ISSN 2078-9181 DOI 10.15622/sp.18.2

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Отделение нанотехнологий и информационных технологий

САНКТ-ПЕТЕРБУРГСКИЙ
ИНСТИТУТ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ РАН

ТРУДЫ СПИИРАН

proceedings.spiiras.nw.ru

ТОМ 18 **№** 2

Санкт-Петербург
2019

18+

SPIIRAS PROCEEDINGS

Volume 18 № 6, 2019

Scientific, educational, and interdisciplinary journal primarily specialized
in computer science, automation, and applied mathematics

Trudy SPIIRAN ♦ Founded in 2002 ♦ Труды СПИИРАН

Founder and Publisher

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences

Editor: A. I. Motienko
Editor: E. P. Miroshnikova
Technical editor: M. S. Avstriyskaya
Translator: N. V. Kashina

Editorial Board's address

14-th line VO, 39, SPIIRAS, St. Petersburg, 199178, Russia,
e-mail: publ@iias.spb.su, web: http://www.proceedings.spiiras.nw.ru/

The journal is indexed in Scopus

ТРУДЫ СПИИРАН

Том 18 № 6, 2019

Научный, научно-образовательный, междисциплинарный журнал с базовой специализацией в области информатики, автоматизации и прикладной математики
Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН)

Адрес редакции

199178, Санкт-Петербург, 14-я линия, д. 39,
e-mail: publ@iias.spb.su, сайт: http://www.proceedings.spiiras.nw.ru/
Журнал индексируется в международной базе данных Scopus
Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук»

CONTENTS

СОДЕРЖАНИЕ

УДК 621.3 **DOI** 10.15622/sp.2019.18.6.1269-1300

Ю.К. Язов, О.С. Авсентьев, А.О. Авсентьев, И.О. Рубцова

МЕТОД ОЦЕНИВАНИЯ ЭФФЕКТИВНОСТИ ЗАЩИТЫ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА С ПРИМЕНЕНИЕМ АППАРАТА СЕТЕЙ ПЕТРИ — МАРКОВА

Язов Ю.К., Авсентьев О.С., Авсентьев А.О., Рубцова И.О. **Метод оценивания эффективности защиты электронного документооборота с применением аппарата сетей Петри — Маркова.**

Аннотация. Традиционные подходы к оцениванию эффективности защиты информации, основанные на сравнении возможностей реализации угроз безопасности информации в условиях отсутствия и применения мер защиты, не позволяют анализировать динамику пресечения мерами защиты процессов реализации угроз. Предложен новый показатель эффективности защиты электронных документов, позволяющий оценивать возможности опережения мерами защиты процесса реализации угроз в системах электронного документооборота и учитывающий вероятностно-временные характеристики динамики применения мер защиты и реализации угроз электронным документам. Разработаны с использованием аппарата сетей Петри — Маркова математические модели и получены аналитические соотношения для расчета предложенного показателя на примере реализации угрозы «туннелирования трафика» (размещение пакетов нарушителя в пакетах доверенного пользователя) и несанкционированного доступа (сетевых атак) к электронным документам, а также угрозы внедрения вредоносной программы путем проведения атаки «неслепой IP-спуфинг» (подмены сетевого адреса). Приведены примеры расчета предложенного показателя и графики его зависимости от вероятности обнаружения сетевых атак системой обнаружения вторжений и от вероятности обнаружения вредоносных программ системой антивирусной защиты. Получены количественные зависимости эффективности защиты электронных документов за счет опережения мерами защиты процессов реализации угроз как от вероятности обнаружения вторжения или вероятности обнаружения вредоносной программы, так и от соотношения времени, затрачиваемого системой защиты на обнаружение попытки реализации угрозы и принятие мер по пресечению процесса ее реализации, и времени реализации угрозы. Модели позволяют не только оценивать эффективность мер защиты электронных документов от угроз уничтожения, копирования, несанкционированных изменений и тому подобное, но и количественно обосновывать требования к времени реакции адаптивных систем защиты на обнаруживаемые действия, направленные на нарушение безопасности электронных документов, а также выявлять слабые места в системах защиты, связанные с динамикой реализации угроз и реакцией на такие угрозы систем защиты электронного документооборота.

Ключевые слова: показатель эффективности, структурно-функциональная модель, сеть Петри — Маркова, система электронного документооборота, угроза безопасности, мера защиты, вероятность обнаружения, система обнаружения вторжений, система антивирусной защиты.

1. Введение. Необходимость оценивания эффективности защиты информации (ЗИ) в информационных системах (ИС) отмечается в целом ряде документов федерального уровня. Так, в Доктрине информационной безопасности Российской Федерации, утвержденной Указом Президента Российской Федерации от 5 декабря 2016 г. №646,

указано, что «планирование, осуществление и оценка эффективности комплекса мер по обеспечению информационной безопасности» являются «задачами государственных органов…». Вместе с тем в настоящее время методики оценивания эффективности ЗИ (или обеспечения безопасности информации) в практическом плане остаются неразвитыми. Это относится и к ИС с системами электронного документооборота (СЭД). Сегодня действия, направленные на копирование и несанкционированное распространение, подделку (модификацию), уничтожение электронных документов (ЭД) в ИС с СЭД, внедрение вредоносных программ (ВП) с целью выполнения таких действий и так далее, становятся важными составляющими информационного противоборства и обусловливают принятие эффективных мер защиты.

В теоретическом плане во многих научных публикациях предлагались весьма разноплановые подходы к оцениванию эффективности ЗИ.

Так, в [1] для оценивания эффективности защиты от отдельных угроз предлагалось в качестве показателя использовать вероятность того, что в условиях применения мер защиты угроза безопасности информации не будет реализована; в [2, 3] были разработаны аналитические модели и методики расчета этой вероятности для различных ИС и угроз. В основе указанных моделей лежит теория риска, понимаемого как произведение вероятности нанесения ущерба определенного уровня (или среднего ущерба) на вероятность реализации угрозы, которая приводит к этому ущербу [4]. При этом, как правило, ущерб от реализации угрозы полагается неприемлемым, и риск понимается только как риск реализации угрозы. Это обусловлено тем, что аналитические методы оценивания ущерба от реализации угроз безопасности информации пока недостаточно развиты.

Указанные показатели используются в интересах выработки решений по управлению рисками [5], при риск-ориентированном выборе мер контроля защищенности информации в ИС [6], при управлении рисками для системно-связанных объектов [7], для оценивания защищенности информации в ИС, построенных на основе «облачных технологий» [8], и так далее.

Вместе с тем разрабатывались и иные подходы. Так, в [9] оценивание эффективности, понимаемой как «степень соответствия результатов защиты информации цели защиты информации», предложено проводить с применением абсолютного, относительного и относительно-разностного показателей, которые рассчитываются путем сравнения вероятностей реализации угроз без мер защиты и в условиях применения мер защиты. Кроме того, наряду с указанными

показателями, рассчитываемыми аналитическими методами, для оценивания эффективности защиты информации в ИС как в России, так и за рубежом использовались и качественные показатели, например при функциональном подходе, который основан на сравнении состава реализуемых мер защиты с составом, заданным нормативными документами [9]; или при подходе, который основан на введении оценочных уровней доверия в соответствии с идеологией международного стандарта ИСО/ МЭК 15408. Широкое применение на практике также нашел подход, реализующий балльный метод, суть которого состоит в экспертной оценке риска реализации угрозы по баллам, определяемым по заранее введенным шкалам. Этот подход используется сегодня в целом ряде программных продуктов, таких как инструментарий стандарта ISO 17799 (стандарт по информационной безопасности, опубликованный в 2005 г. организациями ISO и IEC, в 2013 г. сменил название на ISO/IEC 27002:2005 «Информационные технологии. Методы обеспечения безопасности. Практические правила управления информационной безопасностью»), например, программный продукт COBRA или программный продукт, реализующий метод CRAMM (Central Computer and Telecommunications Agency. Risk Analysis & Management Method — Центральное агентство по компьютерам и телекоммуникациям Великобритании. Метод анализа и контроля рисков), программный продукт RiskWatch и другие. Вместе с тем применение оцениваемых экспертным методом качественных показателей эффективности, в том числе с использованием балльного метода, приводит не только к неполной, но и зачастую к некорректной оценке, поскольку в них практически не учитывается фактор времени, большое разнообразие способов реализации угроз, применяемых мер защиты и так далее.

С учетом изложенного все большее внимание в последнее время уделяется количественным методам оценивания эффективности ЗИ. При этом в качестве цели ЗИ, как правило, рассматривается парирование всех выявленных актуальных угроз безопасности информации, а в качестве показателей используются или вероятностные показатели, или среднестатистические показатели возможностей парирования угроз. Так как рассматриваемые процессы реализации угроз являются случайными, для их аналитического моделирования используется теория случайных процессов, вероятностно-временные характеристики которых определяются с использованием аппарата марковских или полумарковских процессов [10].

Вместе с тем моделирующие возможности указанных аппаратов весьма ограничены и не позволяют учесть важные факторы, которые

определяют динамику реализации многих угроз безопасности информации в информационных системах. Так, в виде марковских моделей невозможно представить разветвленные процессы, характерные для реализации большинства угроз безопасности информации в ИС, и тем более учесть логические условия, которые вводятся для обеспечения адекватности моделей рассматриваемым процессам. Полумарковские модели позволяют представлять процессы разветвленными и анализировать динамику их выполнения, то есть оценивать количественно время выполнения процесса, однако только в том случае, если в нем отсутствуют логические условия его выполнения.

Логику протекания процесса реализации угроз можно представить сетью Петри [11], однако аналитически учитывать временной фактор в моделях, разработанных на основе этого аппарата, практически невозможно. Сегодня аппарат сетей Петри (Е-сети, временные сети Петри и т.д.) применяют для весьма затратного по времени имитационного моделирования. В [12] в интересах реализации требования стандарта функциональной безопасности IEC 61508 предложено использовать для оценивания достигаемой безопасности защищаемой системы при случайных отказах оборудования методы многофазных сетей Маркова и стохастических сетей Петри с предикатами; а в [13] для моделирования и анализа производительности системы аварийно-спасательной логистики — сочетание сетей Петри с марковскими процессами с выводом линейных уравнений для количественного анализа основных показателей эффективности системы. В [14] объединение стохастических сетей Петри с описанием динамики их срабатывания с применением аппарата марковских процессов позволило связать определяемый сетью Петри порядок выполнения моделируемого процесса со случайным временем выполнения парциальных процедур, составляющих этот процесс. Однако предложенные подходы крайне сложно или невозможно применять, когда одновременно нужно учитывать логические условия выполнения моделируемых процессов, их разветвленность, параллельность и время выполнения, что характерно было и для традиционного аппарата марковских и полумарковских процессов.

Для устранения указанных недостатков марковских и полумарковских моделей в [15] был предложен аппарат сетей Петри — Маркова, позволяющий в отличие от сетей Петри аналитически рассчитывать показатели эффективности защиты информации в ИС с учетом фактора времени, а в отличие от аппарата марковских и полумарковских процессов — наряду с разветвленностью моделируемых процессов и параллельностью выполнения во времени

составляющих эти процессы процедур учитывать влияние логических условий на динамику протекания процессов реализации угроз. При этом при помощи сетей Петри — Маркова могут быть получены аналитические соотношения для расчета времени выполнения процесса и указанных выше вероятностных показателей оценки эффективности, рассчитываемых путем сравнения возможностей реализации угроз в условиях отсутствия и применения выбранных мер защиты.

Однако этими показателями оценивается «итоговый эффект» защиты, и они весьма опосредованно учитывают время реакции систем защиты на попытки реализации угроз. В условиях, когда применяются системы адаптивной защиты, к которым относятся, например, системы обнаружения вторжений (СОВ), системы антивирусной защиты (САВЗ), DLP-системы (Data Loss Prevention, системы блокирования попыток несанкционированной передачи данных во внешние сети [16]) и SIEM-системы (объединяют в своем названии аббревиатуры двух терминов: SIM — Security information management, управление информационной безопасностью, и SEM, управление событиями безопасности [17-19]) и так далее, необходимо анализировать зависимости эффективности защиты от времени реакции этих систем на попытки реализации угроз, указанные показатели оказываются мало приемлемыми.

В данной статье предлагается иной подход к оцениванию эффективности ЗИ в ИС с СЭД, основанный на показателях возможности опережения мерами защиты процесса реализации угроз электронному документообороту.

2. Показатели оценки эффективности защиты электронных документов в СЭД на основе определения возможности опережения мерами защиты процесса реализации угроз электронному документообороту. Возможности опережения мерами защиты процесса реализации угрозы электронному документообороту предлагается оценивать вероятностью того, что суммарное время обнаружения факта реализации угрозы и принятия адекватных действий по ее парированию будет меньше времени проникновения в операционную среду СЭД СН до момента начала выполнения несанкционированного действия. Такое опережение оценивается в том случае, когда меры защиты выбираются и применяются в ходе функционирования СЭД СН в зависимости от содержания действий нарушителя или выполняемых функций иного источника угрозы (например, программной закладки, ВП и т.п.), то есть при применении адаптивных мер защиты информации. Применительно к упомянутым мерам защиты ниже приводятся математические модели расчета указанного показателя оценки эффективности защиты электронного документооборота.

Так как реализация большинства угроз и применения мер защиты возможна при выполнении ряда логических условий, а сами процессы реализации, как правило, являются разветвленными и разделяются на параллельно выполняемые подпроцессы, то для построения математических моделей использовался аппарат сетей Петри — Маркова. Ниже такие модели рассматриваются применительно к системам обнаружения вторжений и антивирусной защиты.

Особенность таких моделей заключается в том, что с их использованием оценивается время $\tau_{res.u} = \tau_{det.u} + \tau_{rep.u}$, необходимое СОВ для обнаружения факта вторжения или САВЗ для обнаружения ВП ($\tau_{det.u}$) и пресечения возможности выполнения несанкционированного действия при попытке реализации u-й угрозы ($\tau_{rep.u}$).

Длительность реализации угрозы составляет случайную величину τ_u, включающую в себя продолжительность τ_0 этапа, который предваряет начало функционирования СОВ (или САВЗ), распаковки пакета (или пакетов).

Обозначим разницу $\tau_u - \tau_0$ как $\tau_u^{(0)}$.

Пусть плотности распределения вероятностей величин $\tau_{res.u}$ и $\tau_u^{(0)}$ равны $w_{res}(\tau_{res.u})$ и $w_u(\tau_u^{(0)})$ соответственно.

Угроза не будет реализована, если $\tau_u^{(0)} > \tau_{res.u}$.

Так как рассматриваемые случайные величины являются независимыми, то в соответствии с [20] плотность распределения вероятностей случайной величины $y = \tau_u^{(0)} - \tau_{res.u}$ определяется из соотношения:

$$w_{exc}(y) = \int_0^\infty w_u(y + \tau_{res.u}) \cdot w_{res}(\tau_{res.u}) \cdot d\tau_{res.u}. \quad (1)$$

Тогда вероятность того, что текущее время реализации угрозы в каждой попытке такой реализации будет превышать время принятия адекватных мер защиты, определяется из соотношения:

$$p_{exc}(\tau_u^{(0)} - \tau_{res.u} > 0) = \int_0^\infty w_{exe}(y) \cdot dy, \quad (2)$$

а среднее время реализации угрозы $\overline{\tau_u^{(ЗИ)}}$ с учетом [21, 22] — величину:

$$\overline{\tau_u^{(ЗИ)}} = \overline{\tau_0} + \frac{\overline{\tau_u^0}}{1 - p_{exc}(\tau_u^{(0)} - \tau_{res.u})}, \quad (3)$$

где $\overline{\tau_0}$ и $\overline{\tau_u^{(0)}}$ — математические ожидания величин τ_0 и $\tau_u^{(0)}$ соответственно.

В соответствии с формулой (3) происходит вероятностное прореживание исходного потока событий, описывающего процесс реализации u -й угрозы.

В [21, 22] показано, что для получаемого путем прореживания с вероятностью p потока характеристическая функция $\chi_*(s)$ для интервала времени между соседними событиями в прореженном потоке имеет вид:

$$\chi_*(s) = \frac{p \cdot \chi(s)}{1 - (1-p) \cdot \chi(s)}, \quad (4)$$

где $\chi(s)$ — характеристическая функция для интервала времени между соседними событиями в исходном потоке.

В качестве меры близости потока к пуассоновскому можно использовать коэффициент вариации, определяемый как отношение среднеквадратического отклонения паузы θ между событиями в потоке к его математическому ожиданию:

$$K_{\text{var}} = \frac{\sigma_\theta}{M_\theta} = \frac{\sqrt{\chi^{"}(s) - (\chi^{'}(s))^2}}{\chi^{'}(s)}\Big|_{s=0}, \quad (5)$$

где $\chi^{'}(s)$ и $\chi^{"}(s)$ — первая и вторая производные от характеристической функции $\chi(s)$.

При этом крайними случаями являются детерминированный периодический поток, для которого $K_{\text{var}} = 0$, и пуассоновский поток, для которого $K_{\text{var}} = 1$.

Коэффициент вариации для прореженного с вероятностью p потока K_{var}^* рассчитывается по формуле, аналогичной формуле (5), где вместо характеристической функции исходного потока $\chi(s)$ используется характеристическая функция прореженного потока $\chi_*(s)$, которая определяется по формуле (4).

С учетом изложенного коэффициент вариации для прореженного потока K^*_{var} связывается с коэффициентом вариации исходного потока K_{var} соотношением:

$$K^*_{\text{var}} = \frac{\sqrt{\chi^{''}_*(s) - (\chi^{'}_*(s))^2}}{\chi^{'}_*(s)}\Big|_{s=0} = \sqrt{K_{\text{var}} \cdot p + 1 - p}, \qquad (6)$$

представленном в виде графической зависимости на рисунке 1. Из рисунка 1 видно, что с уменьшением вероятности прореживания результирующий поток быстро приближается к пуассоновскому.

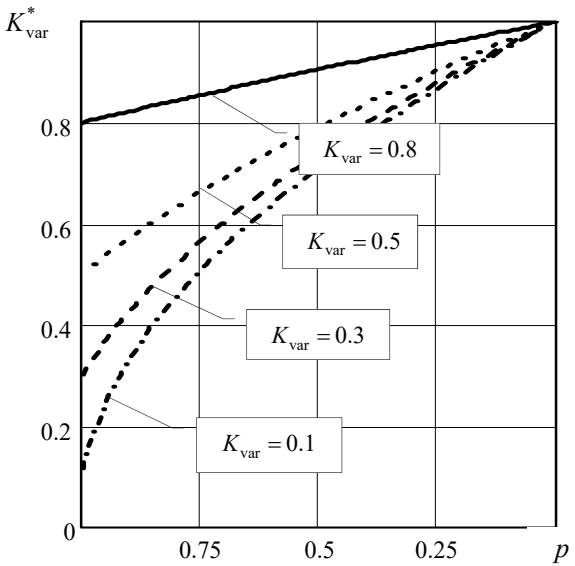


Рис. 1. Зависимость коэффициента вариации прореженного потока от вероятности прореживания и коэффициента вариации исходного потока

При этом даже при вероятности прореживания менее 0.3 для большинства имеющих место на практике потоков ошибка с заменой любой одномодальной плотности распределения на экспоненциальную составляет единицы процентов. Это некритично для результатов оценивания и обусловливает возможность при моделировании процессов, в которых имеет место то или иное вероятностное прореживание исходного потока, использовать только экспоненциальный вид распределения без заметных отклонений в оценках характеристик прореженных потоков.

Представляя плотности распределения $w_{res}(\tau_{res.u})$ и $w_u(\tau_u^{(0)})$ в виде экспонент, нетрудно получить из формул (1) и (2) зависимость:

$$p_{exc}(\tau_u^{(0)} - \tau_{res.u} > 0) = \frac{\overline{\tau_u^{(0)}}}{\overline{\tau_{res.u}} + \overline{\tau_u^{(0)}}}, \tag{7}$$

где $\overline{\tau_{res.u}}$ — среднее время реакции системы защиты на попытку реализации u-й угрозы,

$$\overline{\tau_{res.u}} = \frac{\overline{\tau_{det.u}}}{P_{det.u}} + \frac{\overline{\tau_{rep.u}}}{P_{rep.u}}, \overline{\tau_{det.u}} > 0, \tag{8}$$

P_{det} и P_{rep} — вероятности обнаружения и пресечения процесса реализации u-й угрозы.

С учетом соотношения (8) формула (7) преобразуется к виду:

$$p_{exc.u}(\tau_u^{(0)} - \tau_{res.u} > 0) = \frac{1}{\frac{\overline{\tau_{det}}}{\overline{\tau_u^{(0)}} \cdot P_{det.u}} + \frac{\overline{\tau_{rep.u}}}{\overline{\tau_u^{(0)}} \cdot P_{rep.u}} + 1}. \tag{9}$$

Как правило, если процесс реализации угрозы обнаружен, то пресечение происходит с вероятностью, близкой к единице, то есть $P_{rep.u} \approx 1$. Тогда:

$$\overline{\tau_{res.u}} = \frac{\overline{\tau_{det.u}}}{P_{det.u}} + \overline{\tau_{rep.u}}; \tag{10}$$

$$p_{exc.u}(\tau_u^{(0)} - \tau_{res.u} > 0) = \frac{1}{\frac{\overline{\tau_{det}}}{\overline{\tau_u^{(0)}} \cdot P_{det.u}} + \frac{\overline{\tau_{rep.u}}}{\overline{\tau_u^{(0)}}} + 1}. \tag{11}$$

Вероятность опережения мерой защиты процесса реализации угрозы может быть использована в качестве частного показателя эффективности защиты ЭД. На рисунке 2 в графическом виде

приведена зависимость данного показателя от вероятности обнаружения при различных отношениях $\frac{\overline{\tau_{det.u}}}{\overline{\tau_u^{(0)}}}$ и $\frac{\overline{\tau_{rep.u}}}{\overline{\tau_u^{(0)}}}$.

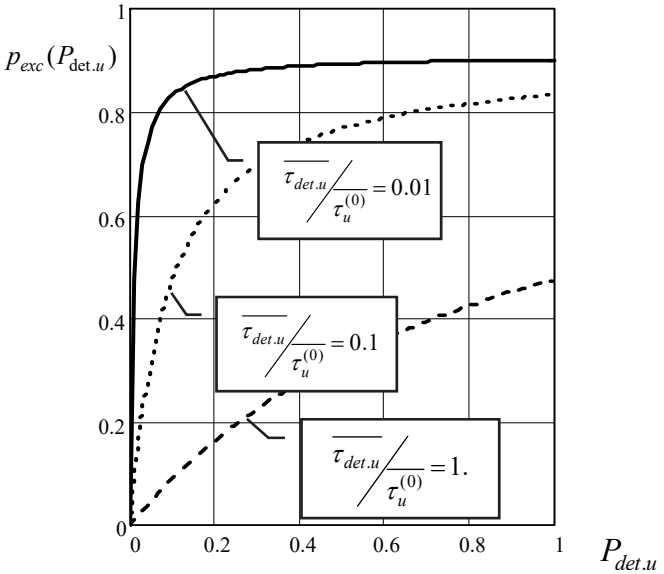


Рис. 2. Зависимость вероятности опережения процесса реализации угрозы мерой защиты

Приведенный показатель позволяет учитывать влияние времени реакции системы защиты на процесс реализации угрозы. Вместе с тем даже при очень высокой эффективности защиты, оцененной по данному показателю, в силу статистического характера процесса реализации угрозы всегда существует вероятность того, что угроза за некоторое конечное время t будет реализована. В связи с этим в качестве основного показателя эффективности защиты ЭД целесообразно использовать вероятность того, что угроза за заданное время не будет реализована. Эта вероятность рассчитывается по формуле:

$$\eta_{exc.u}(t) = \exp\left\{-\frac{t \cdot p_{exc.u}(\overline{\tau_u^{(0)}} - \overline{\tau_{res.u}} > 0)}{\overline{\tau_u^{(0)}}}\right\} = \exp\left(-\frac{t}{\overline{\tau_0} + \overline{\tau_u^{(0)}} + \overline{\tau_{res.u}}}\right). \quad (12)$$

Если необходимо оценить эффективность защиты электронного документооборота в СЭД СН на основе определения возможности

опережения мерами защиты процесса реализации совокупности угроз, то показатель эффективности рассчитывается следующим образом.

Пусть имеется множество U угроз, которые могут быть реализованы в СЭД СН за время t независимо друг от друга относительно как отдельных документов, так и СЭД СН в целом. Для парирования каждой u-й угрозы применяется соответствующая мера защиты, при этом мера защиты должна опередить процесс реализации угрозы.

Тогда показатель эффективности совокупности применяемых мер защиты рассчитывается по формуле:

$$\eta_{exc}^{(U)}(t)=\prod_{u=1}^{U}\exp\left(-\frac{t}{\overline{\tau_u}\cdot\left(1+\frac{\overline{\tau_{res.u}}}{\tau_u}\right)}\right)=\exp\left\{-t\cdot\sum_{u=1}^{U}\left(\frac{1}{\overline{\tau_u}\cdot\left(1+\frac{\overline{\tau_{res.u}}}{\tau_u}\right)}\right)\right\}. \quad (13)$$

Для расчета предложенных показателей необходимо моделирование процесса реализации каждой угрозы с использованием аппарата сетей Петри — Маркова.

3. Структурно-функциональная модель и сети Петри — Маркова, моделирующие процессы реализации угроз электронному документообороту в условиях применения мер защиты. В интересах формирования сетей Петри — Маркова разрабатывались структурно-функциональные модели процессов реализации угроз [23]. Структурно-функциональная модель отражает содержание, взаимосвязь и последовательность выполнения процедур и функций в процессе реализации угрозы в течение всего цикла обработки ЭД [24] и реакцию системы защиты на попытку реализации угрозы.

На рисунке 3 для примера приведена структурно-функциональная модель реализации угрозы сетевой атаки туннелирования трафика в условиях применения СОВ, а на рисунке 4 — угрозы внедрении ВП путем проведения атаки «неслепой IP-спуфинг» в условиях применения САВЗ.

Обозначение и содержание функций, выполняемых в ходе реализации указанных угроз приведены в таблице 1.

На основе этих моделей формировались соответствующие сети Петри — Маркова, моделирующие во времени процессы реализации угроз. Графы сетей Петри — Маркова для указанных выше угроз приведены на рисунках 5 и 6 соответственно, а обозначение и описание позиций и переходов сетей Петри — Маркова для обеих атак приведены в таблице 2.

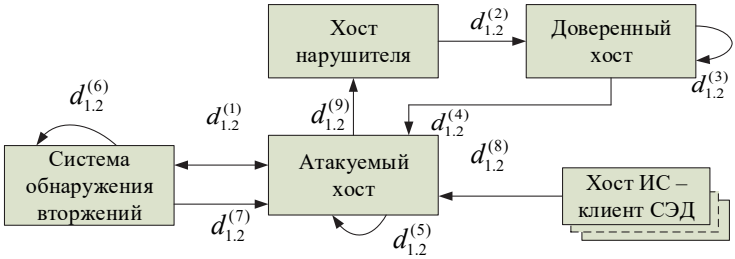


Рис. 3. Структурно-функциональная модель процесса реализации угрозы сетевой атаки с внедрением ВП по сети, реализуемой путем «туннелирования» трафика с использованием протоколов IP или ICMP

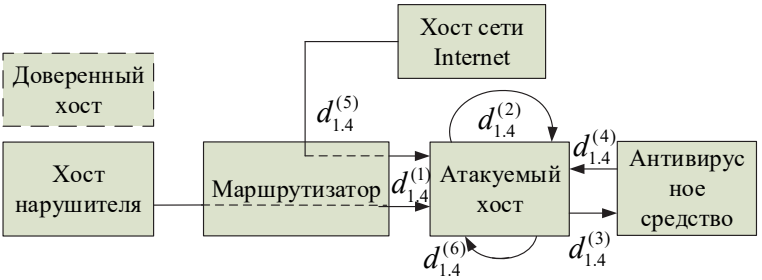


Рис. 4. Структурно-функциональная модель процесса реализации угрозы сетевой атаки с внедрением ВП от имени доверенного хоста путем подмены сетевого адреса в интересах перехвата трафика («неслепой IP-спуффинг») в условиях применения САВЗ

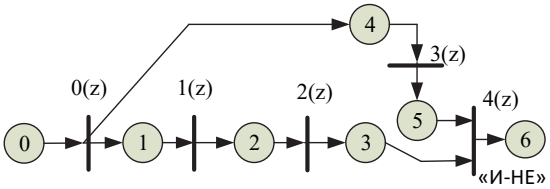


Рис. 5. Сеть Петри – Маркова, моделирующая процесс реализации угрозы сетевой атаки с внедрением ВП путем «туннелирования» трафика с использованием протоколов IP или ICMP в условиях применения СОВ

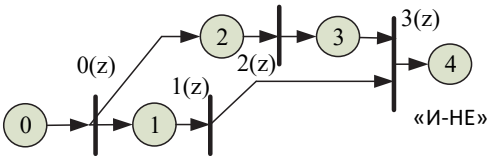


Рис. 6. Сеть Петри – Маркова, моделирующая процесс реализации угрозы внедрения ВП от имени доверенного хоста с подменой сетевого адреса в интересах перехвата трафика («неслепой IP-спуффинг») в условиях применения САВЗ

Таблица 1. Обозначения и содержание функций, выполняемых при реализации угроз туннелирования трафика и проведения атаки «неслепой IP-спуфинг»

Наименование атаки	Содержание функций, выполняемых при реализации угрозы сетевой атаки	Обозначение функции
Сетевая атака туннелирования трафика в условиях применения СОВ	СОВ анализирует входящий трафик на наличие подозрительных команд, ВП и т.п. (по сигнатурам или по аномалиям)	$d_1^{(1)}$
	Инкапсуляция пакетов хоста нарушителя в пакеты, передаваемые на хост, который является доверенным для атакуемого хоста, передача пакетов на доверенный хост	$d_1^{(2)}$
	Распаковка пакета сообщения с хоста нарушителя, обнаружение в поле данных инкапсулированного пакета для атакуемого хоста, в полях адресов которого указан сетевой адрес получателя – атакуемого хоста и адрес отправителя – доверенного хоста, а в поле данных – скрипт, специальная программа-шпион (типа «сниффер», программа наблюдения за операционной средой атакуемого хоста типа Real SPY Monitor и др.)	$d_1^{(3)}$
	Отправка выявленного пакета по адресу атакуемого хоста	$d_1^{(4)}$
	Распаковка полученного пакета и запуск ВП	$d_1^{(5)}$
	СОВ обнаруживает вторжение (вредоносную программу), оповещает пользователя и подготавливает команду на блокирование действий по выполнению программы	$d_1^{(6)}$
	СОВ направляет команду на блокирование действий по выполнению ВП	$d_1^{(7)}$
	Один из хостов ИС-клиентов СЭД направляет ЭД для атакованного хост	$d_1^{(8)}$
	Перехваченный ЭД направляется на хост нарушителя	$d_1^{(9)}$

Продолжение таблицы 1.

Сетевая атаки «неслепой IP-спуфинг» с внедрением вредоносной программы	С хоста нарушителя по протоколу UDP (без установления виртуального канала) от имени доверенного хоста посылается сообщение с ВП через маршрутизатор на атакуемый хост	$d_2^{(1)}$
	Атакуемый хост принимает пакет от «доверенного хоста», распаковывает его. При этом вложенная в него ВП, например, предназначена для уничтожения файлов с ЭД с расширением *.doc и *.docx. ВП запускается операционной системой и готова к поиску документов	$d_2^{(2)}$
	Антивирусное средство просматривает файловую систему хоста на предмет обнаружения ВП, по ее сигнатуре или по аномалиям, режим работы сетевой карты на предмет выявления ВП или следов ее функционирования	$d_2^{(3)}$
	Антивирусное средство обнаруживает и блокирует ВП	$d_2^{(4)}$
	ЭД высылается на атакуемый хост доверенным хостом сети Internet	$d_2^{(5)}$
	ВП на атакуемом хосте обнаруживает появившийся на хосте ЭД и уничтожает его.	$d_2^{(6)}$

При этом учитывалось следующее.

Каждая сеть Петри — Маркова имеет в своем составе набор позиций (на графе обозначены кружочками с номерами 0(a), 1(a) и т.д.) и набор переходов (на графе обозначены жирными линиями с номерами 0(z), 1(z) и т.д.), включающий в себя простые переходы и логические переходы. Позиции и переходы соединены дугами со стрелками, указывающими направления перемещения по графу меток, которые обозначают текущее состояние моделируемого процесса. Простые переходы срабатывают при поступлении в них метки от предстоящих состояний, а логические переходы — при поступлении метки и при выполнении заданных логических условий. Для расчета времени срабатывания сети, то есть перемещения моделируемого процесса из начального в конечное состояние, вся сеть разбивается на участки между начальным состоянием до первого логического перехода, между логическими переходами и от последнего логического перехода до последнего логического состояния (рисунок 7).

В общем случае для такого участка в соответствии с [10] необходимо составить систему уравнений:

$$\Phi_{ij}(d,t) = \pi_{ik}(d) \cdot \int_0^t f_{ik}(d,\tau) \cdot \Phi_{kj}(d,t-\tau) \cdot d\tau. \qquad (14)$$

Таблица 2. Обозначения и описание позиций и переходов сетей Петри-Маркова, моделирующих процессы реализации угроз туннелирования трафика и проведения атаки «неслепой IP-спуфинг»

Наименование атаки	Обозначение элемента сети Петри-Маркова	Описание элемента сети Петри – Маркова
Сетевая атака туннелирования трафика в условиях применения СОВ	Позиции сети	
	0(a)	Начальное состояние процесса, нарушитель готов к проведению атаки, сформировал пакет для передачи на доверенный хост, в который инкапсулировал пакет с программой-шпионом
	1(a)	Сообщение с инкапсулированным пакетом поступило на доверенный хост
	2(a)	Инкапсулированный пакет выделен из сообщения и установлен адрес его пересылки на атакуемый хост; 3(a) — пакет с вредоносной программой от доверенного хоста поступил на атакуемый хост и распакован
	4(a)	На атакуемом хосте включена СОВ
	5(a)	С вероятностью $p_{обн}$ обнаружено вторжение на атакуемом хосте и отправлена команда на пресечение вторжения; 6(a) – угроза не реализована
	Переходы сети	
	0(z)	Передача сообщения с инкапсулированным пакетом на доверенный хост
	1(z)	Распаковка сообщения с инкапсулированным пакетом на доверенном хосте, передача с одного из хостов – клиентов СЭД на атакуемый хост
	2(z)	Инкапсулированный пакет передан на атакуемый хост
	3(z)	Осуществляется обнаружение вторжения по сигнатуре или по аномалиям
	4(z)	Логический переход «И-НЕ», срабатывающий, если к данному моменту времени поступила команда на пресечение вторжений от СОВ, а команда на запуск ВП не поступила, то есть угроза реализована

Продолжение таблицы 2

Сетевая атаки «неслепой IP-спуфинг» с внедрением вредоносной программы	Позиции сети	
	0(a)	Начальное состояние процесса: на атакуемом хосте находится ЭД, нарушитель готов к проведению атаки, сформировал UDP-пакет с адресом доверенного хоста и с программой, предназначенной для уничтожения ЭД с расширением *.doc и *.docx
	1(a)	Пакет с ВП получен на атакуемом хосте и распакован; 2(a) — на атакуемом хосте функционирует система антивирусной защиты
	3(a)	С вероятностью $p_{обн}$ обнаружена ВП и отправлена команда на ее блокирование
	4(a)	Угроза реализована
	Переходы сети	
	0(z)	Передача сообщения с ВП на доверенный хост, запуск на нем антивирусного средства
	1(z)	Запуск ВП
	2(z)	Осуществляется поиск вредоносной программы
	3(z)	Логический переход «И-НЕ», срабатывающий, если не поступила команда на блокирование ВП, а ВП запущена на выполнение, угроза реализуется

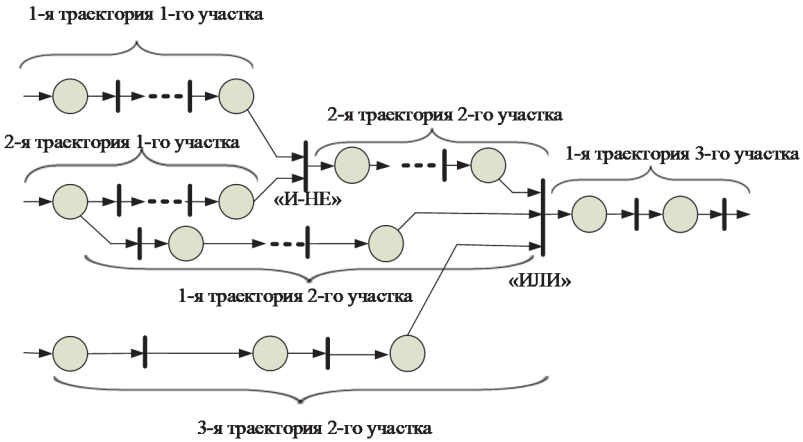


Рис. 7. Пример разбиения сети Петри — Маркова на участки и траектории

$\Phi_{ij}(d,t)$ — вероятность перемещения процесса по траектории d за время t из позиции с номером i в переход с номером j; $f_{ik}(d,\tau)$ — плотность распределения вероятностей времени перемещения процесса по траектории d из i-й позиции в k-й переход; $\pi_{i,k}(d)$ — вероятность того, что процесс пойдет по дуге, соединяющей i-ю позицию с k-м переходом, находящимся на траектории d (при отсутствии разветвлений для инцидентных позиции и перехода эта вероятность равна 1; если позиция не соединяется с переходом по рассматриваемой траектории, то вероятность равна 0).

Если по данной траектории метка из позиции с номером i может достичь перехода с номером k и позиция и переход не являются инцидентными, то имеют место равенства:

$$\pi_{i,j}(d) = \prod_{h_d=1}^{H_d} \pi_{i+h_d,r+h_d}(d); \tag{15}$$

$$\begin{aligned} f_{i,j}(d,t) = f_{i,r}(d,t) * f_{i+1_d,r+1_h}(d,t) * \ldots * \\ * f_{i+h_d,r+h_d}(d,t) * \ldots * f_{i+H_d,r+H_d}(d,t), \end{aligned} \tag{16}$$

где h_d — текущий номер перехода по траектории d при перемещении из позиции с номером i в переход с номером j, $h_d = \overline{1,H_d}$; H_d — общее количество переходов между позицией с номером i и переходом с номером j на траектории d; * — операция свертки [20, 21].

Если на участке имеет место несколько траекторий перемещения, сходящихся на логическом переходе, то время перемещения рассчитывается для каждой траектории.

При этом среднее время выполнения процесса по d-й траектории определяется следующим образом:

$$\overline{\tau_d} = \chi'_{\Sigma_d}(s)\big|_{s=0}, d = \overline{1,D}, \tag{17}$$

где $\chi'_{\Sigma_d}(s)$ — производная от характеристической функции $\chi_{\Sigma_d}(s)$ суммы времен выполнения функций, составляющих процедуры, реализуемые по d-й траектории сети [9, 20]:

$$\chi_{\Sigma_d}(s) = \prod_{r=1}^{R_d} \chi_r(s), r = \overline{1, R_d}, \quad (18)$$

Независимо от того, какому закону подчиняются распределения времен перемещения по дуге (в соответствии с теорией сетей Петри — Маркова [15] время перемещения процесса из позиции в переход считается случайным конечным, а из перехода в позицию — мгновенным), D — общее количество выделенных траекторий; R_d — общее количество ненулевых по времени перемещений на d-й траектории.

Время срабатывания логических переходов существенно зависит не только от количества входящих дуг, но и от того, сколько входящих дуг соответствует логическому условию «И», «ИЛИ», «НЕ» или их сочетаниям (рисунок 8).

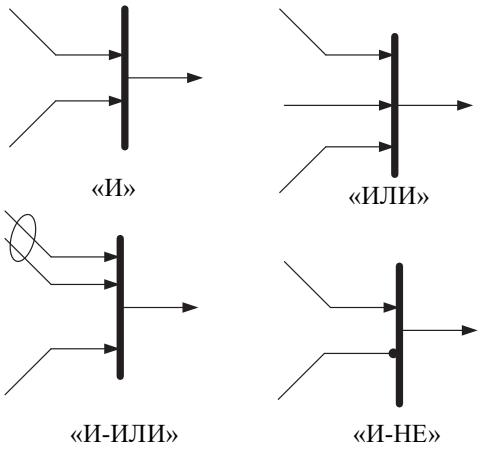


Рис. 8. Некоторые логические переходы, встречающиеся при моделировании процессов реализации угроз в ИС

Соотношения для расчета времен срабатывания логических переходов наиболее часто встречающихся при моделировании процессов реализации угроз безопасности информации в ИС, приведены в соответствии с [25] в таблице 3.

С учетом приведенных соотношений рассчитывались средние времена срабатывания сетей Петри — Маркова, моделирующие угрозу сетевой атаки путем «туннелирования» трафика (рисунок 5) и угрозу «неслепой IP-спуффинг» (рисунок 6).

Таблица 3. Соотношения для расчета времени срабатывания некоторых логических переходов

Логическое условие	Формула для расчета
«И»	Для двух входящих дуг: $\overline{\tau}_{"И"} = \frac{\overline{\tau_1}^2 + \overline{\tau_1}\cdot\overline{\tau_2} + \overline{\tau_2}^2}{\overline{\tau_1} + \overline{\tau_2}}$; для трех входящих дуг: $\overline{\tau}_{"И"} = \overline{\tau_1} + \overline{\tau_2} + \overline{\tau_3} + \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_2}} + \frac{1}{\overline{\tau_3}}} + \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_2}}} + \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_3}}} + \frac{1}{\frac{1}{\overline{\tau_2}} + \frac{1}{\overline{\tau_3}}}$
«ИЛИ»	Для двух входящих дуг: $\overline{\tau}_{"ИЛИ"} = \frac{\overline{\tau_1}\cdot\overline{\tau_2}}{\overline{\tau_1} + \overline{\tau_2}}$; Для трех входящих дуг: $\overline{\tau}_{"ИЛИ"} = \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_2}} + \frac{1}{\overline{\tau_3}}}$
«И-НЕ»	Для двух входящих дуг (первая соответствует условию «И», вторая – условию «НЕ»): $\overline{\tau}_{"И-НЕ"} = \overline{\tau_1}\cdot\left(1 + \frac{\overline{\tau_1}}{\overline{\tau_2}}\right)$; Для трех входящих дуг (первая и вторая соответствует условию «И», третья – условию «НЕ»): $\overline{\tau}_{"И-НЕ"} = \frac{\overline{\tau_1}^2 + \overline{\tau_1}\cdot\overline{\tau_2} + \overline{\tau_2}^2}{\overline{\tau_1} + \overline{\tau_2}}\cdot\left(1 + \frac{\overline{\tau_1}^2 + \overline{\tau_1}\cdot\overline{\tau_2} + \overline{\tau_2}^2}{(\overline{\tau_1} + \overline{\tau_2})\cdot\overline{\tau_3}}\right)$
«И-ИЛИ»	Для трех входящих дуг (первая и вторая дуги или третья дуга): $\overline{\tau_{"И-ИЛИ"}} = \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_2}} + \frac{1}{\overline{\tau_3}}} + \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_2}}} + \frac{1}{\frac{1}{\overline{\tau_1}} + \frac{1}{\overline{\tau_3}}}$

Так, для первой из указанных угроз среднее время ее реализации в условиях применения СОВ определяется из соотношений:

$$\overline{\tau_u^{(ЗИ)}} = \overline{\tau_{0,4}} = \overline{\tau_{0,0}} + \overline{\tau_{"И-НЕ"}};\ \overline{\tau_{"И-НЕ"}} = \overline{\tau_{1,4}}\cdot\left(1 + \frac{\overline{\tau_{1,4}}}{\overline{\tau_{4,4}}}\right), \qquad (19)$$

где $\overline{\tau_{1,4}} = \overline{\tau_{1,1}} + \overline{\tau_{2,2}} + \overline{\tau_{3,4}}$ и $\overline{\tau_{4,4}} = \overline{\tau_{4,3}} + \overline{\tau_{5,4}} = \frac{\overline{\tau_{\det}}}{p_{\det}} + \overline{\tau_{rep}}$.

Здесь $\overline{\tau_{0,0}} \equiv \overline{\tau_0}$ (см. формулу (3)), $\overline{\tau_u^{(0)}} \equiv \overline{\tau_{1,4}}$ и $p_{exc}(\tau_u^{(0)} - \tau_{res.u} > 0) = \dfrac{\overline{\tau_{4,4}}}{\overline{\tau_{4,4}} + \overline{\tau_{1,4}}}$.

Если положить все средние времена переходов, в том числе среднее время пресечения, кроме среднего времени обнаружения вторжения, примерно равными $\bar{\tau}$, то время реализации угрозы определяется из соотношения:

$$\overline{\tau_u^{(ЗИ)}} = \bar{\tau} \cdot \left[4 + \frac{9}{1 + \dfrac{\overline{\tau_{det}}}{P_{det} \cdot \bar{\tau}}} \right]. \qquad (20)$$

Для второй из указанных угроз среднее время ее реализации в условиях применения САВЗ определяется из соотношения:

$$\overline{\tau_u^{(ТЗИ)}} = \overline{\tau_{0,3}} = \overline{\tau_{0,0}} + \overline{\tau_{"И\text{-}НЕ"}}, \qquad (21)$$

где $\overline{\tau_{"И-НЕ"}} = \overline{\tau_{1,3}} \cdot (1 + \dfrac{\overline{\tau_{1,3}}}{\overline{\tau_{2,3}}})$ и $\overline{\tau_{2,3}} = \overline{\tau_{2,2}} + \overline{\tau_{3,3}} = \overline{\tau_{rep}} + \dfrac{\overline{\tau_{det}}}{p_{det}}$.

Здесь $\overline{\tau_0} \equiv \overline{\tau_{0,0}}$ (см. формулу (3)), $\overline{\tau_u^{(0)}} \equiv \overline{\tau_{1,4}}$ и $p_{exc}(\tau_u^{(0)} - \tau_{res.u} > 0) = \dfrac{\overline{\tau_{2,3}}}{\overline{\tau_{2,3}} + \overline{\tau_{1,3}}}$.

Если положить, что все средние времена переходов, в том числе среднее время пресечения, кроме среднего времени обнаружения вредоносной программы, примерно равны $\bar{\tau}$, то среднее время реализации угрозы определяется из соотношения:

$$\overline{\tau_u^{(ЗИ)}} = \bar{\tau} \left\{ 2 + \frac{1}{1 + \dfrac{\overline{\tau_{det}}}{p_{det} \cdot \bar{\tau}}} \right\}. \qquad (22)$$

С учетом приведенных соотношений по формуле (12) рассчитывались зависимости показателя эффективности СОВ от вероятности обнаружения вторжения для угрозы сетевой атаки с внедрением ВП путем «туннелирования» трафика, которые приведены в графическом виде на рисунках 9 и 10. На рисунках 11 и 12 представлены путем проведения атаки «неслепой IP-спуффинг» от вероятности обнаружения вредоносной программы.

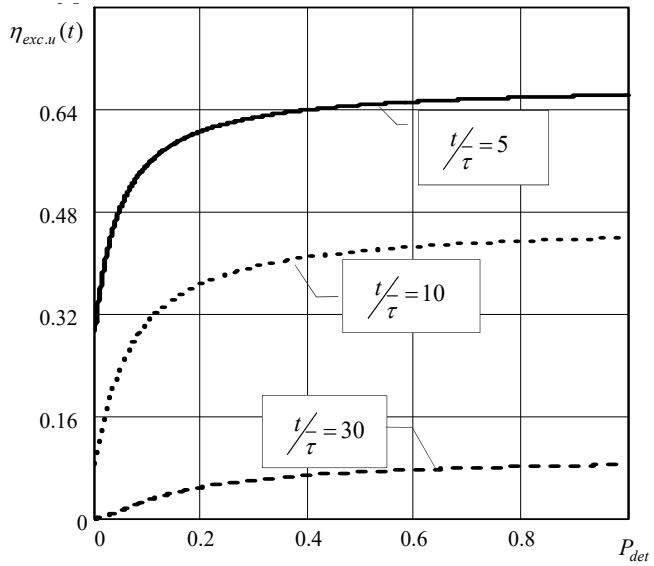


Рис. 9. Зависимость показателя эффективности защиты информации при использовании СОВ от вероятности обнаружения вторжения при $\overline{\tau_{det}}/\tau = 0.1$

Анализ полученных зависимостей показывает, что эффективность защиты ЭД за счет опережения мерами защиты процессов реализации угроз существенно зависит не только от вероятности обнаружения вторжения или вероятности обнаружения вредоносной программы, но и от времени реакции СОВ и САВЗ.

Это обусловливает необходимость формирования требований к СОВ и САВЗ в части ограничения времени, затрачиваемого на обнаружение и пресечение нарушений в СЭД СН, связанных с сетевыми атаками и применением вредоносных программ. Сегодня такие требования в нормативных документах отсутствуют.

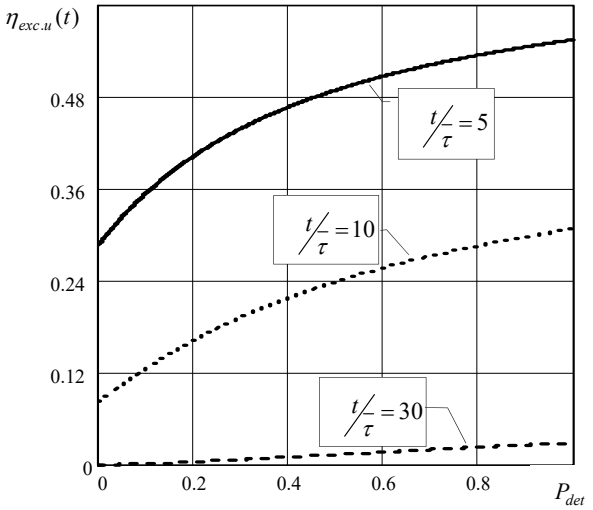


Рис. 10. Зависимость показателя эффективности защиты информации при использовании СОВ от вероятности обнаружения вторжения при $\overline{\tau_{det}}/\tau = 1$

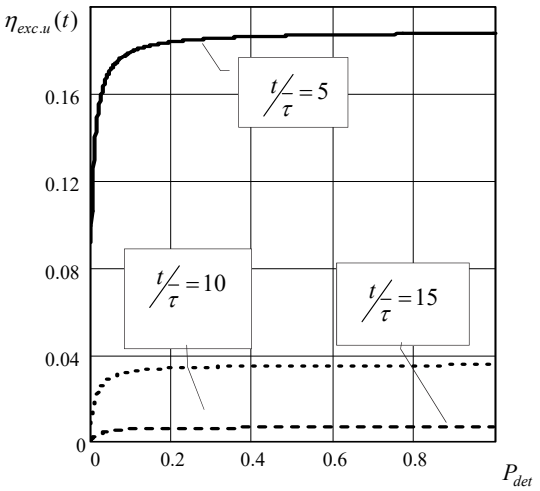


Рис. 11. Зависимость эффективности защиты информации при использовании САВЗ от вероятности обнаружения вредоносной программы при $\overline{\tau_{det}}/\tau = 0.01$

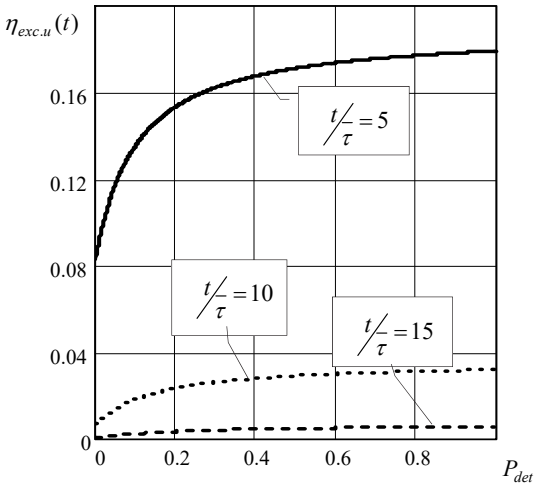


Рис. 12. Зависимость показателя эффективности защиты информации при использовании САВЗ от вероятности обнаружения вредоносной программы при $\overline{\tau_{det}}/\tau = 0.1$

4. Экспериментальные результаты. Для проверки адекватности разработанных моделей и корректности получаемых оценок эффективности защиты ЭД были проведены экспериментальные исследования в форме вычислительного эксперимента с использованием известного методического аппарата функционального моделирования IDEF0 (Integrated Computer Aided Manufacturing DEFinition) в соответствии с Рекомендациями по стандартизации Р 50.1.028-2001 «Информационные технологии поддержки жизненного цикла продукции. Методология функционального моделирования».

Для формирования перечня процедур и функций, выполняемых в ходе электронного документооборота, были использованы сведения о структуре и функционале широко применяемой в органах власти и государственных организациях СЭД типа «Дело» [26].

На рисунке 13 приведены состав программно-аппаратных компонентов и задачи, решаемые с применением такой СЭД, а также перечень и порядок выполнения функций обработки документов в ней на примере входящего ЭД, которые были положены в основу проведения экспериментальных исследований.

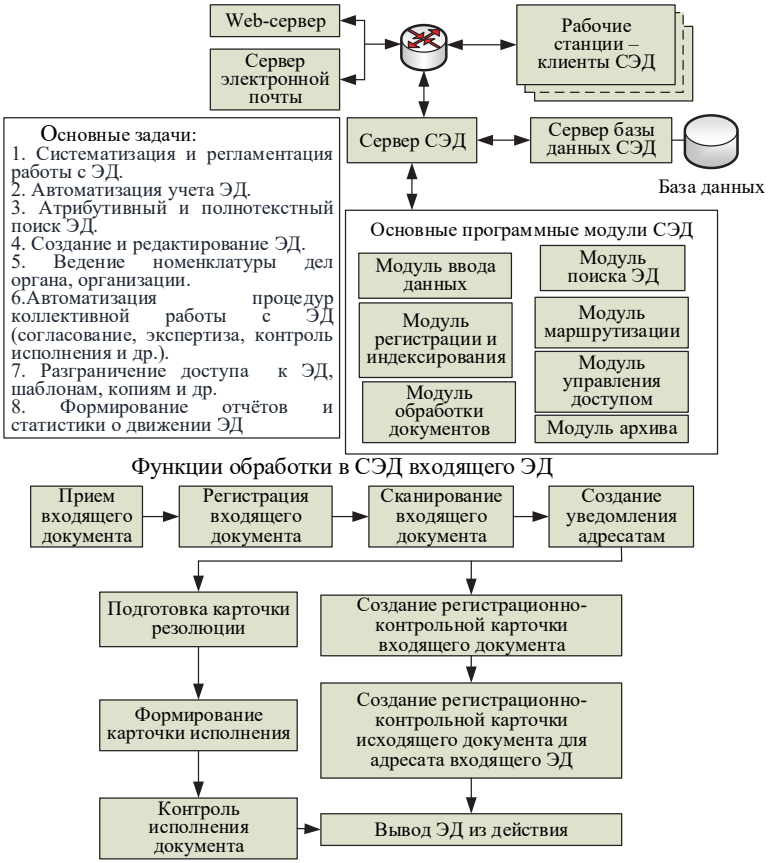


Рис. 13. Состав и задачи СЭД, перечень и порядок выполнения функций обработки документов на примере входящего ЭД, которые учитывались в вычислительном эксперименте

В ходе эксперимента задавалось время реакции системы защиты (СОВ и САВЗ) и сравнивались результаты теоретической оценки времени реализации угроз электронному документообороту с использованием аппарата сетей Петри — Маркова со временем, рассчитываемым по результатам моделирования с использованием IDEF0, а затем по формуле (12) рассчитывался показатель эффективности для модели на сети Петри — Маркова и для модели, разработанной с использованием аппарата IDEF0.

Вычислительный эксперимент показал, что расхождение в результатах оценки эффективности защиты электронного

документооборота по указанным моделям применительно к угрозам туннелирования трафика и «неслепого IP-спуффинга» не превысило 18%.

Это свидетельствует о достаточно высокой адекватности моделей, разрабатываемых с использованием аппарата сетей Петри—Маркова, и корректности получаемых оценок эффективности защиты электронного документооборота.

Важным практическим вопросом применения разработанного подхода является оценивание эффективности адаптивных мер защиты, для чего необходима разработка соответствующей методики и ее реализация в виде программного продукта. Порядок действий по оцениванию эффективности защиты ЭД в СЭД, который может быть положен в основу разработки такой методики, приведен на рисунке 14.

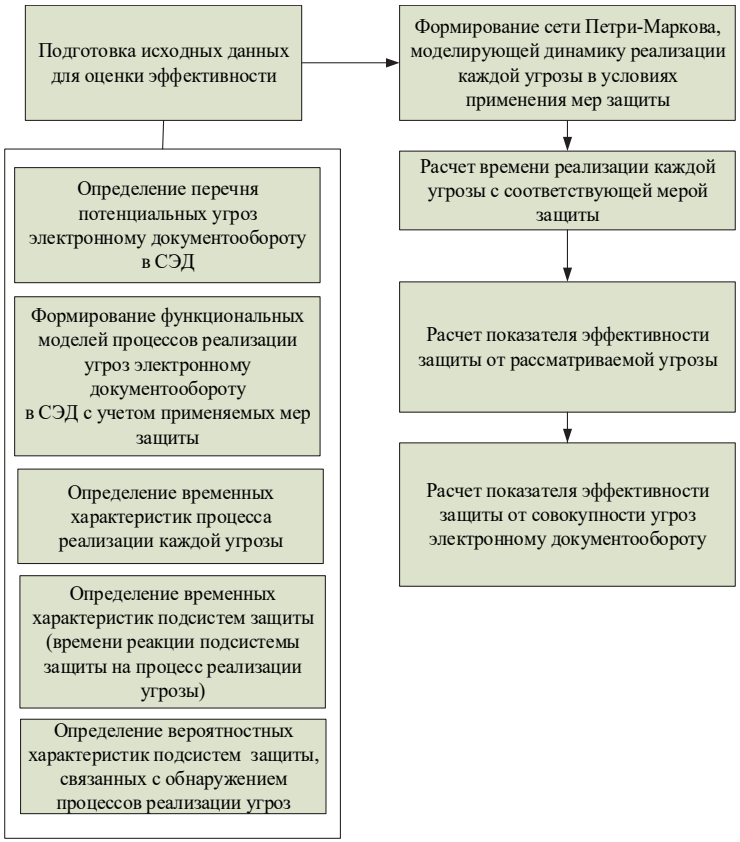


Рис. 14. Порядок действий, составляющий содержание методики оценивания эффективности защиты электронных документов в СЭД на основе определения возможности опережения мерами защиты процесса реализации угроз

4. Заключение. Традиционный подход к оцениванию показателей эффективности защиты электронного документооборота на основе сравнения возможностей реализации угроз безопасности информации без применения и с применением мер защиты в случае, когда применяются адаптивные меры защиты информации, оказывается недостаточным.

Предложенный новый показатель эффективности защиты электронного документооборота, направленный на оценивание возможности опережения мерами защиты процесса реализации угроз документообороту, позволяет учесть время реакции систем защиты на факт обнаружения попытки (процесса) реализации угрозы.

Для расчета указанного показателя в работе предложен подход, во-первых, к формированию структурно-функциональных моделей процессов реализации угроз, во-вторых, к построению на их основе математических моделей оценивания вероятностно-временных характеристик этих процессов с использованием аппарата сетей Петри — Маркова, который позволяет количественно обосновывать требования к временным характеристикам функционирования систем защиты ЭД в СЭД.

Проведенные экспериментальные исследования в форме вычислительного эксперимента путем сравнения результатов расчетов показателей эффективности защиты ЭД по моделям, построенным с применением сетей Петри — Маркова и аппарата функционального моделирования IDEF0, показали их достаточно высокую сходимость с отклонением от 3 до 18 %, что свидетельствует о корректности моделей оценивания эффективности защиты электронного документооборота в СЭД с применением аппарата сетей Петри — Маркова.

Перспективами направлениями дальнейших исследований по данной тематике являются:

– расширение состава логических условий реализации процессов документооборота и угроз безопасности информации в СЭД в сочетании с применением аппарата логических сетей и теории предикатов;

– разработка аналитических моделей расчета показателей оценки влияния различных мер и средств защиты информации на процессы реализации угроз электронному документообороту;

– проведение теоретических и экспериментальных исследований по нормированию значений показателей эффективности защиты электронного документооборота в интересах обоснования требований к системам защиты;

– разработка программных средств автоматизации оценивания эффективности защиты электронного документооборота в СЭД.

Литература

1. *Скрыль С.В. Лаврухин Ю.Н., Курило А.П., Багаев Д.А.* Обоснование показателей для оценки эффективности информационных процессов в информационно-

телекоммуникационных системах в условиях противодействия угрозам информационной безопасности // Информация и безопасность. 2009. № 3. С. 429–432.
2. *Скоредова Ю.В., Гречишников Е.В., Кравченко А.С., Ланкин О.В.* Анализ методик автоматизированной оценки угроз и рисков информационной безопасности информационно-телекоммуникационных систем // Вестник Воронежского института ФСИН России. 2017. № 3. С. 128–133.
3. *Скрыль С.В. и др.* Вероятностные модели информационных процессов в интегрированных системах безопасности в условиях обеспечения защиты информации от несанкционированного доступа // Телекоммуникации. 2015. № 6. С. 26–31.
4. *Muratkhan R., Satybaldina D.Z.* Quantitative method of information security risk assessment by multicomponent threats // Life Science Journal. 2014. vol. 11. no. 11. pp. 372–375.
5. *Al Hadidi M. et al.* Methods of risk assessment for information security management // International Review on Computers and Software. 2016. vol. 11. no. 2. pp. 81–91.
6. *Anisimov V.G., Zegzhda P.D., Anisimov E.G., Bazhin D.A.* A risk-oriented approach to the control arrangement of security protection subsystems of information systems // Automatic Control and Computer Sciences. 2016. vol. 50. no. 8. pp. 717–721.
7. *Кукин С.А., Печерский А.В.* Управление рисками информационных угроз для системно связанных объектов // Сборник статей XV Международной научно-практической конференции «Современные технологии документооборота в бизнесе, производстве и управлении». 2015. С. 41–44.
8. *Makarevich O., Mashkina I., Sentsova A.* The method of the information security risk assessment in cloud computing systems // Proceedings of the 6th International Conference on Security of Information and Networks. 2013. pp. 446–447.
9. *Язов Ю.К., Соловьев С.В.* Защита информации в информационных системах от несанкционированного доступа // Воронеж: Кварта. 2015. 440 с.
10. *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. Издание второе, переработанное и дополненное // М.: Наука. 1987.
11. *Котов В.Е.* Сети Петри // М.: Наука. 1984. 160с.
12. *Brissaud F., Luiz F.* Average probability of a dangerous failure on demand: different modelling methods, similar results // arXiv preprint arXiv: 1501.06487. 2015.
13. *Yang M., Wang M., Qu Y.* Modeling and performance analysis of the emergency rescue logistics system based on Petri nets // Journal of Hebei University of Science and Technology. 2017. vol. 38(3). pp. 269–277.
14. *Vazquez C.R., Silva M.* Stochastic Continuous Petri Nets: An Approximation of Markovian Net Models // IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans. 2011. vol. 42. no. 3. pp. 641–653.
15. *Игнатьев В.М., Ларкин Е.В.* Сети Петри-Маркова // ТулГТУ. 1994.
16. *Mohan L.N., Anjaneyulu G.S.G.N.* A secured digital signature using conjugacy and DLP on non-commutative group over finite field // Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications. 2017. pp. 457–465.
17. *Detken K.O. et al.* SIEM approach for a higher level of it security in enterprise networks // 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2015. vol. 1. pp. 322–327.
18. *Majeed A. et al.* Near-miss situation based visual analysis of SIEM rules for real time network security monitoring // Journal of Ambient Intelligence and Humanized Computing. 2018. vol. 10. no. 4. pp. 1509–1526.
19. *Martinasek Z., Blazek P., Silhavy P., Smekal D.* Methodology for correlations discovery in security logs // 2017 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2017. pp. 294–298.
20. *Тихонов В.И.* Статистическая радиотехника // М.: Сов. радио. 1966.
21. *Климов Г.П.* Стохастические системы массового обслуживания // М.: Наука. 1966.

22. *Тараканов К.В., Овчаров Л.А., Тырышкин А.Н.* Аналитические методы исследований систем // М.: Сов радио. 1974. 240 с.
23. *Авсентьев О.С. Рубцова И.О.* Обобщенное представление информационных процессов в системах электронного документооборота специального назначения в условиях угроз безопасности информации // Вестник Воронежского института МВД России. 2017. № 4. С. 108–115.
24. *Hallé S. et al.* Decentralized enforcement of document lifecycle constraints // Information Systems. 2018. vol. 74. pp. 117–135.
25. *Язов Ю.К., Панфилов В.В.* Моделирование динамики реализации угроз безопасности информации с использованием аппарата сетей Петри-Маркова // Информация и безопасность. 2006. Т. 9. № 1. С. 117–123.
26. Основные функции системы СЭД «ДЕЛО». URL: http://www.interface.ru/home.asp?artid=21844 (дата обращения: 27.07.2019).

Язов Юрий Константинович — д-р техн. наук, профессор, главный научный сотрудник, Государственный научно-исследовательский, испытательный институт проблем технической защиты информации, Федеральная служба по техническому и экспортному контролю России; профессор, кафедра систем информационной безопасности, Воронежский государственный технический университет. Область научных интересов: применение методов математического моделирования в области обеспечения информационной безопасности объектов различных сфер деятельности, разработка, совершенствование и применение методов и средств защиты информации в процессе сбора, хранения, обработки, передачи и распространения информации. Число научных публикаций — 290. Yazoff_1946@mail.ru; ул. 9 Января, 280 а, 394020, Воронеж, Российская Федерация; р.т.: +7(903)-651-42-69.

Авсентьев Олег Сергеевич — д-р техн. наук, профессор, профессор, кафедра информационной безопасности, Воронежский институт Министерства внутренних дел России; профессор, кафедра организации и технологии защиты информации, Белгородский университет кооперации, экономики и права. Область научных интересов: применение методов математического моделирования в области обеспечения информационной безопасности объектов различных сфер деятельности, разработка, совершенствование и применение методов и средств защиты информации в процессе сбора, хранения, обработки, передачи и распространения информации. Число научных публикаций — 101. osaos@mail.ru; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7-(473)-500-52-42; факс: +7(473)-200-52-00.

Авсентьев Александр Олегович — канд. техн. наук, старший преподаватель, кафедра физики, Воронежский институт Министерства внутренних дел России. Область научных интересов: применение методов математического моделирования в области обеспечения информационной безопасности объектов различных сфер деятельности, разработка, совершенствование и применение методов и средств защиты информации в процессе сбора, хранения, обработки, передачи и распространения информации. Число научных публикаций — 40. aoaao8787@mail.ru; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7-(473)-500-52-66; факс: +7-(473)-200-52-00.

Рубцова Ирина Олеговна — аспирант, кафедры организации и технологии защиты информации, Белгородский университет кооперации, экономики и права. Область научных интересов: разработка, совершенствование и применение методов и средств защиты информации в процессе сбора, хранения, обработки, передачи и распространения информации; обеспечения информационной безопасности объектов политической, социально-экономической, оборонной и других сфер деятельности от внешних и внутренних угроз хищения, разрушения и/или модификации информации. Число научных публикаций — 15. irinka23@bk.ru; ул. Садовая, 116 А, 308023, Белгород, Российская Федерация; р.т.: +7-(4722)-26-38-31; факс: +7-(4722)-26-49-65.

DOI 10.15622/sp.2019.18.6.1269-1300

YU.K. YAZOV, O.S. AVSENTEV, A.O. AVSENTEV, I.O. RUBTSOVA

METHOD FOR ASSESSING EFFECTIVENESS OF PROTECTION OF ELECTRONIC DOCUMENT MANAGEMENT USING THE PETRI AND MARKOV NETS APPARATUS

Yazov Yu.K., Avsentev O.S., Avsentev A.O., Rubtsova I.O. **Method for Assessing Effectiveness of Protection of Electronic Document Management using the Petri and Markov Nets Apparatus.**

Abstract. Traditional approaches to assessing the effectiveness of information security, based on a comparison of the possibilities of realizing threats to information security in absence and application of protection measures, do not allow to analyze the dynamics of suppression by security measures of the process of implementing threats. The paper proposes a new indicator of the effectiveness of protection of electronic documents, aimed at assessing the possibility of advancing security measures of the process of implementing threats in electronic document management systems using the probability-time characteristics of the dynamics of the application of protection measures and the implementation of threats to electronic documents. Mathematical models were developed using the Petri-Markov network apparatus and analytical relationships were obtained for calculating the proposed indicator using the example of the "traffic tunneling" threat (placing intruder packets in trusted user packets) and unauthorized access (network attacks) to electronic documents, as well as the threat of intrusion of malicious program by carrying out an "blind IP spoofing" attack (network address spoofing). Examples of calculating the proposed indicator and graphs of its dependence on the probability of detecting network attacks by the intrusion detection system and on the probability of malware detection by the anti-virus protection system are given. Quantitative dependencies are obtained for the effectiveness of protection of electronic documents due to being ahead of protection measures for threat realization processes, both on the probability of detecting an intrusion or the probability of detecting a malicious program, and on the ratio of the time spent by the protection system on detecting an attempt to implement a threat and taking measures to curb its implementation, and threat implementation time. Models allow not only to evaluate the effectiveness of measures to protect electronic documents from threats of destruction, copying, unauthorized changes, etc., but also to quantify the requirements for the response time of adaptive security systems to detectable actions aimed at violating the security of electronic documents, depending on the probability -temporal characteristics of threat realization processes, to identify weaknesses in protection systems related to the dynamics of threat realization and the reaction of defense systems to such threats electronic document.

Keywords: Efficiency Indicator, Functional Model, Petri-Markov Network, Security Threat, Security Measure, Intrusion Detection System, Anti-Virus Protection System.

Yazov Yuri Konstantinovich — Ph.D., Dr.Sci., Professor, Chief Researcher, Testing Institute of Problems of Technical Protection of information, Federal Service for Technical and Export Control of Russia; Professor, Department of Information security systems, Voronezh state technical University. Research interests: application of mathematical modeling methods in the field of information security of objects of various fields of activity, development, improvement and application of methods and means of information security in the process of collection, storage, processing, transmission and dissemination of information. The number of publications — 290. Yazoff_1946@mail.ru; 280 a, 9 January str., 394020, Voronezh, Russian Federation; office phone: +7(903)-651-42-69.

Avsentev Oleg Sergeevich — Ph.D., Dr.Sci., Professor, Professor, Department of Information Security, Voronezh Institute of the Ministry of Interior of Russia; Professor, Department of

Organization and Technology of Information Protection, Belgorod University of Cooperation, Economics and Law. Research interests: application of mathematical modeling methods in the field of information security of objects of various fields of activity, development, improvement and application of methods and means of information security in the process of collection, storage, processing, transmission and dissemination of information. The number of publications — 101. osaos@mail.ru; 53, pr. Patriotov, 394065, Voronezh, Russian Federation; office phone: +7-(473)-500-52-42; fax: +7(473)-200-52-00.

Avsentev Alexander Olegovich — Ph.D., senior lecturer, Department of Physics, Voronezh Institute of the Ministry of Interior of Russia. Research interests: application of mathematical modeling methods in the field of information security of objects of various fields of activity, development, improvement and application of methods and means of information security in the process of collection, storage, processing, transmission and dissemination of information. The number of publications — 40. aoaao8787@mail.ru; 53, pr. Patriotov, 394065, Voronezh, Russian Federation; office phone: +7-(473)-500-52-66; fax: +7-(473)-200-52-00.

Rubtsova Irina Olegovna — Ph.D. Student, Department of Organization and Technology of Information Security, Belgorod University of Cooperation, Economics and Law. Research interests: development, improvement and application of methods and means of information security in the process of collection, storage, processing, transmission and dissemination of information; information security of objects of political, socio-economic, defense and other fields of activity from external and internal threats of theft, destruction and/or modification of information. The number of publications — 15. irinka23@bk.ru; 116 A, Sadovaya str., 308023, Belgorod, Russian Federation; office phone: +7-(4722)-26-38-31; fax: +7-(4722)-26-49-65.

References

1. Skryl' S.V. Lavrukhin YU.N., Kurilo A.P., Bagaev D.A. [Justification of indicators for evaluating the effectiveness of information processes in information and telecommunication systems in the face of threats to information security]. *Informatsiya i bezopasnost' – Information and security.* 2009. vol. 3. pp. 429–432. (In Russ.).
2. Skoredova Yu.V., Grechishnikov E.V., Kravchenko A.S., Lankin O.V. [Analysis of methods for automated evaluation of threats and risks to information security information and telecommunication systems]. *Vestnik Voronezhskogo instituta FSIN Rossii – Vestnik of Voronezh Institute of the Russian Federal Penitentiary Service.* 2017. vol. 3. pp. 128–133. (In Russ.).
3. Skryl' S.V. et al. [Probabilistic models of information processes in integrated security systems in the conditions of information protection from unauthorized access]. *Telekommunikatsii – Telecommunications*. 2015. vol. 6. pp. 26–31. (In Russ.).
4. Muratkhan R., Satybaldina D.Z. Quantitative method of information security risk assessment by multicomponent threats. *Life Science Journal.* 2014. vol. 11. no. 11. pp. 372–375.
5. Al Hadidi M. et al. Methods of risk assessment for information security management. *International Review on Computers and Software*. 2016. vol. 11. no. 2. pp. 81–91.
6. Anisimov V.G., Zegzhda P.D., Anisimov E.G., Bazhin D.A. A risk-oriented approach to the control arrangement of security protection subsystems of information systems. *Automatic Control and Computer Sciences.* 2016. vol. 50. no. 8. pp. 717–721.
7. Kukin S.A., Pecherskij A.V. [Risk management of information threats for systemically connected objects]. *Sbornik statej XV Mezhdunarodnoj nauchno-prakticheskoj konferentsii "Sovremennye tekhnologii dokumentooborota v biznese, proizvodstve i upravlenii"* [Collection of articles XV International scientific and practical conference "Modern technologies of document circulation in business, production and management"]. 2015. pp. 41–44. (In Russ.).

8. Makarevich O., Mashkina I., Sentsova A. The method of the information security risk assessment in cloud computing systems. Proceedings of the 6th International Conference on Security of Information and Networks. 2013. pp. 446–447.
9. Yazov YU.K., Solov'ev S.V. *Zashhita informatsii v informatsionnykh sistemakh ot nesanktsionirovannogo dostupa* [Protection of information in information systems from unauthorized access]. Voronezh: Kvarta. 2015. 440 p. (In Russ.).
10. Gnedenko B.V., Kovalenko I.N. *Vvedenie v teoriyu massovogo obsluzhivaniya. Izdanie vtoroe, pererabotannoe i dopolnennoe* [Introduction to Queuing theory. Second edition, revised and expanded]. M.: Nauka. 1987. (In Russ.).
11. Kotov V.E. *Seti Petri* [Petri Nets]. M.: Nauka. 1984. 160 p. (In Russ.).
12. Brissaud F., Luiz F. Average probability of a dangerous failure on demand: different modelling methods, similar results. arXiv preprint arXiv: 1501.06487. 2015.
13. Yang M., Wang M., Qu Y. Modeling and performance analysis of the emergency rescue logistics system based on Petri nets. *Journal of Hebei University of Science and Technology.* 2017. vol. 38(3). pp. 269–277.
14. Vazquez C.R., Silva M. Stochastic Continuous Petri Nets: An Approximation of Markovian Net Models. *IEEE Transactions on Systems, Man, and Cybernetics* – Part A: Systems and Humans. 2011. vol. 42. no. 3. pp. 641–653.
15. Ignat'ev V.M., Larkin E.V. *Seti Petri-Markova* [Petri-Markov Nets]. TulGTU. 1994. (In Russ.)
16. Mohan L.N., Anjaneyulu G.S.G.N. A secured digital signature using conjugacy and DLP on non-commutative group over finite field. Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications. 2017. pp. 457–465.
17. Detken K.O. et al. SIEM approach for a higher level of it security in enterprise networks. 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2015. vol. 1. pp. 322–327.
18. Majeed A. et al. Near-miss situation based visual analysis of SIEM rules for real time network security monitoring. *Journal of Ambient Intelligence and Humanized Computing.* 2018. vol. 10. no. 4. pp. 1509–1526.
19. Martinasek Z., Blazek P., Silhavy P., Smekal D. Methodology for correlations discovery in security logs. 2017 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2017. pp. 294–298.
20. Tikhonov V.I. *Statisticheskaya radiotekhnika* [Statistical radio engineering]. M.: Sov. radio. 1966. (In Russ.).
21. Klimov G.P. *Stokhasticheskie sistemy massovogo obsluzhivaniya* [Stochastic queueing systems]. M.: Nauka. 1966. (In Russ.).
22. Tarakanov K.V., Ovcharov L.A., Tyryshkin A.N. *Analiticheskie metody issle-dovanij system* [Analytical methods of research systems]. M.: Sov radio. 1974. 240 p. (In Russ.).
23. Rubtsova, I.O. Avsentev O.S. [A consolidated view of information processes in systems of electronic document circulation of special purpose in terms of threats to information security]. *Vestnik Voronezhskogo instituta MVD Rossii – The bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia.* 2017. vol 4. pp. 108–115.
24. Hallé S. et al. Decentralized enforcement of document lifecycle constraints. *Information Systems.* 2018. vol. 74. pp. 117–135.
25. Yazov Yu.K., Tekunov V.V. [Modeling of the dynamics of realization of the threats to information security using the apparatus of Petri nets and Markov]. *Informatsiya i bezopasnost' – Information and security*. 2006. Issue 9. vol. 1. pp. 117–123.
26. Osnovnye funkcii sistemy SED "DELO" [The main functions of the system EDMS "CASE"]. Available at: http://www.interface.ru/home.asp?artid=21844 (accessed: 27.07.2019).

УДК 004.056.53 **DOI** 10.15622/sp.2019.18.6.1301-1332

О.И. Бокова, И.Г. Дровникова, А.С. Етепнев, Е.А. Рогозин, В.А. Хвостов

МЕТОДИКИ ОЦЕНИВАНИЯ НАДЕЖНОСТИ СИСТЕМ ЗАЩИТЫ ИНФОРМАЦИИ ОТ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ

Бокова О.И., Дровникова И.Г., Етепнев А.С., Рогозин Е.А., Хвостов В.А. **Методики оценивания надежности систем защиты информации от несанкционированного доступа в автоматизированных системах.**

Аннотация. Современные методы защиты информации от несанкционированного доступа в обязательном порядке включаются в виде дополнительных модулей в программное обеспечение автоматизированных систем в защищенном исполнении. Применение систем защиты информации от несанкционированного доступа может снизить надежность автоматизированных систем, если они содержат ошибки, не обнаруживаемые при отладке.

Методической основой при формировании облика систем защиты информации как в процессе разработки, так и в процессе модернизации являются руководящие документы Федеральной службы по техническому и экспортному контролю (ФСТЭК) России. Руководящие документы ФСТЭК России не содержат методических подходов к оценке надежности указанных программных систем. В этой связи актуальна разработка методик оценивания надежности систем защиты информации от несанкционированного доступа, структурная сложность и значительное количество выполняемых функций которых обусловили необходимость использования трех показателей надежности, характеризующих систему при решении задач обеспечения конфиденциальности, целостности и доступности информации. Для разработки методик использованы известные методы оценивания надежности сложных систем, не допускающие их разложение на последовательное и параллельное соединение. Разработанные методики апробированы при оценивании надежности систем защиты информации от несанкционированного доступа, имеющих типовые показатели исходных характеристик. Результаты расчетов и перспективы использования разработанных методик представлены в статье.

Ключевые слова: несанкционированный доступ, система защиты информации, надежность, отказ, автоматизированная система, конфиденциальность информации, целостность информации, доступность информации.

1. Введение. Проникновение современных информационных технологий (ИТ) во все сферы человеческой деятельности обуславливает необходимость научного осмысления последствий их внедрения и практического применения, а также требует содержательного анализа проблемы информационной безопасности (ИБ).

Необходимость осмысления результатов внедрения ИТ в значительной степени связана с объектами информатизации критического применения (ОИКП) (автоматизированными

системами (АС), используемыми в военной сфере, органах государственной безопасности и охраны, органах внутренних дел и т.д.), выход из строя которых может привести к существенным финансовым, человеческим и другим потерям, что является неприемлемым для общества.

Трудности правового регулирования ИТ и недостатки организационного регулирования процесса обеспечения ИБ, отсутствие методик оценивания и обоснования требований к ИБ, проблемы с кадровым обеспечением ОИКП могут привести к тому, что конфиденциальная информация, используемая узким кругом потребителей, становится объектом неправомерного доступа для злоумышленников.

Исключительная важность задач, решаемых в АС на ОИКП с использованием средств вычислительной техники, а также существенный ущерб интересам личности и государства, который возникает в результате снижения уровня безопасности информации в АС, подчеркивают значимость задачи обеспечения надежности функционирования АС в защищенном исполнении и в особенности безопасности хранимой, обрабатываемой и передаваемой конфиденциальной информации.

Классификационная схема угроз ИБ АС представлена как в нормативных документах ФСТЭК России [1, 2], так и в общедоступной научно-технической литературе [3-6]. В указанных источниках рассматривается основное содержание угроз ИБ, точки их приложения, оценки ущербов от возникновения и реализации угроз. Формализованы угрозы ИБ в виде базы знаний угроз ФСТЭК России [4-6], включающей как описательную, так и расчетную части в виде калькулятора оценки опасности угрозы. База знаний содержит информацию об уязвимостях операционных систем (ОС) и систем управления базами данных, прикладных программ, методов защиты информации, а также о связанных с этими уязвимостями и методами защиты рисках.

Международные стандарты, регламентирующие область обеспечения БИ и созданные для развития и углубления методологии ГОСТ «Общие критерии оценки безопасности информационных технологий» ISO/IEC 15408: 2013. «Информационная технология — Методы и средства защиты информации — Критерии оценки безопасности информационных технологий» (ОК), содержат классификацию угроз ИБ АС. Подробный перечень угроз ИБ, предназначенный для обоснования и выбора мер по защите информации от несанкционированного доступа АС, приведен в стандарте Национального института стандартов США (NIST)

ISO/IEC13335:2004 «Information technology — Security techniques — Management of information and communication technology security» (информационный портал национального института стандартов США http://www.nist.org) [7, 8].

В целях обнаружения и противодействия наиболее опасным видам угроз ИБ, связанных с НСД к информации АС, традиционно применяют системы защиты информации (СЗИ) от НСД, которые в обязательном порядке включаюися в виде дополнительных программных систем в состав ОС АС критического применения. Системы защиты могут разрабатываться совместно с АС в ходе ее проектирования или устанавливаться в общесистемное программное обеспечение готовой системы.

Использование систем защиты снижает надежность АС, поскольку они, как и большинство программ, могут содержать не обнаруженные при отладке ошибки. В процессе эксплуатации ошибки СЗИ приводят к снижению интегральной надежности АС. Надежность, в свою очередь, оказывает влияние на эффективность защиты информации (обеспечение конфиденциальности, целостности и доступности информации).

Методической основой обоснования требований к системам защиты является ряд руководящих документов ФСТЭК России, в которых установление требований к защите информации определяется требуемым классом защищенности [9, 10]. Для СЗИ, разрабатываемых в соответствии с международным стандартом ISO/IEC 15408 ОК [7], установление требований состоит в выполнении профиля защиты. Замена понятия класса защищенности понятиями профиля защиты и задания по безопасности является отличительным признаком стандарта ОК.

Профиль защиты включает в себя совокупность функций защиты, применяемых в конкретном профиле, и элементы доверия для конкретного изделия ИТ.

Результаты анализа отечественной и международной нормативной документации показали, что в нормативных документах в области ИБ фактически задается совокупность функций защиты, которые требуется реализовать в СЗИ от НСД. Совокупность функций защиты является признаком, определяющим соответствие СЗИ от НСД как классу защищенности, так и профилю защиты.

Качеству программных систем в Российской федерации посвящены ГОСТ 28195 — 89 и ГОСТ28806 — 90 [11, 12]. Анализ их содержания позволяет сделать вывод об ограниченной применимости методик, содержащихся в стандартах для оценивания надежности СЗИ

от НСД. В представленных стандартах СЗИ не рассматривается как объект оценки качества программных систем. Данное обстоятельство требует творческой доработки методического обеспечения указанных стандартов как в части показателей надежности СЗИ от НСД, так и в части методик их оценивания.

Аналогичные утверждения применимы и в отношении известной научно-технической литературы, посвященной вопросам качества программных систем [13-18].

Оценивание показателей надежности технических систем в России регламентировано рядом ГОСТ, составляющих целую группу стандартов. В частности, стандарты группы 27.XXX посвящены общим проблемам надежности в технике. Несмотря на методическую полноту системы стандартов указанной серии, вопросы оценивания надежности СЗИ от НСД в них напрямую не проработаны. Существует лишь возможность применения отдельных теоретических положений и подходов к оцениванию надежности при разработке математических зависимостей, используемых в методиках.

Известная научно-техническая литература, посвященная проблеме надежности, в частности [19-28], также не рассматривает вопросы оценивания надежности СЗИ от НСД.

При создании АС в защищенном исполнении согласно нормативной документации [29], в которой определены имеющиеся недостатки существующих систем обеспечения ИБ на ОИКП и направления совершенствования АС в части обеспечения защиты информации, оценивание надежности СЗИ от НСД является неотъемлемой составляющей процессов проектирования и эксплуатации указанных АС. При этом вопросам оценивания надежности посвящена конструкторская документация («Техническое задание...», «Эскизный проект...», «Технический проект...», «Пояснительная записка...» и др.) и разрабатываемая на ее основе эксплуатационная документация («Руководство пользователя», «Руководство администратора безопасности», «План мероприятий по обеспечению защиты информации», «Руководство по резервному копированию и восстановлению» и др.).

Таким образом, разработка методического обеспечения, включающего конкретные методики оценивания надежности функционирования СЗИ от НСД при решении задач обеспечения конфиденциальности, целостности и доступности информации, является актуальной проблемой для обеспечения ИБ АС в защищенном исполнении на ОИКП.

Поставлена научная задача разработать и апробировать методики оценивания надежности СЗИ от НСД, включающих показатели вероятности безотказной работы СЗИ от НСД на этапах проектирования и эксплуатации АС в защищенном исполнении.

2. Методика оценивания вероятности безотказной работы систем защиты информации от несанкционированного доступа при обеспечении конфиденциальности информации. Анализ научно-технической литературы [26-28, 30-32] показал, что наиболее целесообразным методом расчета вероятности безотказной работы СЗИ от НСД в АС является метод эквивалентных схем. В основе указанного метода используется следующая формула для расчета полной вероятности [26]:

$$P_c = f(p_1, p_2, ...p_n) = p_1 P(t \mid p_1 = 1) + q_1 P(t \mid p_1 = 0), \quad (1)$$

где: p_i — вероятность безотказной работы i-го элемента; $q_i = 1 - p_i$ — вероятность отказа i-го элемента; $P(t \mid p_1 = 1)$, $P(t \mid p_1 = 0)$ — условные вероятности работоспособного состояния системы при работоспособном состоянии (отказе) первого элемента.

Математические выражения $P(t \mid p_1 = 1)$, $P(t \mid p_1 = 0)$ предназначены для расчета вероятностей безотказной работы структурных схем, эквивалентных исходной схеме, при условии, что первый элемент первой схемы является абсолютно надежным, а первый элемент второй схемы отказал. Отражаемая выражением (1) операция называется операцией разрезания по элементу 1 [31]. В источниках [26, 27, 31-36] показано, что операция разрезания может проводиться по любому элементу структурной схемы надежности системы. При этом выполняется следующая последовательность действий.

1. Определяются простые точки соединения исходной структурной схемы и в соответствии с этим правилом выбирается элемент, по которому будет реализовываться операция разрезания.

2. Исходная структура преобразуется в две эквивалентные схемы. В первой из них элемент разрезания заменяется абсолютно надежной перемычкой. Во второй схеме элемент полностью удаляется.

3. Проводится расчет надежности по каждой эквивалентной схеме и используется математическое выражение (1) для получения результирующего выражения расчета показателя надежности.

Исходная структурная схема по надежности СЗИ от НСД, разработанная в [37], представлена на рисунке 1.

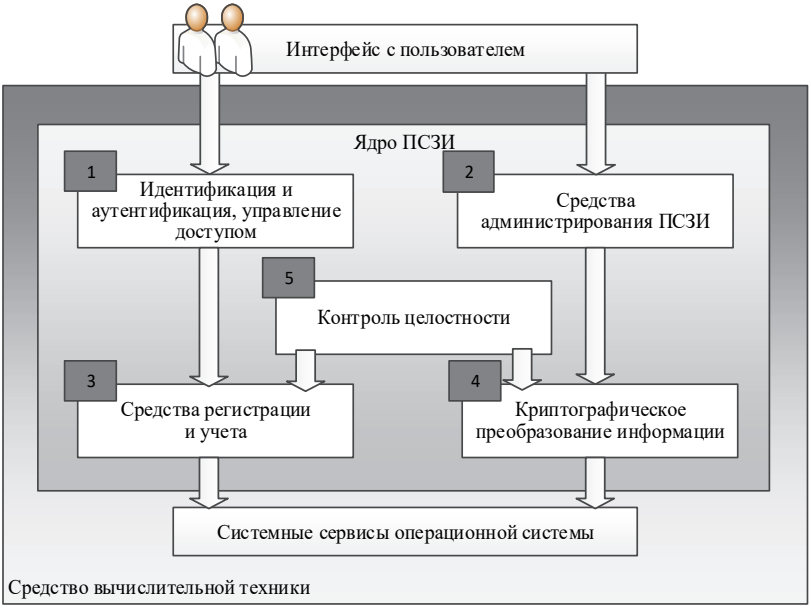


Рис. 1. Исходная структурная схема по надежности СЗИ от НСД

При оценивании вероятности безотказной работы данную структурную схему СЗИ от НСД можно представить в виде двух эквивалентных схем.

Разрезание структурной схемы СЗИ целесообразно провести по элементу «Контроль целостности».

При этом в результате операции разрезания получим две эквивалентные схемы СЗИ от НСД, представленные на рисунках 2 и 3.

Используя формулу (1), с учетом разрезания по элементу «Контроль целостности» (элемент идеально работоспособен или полностью отказал) результирующий показатель надежности СЗИ можно получить в виде:

$$P_c = p_5(1-q_1q_2)(1-q_3q_4)+q_5(1-(1-p_1p_3)(1-p_2p_4)), \qquad (2)$$

где $p_i = 1-q_i$ — вероятность безотказной работы i-го элемента [26].

Математическое выражение (2) позволит оценивать вероятность безотказной работы СЗИ от НСД в АС.

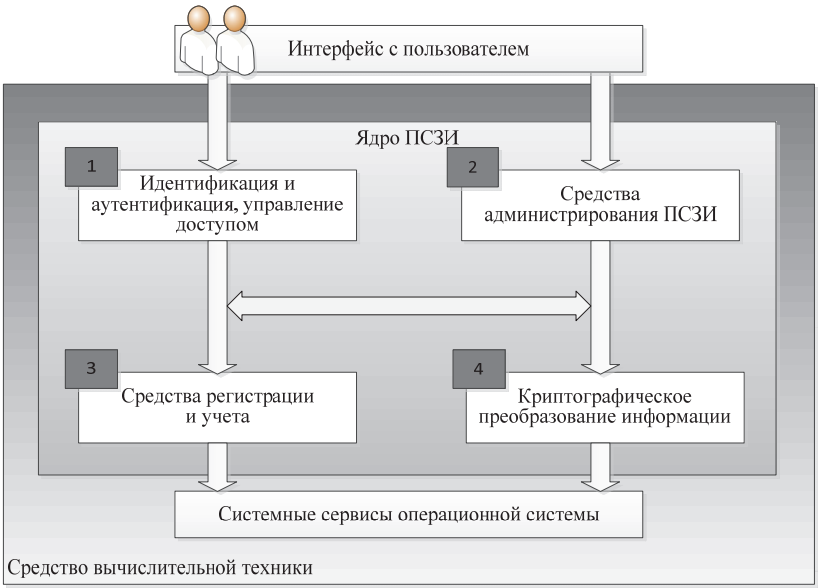


Рис. 2. Эквивалентная схема структуры СЗИ от НСД с заменой элемента «Контроль целостности» абсолютно надежной перемычкой

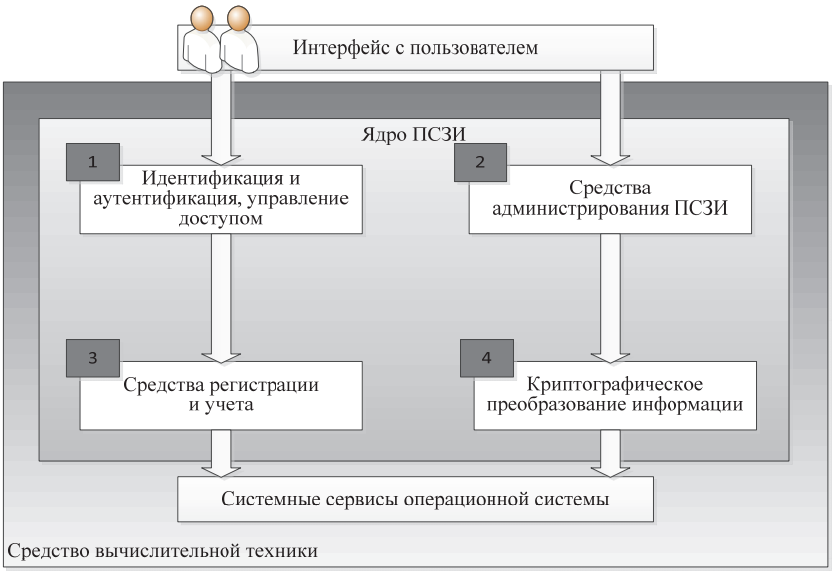


Рис. 3. Эквивалентная схема структуры СЗИ от НСД с заменой элемента «Контроль целостности» разрывом

3. Методика оценивания коэффициента готовности систем защиты информации от несанкционированного доступа при обеспечении доступности информации. При расчетах показателя надежности СЗИ от НСД «Коэффициент готовности» целесообразно использовать метод построения логико-вероятностной функции надежности СЗИ. Основой методики является математический аппарат булевой алгебры. На начальной стадии оценивания надежности структурно-сложных систем при построении расчетной зависимости необходимо последовательно реализовать ряд этапов [37-41].

Элементам сопоставляются логические переменные x_i, принимающие значения единицы при работоспособности элемента и нуля — при его неработоспособности. По результатам логического анализа работоспособности системы конструируется логическая функция работоспособности (ЛФР), имеющая следующий вид: $f(X)$, где многомерный аргумент ЛФР $X = (x_1, x_2, ...x_n)$ является вектором логических переменных.

Наличие хотя бы единственного пути от исходного структурного элемента схемы надежности до ее конечного элемента позволяет считать данную схему частично работоспособной $(f(X) = 1)$. Такой путь может считаться работоспособным при условии работоспособности всех входящих в него элементов СЗИ. Формализация всех возможных путей в ЛФР осуществляется элементарными конъюнкциями булевых переменных, которые соответствуют входящим в данные пути элементам. Функция $f(X)$ выражается в виде дизъюнкции элементарных конъюнкций, соответствующих всем работоспособным путям ЛФР. Форма ЛФР, полученная в ходе составления дизъюнкций элементарных конъюнкций, будет являться исходной для последующего построения математического выражения расчета коэффициента готовности СЗИ от НСД.

Исходная форма ЛФР может быть преобразована к любой форме полного замещения (ФПЗ). ЛФР СЗИ от НСД в виде ФПЗ используется для построения аналитической зависимости для оценивания надежности. При этом проводится замена логических переменных вероятностями, а логических операций — арифметическими.

В соответствии с выше изложенным осуществим ряд замещений: переменная x_i замещается вероятностью $p_i = P(x_i = 1)$, операция инверсии логической переменной $\overline{x_i}$ — вероятностью $q_i = P(x_i = 0)$,

операция дизъюнкции $\vee$ — сложением $+$, операция конъюнкции $\wedge$ — умножением $\times$. Операция логического отрицания $\neg y$ замещается вычитанием вероятности из единицы: $1 - P(y = 1)$.

Для проведения оценки коэффициента готовности СЗИ от НСД целесообразно использовать ФПЗ в базисе «конъюнкция-отрицание» [38, 41].

ЛФР СЗИ от НСД, представленной на рисунках 1 и 2, имеет следующий вид:

$$f = x_1(x_3 \vee x_4 x_5) \vee x_2(x_4 \vee x_3 x_5). \quad (3)$$

Число вхождений математического выражения (3) определяется как двойное. Дважды входят переменные x_3, x_4, x_5. Выбрав x_5 для операции разрезания, получим из исходной формы следующее математическое выражение:

$$f = x_5(x_1(x_3 \vee x_4)) \vee x_2(x_3 \vee x_4) \vee x_5'(x_1 x_3 \vee x_4 x_2). \quad (4)$$

Преобразуя математическое выражение (4) к ФПЗ, получим следующее математическое выражение:

$$f = x_5((x_1'x_2')'(x_3'x_4')')' \vee x_5'((x_1'x_3')'(x_2'x_4')')'. \quad (5)$$

Заменив логические переменные арифметическими в математическом выражении (5), получим:

$$P(f = 1) = p_5(1 - q_1 q_2)(1 - q_3 q_4) + q_5(1 - (1 - p_1 p_3)(1 - p_2 p_4)). \quad (6)$$

Коэффициент готовности СЗИ от НСД ($K_{ГСЗИ}$) получим путем замены вероятности безотказной работы программных модулей СЗИ коэффициентами готовности этих модулей ($K_{Гi}$) [38]. Вероятности отказа заменим коэффициентами простоя ($K_{ПРi}$):

$$K_{ГСЗИ} = p_5(1 - q_1 q_2)(1 - q_3 q_4) + q_5(1 - (1 - p_1 p_3)(1 - p_2 p_4)), \quad (7)$$

где: $p_i = K_{Гi}$, $q_i = 1 - K_{Гi}$.

Для марковской модели надежности имеем:

$$K_{Гi} = \frac{\mu_i}{\mu_i + \lambda_i},$$

$$K_{ПРi} = \frac{\lambda_i}{\mu_i + \lambda_i}.$$

При одинаковых значениях характеристик надежности программных модулей СЗИ от НСД $\lambda_i = \lambda$, $\mu_i = \mu$, приняв $\rho = \frac{\lambda}{\mu}$, математическое выражение (7) можно преобразовать к следующему виду:

$$K_{ГСЗИ} = \frac{(1 + 5\rho + 8\rho^2 + 2\rho^3)}{(1+\rho)^5}. \quad (8)$$

Математическое выражение (8) позволит рассчитать коэффициент готовности СЗИ от НСД в АС [42, 43].

4. Методика оценивания средней наработки систем защиты информации от несанкционированного доступа на отказ при обеспечении целостности и доступности информации. В основе методики проведения оценки средней наработки СЗИ от НСД на отказ использована система алгебраических уравнений полумарковской модели оценки надежности [40, 41, 43]. Для конструирования расчетной формулы оценки средней наработки на отказ применим следующий алгоритм.

В соответствии с исходной структурой СЗИ от НСД S_0 определим минимальный d и максимальный m запасы живучести. При этом согласно [38] под минимальным запасом живучести будем понимать число элементов СЗИ от НСД в минимальном замыкающем множестве s (s — минимальное количество модулей, при которых СЗИ от НСД остается частично работоспособной), уменьшенное на единицу.

Для максимального запаса живучести необходимо определить максимальное количество отказавших элементов, при котором остается хотя бы один работоспособный путь от входного элемента структуры по надежности СЗИ до ее выходного элемента.

Выделим m уровней деградации СЗИ от НСД и зададим состояния S_{ij}, при которых структурная схема по надежности СЗИ считается частично работоспособной. В качестве номера уровня деградации СЗИ выберем количество отказавших элементов.

Каждому типу структуры СЗИ от НСД сопоставим состояние S_{ij}. Назначим для каждого состояния номер соответствующей структуры, используя сквозную нумерацию состояний, а также зададим обобщенное неработоспособное состояние СЗИ от НСД.

Зададим возможные переходы между различными работоспособными состояниями, а также возможные переходы из работоспособных состояний в обобщенное неработоспособное состояние.

Установив интенсивности переходов между состояниями работоспособности и обобщенным состоянием неработоспособности, получим взвешенный граф состояний. Далее необходимо составить систему алгебраических уравнений для полумарковской модели расчета надежности системы со сложной структурой [39] и решить ее относительно состояния СЗИ $S_0(\overline{T_0})$.

Анализ технической документации основных типов СЗИ от НСД в АС [41] показал, что интенсивности отказов отдельных программных модулей можно принять равными, то есть $\lambda_i = \lambda$, интенсивности восстановления работоспособности отдельных программных модулей — также равными друг другу, то есть $\mu_i = \mu$, а число операций по восстановлению работоспособности — равным 3.

Анализ структурной схемы по надежности СЗИ от НСД в АС [40], представленной на рисунке 1, позволяет выбрать следующие параметры: $d = 1$, $m = 3$. В соответствии с принятыми значениями минимального и максимального запасов живучести СЗИ выделим три уровня их деградации.

Для первого уровня деградации можно выделить две различные структуры СЗИ от НСД — S_{11} и S_{12}, остающиеся частично работоспособными. Структурные схемы СЗИ от НСД при первом уровне деградации представлены на рисунках 4 и 5 соответственно.

Для второго уровня деградации можно выделить три различные структуры СЗИ от НСД — S_{21}, S_{22}, S_{23}, остающиеся частично работоспособными. Структурные схемы СЗИ от НСД при втором уровне деградации представлены на рисунках 6-8 соответственно.

Для третьего уровня деградации можно выделить структуру СЗИ от НСД S_{31}, остающуюся частично работоспособной. Структурная схема СЗИ от НСД при третьем уровне деградации представлена на рисунке 9.

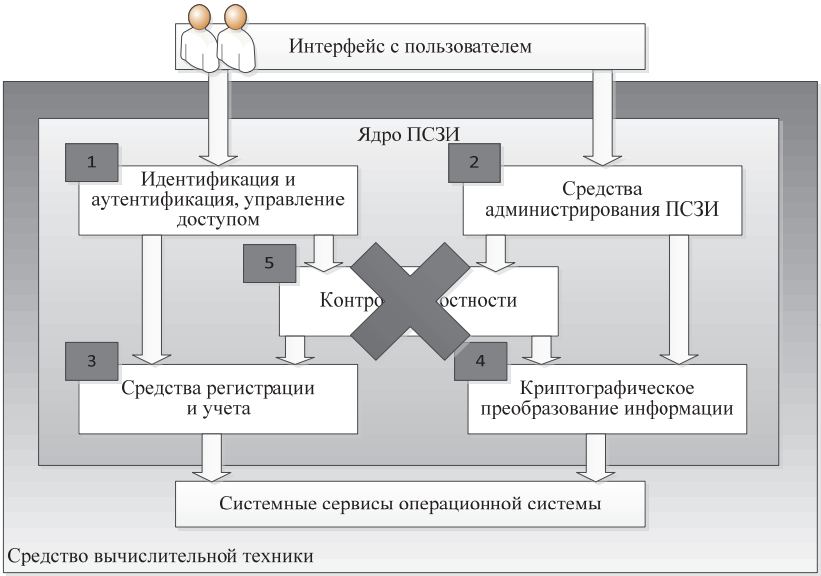


Рис. 4. Структурная схема S_{11} СЗИ от НСД для первого уровня деградации

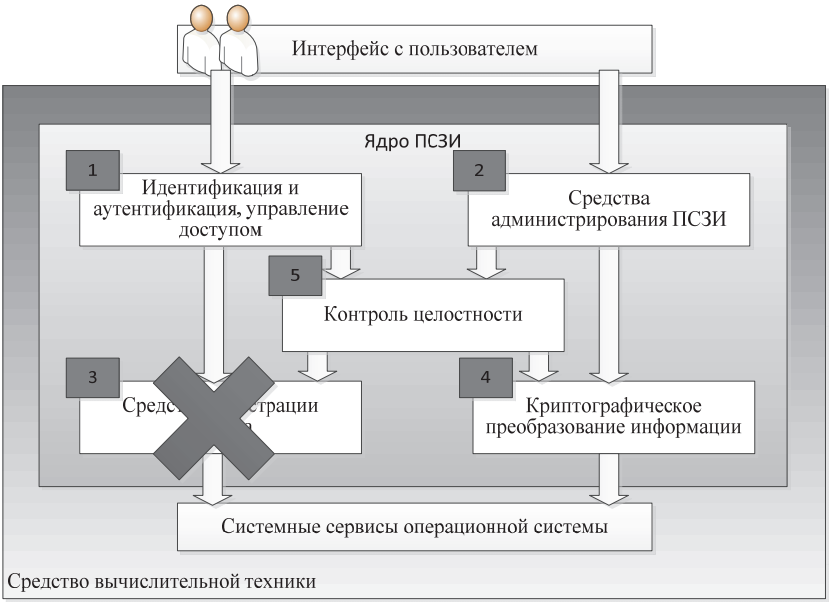


Рис. 5. Структурная схема S_{12} СЗИ от НСД для первого уровня деградации

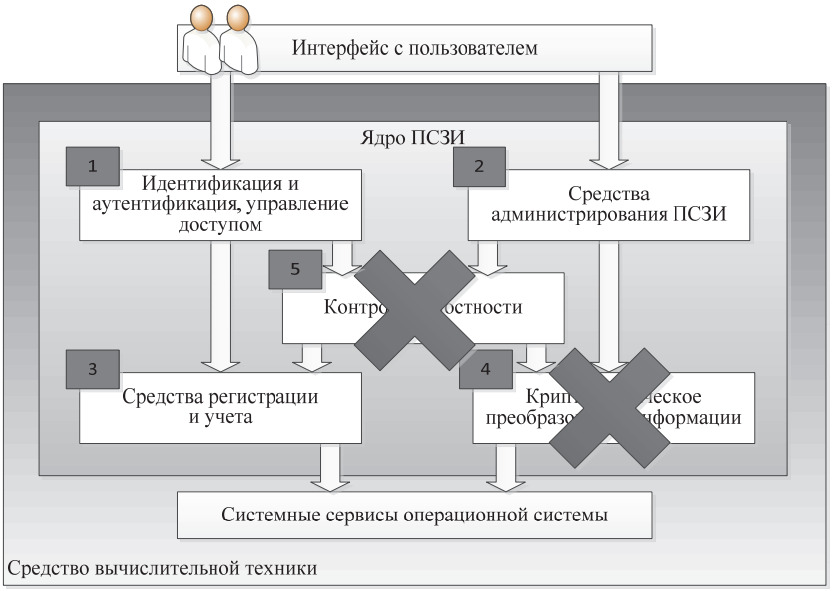


Рис. 6. Структурная схема S_{21} СЗИ от НСД для второго уровня деградации

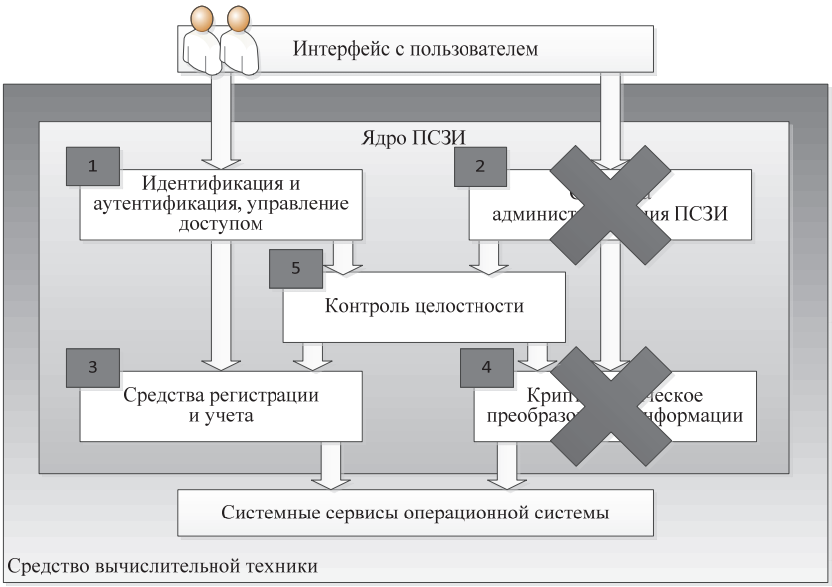


Рис. 7. Структурная схема S_{22} СЗИ от НСД для второго уровня деградации

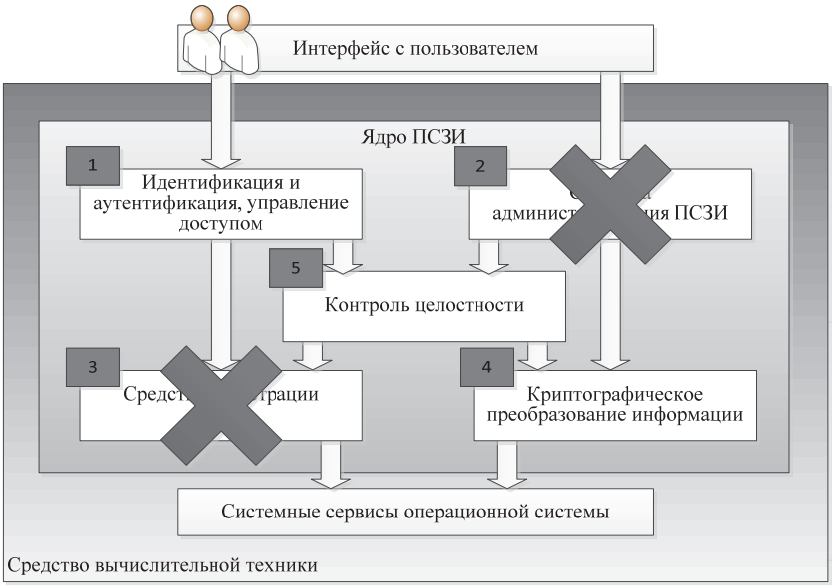


Рис. 8. Структурная схема S_{23} СЗИ от НСД для второго уровня деградации

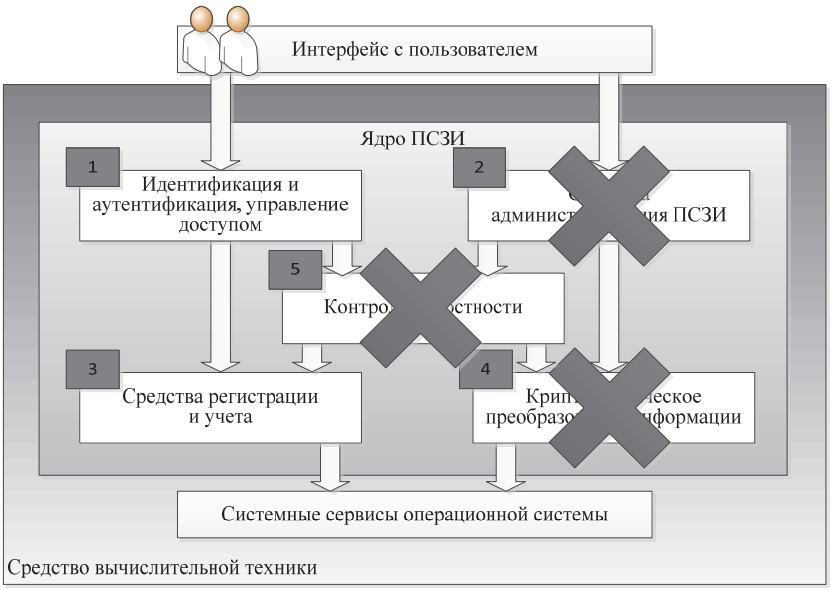


Рис. 9. Структурная схема S_{31} СЗИ от НСД для третьего уровня деградации

Составим граф смены состояний структурной схемы по надежности СЗИ от НСД. При этом можно выделить 6 состояний с частичной работоспособностью, одно состояние определить как полностью работоспособное, а другое — как обобщенное неработоспособное состояние.

Граф смены состояний работоспособности программных модулей СЗИ представлен на рисунке 10.

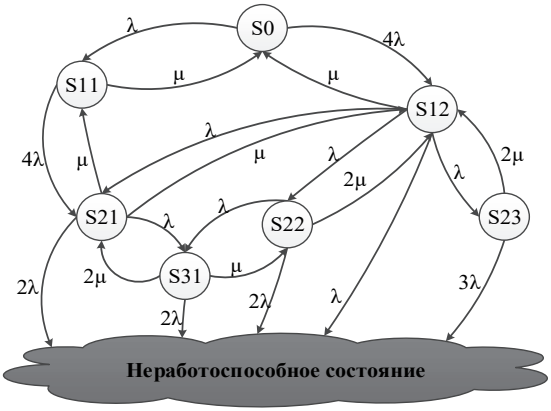


Рис. 10. Граф смены состояний работоспособности СЗИ от НСД

Система алгебраических уравнений имеет следующий вид:

$$\begin{cases} \overline{T}_0 = \dfrac{1}{5\lambda} + \dfrac{1}{5}\overline{T}_1 + \dfrac{4}{5}\overline{T}_2, \\ \overline{T}_1 = \dfrac{1+\mu\overline{T}_0 + 4\lambda\overline{T}_3}{\mu+4\lambda}, \\ \overline{T}_2 = \dfrac{1+\mu\overline{T}_0 + \lambda\overline{T}_3 + \lambda\overline{T}_4 + \lambda\overline{T}_5}{\mu+4\lambda}, \\ \overline{T}_3 = \dfrac{1+\mu\overline{T}_1 + \mu\overline{T}_2 + \lambda\overline{T}_6}{2\mu+3\lambda}, \\ \overline{T}_4 = \dfrac{1+2\mu\overline{T}_2 + \lambda\overline{T}_6}{2\mu+3\lambda}, \\ \overline{T}_5 = \dfrac{1+2\mu\overline{T}_2}{2\mu+3\lambda}, \\ \overline{T}_6 = \dfrac{1+2\mu\overline{T}_1 + \mu\overline{T}_4}{3\mu+2\lambda}, \end{cases} \quad (9)$$

где $\overline{T_i}$ — среднее время нахождения процесса, формализуемого графом (рисунок 10), в *i*-ом состоянии.

Для решения представленной системы алгебраических уравнений перейдем к безразмерным переменным $a_i = \lambda \overline{T_i}$ и $\rho = \frac{\lambda}{\mu}$. Тогда систему уравнений (9) можно преобразовать к виду:

$$\begin{cases} a_o = \dfrac{(1 + a_1 + 4a_2)}{5}, \\ a_1 = \dfrac{(\rho + a_0 + 4\rho a_3)}{1 + 4\rho}, \\ a_2 = \dfrac{(\rho + a_0 + \rho(a_3 + a_4 + a_5)}{1 + 4\rho}, \\ a_3 = \dfrac{(\rho + a_1 + a_2 + \rho a_6)}{(2 + 3\rho)}, \\ a_4 = \dfrac{(\rho + 2a_2 + \rho a_6)}{(2 + 3\rho)}, \\ a_5 = \dfrac{(\rho + 2a_2)}{(2 + 3\rho)}, \\ a_6 = \dfrac{(\rho + 2a_2 + a_4)}{(3 + 2\rho)}. \end{cases} \tag{10}$$

Совершив эквивалентные преобразования, упростим систему алгебраических уравнений (10) и приведем ее к виду:

$$\begin{cases} a_o = \dfrac{(1 + a_1 + 4a_2)}{5}, \\ a_1 = \dfrac{(\rho + a_0 + 4\rho a_3)}{1 + 4\rho}, \\ a_2 = \dfrac{(2\rho(1 + 2\rho) + (2 + 3\rho)(a_0 + \rho a_3 + \rho a_4)}{(2 + 9\rho + 12\rho^2)}, \\ a_3 = \dfrac{(3\rho(1 + \rho) + (3 + 2\rho)(a_1 + a_2) + \rho a_4)}{(6 + 11\rho + 6\rho^2)}, \\ a_4 = \dfrac{(3\rho(1 + \rho) + 2(3 + 2\rho)a_2 + 2\rho a_3)}{(6 + 10\rho + 6\rho^2)}. \end{cases} \tag{11}$$

Для решения полученной системы алгебраических уравнений (11) относительно состояния S_0 целесообразно использовать метод определителей:

$$P_0 = a_0 = \lambda \overline{T}_0 = \Delta_0 / \Delta, \tag{12}$$

где:

$$\Delta = \begin{vmatrix} 1 & -0.2 & -0.8 & 0 & 0 \\ \frac{-1}{1+4\rho} & 1 & 0 & -\frac{4\rho}{1+4\rho} & 0 \\ \frac{-2-2\rho}{2+9\rho+12\rho^2} & 0 & 1 & \frac{-\rho(2+3\rho)}{2+9\rho+12\rho^2} & \frac{-\rho(2+3\rho)}{2+9\rho+12\rho^2} \\ 0 & \frac{-3-2\rho}{6+11\rho+6\rho^2} & \frac{-3-2\rho}{6+11\rho+6\rho^2} & 1 & \frac{-\rho}{6+11\rho+6\rho^2} \\ 0 & 0 & \frac{-2(3+2\rho)}{6+10\rho+6\rho^2} & \frac{-2\rho}{6+10\rho+6\rho^2} & 1 \end{vmatrix},$$

$$\Delta_0 = \begin{vmatrix} 0.2 & -0.2 & -0.8 & 0 & 0 \\ \frac{\rho}{1+4\rho} & 1 & 0 & -\frac{4\rho}{1+4\rho} & 0 \\ \frac{2\rho(1+2\rho)}{2+9\rho+12\rho^2} & 0 & 1 & \frac{-\rho(2+3\rho)}{2+9\rho+12\rho^2} & \frac{-\rho(2+3\rho)}{2+9\rho+12\rho^2} \\ \frac{3\rho(1+\rho)}{6+11\rho+6\rho^2} & \frac{-3-2\rho}{6+11\rho+6\rho^2} & \frac{-3-2\rho}{6+11\rho+6\rho^2} & 1 & \frac{-\rho}{6+11\rho+6\rho^2} \\ \frac{3\rho(1+\rho)}{6+10\rho+6\rho^2} & 0 & \frac{-2(3+2\rho)}{6+10\rho+6\rho^2} & \frac{-2\rho}{6+10\rho+6\rho^2} & 1 \end{vmatrix}.$$

Выражение (12) является результирующим аналитическим выражением оценивания среднего времени наработки на отказ СЗИ от НСД в АС.

5. Результаты оценивания надежности систем защиты информации от несанкционированного доступа. При использовании разработанных методик и автоматизированных средств оценивания показателей надежности СЗИ от НСД в АС выбрана распространенная технологическая схема обработки конфиденциальной информации, рекомендованная технической документацией [38, 41]. При этом приняты следующие допущения: АС представляет собой отдельное автоматизированное рабочее

место; объектом защиты является конфиденциальная информация; защита информации реализуется СЗИ от НСД в соответствии с 3 классом защищенности [9, 10]; ЭВМ работает под управлением ОС Windows на ПЭВМ класса Pentium.

Расчеты по оцениванию надежности СЗИ проведены в среде Mathcad 15.

5.1. Результаты оценивания вероятности безотказной работы систем защиты информации от несанкционированного доступа в течение рабочей смены. На рисунках 11-19 представлены значения вероятностей работоспособного состояния СЗИ от НСД с исходным уровнем первичных ошибок $\lambda_i = 10^{-7}, 10^{-6}, 10^{-5}$ и временем работы СЗИ, равным 5, 10 и 20 часов соответственно, рассчитанные с использованием формулы (1) расчета вероятности безотказной работы СЗИ от НСД для следующих возможных размеров исходного кода программы, реализующей СЗИ от НСД [11]:

——— — размер исходного кода программы — 100 Мб;

– – – – – — размер исходного кода программы — 150 Мб;

..... — размер исходного кода программы — 200 Мб;

— · — · — размер исходного кода программы — 250 Мб.

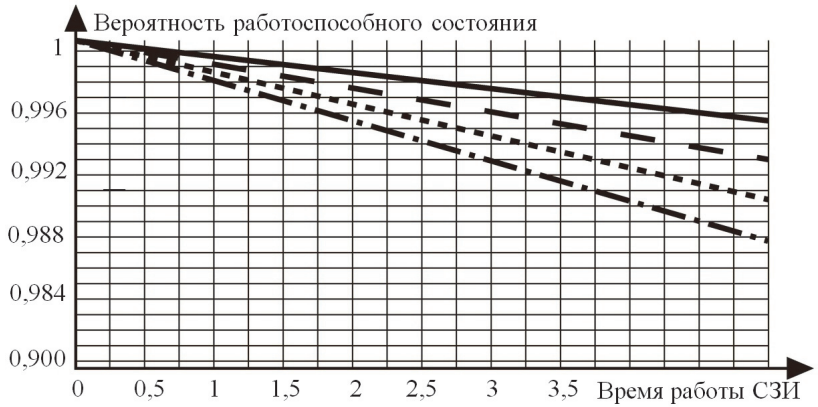


Рис. 11. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-7}$, время работы СЗИ — 5 часов

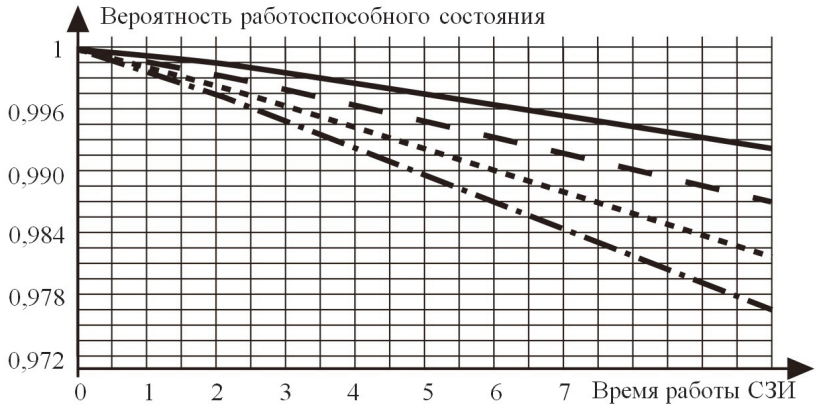


Рис. 12. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-7}$, время работы СЗИ — 10 часов

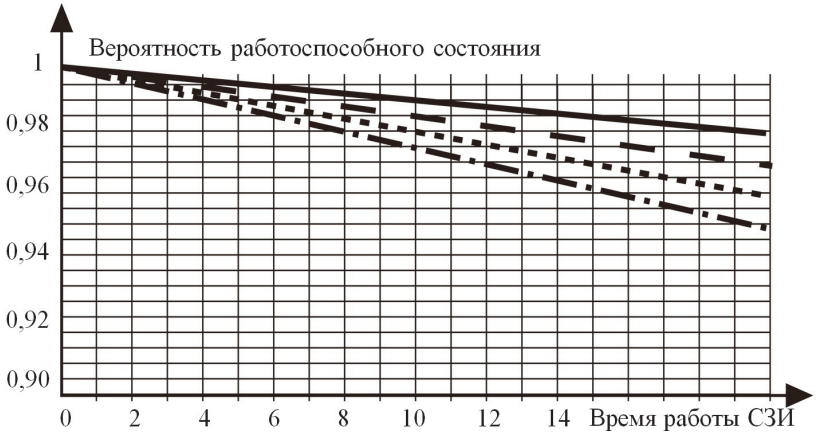


Рис. 13. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-7}$, время работы СЗИ — 20 часов

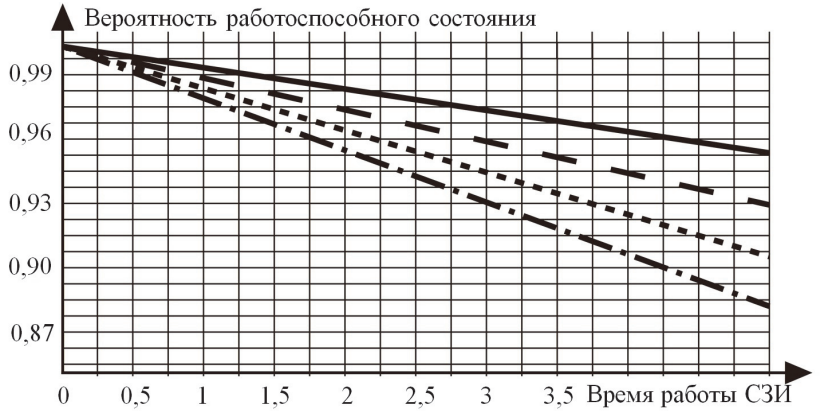


Рис. 14. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-6}$, время работы СЗИ — 5 часов

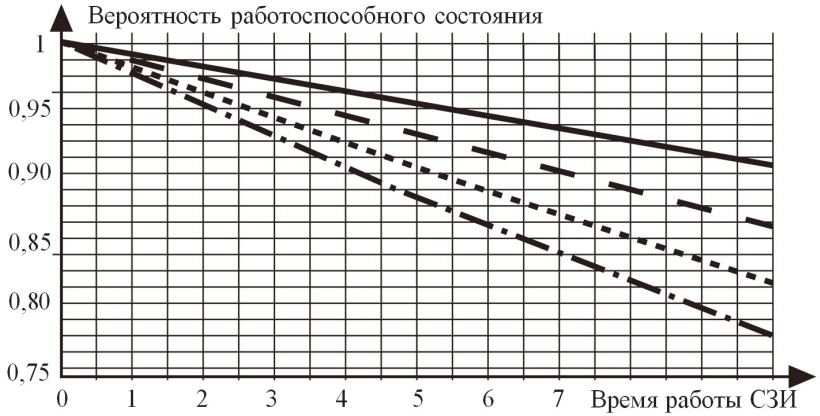


Рис. 15. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-6}$, время работы СЗИ — 10 часов

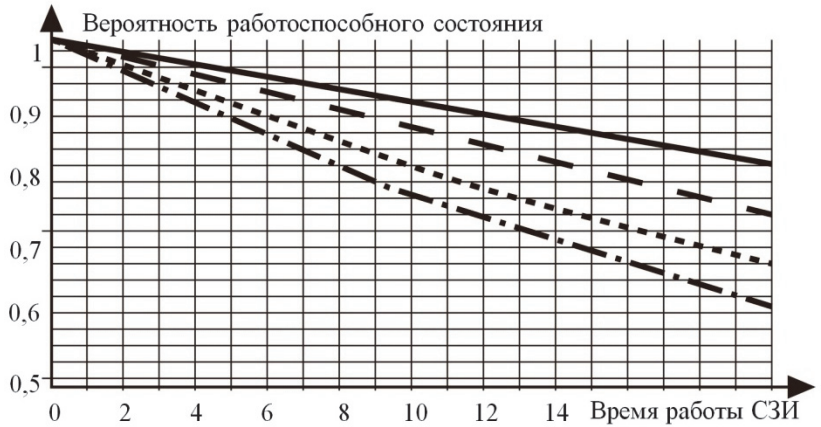


Рис. 16. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-6}$, время работы СЗИ — 20 часов

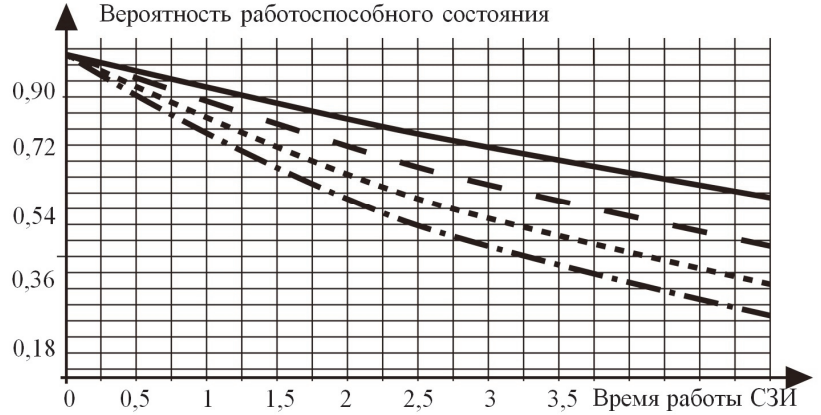


Рис. 17. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-5}$, время работы СЗИ — 5 часов

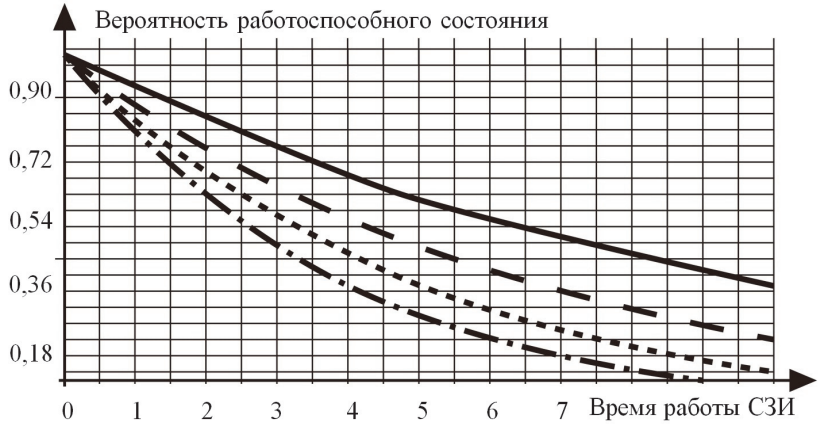


Рис. 18. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-5}$, время работы СЗИ — 10 часов

Рис. 19. Зависимость вероятности работоспособного состояния СЗИ от НСД от времени с исходным уровнем первичных ошибок $\lambda_i = 10^{-5}$, время работы СЗИ — 20 часов

5.2. Результаты оценивания коэффициента готовности систем защиты информации от несанкционированного доступа. На рисунке 20 представлены значения коэффициента готовности СЗИ от НСД, рассчитанные с использованием методики расчета коэффициента готовности для следующих возможных размеров исходного кода программы (из всего диапазона возможных значений), реализующей СЗИ от НСД, с исходным уровнем первичных ошибок $\lambda_i = 10^{-7}$ [11]:

— размер исходного кода программы — 100 Мб:

— размер исходного кода программы — 200 Мб;

— размер исходного кода программы — 300 Мб:

— размер исходного кода программы — 1000 Мб.

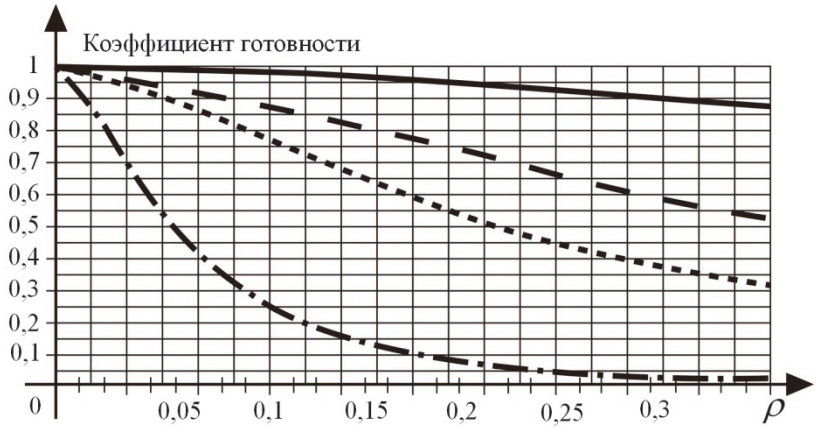


Рис. 20. Зависимость коэффициента готовности СЗИ от обобщенного параметра потоков отказов восстановлений ρ

5.3. Результаты оценивания показателя средней наработки систем защиты информации от несанкционированного доступа на отказ. В таблице 1 представлены результаты расчетов средней наработки СЗИ от НСД на отказ. Расчеты проведены в соответствии с методикой оценивания показателя средней наработки СЗИ от НСД на отказ, разработанной в статье. В таблице приведены средние значения наступления событий, связанных с деградацией структуры СЗИ от НСД, в соответствии с рисунком 10 (шесть состояний СЗИ от НСД с частичной работоспособностью).

Таблица 1. Результаты расчётов средней наработки СЗИ от НСД на отказ

ρ	$\lambda_0 \overline{T}_0$	$\lambda_0 \overline{T}_{11}$	$\lambda_0 \overline{T}_{21}$	$\lambda_0 \overline{T}_{22}$	$\lambda_0 \overline{T}_{23}$	$\lambda_0 \overline{T}_{31}$
0,01	25,47	35,46	25,22	25,21	25,10	25,06
0,05	5,52	5,50	5,47	5,27	5,15	5,12
0,10	3,07	3,03	3,03	2,82	2,69	2,68
0,20	1,87	1,83	1,83	1,64	1,51	1,51
0,30	1,49	1,44	1,44	1,26	1,14	1,14

6. Заключение. В статье впервые предложены методики оценивания надежности СЗИ от НСД АС в защищенном исполнении, учитывающие особенности архитектурного построения СЗИ от НСД и принципы их функционирования при решении задач защиты информации. При решении триединой задачи обеспечения конфиденциальности, целостности и доступности информации применены разные критерии надежности. Разработанные методики апробированы путем проведения расчетов показателей надежности СЗИ от НСД с типовыми структурными схемами построения и типовыми значениями характеристик остаточных первичных ошибок в программном коде и объеме программы.

Применение предложенных методик позволит разработать разделы конструкторской документации («Пояснительная записка к техническому проекту») при создании СЗИ от НСД, содержащие расчеты надежности. Полученные оценки показателей надежности для усредненных характеристик СЗИ могут служить основой для обоснования требований к системе эксплуатации СЗИ от НСД АС подразделениями по обеспечению ИБ, а также использоваться как требования по надежности, содержащиеся в нормативной документации регуляторов совместно с техническими требованиями по защите информации.

Литература

1. ФСТЭК РФ. Руководящий документ. Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка). URL: https://azanpa.ru/fstek-rossii-vypiska-ot15022008-h1468056/6/6.3/ (дата обращения: 10.07.2019).
2. ФСТЭК РФ. Руководящий документ. Методика определения угроз безопасности информации в информационных системах. URL: https://fstec.ru/component/attachments/download/ 812 (дата обращения: 10.07.2019).
3. *Герасименко В.А., Малюк А.А.* Основы защиты информации // М.: МИФИ. 1997. 537 с.
4. *Zhu R., Zeng Y., Xu L., Yi X.* Lightweight Privacy Preservation for Securing Large-Scale Database-Driven Cognitive Radio Networks with Location Verification // Security and Communication Networks. 2019. vol. 10. pp. 1–12.

5. *Nia M.A., Bahrak B., Kargahi M., Fabian B.* Detecting New Generations of Threats Using Attribute-Based Attack Graphs // IET Information Security. 2019. vol. 13. no. 4. pp. 293–303.
6. *Яковина В.С., Федасюк Д.В., Мамроха Н.М.* Аналіз використання аспектно-орієнтованого програмування як засобу підвищення надійності програмного забезпечення // Інженерія програмного забезпечення. 2010. Т. 2. №. 2. С. 24–29.
7. ГОСТ Р ИСО/МЭК 15408-1-2013. Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель. Часть 2. Функциональные компоненты безопасности. Часть 3. Компоненты доверия к безопасности // М.: Стандартинформ. 2014.
8. National vulnerability database. URL: https://nvd.nist.gov (дата обращения: 22.06.2019).
9. ФСТЭК РФ. Руководящий документ. Средства вычислительной техники. Защита от несанкционированного доступа к информации. Показатели защищённости от несанкционированного доступа к информации. URL:https://dokipedia.ru/document/5326599 (дата обращения: 10.07.2019).
10. ФСТЭК РФ. Руководящий документ. Автоматизированные системы. Защита от несанкционированного доступа к информации. Классификация автоматизированных систем и требования по защите информации. URL:https://dokipedia.ru/document/ 5182727 (дата обращения: 10.05.2019).
11. ГОСТ 28195. Оценка качества программных средств. Общие положения. URL:http://www.a-podkidyshev.ru/GOST/28195-89.pdf (дата обращения: 23.04.2019).
12. ГОСТ 28806-89. Качество программных средств. Термины и определения. URL: http://www.kimmeria.nw.ru/standart/glosys/gost_ 28806_90.pdf (дата обращения: 23.04.2019).
13. *Dordevic N.* Software quality standards // Military Technical Courier. 2017. vol. 65. no. 1. pp. 102–124.
14. *Аббасов А.Э., Аббасов Т.Э.* Оценка качества программного обеспечения для современных систем обработки информации // Информационно-технологический вестник. 2015. № 5(3). С. 15–28.
15. *Pandian P.S.* Adopting security checks in business transactions using formal-oriented analysis processes for entrepreneurial students // International Journal of Electrical Engineering & Education. 2019. pp. 101–112.
16. *Arsanjani A.* Empowering the business analyst for on demand computing // IBM Systems Journal. 2005. vol. 44. no. 1. pp. 67–80.
17. *Пак В.О., Абраров Р.Д., Курязов Д.А.* Software testing as integral part of software quality // Молодой учёный. 2016. № 9-5. С. 29–32.
18. *Щенников А.Н.* Качество информационных систем // ИТНОУ. 2018. № 1(5). С. 53–62.
19. *Ayyub B.M., McCuen R.H.* Probability, Statistics and Reliability for Engineers and Scientists // CRC Press. 2016. 656 p.
20. *Тимашев С.А., Похабов Ю.П.* Проблемы комплексного анализа и оценки индивидуальной конструкционной надёжности космических аппаратов (на примере поворотных конструкций) // Екатеринбург: АМБ. 2018. 38 с.
21. *Shubinsky I.B., Rozenberg I.N., Papic L.* Adaptive fault tolerance in real time information systems // Reliability: Theory & Applications. 2017. vol. 12. no. 1(44). pp. 18–25.
22. *Levitin G., Finkelstein M., Huang H.Z.* Scheduling of imperfect inspections for reliability critical systems with shock-driven delayed defects and failures // Reliability Engineering & System Safety. 2019. vol. 189. . pp. 89–98.

23. *Paredes R., Dueñas-Osorio L., Meel K.S., Vardi M.Y.* Principled network reliability approximation: A counting-based approach // Reliability Engineering & System Safety. 2019. vol. 191. pp. 93–110.
24. *Jones C.* Applied software measurement: Assuring // Productivity and Quality. 1997.
25. *Kit E.* Software Testing in the Real World: Improving the Process // Addison-Wesley. 1996.
26. *Гнеденко Б.В., Беляев Ю.К., Соловьёв А.Д.* Математические методы в теории надёжности // М.: КД Либроком. 2019. 584 с.
27. *Казарин О.В., Шубинский И.Б.* Надёжность и безопасность программного обеспечения: учеб. пособие для бакалавриата и магистратуры // М.: МГУ им. М.В. Ломоносова. 2018. 342 с.
28. *Londeix B.* Cost estimation for software development // Addison-Wesley Longman Publishing Co. 1987.
29. Об утверждении Доктрины информационной безопасности Российской Федерации: указ Президента РФ от 05.12.2016 № 646. URL: http://www.consultant.ru/document/cons_doc_LAW_208191/ (дата обращения: 07.06.2019).
30. *Малафеев С.И., Копейкин А.И.* Надёжность технических систем: примеры и задачи // СПб.: Лань. 2016. 320 с.
31. *Зуб А.Т.* Принятие управленческих решений: учебник и практикум. 2-е изд. испр. и доп. // М.: Юрайт. 2018. 332 с.
32. *Гулов В.П. и др.* Методика оценки надёжности системы защиты информации от несанкционированного доступа медицинской информационной системы // Прикладные информационные аспекты медицины. № 1. 2018. С. 202–209.
33. *Скрыпников А.В. и др.* Нормирование требований к характеристикам программных систем защиты информации // Вестник Воронежского государственного университета инженерных технологий. 2018. Т. 80. № 4. С. 96–110.
34. *Филяк П.Ю., Данилова Ю.Н., Гришина Н.В., Мухаммед Н.А.* Обеспечение безопасности в сети интернет на основе сертифицированных решений для обнаружения и предотвращения вторжений/атак // Информация и безопасность. 2018. Т. 21. № 4. С. 510–515.
35. *Оленева Н.Р., Семяшкина Д.С.* Российские и зарубежные разработки в области информационной безопасности // Информация и безопасность. 2018. Т. 21 № 3. С. 380–383.
36. *Samaan N.A. et al.*. Dynamic Contingency Analysis Tool — Phase 1, PNNL-24843, Pacific Northwest National Laboratory, Richland, WA, 2015. URL: http://www.pnnl.gov/main/ publications/external/technical_reports/PNNL-24843.pdf (дата обращения: 28.05.2019).
37. *Дровникова И.Г., Етеппев А.С., Рогозин Е.А.* Основные виды уязвимостей и взаимосвязь компонентов безопасности при обосновании показателей надёжности системы защиты информации от несанкционированного доступа в автоматизированных системах // Приборы и системы. Управление, контроль, диагностика. 2019. № 3. С. 59–64.
38. *Черкесов Г.Н., Воропай Н.И., Сухарев М.Г., Чельцов М.Б.* Надёжность систем энергетики // Новосибирск: Наука. 1999. 434 с.
39. *Дровникова И.Г., Етеппев А.С., Рогозин Е.А.* Формирование критериев работоспособности и отказов системы защиты информации от несанкционированного доступа автоматизированной системы // Приборы и системы. Управление, контроль, диагностика. М.: Научтехлитиздат. 2019. № 5. С. 18–24.
40. *Рогозин Е.А. и др.* Методы и средства оценки защищённости автоматизированных систем органов внутренних дел: монография // Воронежский институт МВД России. 2017. 88 с.

41. *Conto J.* MPjobs — a tool to run PSSe scripts in parallel // ERCOT. 2015.
42. *Змеев А.А. и др.* Методы и средства эволюционного и структурного моделирования при обосновании требований к программным системам защиты информации // Воронежский институт МВД России. 2015. 91 с.

Бокова Оксана Игоревна — д-р техн. наук, профессор, заместитель начальника, институт по научной работе, Воронежский институт Министерства внутренних дел России. Область научных интересов: системы защиты информации, оптимальное управление безопасностью территориальных сегментов информационно-телекоммуникационных систем, управление в социально-экономических системах, математическое моделирование информационных процессов в условиях информационного конфликта. Число научных публикаций — 220. o.i.bokova@gmail.com; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7(473) 200-50-03; факс: +7(473) 200-55-00.

Дровникова Ирина Григорьевна — д-р техн. наук, доцент, профессор, кафедра автоматизированных информационных систем органов внутренних дел, Воронежский институт Министерства внутренних дел России. Область научных интересов: проектирование систем защиты информации от несанкционированного доступа в автоматизированных системах, эволюционное моделирование, теория вероятности, прикладная информатика, управление в социально-экономических системах. Число научных публикаций — 215. idrovnikova@mail.ru; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7(473) 247-67-07; факс: +7(473) 200-55-00.

Етеннев Андрей Сергеевич — адъюнкт, кафедра автоматизированных информационных систем органов внутренних дел, Воронежский институт Министерства внутренних дел России. Область научных интересов: надёжность функционирования систем защиты информации от несанкционированного доступа в автоматизированных системах, прикладная информатика. Число научных публикаций — 4. electronag@gmail.com; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7(473) 247-67-07; факс: +7(473) 200-55-00.

Рогозин Евгений Алексеевич — д-р техн. наук, профессор, профессор, кафедра автоматизированных информационных систем органов внутренних дел, Воронежский институт Министерства внутренних дел России. Область научных интересов: защита информации от несанкционированного доступа в автоматизированных системах, проектирование и управление процессами защиты информации на основе количественной оценки систем защиты информации, прикладная информатика. Число научных публикаций — 250. evgenirogozin@yandex.ru; пр. Патриотов, 53, 394065, Воронеж, Российская Федерация; р.т.: +7(473) 247-67-07; факс: +7(473) 200-55-00.

Хвостов Виктор Анатольевич — канд. техн. наук, доцент, кафедра информационной безопасности, Воронежский государственный университет инженерных технологий России. Область научных интересов: надёжность систем защиты информации, применение методов системного анализа в защите информации, технология разработки программных систем защиты информации. Число научных публикаций. Число научных публикаций — 45. hvahva1@mail.ru; пр. Революции, 19, 394036, Воронеж, Российская Федерация; р.т.: +7(473) 255-42-67; факс: +7(473) 255-42-67.

DOI 10.15622/sp.2019.18.6.1301-1332

O.I. BOKOVA, I.G. DROVNIKOVA, A.S. ETEPNEV, E.A. ROGOZIN,
V.A. KHVOSTOV

METHODS OF ESTIMATING RELIABILITY OF INFORMATION SECURITY SYSTEMS WHICH PROTECT FROM UNAUTHORIZED ACCESS IN AUTOMATED SYSTEMS

Bokova O.I., Drovnikova I.G., Etepnev A.S., Rogozin E.A., Khvostov V.A. **Methods of Estimating Reliability of Information Security Systems which Protect from Unauthorized Access in Automated Systems.**

Abstract. Modern methods of protecting information from unauthorized access in automated systems are based on the use of specialized information security systems from unauthorized access. Security systems are necessarily included in the form of additional software systems in the software as in a secure execution. Information security systems from unauthorized access can be developed not only in a process of automated systems design, but also complement the system-wide software of functioning systems. The use of the information security systems from unauthorized access can reduce a overall reliability of the automated systems, if they contain errors that are not detected during debugging. The reliability of the information security systems affects effectiveness of information security (confidentiality, integrity and availability). Guidelines of the Federal Service for Technical and Export Control (FSTEC) of Russia are a methodological basis for the formation of the information security systems' image both in the process of development and in the process of modernization of the automated systems. The guidance documents of FSTEC of Russia do not contain methodological approaches to assessing the reliability of these program systems. In this regard, the actual design of techniques of estimating reliability of the information security systems from unauthorized access in automated systems in a secure execution. The structural complexity of the information security systems from unauthorized access and large number of functions performed necessitates the use of three reliability indicators that characterize the system in solving problems of confidentiality, integrity and availability of information. To develop the technique, the known methods of evaluating the reliability of complex systems are used, which do not allow their decomposition into serial and parallel connection. The developed methods were tested in assessing the reliability of the information security systems from unauthorized access with typical indicators of initial characteristics. The results of calculations and prospects of using the developed methods are presented in the paper.

Keywords: Information Security System, Unauthorized Access, Automated System, Reliability, Refusal, Information Confidentiality, Information Integrity, Information Availability.

Bokova Oksana Igorevna — Ph.D., Dr.Sci., Professor, Deputy Head, Institute of Scientific Work, Voronezh Institute of the Ministry of Interior of Russia. Research interests: information systems security, optimal security management of territorial segments of information and telecommunication systems, management in socio-economic systems, mathematical modeling of information processes in the context of information conflict. The number of publications — 220. o.i.bokova@gmail.com; 53, pr. Patriotov, 394065, Voronezh, Russian Federation; office phone: +7(473) 200-50-03; fax: +7(473) 200-55-00.

Drovnikova Irina Grigorevna — Ph.D., Dr.Sci., Associate Professor, Professor, Department of Automated Information Systems in Interior Affairs, Voronezh Institute of the Ministry of Interior of Russia. Research interests: design of information security systems against unauthorized access in automated systems, evolutionary modeling, probability theory, applied

informatics, social-and-economic system management. The number of publications — 215. idrovnikova@mail.ru; 53, pr. Patriotov, 394065, Voronezh, Russian Federation; office phone: +7(473) 247-67-07; fax: +7(473) 200-55-00.

Etepnev Andrei Sergeevich — Ph.D. student, Department of Automated Information Systems in Interior Affairs, Voronezh Institute of the Ministry of Interior of Russia. Research interests: reliability of information security systems against unauthorized access in automated systems, applied informatics. The number of publications — 4. electronag@gmail.com; 53, pr. Patriotov, 394065, Voronezh, Russian Federation; office phone: +7(473) 247-67-07; fax: +7(473) 200-55-00.

Rogozin Evgeniy Alekseevich — Ph.D., Dr.Sci., Professor, Professor, Department of Automated Information Systems in Interior Affairs, Voronezh Institute of the Ministry of Interior of Russia. Research interests: information security from unauthorized access in automated systems, design and management of information security processes based on the quantitative assessment of information security systems, applied informatics. The number of publications — 250. evgenirogozin@yandex.ru; 53, 394065, Russian Federation; office phone: +7(473) 247-67-07; fax: +7(473) 200-55-00.

Khvostov Victor Anatolevich — Ph.D., Associate Professor, Department of Information Security, Voronezh State University of Engineering Technologies. Research interests: reliability of information security systems, application of system analysis methods in information security, technology for developing software systems for information security. The number of publications — 45. hvahva1@mail.ru; 19, pr. Revolucii, 394036, Voronezh, Russian Federation; office phone: +7(473) 255-42-67; fax: +7(473) 255-42-67.

References

1. FSTEK RF. Rukovodyashchij dokument. Bazovaya model' ugroz bezopasnosti personal'nyh dannyh pri ih obrabotke v informacionnyh sistemah per-sonal'nyh dannyh (vypiska) [FSTEC RF. Guidance document. The basic model of threats to the security of personal data when they are processed in personal data information systems (extract)]. Available at: https:azanpa.ru.fstek-rossii-vypiska-ot15022008-h1468056-6-6.3 (accessed: 10.07.2019). (In Russ.).
2. FSTEK RF. Rukovodyashchij dokument. Metodika opredeleniya ugroz bez-opasnosti informacii v informacionnyh sistemah [FSTEC RF. [Guidance document. Methodology for identifying information security threats in information systems]. Available at: fstec.ru-component-attachments-download-812 (accessed: 10.07.2019). (In Russ.).
3. Gerasimenko V.A., Maljuk A.A. *Osnovy zashhity informacii* [Basis of information security]. Moscow: MIPI. 1997. 537 p. (In Russ.).
4. Zhu R., Zeng Y., Xu L., Yi X. Lightweight Privacy Preservation for Securing Large-Scale Database-Driven Cognitive Radio Networks with Location Verification. *Security and Communication Networks*. 2019. vol. 10. pp. 1–12.
5. Nia M.A., Bahrak B., Kargahi M., Fabian B. Detecting New Generations of Threats Using Attribute-Based Attack Graphs. *IET Information Security*. 2019. vol. 13. no. 4. pp. 293–303.
6. Jakovina V.S., Fedasjuk D.V., Mamroha N.M. [Analysis of the use of aspect-oriented programming as a means of improving software reliability]. *Inzhenerija programnogo zabezpechennja — Software Engineering*. 2010. Issue 2. vol. 2. pp. 24–29. (In Ukr.).
7. GOST R ISO/IEC 15408-1-2013. [Information technology. Methods and means of security. Criteria for assessing the security of information technology. Part 1. Introduction and general model. Part 2. Functional safety components. Part 3. Components of security confidence]. M.: Standardinform. 2014. (In Russ.).

8. National vulnerability database. Available at: https:nvd.nist.gov (accessed: 22.06.2019).
9. FSTEK RF. Rukovodyashchij dokument. Sredstva vychislitel'noj tekhniki. Zashchita ot nesankcionirovannogo dostupa k informacii. Pokazateli zashchishchyonnosti ot nesankcionirovannogo dostupa k informacii [FSTEC RF. Guidance document. Computing facilities. Protection against unauthorized access to information. Indicators of security against unauthorized access to information]. Available at: https:dokipedia.ru.document.5326599 (accessed: 10.07.2019). (In Russ.).
10. FSTEK RF. Rukovodyashchij dokument. Avtomatizirovannye sistemy. Zashchita ot nesankcionirovannogo dostupa k informacii. Klassifikaciya avtomatizirovannyh sistem i trebovaniya po zashchite informacii [FSTEC RF. Guidance document. Automated systems. Protection against unauthorized access to information. Classification of automated systems and information security requirements]. Available at: https:dokipedia.ru.document. 5182727. (accessed: 10.05.2019). (In Russ.).
11. GOST 28195. Ocenka kachestva programmnyh sredstv. Obshchie polozheniya [GOST 28195. Software quality assessment. General provisions]. Available at: http:www.a-podkidyshev.ru.GOST.28195-89.pdf (accessed: 23.04.2019). (In Russ.).
12. GOST 28806-89. Kachestvo programmnyh sredstv. Terminy i opredeleniya. [GOST 28806-89. Quality of software. Terms and definitions]. Available at: http:www.kimmeria.nw.ru.standart.glosys.gost_28806_90.pdf. (accessed: 23.04.2019). (In Russ.).
13. Dordevic N. Software quality standards. *Military Technical Courier*. 2017. vol. 65. no. 1. pp. 102–124.
14. Abbasov A.E., Abbasov T.E. [Quality evaluation software for modern information processing systems]. *Informacionno-tekhnologicheskij vestnik – Information Technology Bulletin.* 2015. vol. 5(3). pp. 15–28. (In Russ.).
15. Pandian P.S. Adopting security checks in business transactions using formal-oriented analysis processes for entrepreneurial students. *International Journal of Electrical Engineering & Education.* 2019. pp. 101–112.
16. Arsanjani A. Empowering the business analyst for on demand computing. *IBM Systems Journal.* 2005. vol. 44. no. 1. pp. 67–80.
17. Pak V.O., Abrarov R.D., Kuryazov D.A. [Software testing as integral part of software quality].*Young scientist*. 2016. vol. 9.5. pp. 29–32. (In Russ).
18. Shchennikov A.N. [Quality of information systems]. *ITNOU – ITNOU*. 2018. vol. 1(5). pp. 53–62. (In Russ.).
19. Ayyub B.M., McCuen R.H. Probability, Statistics and Reliability for Engineers and Scientists. CRC Press. 2016. 656 p.
20. Timashev S.A., Pohabov Ju.P. *Problemy kompleksnogo analiza i ocenki individual'noj konstrukcionnoj nadjozhnosti kosmicheskih apparatov* (*na primere povorotnyh konstrukcij*) [Problems of complex analysis and evaluation of individual structural reliability of spacecraft (on the example of rotary structures)]. Yekaterinburg: AMB. 2018. 38 p. (In Russ.).
21. Shubinsky I.B., Rozenberg I.N., Papic L. Adaptive fault tolerance in real time information systems. *Reliability: Theory & Applications.* 2017. vol. 12. no. 1(44). pp. 18–25.
22. Levitin G., Finkelstein M., Huang H.Z. Scheduling of imperfect inspections for reliability critical systems with shock-driven delayed defects and failures. *Reliability Engineering & System Safety.* 2019. vol. 189. . pp. 89–98.
23. Paredes R., Dueñas-Osorio L., Meel K.S., Vardi M.Y. Principled network reliability approximation: A counting-based approach. *Reliability Engineering & System Safety*. 2019. vol. 191. pp. 93–110.
24. Jones C. Applied software measurement: Assuring. Productivity and Quality. 1997.

25. Kit E. Software Testing in the Real World: Improving the Process. Addison-Wesley. 1996.
26. Gnedenko B.V., Belyaev Yu.K., Solov'yov A.D. *Matematicheskie metody v teorii nadyozhnosti* [Mathematical methods in the theory of reliability]. Moscow: KD Librokom. 2019. 584 p. (In Russ.).
27. Kazarin O.V., Shubinskij I.B. *Nadyozhnost' i bezopasnost' programmnogo obespecheniya: ucheb. posobie dlya bakalavriata i magistratury* [Software reliability and security: studies' manual for bachelor's and master's degrees]. M.: MGU im. M.V. Lomonosova. 2018. 342 p. (In Russ.).
28. Londeix B. Cost estimation for software development. Addison-Wesley Longman Publishing Co. 1987.
29. Ob utverzhdenii Doktriny informacionnoj bezopasnosti Rossijskoj Federacii: ukaz Prezidenta RF ot 05.12.2016 № 646. [On of the Information Security Doctrine of the Russian Federation: Decree of the President of the Russian Federation of December 5. 2016. №. 646]. Available at: http://www.consultant.ru/document/cons_doc_LAW_208191/ (accessed: 07.06.2019). (In Russ.).
30. Malafeev S.I., Kopejkin A.I. *Nadyozhnost' tekhnicheskih sistem: primery i zadachi* [Reliability of technical systems: examples and tasks]. SPb.: Lan'. 2016. 320 p. (In Russ.).
31. Zub A.T. *Prinyatie upravlencheskih reshenij*: uchebnik i praktikum. 2-e izd. ispr. i dop. [Management decision-making: textbook and workshop]. M.: Yurajt. 2018. 332 p. (In Russ.).
32. Gulov V.P. et al. [Methods of assessing the reliability of the information protection system from unauthorized access to medical information system]. *Prikladnye informacionnye aspekty mediciny —Applied information aspects of medicine*. vol. 1. 2018. pp. 202–209. (In Russ.).
33. Skrypnikov A.V. et al. [Regulation of requirements to characteristics of information security software systems]. *Vestnik Voronezhskogo gosudarstvennogo universiteta inzhenernyh tekhnologij – Proceedings of the Voronezh State University of Engineering Technologies.* 2018. Issue 80. vol. 4. pp. 96–110. (In Russ.).
34. Filyak P.Yu., Danilova Yu.N., Grishina N.V., Muhammed N.A. [Internet security based on certified intrusion/attack detection and prevention solutions]. *Informaciya i bezopasnost' – Information and security*. 2018. Issue 21. vol. 4. pp. 510–515. (In Russ.).
35. Oleneva N.R., Semyashkina D.S. [Russian and foreign developments in the field of information security]. *Informaciya i bezopasnost' – Information and security*. 2018. Issue 21. vol. 3. pp. 380–383. (In Russ.).
36. Samaan N.A. et al. Dynamic Contingency Analysis Tool — Phase 1, PNNL-24843, Pacific Northwest National Laboratory, Richland, WA, 2015. Available at: http://www.pnnl.gov/main/ publications/external/technical_reports/PNNL-24843.pdf (accessed: 28.05.2019).
37. Drovnikova I.G., Etepnev A.S., Rogozin E.A. [The main types of vulnerabilities and the relationship of security components in justifying the reliability of the information protection system from unauthorized access in automated systems]. *Pribory i sistemy. Upravlenie, kontrol', diagnostika – Instruments and systems. Monitoring, control and diagnostics.* 2019. vol. 3. pp. 59–64. (In Russ.).
38. Cherkesov G.N., Voropaj N.I., Suharev M.G., Chel'cov M.B. *Nadyozhnost' sistem energetiki* [Reliability of energy systems]. Novosibirsk:Nauka. 1999. 434 p. (In Russ.).
39. Drovnikova I.G., Etepnev A.S., Rogozin E.A. [Formation of performance criteria and failures of the system for protecting information from unauthorized access of the automated system]. *Pribory i sistemy. Upravlenie, kontrol', diagnostika – Instruments and systems. Monitoring, control and diagnostics.* Moscow: Nauchtechlitizdat. 2019. vol. 5. pp. 18–24. (In Russ.).

40. Rogozin E.A. et al. *Metody i sredstva ocenki zashchishchyonnosti avtomatizirovannyh sistem organov vnutrennih del: monografiya* [Methods and means of assessing the security of automated systems of internal Affairs bodies: monograph]. Voronezhskij institut MVD Rossii. 2017. 88 p. (In Russ.).
41. Conto J. MPjobs — a tool to run PSSe scripts in parallel. ERCOT. 2015.
42. Zmeev A.A. et al. *Metody i sredstva evolyucionnogo i strukturnogo modelirovaniya pri obosnovanii trebovanij k programmnym sistemam zashchity informacii* [Methods and means of evolutionary and structural modeling in justifying the requirements for software systems to protect information]. Voronezhskij institut MVD Rossii. 2015. 91 p. (In Russ.).

DOI 10.15622/sp.2019.18.6.1333-1356

G.R. TSOCHEV, R.D. YOSHINOV, O.P. ILIEV

KEY PROBLEMS OF THE CRITICAL INFORMATION INFRASTRUCTURE THROUGH SCADA SYSTEMS RESEARCH

Tsochev G.R., Yoshinov R.D., Iliev O.P. **Key Problems of the Critical Information Infrastructure through SCADA Systems Research.**

Abstract. In the current age, cyber security is an essential element of any information system. A key aspect is in the critical information infrastructure, where information security has become a top priority for information and network security experts. The interoperability of an ICT infrastructure with other components of it is an important aspect of its life cycle. Because Supervisory Control and Data Acquisition (SCADA) systems form part of the critical infrastructure, their cyber protection is particularly important in strategically important industrial and infrastructure sites — power plants, refineries, oil pipelines, treatment plants, manufacturing facilities, communications and transportation infrastructures. Along with the advancement of technology, the increasing number of Scada devices available online, the vulnerability of the sectors controlled by them has also increased. In the world of Internet of things (everything), the end devices cause a new wave of possible vulnerabilities in SCADA. They become the new places for attacks and breaches trough wich the system may be accessed or even compromised. There are a number of critical infrastructures in the Community whose disruption or destruction would have significant cross-border implications for more than one sector as a result of the interdependence of interconnected infrastructures. Such European critical infrastructures have been established and launched under a common procedure developed by the European Commission, with security requirements assessed according to a common minimum approach.

The present article exposes and examines the critical infrastructures of the European Union and Bulgaria. Trough presenting the structure of a Scada system the vulnerabilities and the various possibilities of attacking it were analised. As an example, a specific case based on trees has been considered, and the obtained results were summarized and visualized. The consequences were analized and respecctive conclusion was done.

Keywords: Critical Infrastructure, SCADA, Attack Tree, Cyber Security, Network and Information Security.

1. Introduction. The last decade was filled with great dynamics in the field of information technology. Increasingly, terms such as cyber attacks, espionage, stolen personal information from hacked profiles in social networks or elsewhere have become more common.

In times of high technological growth and multiple directions of development, everyone is looking for their own way to succeed. Both the innovativeness and knowledge of the factors that lead to the successful implementation of the steps for implementing the final decision are decisive. For the most part, the information is contained in the shared network space. Its value is greater than ever, and may be increased or decreased by some of the characteristics it possesses [1]:

– availability allows authorized users — physical persons or computer systems – to access information without hindrance;

– information accuracy exists when it is free of errors and inaccuracies, and has the value that end users expect to have. If the information has been intentionally or inadvertently modified, it is no longer reliable and accurate;

– of major importance is the authenticity of information, and its quality and condition to be original. It can be considered authentic when it is on the same level at which it was created, stored or transmitted;

– information is confidential when it is protected from the disclosure or exposure of unauthorized users or systems.

For most organizations, the security of information and the systems that process, transmit, and store it is crucial. In many more cases, information is also business.

Creating an information security program that adheres to the principles of security as a business factor is the first step in an association's efforts in building an effective security strategy. Continuous risk assessment is necessary, as well as evaluating and implementing sound policies, standards, and controls in order to reduce them.

Computer systems and networks are one of the highest technological products of humanity. Apart from all the advantages they offer, they have a number of disadvantages. Security problems in the form of malware, loss of privacy, reception of unwanted advertising, commercial or spam messages affect almost every computer user. One of the many definitions of security is that it represents the ability of a system to withstand external or internal destabilizing factors that can lead to its undesirable state or behavior.

The purpose of information security is to protect the valuable resources (information, computer hardware and software) of an organization. By selecting and implementing appropriate safeguards, security assists the mission of the organization by protecting its physical and financial resources, reputation, legal position, employees and other tangible and intangible assets.

The endless development of technology has undoubtedly improved the efficiency of industrial processes, and the delivery of health services. Jobs that would take more time and work force have already been taken up by devices that are connected to each other to achieve the same goal and even make it better. Healthcare people now rely on information systems, including mobile devices (implantable or external), to facilitate patient monitoring, so delivery of services has become more automated [7]. These connected devices, called the Internet of Things, have grown significantly when it comes to meaning, number and value over the years.

Despite the enormous advantages of these systems, it is likely that these devices can be used by malicious cyber attackers as a means of endangering the system itself, and human existence in general. A highly motivated person or group of people for cyber-terrorism will endanger human lives and property by disrupting services provided by critical infrastructures, e.g. healthcare IT infrastructure. By endangering human life, the goals might be panic, personal injury, health risk or death. The satisfaction for the people

comes from the use of connected computer infrastructure to cause some physical impact on the environment.

Although various cyber attacks are already targeting IoT health services, pumps that overflow drugs are known to be very reliable and useful for doctors and patients as they provide safe and accurate administration of drugs and fluids. It is suggested that the attacker's motive should change from financial gain to altering the overflow rate of the overflow pump, changing the configuration of the device, sending malicious commands, or simply interfering with the communication of the device. The attacker may also decide to gain physical or remote access to the electronic health records (HER), and to modify the patient's files, such as blood type, type of dosage, therapy session, etc. All this can lead to catastrophic events and can also cause panic in society, especially if the event is unexpected — or there are no appropriate control measures in place.

An alternative industry relevant to this comparative study is the one that drives the industrial SCADA system [3]. No rigorous security research has been done when SCADA systems were introduced decades ago. This has led to many of the security problems that they face today.

2. Critical Infrastructure. The term **c**ritical infrastructure refers to an element, system or parts thereof located in Member States which is essential for the maintenance of vital public functions, health, safety, security, economic or social well-being of the population, and whose disruption or destruction would have significant consequences in a given Member State as a result of the inability to retain these functions [4].

European critical infrastructures are the critical infrastructures located in Member States whose disruption or destruction would have significant consequences for two or more Member States.

The Council of the European Union has adopted Directive 2008 / 114 / EC of 8 December 2008 [8] about the establishment and designation of European critical infrastructures, and assessing the need to improve their protection. This Directive mainly covers the energy and transportation sectors. The requirements of Directive 2008/114/EC have been incorporated into the national legislation of Bulgaria by Decree No.18 of the Council of Ministers of 01.02.2011 on the establishment and designation of European Critical Infrastructures in the Republic of Bulgaria, and the measures for their protection [9].

Critical infrastructures are taken into account when protecting the critical infrastructures resulting from human activity, technological threats and natural disasters, but priority is given to terrorist threats. Member States are ultimately responsible for the management of critical infrastructure protection mechanisms within their national borders.

There are a number of critical infrastructures in the Community whose disruption or destruction would have significant cross-border implications for more than one sector as a result of the interdependence of interconnected infrastructures. Such European critical infrastructures have been established and launched under a common procedure developed by the European Commission, with security requirements assessed according to a common minimum approach.

The critical infrastructure of Europe is divided into several sectors:

– Energy;
– Information and Telecommunication Systems;
– Water Resources;
– Food Industry;
– Healthcare;
– Financial Services;
– Law and Order;
– Public Services;
– Transportation Systems;
– Chemical and Nuclear Industries;
– Space Systems and Exploration.

– However, in Bulgaria, 19 sectors have been identified as critical, as follows [9]:

– Energy;
– Transportation;
– Telecommunications, including electronic communication networks, and information and communication infrastructure;
– Healthcare;
– Agriculture and Food;
– Environment;
– Finance and Banking;
– Defense;
– Justice, Public Order and Security;
– Economics;
– Education, Science and Technology;
– Natural Resources;
– Regional Development and Public Works;
– Tourism;
– Government and Social Governance;
– Disaster Protection;
– Cultural Heritage;
– Postal and Courier Services;
– Sports Facilities.

All of these sectors which have been identified as critical to the country are governed by the strategic documents on which the models for guaranteeing their security are based, and developed. The European Council meeting of June 2004 invited the Commission to prepare a common strategy for the protection of its critical infrastructure. On 20 October 2004, the Commission adopted the document "Protecting the Critical Infrastructure in the Fight against Terrorism" [8] with clear proposals for measures needed to improve the prevention, preparedness and response in Europe of terrorist attacks affecting the critical infrastructure. In its conclusions on the "Prevention, Preparedness and Response to Terrorist Attacks" and the "EU Solidarity Program on the Effects of Terrorist Threats and Attacks", adopted in December 2004, the Council endorses the Commission's plan to propose a European Critical Infrastructure Protection Program (EPCIP) [9], and approved the creation of a Critical Infrastructure Warning Information Network (CIWIN) [10].

Taking these two measures defines the framework for the protection of EU infrastructure and accordingly defines the horizontal framework for the protection of EU critical infrastructure, explaining how EPCIP [9,10] (including CIWIN) can be put into effect.

The CIWIN initiative [9,10] is part of EPCIP, focusing in particular on the information sharing process between EU Member States, and addressing the IT system that supports this process.

Essentially, in the course of protecting the critical infrastructure, this means the identification of risk factors in the sector concerned. Critical infrastructure protection encompasses a set of activities designed to ensure the normal functioning, continuity and integrity of critical infrastructures to deter, reduce, mitigate or counteract threats, risks or vulnerabilities. The availability, integrity, and confidentiality model of security, integrity, and confidentiality is fully applicable to the security of critical information infrastructure.

In order to maximize the use of ICT infrastructures, and thus to make full use of the economic and social opportunities provided by the information society, all actors need to have high confidence in these infrastructures. This depends on various factors, the most important of which is to guarantee their high level of security and resilience. Diversity, openness, interoperability, usability, transparency, accountability, verifiability of different components and competition — these are key factors for improving security, and promoting the introduction of security products, processes and services.

The rapid pace of industrialization, as well as the use of artificial intelligence in the management and control of the production, transportation and storage of products and raw materials, necessitates an appropriate class of connectivity technologies.

These technologies have requirements for reliability, operation under extreme conditions, and protection that are many times higher than normal. This is due to the risks associated with the collapse of the systems that control critical infrastructures, such as nuclear power plants, gas transmission network, water supply and sewerage, and electricity transmission networks.

Typically, such networks have a mix of Scada [10] devices and Ethernet transmission. These two technologies must work seamlessly together, and must be well protected against cyber attacks aimed at disruption, theft, or manipulation of traffic.

The security of Scada systems is of particular importance in strategically important industrial and infrastructure sites — power plants, refineries, oil pipelines, treatment plants, manufacturing facilities, communication and transport infrastructures. With the advancement of technology, and the increasing number of Scada devices available online, the vulnerability of the sectors controlled by them has increased.

3. Industrial Control Systems. The industrial management system (ICS) is a generic term that covers several types of management systems, including Supervisory Management and Data Acquisition Systems (Scada) [11] (Figure 1), control systems (DCS), and other control systems such as PLCs, common in the industrial sectors and critical infrastructures. An ICS consists of combinations of components with specific controls (electrical, mechanical, hydraulic, pneumatic) that work together to achieve their purpose (e.g. production, material or energy transportation). The part of the system that mainly deals with the production of industrial products is called technological process. The part that controls the system includes a specification of the desired output or performance [12]. The control may be fully automated, or it may involve a person. Systems can be configured to work within an open cycle, closed cycle, and manual mode. For open-loop control systems, output is controlled by the settings set. In the closed loop control systems, output has such an effect on the input that it supports the desired purpose. In the manual mode, the system is entirely controlled by humans. The part of the system that is mainly concerned with maintaining the specification is called controller. A typical ICS system may contain multiple control cycles, human machine interfaces (HMI) [11], and remote diagnostics and support tools built using multiple network protocols. ICS industrial control processes are commonly used in sectors, such as electricity, dams, oil and natural gas, chemicals, transportation, pharmaceutical, pulp and paper, food and beverage industries (including automotive, aerospace, and durables). ICSs can largely be grouped by function into one or more of these three categories: overview, monitoring and management.

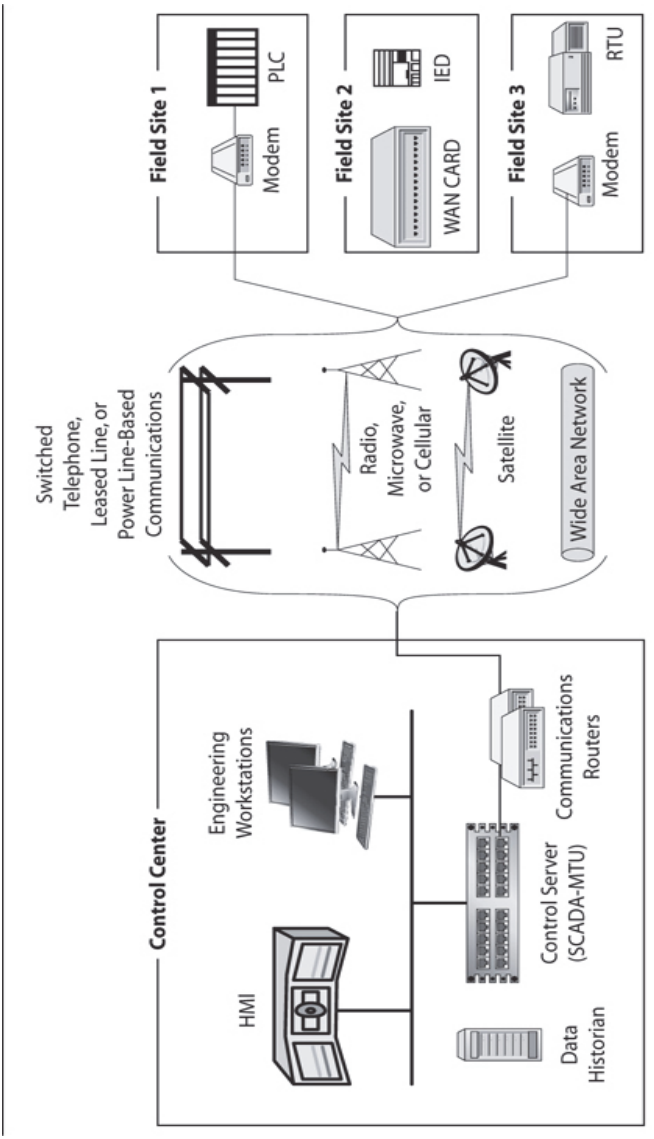


Fig. 1. General view of SCADA system

Труды СПИИРАН. 2019. Том 18 № 6. ISSN 2078-9181 (печ.)
ISSN 2078-9599 (онлайн) www.proceedings.spiiras.nw.ru

The Scada components [13] (Figure 2) system consist of sensors and actuators that are responsible for the collection of physical parametric data from field devices. These signals are typically stored in analogue format and converted through a remote terminal device, programmable logic controllers or an intelligent electronic device (IED). Once the data is converted, it is transmitted via a communication channel to the Scada control unit where the collected data is processed, and the operations are transmitted back to the field devices.

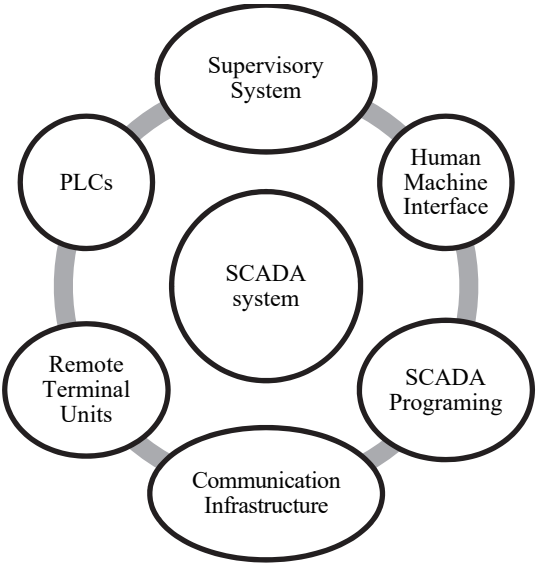


Fig. 2. Components of SCADA architecture

The Scada controller uses the HMI [14], which is responsible for submitting the collected data to operator in readable form. Further expansion of the architecture exists in situations where there is an operator outside the area of the industrial network, trying to control devices from remote locations. Communication between the field devices and the Scada host can be accomplished via dial-up, satellite, serial, radio, telephone, or WLAN. Also, specialized communication protocols such as DNP, DNP3, ICCP, Modbus and ProfiBus etc. have been integrated into the Scada network. As some of these protocols were in existence more than 20 years ago, when security was not a key factor in their construction, the main focus was on their effective operation, and not security.

With up to 4 layers (collection, conversion, communication and control) in the Scada system, it can be seen that each of the layers can be

used as an entry point to attack the system. A physical attack where the on-site devices are corrected by the attacker or the malicious person alters the data that is sent to the HMI. Some of the popular protocols that are still in use do not include authentication and encryption during their operation, and these shortcomings can be used to capture data that is transferred between devices. Insider exploitation errors or lack of proper access control can make HMI vulnerable, this can be used by a malicious attacker to interfere with industrial processes.

4. Security Issues within Scada Architecture.

4.1. HMI Vulnerability. An attacker can exploit the industrial Scada system by exploiting vulnerabilities that exist in the system's HMI [14, 15].

– Preset Name and Password — HMI component hacking can allow a remote attacker to control field devices. Exploitable vulnerabilities include embedded username and password in plain text in Java code used to design Web HMI;

– Validation of Field for Incorrect Input — also, incorrect encoding techniques can allow an attacker to execute requests through the Web HMI input field, leading to SQL injection attacks;

– Incorrect Authentication and Permission — bad authentication techniques are a door in the system. For example, the lack of a two-factor system for certification in critical infrastructure leaves an attacker unable to gain access to the system without much effort.

4.2. Zero Day Exploits. Unidentified vulnerabilities by system software vendors are one of the most common threats exploited by Scada network attackers. Due to the nature of the vulnerability, an attacker can infiltrate the system by performing intelligence, scanning and logging without being detected.

4.3. PLC Vulnerabilities. PLCs, sensors, and actuators serve as aggregators of data in the industrial field. It is possible for the attacker to take control of the PLC by exploiting vulnerabilities, and then directly or indirectly intervening in industrial processes. Attacking PLC firmware vulnerabilities by an attacker can provide direct access to the sensors and actuators of the field.

4.4. Social Engineering. Cyber-attackers can use social engineering methods to break into the web. Checking IoT databases such as shodan.io for device username and passwords is one way to learn the web. Email spear phishing methods can also be considered an entry point into the corporate network if the attacker's intention is to bring malware into the network. Hijacking an insider to physically connect a device (such as USB, disk drive) to distribute malware is also possible.

4.5. Inadequate Physical Security. The field devices themselves are responsible for the industrial process, and inappropriate physical security procedures can allow an unauthorized attacker to gain physical access to the devices. As a result — although Scada monitoring includes an alarm due to the infectious data it has received — it will be quite difficult for the situation to be remedied except by physical means.

4.6. Vulnerabilities of the Scada Protocol. Most of the common protocols used on Scada networks are for operational efficiency, not security. They do not include authentication mechanisms used in traditional IT systems that are used to authenticate the sender or recipient of the data, thereby allowing attackers to compromise the integrity and confidentiality of sensor values.

4.7. Corporate Network Connection. Earlier Scada systems relied on point-to-point networks. To scale the Scada networks to fit current organizational needs, Scada systems have been connected to the corporate network through a secure gateway. Corporate networks operate as computer networks; they are susceptible to attacks such as SQL injection, phishing, spear phishing, and other vulnerability exploits.

5. Tree-based Security Breach Risk Analysis. Attack trees were introduced by Schneier [29] as a way of officially analyzing the security of systems and subsystems based on various attacks. Schneier's work is remarkable, as it is the first time that this information security approach has been implemented. The purpose of the attack is the root of the tree, and the various ways of carrying out the attack are the leaves, with connections through nodes AND and OR.

Figure 3 shows an image of a target tree. The attack tree consists of OR nodes, AND nodes, and Leaf nodes. The topmost node is the root node, which shows the overall purpose of the hacker. The top node is decomposed into several sub-goals that consist of other nodes and leafs. The OR node indicates that the attack can be performed by completing 1 or more sub-targets. For example, the OR target #1 can be achieved either through subgroup #1a, or subgroup #1b. The AND node indicates that the attack can be carried out by completing all the sub-goals. For example, goal AND #2 can be achieved by performing subgroup #2a and subgroup #2b.

Moore et al. [19] describe and illustrate an approach for documenting attacks against software systems that use structured and reusable attack tree information. Analysts can then use the approach to document and identify common attack patterns, and then modify attack trees to improve security development.

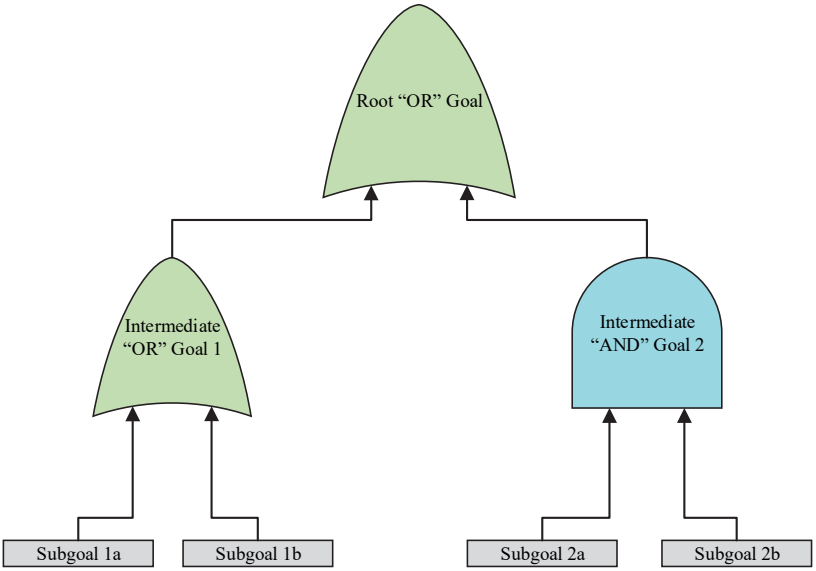


Fig. 3. Example Tree

Recently, attack trees have been applied to the Scada communication system[29]. The authors have identified eleven targets of attackers and related security vulnerabilities in the specification and development of a typical Scada system:

1. Easing the access to the Scada system;
2. Identifying the Modbus device;
3. Disconnecting the master and slave modules, or breaking the Master-Slave rule;
4. Disabling subordinates;
5. Reading data from subordinates;
6. Recording data in subordinates;
7. Programming of subordinates;
8. Compensation of subordinates;
9. Disabling the Main Module;
10. Main Module Data Recording;
11. Modifying the Master Module.

6. Methodology of the Analysis. Figure 4 shows a schematic overview of the methodology for generating attacking trees, and a description of the workflow from the adoption of a system model to the categorization of opportunities for threat level. Secure-Tree Modeling Tool [2] by Amenaza Technologies Ltd. used in developing the attack tree.

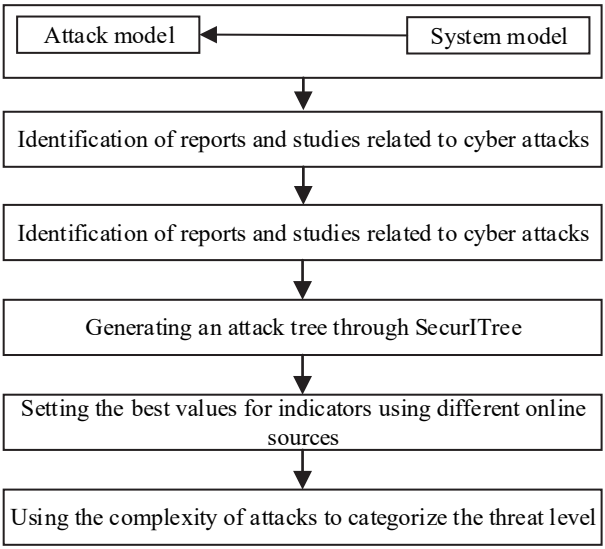


Fig. 4. Methodology of the Analysis

7. Scada System Attack. The purpose of the attack is to terminate fuel operations [20]. Upon successful attack, the attacker can change the Fuel Manager Database (FMD) protection data, change the real-time HMI data, crash the FMD hard drive, submit a fake fuel company report, and violate the FMD communication.

7.1. Indicators. An attacker needs a variety of resources to attack. Four different indicators are considered below:

– accessibility — accessibility is measured by how easily an attacker can access a restricted system;

– sight — sight is the measure of how secretly an attacker can perform an attack without being noticed. Visibility measurement comes from the mystery of the attack from the planning stage to intelligence, until the attack is successfully completed;

– technical ability — technical ability is the measure of how easily an attack can be done. This is related to the level of the attacker, the level of expertise and specialization that are taken into account in determining the technical capabilities of the attacker;

– breakthrough time — breakthrough time is a measure of how long it will take from planning to the final execution of an attack.

The attack tree is built from the end leaves to the nodes until the attacker reaches his target, which is the root node. Table 1 shows how to

select Leaf attack tree nodes and how OR and AND nodes are calculated. For OR nodes, all Leaf nodes are considered in the attack. This means that the number of Leaf nodes added to the OR nodes affects the number of attack scenarios generated.

Table 1. Node range

Indicator name	LOGICAL OR	LOGICAL AND	Range
Accessibility	Minimum Peaks: The tree selects and examines each of the nodes for each attack.	Minimum Peak Value: The value of Leaf nodes with the highest system access restriction is selected to move to the root node. For example, in an attack consisting of a high-access (unrestricted) node and a medium-access (moderate-access) node, the value of a moderate node rises on the tree, being a more complex node.	1-3
Landmark	Minimum Peaks: The tree selects and examines each of the nodes for each attack.	Minimum Peak Value: The value of the leaf node with the most inconspicuous attack is selected to move to the root node. This is because the lower the value of the leaf node, the less noticeable the attack.	1-5
Technical ability	Minimum Peaks: The tree selects and examines each of the nodes for each attack.	Maximum Peak Value: The value of Leaf nodes with the highest technical ability is selected to move to the primary node. This is because the higher the score, the higher the technical ability required to perform the attack. If one of the leaf attack nodes is a combination of very difficult attacks, then the attack must also have high technical ability.	1-5
Breakthrough time	Minimum Peaks: The tree selects and examines each of the nodes for each attack.	Maximum Peak Value: The value of the Leaf node with the most time to perform the attack is selected to move to the root node.	1-5

7.2. Scada System Attack Scenario. Many attack nodes describe elements of an advanced persistent threat (APT). The anatomy of APT attacks describes social engineering, spear phishing, malware, mapping, escalation of privileges, and networking. APTs are designed to be difficult to see in network traffic as they move from one host to another. For example, Trojan, Rootkits, and Backdoors are designed in such a way that they can maintain a low profile in the system. Figure 5 gives an overview to all nodes of the attack.

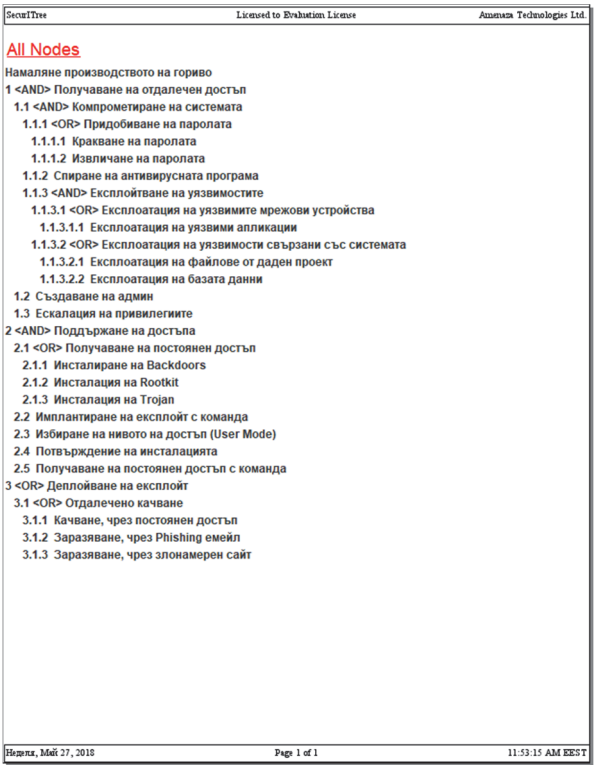


Fig. 5. All nodes of the attack in full view

7.3. Scada System Attack Scenario — analysis using complexity index SecurITree software provides a tool that allows you to identify threat profiles. Attack scenarios that fall into threat level 1 have the highest level of attack complexity (Table 2). The level of complexity of attacks decreases from threat level 1 to threat level 5. While attacks under threat level 1 are the most complex, threat level 5 can lead to an attack against healthcare infrastructure with less complexity and better results. In the Scada attack scenario, it can be seen that only attackers under threat 1 and 3 can carry out the attack.

Table 2. Assigning weight to each of the nodes

№	Attack nodes	Technical ability	Accessibility	Landmark	Attack time
1	Create admin	4	3	4	1
2	Escalation of privileges	4	3	2	1
3	Password cracking	3	3	3	2
4	Password retrieval	4	3	4	4
5	Stopping the antivirus program	4	2	4	2
6	Operation of vulnerable applications	4	3	3	2
7	Operation of files from a project	4	2	3	2
8	Operation of the database	5	2	4	3
9	Installing Backdoors	3	2	2	2
10	Rootkit installation	3	2	2	2
11	Trojan installation	3	2	2	3
12	Implanting an exploit with a command	2	3	3	1
13	Selecting the Access Level (User Mode)	3	3	3	2
14	Selecting the Access Level (User Mode)	2	2	3	1
15	Get permanent access with a command	3	3	4	2
16	Upload, with permanent access	2	3	4	3
17	Contamination by Phishing Email	4	1	2	2
18	Infection through a malicious site	4	1	2	2

An attack causing reduced WPAFB combustion processing utilizes all the characteristics of a sophisticated attack. There are elements of social engineering, APT, insubordination, remote administration, and exploitation of Zero-Day vulnerabilities in the attack being analyzed. This gives the total SI score for this attack to be 5. Which is the highest SI score obtained. An additional illustration is shown in Table 3.

Table 3. Determination of SI coefficient for SCADA system

Types of features	Attack nodes	SI
Social Engineering	Infection by Phishing email	1
Remote Administration	Backdoors installation, Rootkit installation, command implant implantation, installation confirmation	1
Landmark	Antivirus suspension, escalation of privileges, exploitation of vulnerable applications, installation of Rootkit	1
Zero-Day Vulnerabilities	Database vulnerability	1
APT	Backdoors installation, Rootkit installation, admin creation	1
		All = 5

8. Results of the Attack on Scada. The minimum level of threat required to attack an industrial Scada infrastructure is threat level 3. Due to the AND function of the root node, all attacks must be made to achieve an attack from any of the threat levels. Threat Level 3 has 36 scenarios that are the same as the total number of scenarios generated by the attack tree. This means that an attacker with high technical skills, limited access to the system, an increased level of insubordination, and ready to devote months to years to planning and executing the attack, can make a successful attack. The maximum level of threat required to attack an industrial Scada system for generating a physical result is threat level 1. An attacker with a level 1 threat

will also be able to execute all 36 attacks on a Scada system attack, and he does so with unlimited access to the system.

Although a third-level attacker can execute an attack, the result of the analysis indicates that an attacker with a threat level 2 cannot execute the attack. This is because the attacker with threat level 2 has only an average technical level of cyber knowledge, although the indicators of incompleteness and access are high. For this reason, it can be concluded that the attacker's technical ability is one of the most important indicators when analyzing an attacker's capabilities.

There are two types of errors for evaluating IDS, for each possible test value: false positive (FP) and false negative (FN). FP occurs when an event is intended to be intrusive, but is actually normal, whereas FN occurs when a truly intrusive event occurs without being recognized as one. On the other hand, the true positive (TP) measures the proportion of true positives that are correctly identified as such, while the true negative (TN) measures the proportion of negatives that are correctly identified as such.

The performance of each classifier can be quantified using measures of measurement of detection rate (DR) and overall accuracy (OA). DR shows the percentage of true breaks that were successfully detected:

$$DR = \frac{TP}{TP + FN} \times 100\%.$$

OA is calculated as the total number of correctly classified infiltrations divided by the total number of observations:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%.$$

Effective IDS requires a high degree of DR and OA while maintaining low levels of false alarms. Accuracy is critical to developing an effective IDS, since high FP speeds or low DRs will render it practically useless.

The Table 4 below lists data from a standard IDS system that is used to protect the Scada information infrastructure under study.

Table 4. Simulation results for DDoS attack detection

Attack (% from affected endpoints)	FP %	FN %	OA
1	1.2	1.1	60.1
5	1.3	1.2	60.05
10	1.5	1.6	60.13

Table 4 (continued)

15	1.7	1.8	60.7
20	1.9	2	62.03
25	2.1	2.3	61.83
30	2.4	2.5	62.38
35	2.6	2.7	62.58
40	3.1	3	62.68
45	3.2	3.4	62.7
50	3.3	3.7	63.65
55	3.5	3.8	65.08
60	3.7	3.9	66.05

From the graph above (Figure 6), it is visible that the system becomes more accurate with the increase of the relevant percentage. More affected endpoints lead to more accurate results. If the attack is smaller and targeted at a specific endpoint, it will be very difficult to detect it. In the course of this experiment, the gateway capability of the Scada system in/out point of the network was also recorded (Figure 7).

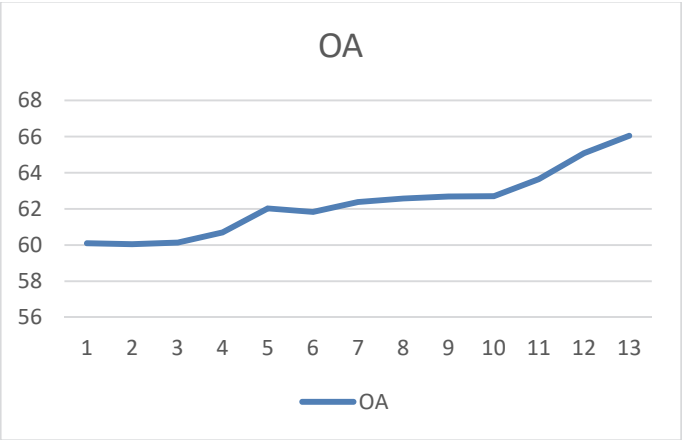


Fig. 6. Overall accuracy

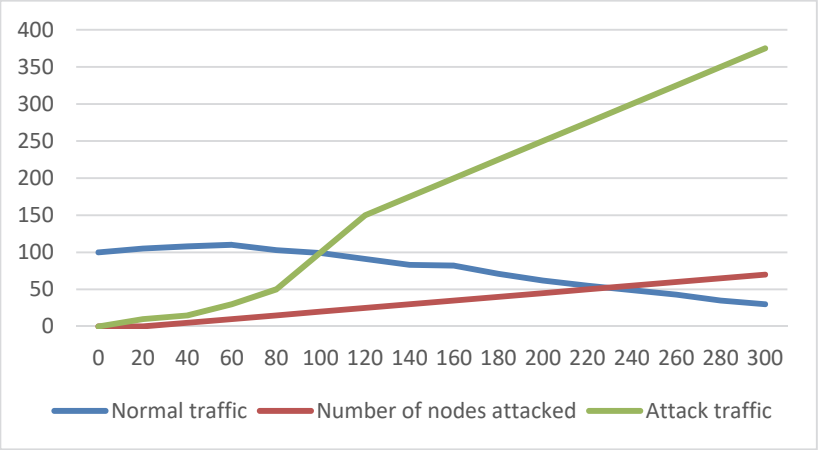


Fig. 7. Network traffic

9. Conclusion. Strategic planning for critical infrastructure protection requires the coordination and coordination of actions taken across time and space. Critical infrastructure protection planning needs to be based on a specific platform. It is created with the help of relevant information, mainly of a regulatory nature — laws, decrees, directives, regulations, instructions and other documents.

The real-time cybersecurity of continuous systems requires an overall view and comprehensive understanding of network security, control theory, and the physical system. Ultimately, all viable technical solutions and research guidelines for the provision of Scada systems must be related to computer security, communications network and control engineering. The idea itself of looking into the problem within the context of performance control has its foundations. There are two types of errors for evaluating by the Intruder Detection System, for each possible test value: false positive and false negative. Effective Intruder Detection System requires a high degree of detection rate and overall accuracy, while maintaining low levels of false alarms. Accuracy is critical to developing an effective Intruder Detection System, since high false positive speeds or low detection rates will render it practically useless. The experiments show that with increasing attack traffic at multiple entry points the inrtuder detection system becomes more accurate. More affected endpoints lead to more accurate results. If the attack is smaller and targeted at a specific endpoint, it is very difficult to detect it. The damage increases when adequate attention to the various devices involved in the detection process is not provided. Each of these devices has a specific structure (software and hardware) and thus requires individual and specific protection.

The attacker's technical ability is one of the most important indicators when analyzing attacker's capabilities. This means that an attacker with high technical skills, even with limited access to the system, but ready to devote months to years in planning and executing the attack, can make a successful attack.

The future work of the team will include a more in-depth study of the cyber-physical connections of the individual elements and the various participants in a critical information infrastructure, as well as the application of active polling in order to reduce the rate of cyber attacks.

References

1. Almalawi A. et al. An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems. Computers & Security. 2014. vol. 46. pp. 94–110.
2. Amenaza Technologies Limited. Available at: https://www.amenaza.com/faq.php (accessed: 25.08.2019).
3. Byres E.J, Franz M., Miller D. The Use of Attack Trees in Assessing Vulnerabilities in SCADA System. Proceedings of the International Infrastructure Survivability Workshop (IISW'04). 2004. pp. 3–10.
4. Dochev D., Pavlov R., Paneva-Marinova D., Pavlova L. Towards Modeling of Digital Ecosystems for Cultural Heritage. *Digital Presentation and Preservation of Cultural and Scientific Heritage*. 2019. vol. 9. pp. 77–88.
5. IEEE Standard Definition, Specification, and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control. Institute of Electrical and Electronics Engineers. 1987.
6. IoT And SCADA: Is One Going To Replace The Other? Available at: https://www.iotnxt.com/2018/07/18/iot-or-scada/ (accessed: 25.08.2019).
7. Oficialen sajt Microsoft Story Lab [Official web site of Microsoft Story Lab]. Available at: https://news.microsoft.com/stories/cloud-security (accessed: 25.08.2019). (In Bulg.).
8. EU Directive 2008/114/EC of 8 December 2008. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2008.345.01.0075.01.ENG (accessed: 25.08.2019).
9. Naredba za reda za ustanovyavaneto i oznachavaneto na evropejski kritichni infrastrukturi v republika B"lgariya i merkite za tyahnata zashchita [Ordinance on the procedure for the establishment and designation of European Critical Infrastructures in the Republic of Bulgaria and the measures for their protection]. Available at: https://www.lex.bg/laws/ldoc/2135839567 (accessed: 25.08.2019). (In Bulg).
10. Communication from the Commission on a European Programme for Critical Infrastructure Protection. Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52006DC0786 (accessed: 25.08.2019).
11. Industrialen Ethernet i SCADA [Industrial and SCADA]. Available at: https://radesol.com/bg/industrialen-ethernet-scada/ (accessed: 25.08.2019).
12. Xing L., Demertzis K., Yang J. Identifying data streams anomalies by evolving spiking restricted Boltzmann machines. *Neural Computing and Applications*. 2019. pp. 1–15.
13. Lecture notes. Available at: https://www.cs.fsu.edu/~redwood/OffensiveComputerSecurity/lectures.html (accessed: 25.08.2019).
14. Guide for Applying the Risk Management Framework to Federal Information Systems: a Security Life Cycle Approach, NIST 2018. Available at: https://csrc.nist.gov/publications/detail/sp/800-37/rev-1/final (accessed: 25.08.2019).
15. Lee M.R., Assante M.J., Conway T. ICS CP/PE (Cyber-to-Physical or Process Effects) case study paper – German Steel Cyber Attack. SANS ICS. 2014. 15 p. Available at: https://ics.sans.org/media/ICS-CPPE-case-Study-2-German-Steelworks_Facility.pdf (accessed: 25.08.2019).
16. Malaviya S. SCADA Cybersecurity Framework. *ISACA Journal.* 2014. vol. 1. 5 p.

17. Eriksson M., Johansson T.T. An Example of a Man-in-the-Middle Attack Against Server Authenticated SSL-Sessions. International conference on applied cryptography and network security. 2003.
18. Moore A.P., Ellison R.J., Linger R.C. Attack Modeling for Information Security and Survivability. Carnegie-Mellon University. 2001. 32 p.
19. Iliev O. Radar Charts: A Novel Means to Explore the Relationship Between QoS and QoE.
20. Paneva-Marinova D., Stoikov J.S., Pavlova L.R., Luchev D.M. [System Architecture and Intelligent Data Curation of Virtual Museum for Ancient History]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2019. vol. 18(2). pp. 444–470.
21. Radvanovsky R., Brodsky J. Handbook of SCADA/Control Systems Security. CRC Press. 2013.
22. Trifonov R. et al. Increasing the level of network and information security using artificial intelligence. Fifth International Conference on Advances in Computing, Communication and Information Technology. 2017. pp. 83–88.
23. Trifonov R., Yoshinov R., Jekov B., Pavlova G. Methodology for Assessment of Open Data. *International Journal of Computers*. 2017. Issue 2. pp. 28–37.
24. Trifonov R., Yoshinov R., Jekov B., Pavlova G. E-government assessment. *International Journal of Development Research*. 2017. vol. 07. pp. 14874–14881.
25. Trifonov R., Yoshinov R., Pavlova G., Tsochev G. Artificial neural network intelligent method for prediction. AIP Conference Proceedings. vol. 1872. no. 1. pp. 020021.
26. Trifonov R. et al. A Survey of Artificial Intelligence for Enhancing the Information Security. International Journal of Development Research. 2017. vol. 7. no. 11. pp. 16866–16872.
27. Trifonov R. et al. Conceptual model for cyber intelligence network security system. *International Journal of Computers*. 2017. vol. 11. pp. 85–92.
28. Krutz R.L. Securing SCADA systems. Wiley Publishing, Inc. 2006. 240 p.
29. Schneider Electric. Available at: https://www.se.com/ww/en/product-category/6000-telemetry-and-remote-scada-systems/ (accessed: 25.08.2019).
30. Morris T., Gao W. Industrial control system traffic data sets for intrusion detection research. International Conference on Critical Infrastructure Protection. 2014. pp. 65–78.
31. Fillatre L., Nikiforov I. A statistical method for detecting cyber/physical attacks on SCADA systems. 2014 IEEE Conference on Control Applications (CCA). 2014. pp. 364–369.

Tsochev Georgi Rumenov — Ph.D., Chief Assistant Professor, Laboratory of Telematics, Bulgarian Academy of Sciences (BAS). Research interests: computer science, computer networks and communication, neural networks, deep learning, application of mathematics and informatics in cybersecurity. The number of publications — 25. gtsochev@gmail.com; 8 bl., Acad. Georgi Bonchev Str., 1113, Sofia, Bulgaria; office phone: +359895589861.

Yoshinov Radoslav Dakov — Ph.D., Professor, Head of Laboratory, Laboratory of Telematics, Bulgarian Academy of Sciences (BAS). Research interests: computer science, medical systems, computer networks and communication, deep learning, cybersecurity, E-Government cybersecurity of computer networks. The number of publications — 191. yoshinov@cc.bas.bg; 8 bl., Akad. G. Bonchev Str., 1113, Sofia, Bulgaria; office phone: +359888627190.

Iliev Oleg Petrov — Junior Researcher, Laboratory of Telematics, Bulgarian Academy of Sciences (BAS). Research interests: information technologies, computer science, IT components to support education process. The number of publications — 6. iliev.oleg@gmail.com; 8 bl., Akad. G. Bonchev Str., 1113, Sofia, Bulgaria; office phone: +359884381052.

Acknowledgements. This research is supported through IKT in NOS program of MES.

УДК 006.72 **DOI** 10.15622/sp.2019.18.6.1333-1356

Г.Р. Цочев, Р.Д. Йошинов, О.П. Илиев

ОСНОВНЫЕ ПРОБЛЕМЫ ЗАЩИТЫ КРИТИЧЕСКОЙ ИНФРАСТРУКТУРЫ СЕТЕЙ НА ОСНОВЕ СИСТЕМ SCADA

Цочев Г.Р., Йошинов Р.Д., Илиев О.П. **Основные проблемы защиты критической инфраструктуры сетей на основе систем SCADA.**

Аннотация. Критические инфраструктуры и оперативная совместимость составляющих ее информационно-ресурсных компонентов — главная составляющая жизненного цикла инфраструктуры. Поскольку системы диспетчерского контроля и сбора данных ((анг. Supervisory Control and Data Acquisition — SCADA) являются частью критической инфраструктуры, их киберзащита особенно важна на стратегических объектах, таких как электростанции, нефтеперерабатывающие заводы, нефтепроводы, очистные сооружения, производственные объекты, транспорт и так далее. Наряду с развитием технологий и онлайн доступности устройств систем SCADA, также увеличилась уязвимость подконтрольных им секторов. В мире Интернета вещей конечные устройства вызывают новую волну возможных уязвимостей в SCADA, так как они подвержены атакам и взломам и через них можно получить доступ к системе. В Европейском сообществе существует ряд критически важных инфраструктур, нарушение или разрушение которых может иметь значительные по масштабу трансграничные последствия для более чем одного сектора как результатвзаимозависимости взаимосвязанных инфраструктур. Такие европейские критические инфраструктуры были созданы и запущены в соответствии с разработанной Европейской комиссией общей процедурой, включающей в себя оценку требований безопасности, с учетом общего минимального подхода.

Рассматриваются критические инфраструктуры Европейского Союза и Болгарии. Посредством структуры системы SCADA были проанализированы уязвимости и различные возможности для ее атаки. В качестве примера рассмотрен конкретный случай, на премере деревьев атак, и полученные результаты были обобщены и визуализированы. Проанализированы последствия и сделаны соотвествующие выводы.

Ключевые слова: критическая инфраструктура, SCADA, дерево атаки, киберзащита, сетевая информационная защита.

Цочев Георги Руменов — Ph.D., главный ассистент, лаборатория телематики, Болгарская академия наук (БАН). Область научных интересов: информатика, компьютерные сети и связь, нейронные сети, глубинное обучение, применение математики и информатики в кибербезопасности. Число научных публикаций — 25. gtsochev@gmail.com; ул. Академика Георги Бончев, 8 bl., 1113, София, Болгария; р.т.: +359895589861.

Йошинов Радослав Даков — Ph.D., профессор, заведующий лабораторией, лаборатория телематики, Болгарская академия наук (БАН). Область научных интересов: информатика, медицинские системы, компьютерные сети и связь, глубинное обучение, кибербезопасность, кибербезопасность компьютерных сетей электронного правительства. Число научных публикаций — 191. yoshinov@cc.bas.bg; ул. Академика Георги Бончев, 8 bl., 1113, София, Болгария; р.т.: +359888627190.

Илиев Олег Петров — младший научный сотрудник, лаборатория телематики, Болгарская академия наук (БАН). Область научных интересов: информационные технологии, информатика, ИТ-компоненты для поддержки учебного процесса. Число научных публикаций — 6. iliev.oleg@gmail.com; ул. Академика Георги Бончев, 8 bl., 1113, София, Болгария; р.т.: +359884381052.

Поддержка исследований. Работа выполнена при финансовой поддержке программы ИКТ в НОС, МОН.

Литература

1. *Almalawi A. et al.* An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems // Computers & Security. 2014. vol. 46. pp. 94–110.
2. Amenaza Technologies Limited. URL: https://www.amenaza.com/faq.php (дата обращения: 25.08.2019).
3. *Byres E.J, Franz M., Miller D.* The Use of Attack Trees in Assessing Vulnerabilities in SCADA System // Proceedings of the International Infrastructure Survivability Workshop (IISW'04). 2004. pp. 3–10.
4. *Dochev D., Pavlov R., Paneva-Marinova D., Pavlova L.* Towards Modeling of Digital Ecosystems for Cultural Heritage // Digital Presentation and Preservation of Cultural and Scientific Heritage. 2019. vol. 9. pp. 77–88.
5. IEEE Standard Definition, Specification, and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control // Institute of Electrical and Electronics Engineers. 1987.
6. IoT And SCADA: Is One Going To Replace The Other? URL: https://www.iotnxt.com/2018/07/18/iot-or-scada/ (дата обращения: 25.08.2019).
7. Официален сайт Microsoft Story Lab. URL: https://news.microsoft.com/stories/cloud-security (дата обращения: 25.08.2019).
8. EU Directive 2008/114/EC of 8 December 2008. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2008.345.01.0075.01.ENG (дата обращения: 25.08.2019).
9. Наредба за реда за установяването и означаването на европейски критични инфраструктури в република България и мерките за тяхната защита. URL: https://www.lex.bg/laws/ldoc/2135839567 (дата обращения: 25.08.2019).
10. Communication from the Commission on a European Programme for Critical Infrastructure Protection. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52006DC0786 (дата обращения: 25.08.2019).
11. Индустриален Ethernet и SCADA. URL: https://radesol.com/bg/industrialen-ethernet-scada/ (дата обращения: 25.08.2019).
12. *Xing L., Demertzis K., Yang J.* Identifying data streams anomalies by evolving spiking restricted Boltzmann machines // Neural Computing and Applications. 2019. pp. 1–15.
13. Lecture notes. URL: https://www.cs.fsu.edu/~redwood/OffensiveComputerSecurity/lectures.html (дата обращения: 25.08.2019).
14. Guide for Applying the Risk Management Framework to Federal Information Systems: a Security Life Cycle Approach, NIST 2018. URL: https://csrc.nist.gov/publications/detail/sp/800-37/rev-1/final (дата обращения: 25.08.2019).
15. *Lee M.R., Assante M.J., Conway T.* ICS CP/PE (Cyber-to-Physical or Process Effects) case study paper – German Steel Cyber Attack // SANS ICS. 2014. 15 p. URL: https://ics.sans.org/media/ICS-CPPE-case-Study-2-German-Steelworks_Facility.pdf (дата обращения: 25.08.2019).
16. *Malaviya S.* SCADA Cybersecurity Framework // ISACA Journal. 2014. vol. 1. 5 p.
17. *Eriksson M., Johansson T.T.* An Example of a Man-in-the-Middle Attack Against Server Authenticated SSL-Sessions // International conference on applied cryptography and network security. 2003.
18. *Moore A.P., Ellison R.J., Linger R.C.* Attack Modeling for Information Security and Survivability // Carnegie-Mellon University. 2001. 32 p.
19. *Iliev O.* Radar Charts: A Novel Means to Explore the Relationship Between QoS and QoE.
20. *Paneva-Marinova D., Stoikov J.S., Pavlova L.R., Luchev D.M.* System Architecture and Intelligent Data Curation of Virtual Museum for Ancient History // Труды СПИИРАН. 2019. vol. 18(2). pp. 444–470

21. *Radvanovsky R., Brodsky J.* Handbook of SCADA/Control Systems Security // CRC Press. 2013
22. *Trifonov R. et al.* Increasing the level of network and information security using artificial intelligence // Fifth International Conference on Advances in Computing, Communication and Information Technology. 2017. pp. 83–88.
23. *Trifonov R., Yoshinov R., Jekov B., Pavlova G.* Methodology for Assessment of Open Data // International Journal of Computers. 2017. Issue 2. pp. 28–37.
24. *Trifonov R., Yoshinov R., Jekov B., Pavlova G.* E-government assessment // International Journal of Development Research. 2017. vol. 07. pp. 14874–14881.
25. *Trifonov R., Yoshinov R., Pavlova G., Tsochev G.* Artificial neural network intelligent method for prediction // AIP Conference Proceedings. vol. 1872. no. 1. pp. 020021.
26. *Trifonov R. et al.* A Survey of Artificial Intelligence for Enhancing the Information Security // International Journal of Development Research. 2017. vol. 7. no. 11. pp. 16866–16872.
27. *Trifonov R. et al.* Conceptual model for cyber intelligence network security system // International Journal of Computers. 2017. vol. 11. pp. 85–92.
28. *Krutz R.L.* Securing SCADA systems // Wiley Publishing, Inc. 2006. 240 p.
29. Schneider Electric. URL: https://www.se.com/ww/en/product-category/6000-telemetry-and-remote-scada-systems/ (дата обращения: 25.08.2019).
30. *Morris T., Gao W.* Industrial control system traffic data sets for intrusion detection research // International Conference on Critical Infrastructure Protection. 2014. pp. 65–78.
31. *Fillatre L., Nikiforov I.* A statistical method for detecting cyber/physical attacks on SCADA systems // 2014 IEEE Conference on Control Applications (CCA). 2014. pp. 364–369.

УДК 621.391.1 **DOI** 10.15622/sp.2019.18.6.1357-1380

О.О. Басов, И.А. Саитов, А.И. Мотиенко, С.С. Астапов

СИНТЕЗ ТОПОЛОГИЧЕСКОЙ СТРУКТУРЫ РАСПРЕДЕЛЕННОЙ ТЕРМИНАЛЬНОЙ СИСТЕМЫ ДЛЯ АУДИОМОНИТОРИНГА ПОЛЬЗОВАТЕЛЕЙ ЛОКАЛЬНЫХ ИНФОРМАЦИОННЫХ ПРОСТРАНСТВ

Басов О.О., Саитов И.А., Мотиенко А.И., Астапов С.С. **Синтез топологической структуры распределенной терминальной системы для аудиомониторинга пользователей локальных информационных пространств.**

Аннотация. Широкое использование многомодальных интерфейсов, обусловленное многомодальностью традиционного межличностного общения, переход к полимодальному представлению информации и систем ее обработки позволили по-новому взглянуть на ряд инфокоммуникационных технологий и сервисов и предложить подход к их реализации на основе распределенных терминальных систем. Предложен подход к синтезу топологической структуры таких систем, реализуемый в два этапа: на первом определяется минимальная совокупность узлов коммуникации и их размещение на основе требований к доступности узлов коммуникации для различных категорий пользователей и глобальности распределенной терминальной системы, на втором — варианты построения узлов коммуникации и связей между ними, которые обеспечивают выполнение функций коммуникации пользователей локальных информационных пространств при обеспечении непрерываемости связи для различных категорий пользователей. Представлен модельный пример задачи синтеза распределенной терминальной системы для аудиомониторинга двух категорий пользователей (взрослые и дети) в локальном информационном пространстве (доме), голосового управления подсистемами «умного» дома. Для ее решения на каждом этапе синтеза определены исходные данные, осуществлена формальная постановка задачи синтеза, представлен алгоритм решения и получаемые результаты. Так задача первого этапа синтеза представляет собой линейную целочисленную задачу математического программирования, решенную в модельном примере симплекс-методом. Решение задачи второго этапа основано на альтернативно-графовой формализации и методе «ветвей и границ». Полученные результаты демонстрируют возможности предложенного научно-методического инструментария синтеза топологической структуры распределенных терминальных систем и перспективность его использования во вновь возникающих задачах технической реализации новых инфокоммуникационных технологий и сервисов.

Ключевые слова: инфокоммуникационная система, распределенная терминальная система, топологическая структура, доступность узлов, непрерываемость связи.

Введение. Использование в пределах локальных пространств (помещений, офисов, квартир, домов, зданий) многомодальных устройств ввода/вывода [1] и систем на их основе, объединяющих различные комбинации входных и выходных модальностей и названных в [2] распределенными терминальными системами (РТС), обеспечивает их пользователей рядом инфокоммуникационных технологий (таблица 1), придавая таким пространствам свойство интеллектуальности. При этом исходя из

личных предпочтений и потребностей в коммуникативных каналах пользователям РТС предоставляются различные функции коммуникации на основе необходимой комбинации доступных модальностей. Так, например, для реализации распределенной телефонии (телефонной связи) в локальном пространстве возможно использовать микрофонные и акустические системы, а для голосового управления системами «умного» дома или интеллектуального зала [3] или аудиомониторинга пользователей [4] такие локальные информационные пространства должны содержать распределенную сеть микрофонов [5].

Таблица 1. Инфокоммуникационные технологии локальных пространств

№ п/п	Акустический канал коммуникации		Визуальный канал коммуникации		Инфокоммуникационная технология
	входные модальности	выходные модальности	входные модальности	выходные модальности	
1	+				Многоканальная аудиозапись, голосовое управление, аудиомониторинг пользователей
2			+		Видеонаблюдение, управление жестами, видеомониторинг пользователей
3	+		+		Многоканальная аудиовидеозапись, многомодальное управление, бимодальный мониторинг пользователей
4		+			«Звук вокруг», сонификация
5				+	Масштабная визуализация, дополненная реальность
6	+	+			Распределенная телефония
7			+	+	Виртуальная реальность, усиленная сонификация
8	+	+	+	+	Распределенная видеосвязь

2. Обзор существующих систем мониторинга пользователей локальных информационных пространств. Существующие системы мониторинга пользователей предназначены, как правило, для анализа физического (физиологического) состояния

различных групп населения, в первую очередь — пожилых людей [6-12]. Так, в [6] предлагается система определения факта падения человека на основе многомодальных беспроводных сенсорных сетей, включающих датчики температуры, движения и биосенсоры, видеокамеры, аудиодатчики и RFID-метки. Для повышения точности обнаружения падений и минимизации ложных срабатываний сенсоры используются совместно. RFID-метки применяются для отслеживания местоположения людей и активации видеокамер в помещениях, где последние находятся. Видеокамеры обеспечивают наблюдение и распознавание действий людей. Биосенсоры, например, измеряют артериальное давление или частоту сердечных сокращений, контролируя состояние здоровья. Датчики окружающей среды используются для предоставления контекстной информации об окружающей среде. Система позволяет осуществлять удаленный мониторинг ее пользователей посредством мобильных и локальных вычислительных сетей.

В системах предотвращения падений основное внимание уделяется их обнаружению. Как таковые падения неизбежны, и их определение для предотвращения травм имеет большое значение, однако другие области предотвращения риска падения представляют серьезную проблему. Предполагается, что будущие системы предотвращения падений будут более эффективны для устранения внешних рисков, в частности от того, насколько успешно врачи, проводящие домашние обследования, могут использовать оборудование, а пожилые люди — самостоятельно оценить свои потребности во вспомогательном оборудовании при отсутствии персонала в доме [7]. С этой целью изучение расположения мебели и точности ее измерения могло бы обеспечить правильное размещение оборудования в доме, что привело бы к более эффективному его использованию.

В [8] исследуется преимущества использования мобильной трехмерной (3D) технологии визуализации для расширения мероприятий по оценке состояния окружающей среды, направленных на преодоление внешних факторов риска падения в домашних условиях.

Беспроводная домашняя система мониторинга [9] включает в себя монитор состояния человека, носимый на запястье, и опционально пульсоксиметры или мониторы артериального давления, подключаемые по беспроводной сети, и позволяет собирать данные о деятельности и местоположении человека, распознавать падения, выдавать оповещения о панике.

В [10] предлагается система удаленного мониторинга здоровья пожилых людей на основе «умного дома», состоящая из «умной одежды», считывающей параметры сердцебиения и движения человека, шлюза «умного дома», предназначенного для мультиплексирования и передачи данных на сервер здравоохранения. Предлагаемая система обладает хорошей масштабируемостью и простотой эксплуатации.

В России аналогом рассмотренных систем является автоматизированная система дистанционного мониторинга здоровья человека «Монитор здоровья» [11], позволяющая автоматически считывать данные с различных специализированных приборов анализа параметров здоровья человека. При необходимости пользователь может предоставить врачу доступ к хранящимся в системе измерениям.

В подавляющем большинстве систем, за исключением [6], мало внимания уделено анализу всей совокупности акустических данных, что не позволяет полно оценивать текущую ситуацию в локальном информационном пространстве, оперативно реагировать на ее изменение и вырабатывать управляющие воздействия при возникновении нештатных (критических) ситуаций.

Недостатками системы мониторинга [6] являются необходимость размещения на пользователях RFID-меток, использование для их локализации встроенных видеокамер и, как следствие, сложность и дороговизна такой системы, а также отсутствие рекомендаций по размещению аудиодатчиков и возможности их использования для локализации пользователя в некотором пространстве. Для их преодоления должна быть решена задача синтеза топологической структуры РТС (сенсорной сети или более конкретно — распределенной сети микрофонов), позволяющая выбрать наилучший по заданным критериям оптимальности вариант размещения аудиодатчиков в локальном информационном пространстве для локализации пользователя в нем, считывания акустического сигнала с требуемым качеством, его предварительной обработки и передачей для последующего анализа физического (физиологического) состояния человека либо для распознавания команд управления системами «умного» дома.

Следует отметить, что последовательный синтез РТС должен основываться на разработке ее физической (топологической и функциональной) структуры, начиная с самого простого случая — одной входной (выходной) модальности одного из доступных каналов коммуникации с последовательным наращиванием числа модальностей и каналов.

3. Обзор существующих подходов к постановкам и решению задач анализа и синтеза (оптимизации характеристик) РТС и их элементов. В современной науке сложилось три принципиально различающихся подхода к формализации задач синтеза, отражающих взаимосвязь между исходными данными и представлениями о структуре и функционировании сложной системы:

– синтез структуры системы при заданных алгоритмах ее функционирования (1-й класс задач);

– синтез оптимального поведения и алгоритмов функционирования системы при известной структуре (2-й класс задач);

– синтез структуры и алгоритмов (поведения) функционирования системы, распределения их оптимального состава (3-й класс задач).

Исследования показывают, что при наблюдающемся прогрессе в решении 1 и 2 класса задач практически отсутствуют подходы к формализации и решению 3 класса задач, к которым в методологическом плане примыкают задачи синтеза РТС. Их решение может быть найдено при комплексировании структурно-функционального и функционально-структурного подходов.

В структурно-функциональном подходе системотехнические задачи представляют собой взаимоувязанный комплекс задач анализа и оптимизации характеристик РТС или ее подсистем через параметры ее физической структуры. Такие задачи далее будем называть задачами по оптимизации структуры (ЗОС). Классическая ЗОС включает в свой состав частные задачи по синтезу (оптимизации) топологической и функциональной структур, в совокупности обеспечивающих минимум критерия экономической природы при выполнении ограничений на вероятностно-временные характеристики. При использовании графо-матричных моделей вычислительная сложность ЗОС является полиномиальной. Для решения таких задач на сегодняшний день в практике проектирования ИКС установились декомпозиционный и глобальный подходы.

В первом случае РТС разбивается, например, на центральный узел коммутации (ЦУК), узлы коммутации (УК) и вспомогательные УК (ВУК), структуры которых оптимизируются раздельно (рисунок 1). Очевидно, что декомпозиционному подходу свойственен такой недостаток, как неконтролируемая ошибка декомпозиции, поскольку часто открытым остается вопрос выбора подсистем и распределения общесистемных норм между ними.

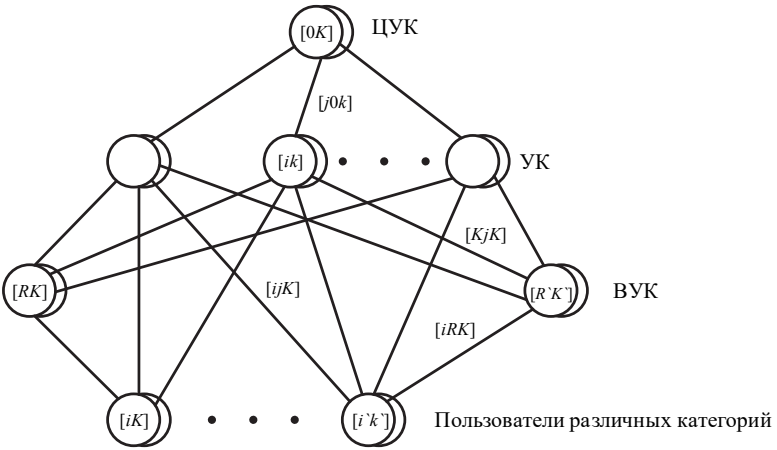


Рис. 1. Обобщенная схема распределенной терминальной системы для аудиомониторинга пользователей локальных информационных пространств

Второй подход подразумевает работу со структурой РТС в целом и на первый план выступает проблема снижения вычислительной сложности ЗОС. Снижение вычислительной сложности ЗОС может быть получено за счет огрубления топологической, функциональной и/или конструктивной структуры РТС.

В ряде случаев для решения ЗОС высокой вычислительной сложности применяется многоэтапная процедура [7]. Такой подход базируется на применении серии различающихся степенью адекватности (и соответственно размерностью) моделей, для которых выходные результаты очередного этапа оптимизации являются начальными условиями поиска для последующего этапа. Исследования показали, что наблюдаемые в реальных условиях среднестатистическая точность исходных данных и погрешности, вносимые модельными допущениями, позволяют с достаточной для практики точностью ограничиться двумя этапами решения типовых ЗОС.

Независимо от применения декомпозиционного или глобального подхода выбор алгоритма решения конкретной ЗОС определяется особенностями ее постановки, например спецификой предпроектной ситуации. Таким образом, ЗОС может быть сведена к той или иной задаче математического программирования или их последовательности.

При синтезе структур систем, аналогичных РТС, например инфокоммуникационных систем, уже многие годы широко используются методы «ветвей и границ» и «устранения ребер», адаптированных под цели оптимизации структуры. Однако из-за высокой вычислительной сложности вне зависимости от применяемого поискового алгоритма возможности этих методов решения ЗОС на практике ограничиваются числом узлов соответствующего графа.

Недостатками применения эвристических (в том числе переборных) алгоритмов являются трудность выбора стартовой топологии, во многом определяющей эффективность оптимизации, и необходимость периодической проверки ограничений по связности структуры. Им также свойственны существенные временные затраты на поиск допустимого начального плана распределения потоков.

Выявленные недостатки существующих подходов к решению ЗОС РТС стимулируют развитие теоретико-методологической базы проектирования. Перспективным в этом смысле является комплексный подход, базирующийся одновременно на структурно-функциональной и функционально-структурной методологиях. В данном случае задача синтеза топологической структуры РТС для аудиомониторинга пользователей локальных информационных пространств состоит в рациональном отображении множества взаимосвязанных функций коммуникации на множество взаимосвязанных узлов коммуникации, обслуживающих различные категории пользователей с учетом их числа, расположения и состава аппаратно-программными средств. После ее решения оптимизируется распределение выполняемых функций и задач коммуникации по узлам РТС — определяется ее функциональная структура.

Задачи синтеза топологической и функциональной структур РТС, по сути, представляют собой задачи оптимизации характеристик сложных систем, функционирование которых во многом определяется принятыми принципами и алгоритмами обработки многомодальной информации, степенью централизации управления, согласованностью целей узлов коммуникации различного уровня. Моделям оптимизации функционирования РТС посвящены работы [1, 2], здесь же предложен научно-методический аппарат двухэтапной процедуры синтеза топологической структуры РТС, рассмотрен модельный пример его применения.

4. Постановка и решение задачи определения минимальной совокупности узлов коммуникации и их размещения. На первом этапе синтеза топологической структуры РТС решается задача определения минимальной совокупности узлов коммуникации (УК) и их размещения, исходя из требований к доступности УК для различных

категорий пользователей и глобальности РТС. Под различными категориями пользователей понимаются их группы, отличительной особенностью которых является использование для взаимодействия с РТС одинаковых каналов коммуникации (модальностей). *Доступность узлов* определяется временем τ_{ij} максимальной продолжительности взаимодействия пользователя *i*-й категории с *j*-м УК $\left(j=\overline{1,J}\right)$, а *глобальность*

$$H_i=\frac{\tau_i}{\tau},$$

где $\tau_i\left(i=\overline{1,I}\right)$ — общее время доступности РТС *i*-й категории пользователей; τ — общее время нахождения пользователя в заданном локальном информационном пространстве.

Рассмотрим модельный пример решения задачи синтеза топологической структуры РТС для аудиомониторинга пользователей в заданном локальном пространстве (рисунок 2). В соответствии с определенными в [2] обозначениями такая система включает в себя:

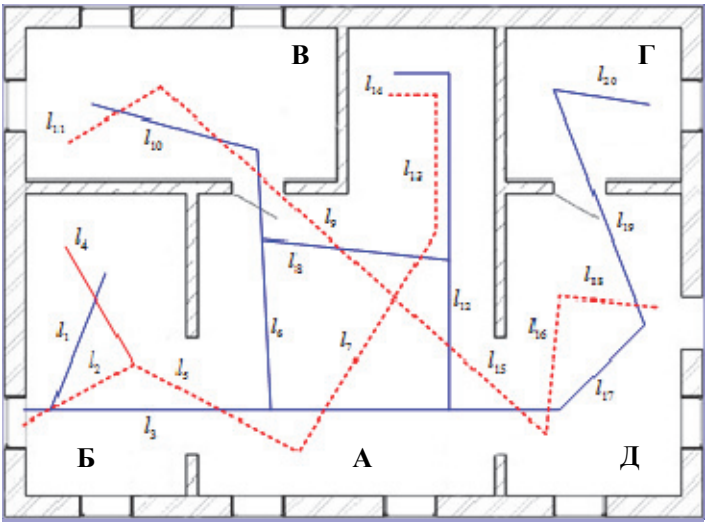
– ЦУК, предназначенный для обработки и передачи аудиоинформации от УК в систему мониторинга;

– непосредственно УК, представляющие в рассматриваемом случае комплексы аппаратно-программных средств сбора, обработки и передачи ЦУК аудиоинформации в цифровом виде;

– ВУК, представляющие собой отдельные микрофоны.

В модельном примере микрофоны УК и ВУК размещаются на потолке помещений и имеют направленную вниз диаграмму направленности, а пользователи делятся на две категории по возрастному признаку: первой категории $\left(i=1\right)$ — взрослые, для которых $k_1=2$, — доступны все помещения локального информационного пространства; второй категории $\left(i=2\right)$ — дети, для которых $k_2=1$, — доступны все помещения, за исключением Г, усредненные маршруты их передвижения отличны в связи с отсутствием необходимости пользования некоторыми предметами интерьера (рисунок 2 б).

а)



б)

Рис. 2. 3D-проекция (а) и план локального информационного пространства (б) с указанием маршрутов движения пользователей двух категорий (маршрут передвижения пользователей первой категории обозначен сплошной линией, пользователей второй категории – прерывистой линией)

Пользователи каждой из категорий предъявляют идентичные требования к номенклатуре и качеству предоставляемых функций коммуникации (услугам) и имеют близкие параметры движения (скорость, направление и т.п.).

Исходными данными для решения задачи определения состава узлов коммуникации и их размещения (рисунок 3) являются (таблицы 2-4):

– $l = \overline{1, L}$ — участки, на которых пользователям доступны УК [2];

– n_l — множество узлов, которые доступны пользователям соответствующей категории на l-м участке;

– k_i — число узлов, необходимых для обслуживания пользователей i-й категории;

– L_i — множество индексов участков, на которых пользователям i-й категории $\left(i = \overline{1, I}\right)$ доступна РТС;

– θ_l — время доступности УК для пользователя на l-м участке;

– c_j — затраты на создание j-го УК $\left(j = \overline{1, J}\right)$.

Начальная (до оптимизации) схема возможного размещения узлов коммуникации в локальном информационном пространстве предполагает размещение центрального узла коммуникации в техническом помещении Г, по одному узлу коммуникации — в каждом из помещений, вспомогательные узлы коммуникации — по помещениям с целью увеличения доступности узлов коммуникации.

Задача определения состава и размещения УК в общем случае состоит в минимизации затрат на создание УК:

$$\min \sum_{j=1}^{J} c_j x_j, \tag{1}$$

$$x_j = \begin{cases} 1, \text{ если в системе используется узел } j\left(j = \overline{1, J}\right); \\ 0 \ - \text{ в противном случае;} \end{cases}$$

при ограничениях на глобальность системы связи, предоставляемой пользователям различных категорий:

$$\sum_{l\in L_i}\theta_l y_l \geq H_i\tau_i,\ i=\overline{1,I}\ , \tag{2}$$

где

$$y_l=\begin{cases}1, & \text{если } l\text{-й участок доступности используется в РТС;}\\ 0 & -\text{ в противном случае.}\end{cases}$$

при ограничениях, например:

$$\sum_{j\in m_i} x_j \geq k_i,\ \ i=\overline{1,I}\ , \tag{3}$$

при этом взаимосвязь между переменными x_j и y_l задается условиями:

$$\sum_{j\in n_l} x_j \geq y_l \geq \frac{1}{|n_l|}\sum_{j\in n_l} x_j,\ \ l=\overline{1,I}\ , \tag{4}$$

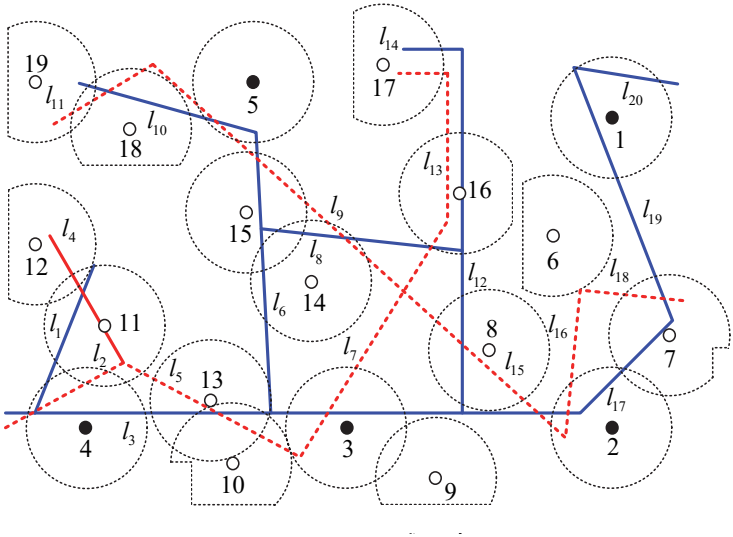


Рис. 3. Исходные данные для решения задачи определения минимальной совокупности узлов коммуникации и их размещения

Таблица 2. Время доступности УК (ВУК) для пользователей первой категорий

L_1	1	3	6	9	10	12	14	17	19	20
n_l	4, 11, 12	2, 3, 4, 10, 13	5, 13, 14, 15	14, 15, 16	5, 18, 19	8, 16	17	2, 7	1, 7	1
θ_l, с	7	10	3	2	3	6	1	4	2	1

Таблица 3. Время доступности УК (ВУК) для пользователей второй категорий

L_2	2	4	5	7	9	11	13	14	15	16	18
n_l	4, 11	11, 12	3, 10, 11, 13	3, 16	5, 14, 15	18, 19	16	17	2, 8	2	7
θ_l, с	4	4	5	1	6	2	1	1	2	1	1

Таблица 4. Временные и стоимостные параметры узлов коммуникации

j	1	2	3	4	5	6	7	8	9	10
τ_{1j}, с	3	1	2	2	0,5	0	2	3	0	1
τ_{2j}, с	0	1	1	1	0,5	0	1	2	0	2
c_j, у.е	4	4	4	4	4	2	2	2	2	2
j	11	12	13	14	15	16	17	18	19	
τ_{1j}, с	2	1	2	4	3	2	0,5	2	0,5	
τ_{2j}, с	2	0	2	1	1	2	0,5	2	1	
c_j, у.е.	2	2	2	2	2	2	2	2	2	

Задача (1)–(4) представляет собой линейную целочисленную задачу математического программирования, результаты решения которой симплекс-методом позволили определить в рассматриваемом модельном примере, какие узлы необходимы в системе аудиомониторинга пользователей локальных информационных пространств, их размещение, а также состав узлов, обслуживающих пользователей двух категорий (рисунок 4) [13-15].

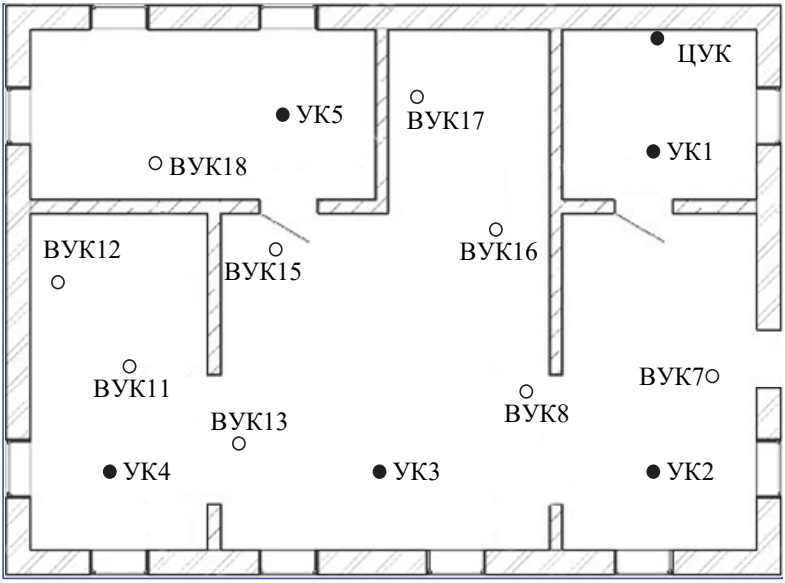


Рис. 4. Результаты решения задачи определения минимальной совокупности узлов коммуникации и их размещения

Для полученного решения глобальность РТС для различных категорий пользователей составила:

$$H_1 = \sum_j \tau_{1j} \Big/ \tau = 0,95;$$

$$H_2 = \sum_j \tau_{2j} \Big/ \tau = 0,93,$$

где τ_{1j}, τ_{2j} — общее время доступности j-го узла РТС пользователям первой и второй категории соответственно.

Следовательно, пользователям первой категории РТС будет доступна в 95 % от общего времени нахождения в заданном локальном информационном пространстве (рисунок 2), пользователям второй категории — в 93 % от общего времени нахождения.

5. Постановка и решение задачи определения вариантов построения узлов коммуникации и связей между ними. На втором этапе синтеза топологической структуры РТС для выбранной

совокупности узлов РТС и их взаимосвязей (рисунок 5) определяются варианты построения УК (ВУК) и связей между ними, которые обеспечивают выполнение функций аудиомониторинга пользователей локального информационного пространства при обеспечении непрерываемости связи для заданных категорий пользователей.

С учетом возможных внешних неблагоприятных воздействий, определяющих вероятности надежной работы (выполнения функций коммуникации) РТС и ее элементов, исходными данными для решения задачи определения вариантов построения УК и связей между ними являются:

– $l = \overline{1, L}$ — варианты построения узла (связи между узлами). Для полученных результатов решения задачи первого этапа синтеза (рисунок 4) выбраны два варианта реализации программно-аппаратных комплексов для УК и три варианта микрофонов для ВУК;

– P_{jl} — вероятности надежной работы l-го варианта построения j-го УК за рассматриваемый интервал времени;

– P_{ijl} — вероятности надежной работы акустических каналов между пользователем и УК;

– P_{j0l} — вероятности надежной работы каналов связи между УК и ЦУК за рассматриваемый интервал времени для различных вариантов их реализации;

– P_{Rl} — вероятности надежной работы l-го варианта R-го ВУК за рассматриваемый интервал времени $\left(l = \overline{1, L_r},\, r = \overline{1, R}\right)$;

– P_{iRl} — вероятности надежной работы акустических каналов между пользователем и ВУК;

– P_{Rjl} — вероятности надежной работы каналов связи между ВУК и УК за рассматриваемый интервал времени для различных вариантов их реализации;

– $P_i^{\text{ДОП}}$ — допустимая вероятность надежной работы РТС для пользователей i-й категории;

– c_{jl}, c_{Rl}, c_{ijl}, c_{iRl}, c_{Rjl}, c_{j0l} — затраты на организацию основных и вспомогательных УК, каналов связи и акустических каналов для различных вариантов их реализации;

В рассматриваемом примере индекс $j = 0$ соответствует ЦУК, индексы $j = \overline{1, J}$ — совокупности УК $\left(l = \overline{1, L_j},\, j = \overline{1, J}\right)$, а под неблагоприятными внешними воздействиями, влияющими на

надежность работы различных вариантов построения УК и ВУК, понимаются сложные акустические условия, возникающие в следствие появления нескольких пользователей в определенных местах локального информационного пространства, открывания окон и дверей (или их отсутствия). Каналы связи между ВУК и УК, УК и ЦУК за рассматриваемый интервал времени считаются абсолютно надежными, в связи с чем, вероятности $P_{Rjl}=1$, $P_{j0l}=1$.

Допустимые вероятности успешной работы РТС для выделенных категорий пользователей $P_1^{\text{ДОП}}=0,9$, $P_2^{\text{ДОП}}=0,8$. Затраты на организацию вариантов построения каналов коммуникации (связи) между основными и вспомогательными узлами коммуникации, центральным узлом коммуникации приняты равными $c_{ijl}=c_{iRl}=c_{Rjl}=c_{j0l}=0,25$ у.е. ввиду относительно простоты и дешевизны реализации проводной связи между указанными узлами коммуникации в локальном информационном пространстве. Остальные параметры для определения вариантов построения УК и связей между ними в рассматриваемом модельном примере определены случайным образом (таблицы 5, 6).

Таблица 5. Характеристики надежности и стоимости УК и каналов коммуникации

j			1	2	3	4	5
P_{jl}	$l=1$		0,90	0,90	0,80	0,90	0,90
	$l=2$		0,91	0,91	0,90	0,91	0,91
P_{ijl}	$l=1$	$i=1$	0,80	0,81	0,80	0,78	0,80
		$i=2$	0,70	0,65	0,70	0,75	0,70
	$l=2$	$i=1$	0,85	0,85	0,85	0,80	0,85
		$i=2$	0,75	0,70	0,70	0,80	0,75
c_{jl}, у.е.	$l=1$		7	7	7	7	7
	$l=2$		6	6	6	6	6

Задача определения вариантов построения УК и связей между ними в общем случае сводится к минимизации затрат

$$\min\left(\sum_k c_{jk}x_{jk}+\sum_k c_{iRk}x_{iRk}+\sum_k c_{ijk}x_{ijk}+\sum_k c_{jRk}x_{jRk}+\sum_k c_{j0k}x_{j0k}\right), \qquad (5)$$

при ограничениях:

$$P_{Rk}\left(x_{Rk},x_{iRk}\right)\left(1-P_{ijk}\left(x_{ijk},x_{Rjk},x_{jk}\right)\right)+$$
$$+\left(1-P_{Rk}\left(x_{Rk},x_{iRk}\right)\right)\left(1-P_{ij}\left(x_{ijk},x_{jk}\right)\right)\geq P_i^{\text{ДОП}}, \quad (6)$$

$$\sum_k x_{ik}=1,\ \sum_k x_{iRk}=1,\ \sum_k x_{ijk}=1,\ \sum_k x_{jRk}=1,\ \sum_k x_{i0k}=1,$$

где

$$x_{jk}=\begin{cases}1, \text{ если выбирается } k\text{-й вариант построения } j\text{-го узла;}\\ 0- \text{ в противном случае;}\end{cases}$$

$$x_{ijk}=\begin{cases}1, \text{ если выбирается } k\text{-й вариант связи между элементами } i \text{ и } j;\\ 0- \text{ в противном случае;}\end{cases}$$

$$x_{j0k}=\begin{cases}1, \text{ если выбирается } k\text{-й вариант связи между элементами } j \text{ и } 0;\\ 0- \text{ в противном случае.}\end{cases}$$

$$P_{Rk}\left(x_{Rk},x_{iRk}\right)=\left(\sum_k P_{Rk}x_{Rk}\right)\left(\sum_k P_{iRk}x_{iRk}\right);$$

$$P_{ijk}\left(x_{ijk},x_{Rjk},x_{jk}\right)=\prod_j\left[1-\left(1-\sum_k P_{ijk}x_{ijk}\right)\left(1-\sum_k P_{Rjk}x_{Rjk}\right)\left(\sum_k P_{jk}x_{jk}\right)\right];$$

$$P_{ij}\left(x_{ijk},x_{jk}\right)=\prod_j\left[1-\left(\sum_k P_{ijk}x_{ijk}\right)\left(\sum_k P_{jk}x_{jk}\right)\right].$$

Для решения задачи (5), (6) в модельном примере использован вариант переборного решения, тогда как для большего числа категорий пользователей узлов коммуникации и альтернатив может быть использована альтеративно-графовая формализация и предложенный в [16, 17] алгоритм, основанный на схеме «ветвей и границ». Для задания возможных вариантов построения узлов РТС и взаимосвязей между ними введен граф G_J, а для определения одного из возможных вариантов реализации узлов РТС с их взаимосвязями подграф $G_J^* \in G_J$ [18, 19].

Таблица 6. Характеристики надежности и стоимости ВУК и каналов коммуникации

j			6	7	8	9	10	11	12
P_{Rl}	$l=1$		0,90	0,80	0,90	0,90	0,90	0,80	0,90
	$l=2$		0,91	0,90	0,91	0,91	0,91	0,90	0,91
	$l=3$		0,81	0,80	0,78	0,80	0,81	0,80	0,78
P_{iRl}	$l=1$	0,65	0,70	0,75	0,70	0,65	0,70	0,75	0,70
		0,85	0,85	0,80	0,85	0,85	0,85	0,80	0,85
	$l=2$	0,70	0,70	0,80	0,75	0,70	0,70	0,80	0,70
		0,90	0,80	0,90	0,90	0,90	0,80	0,90	0,90
	$l=3$	0,91	0,90	0,91	0,91	0,91	0,90	0,91	0,88
		0,81	0,80	0,78	0,80	0,81	0,80	0,78	0,75
c_{Rl}, у.е.	$l=1$		5	3	4	3	2	3	3
	$l=2$		3	4	4	5	4	4	5
	$l=3$		3	3	2	3	4	5	3
j			13	14	15	16	17	18	19
P_{Rl}	$l=1$		0,85	0,85	0,85	0,85	0,85	0,85	0,85
	$l=2$		0,70	0,70	0,70	0,70	0,70	0,70	0,70
	$l=3$		0,90	0,90	0,90	0,90	0,90	0,90	0,90
P_{iRl}	$l=1$	$i=1$	0,85	0,85	0,85	0,85	0,85	0,85	0,85
		$i=2$	0,70	0,70	0,70	0,70	0,70	0,70	0,70
	$l=2$	$i=1$	0,90	0,90	0,90	0,90	0,90	0,90	0,90
		$i=2$	0,85	0,85	0,85	0,85	0,85	0,85	0,85
	$l=3$	$i=1$	0,70	0,70	0,70	0,70	0,70	0,70	0,70
		$i=2$	0,90	0,90	0,90	0,90	0,90	0,90	0,90
c_{Rl}, у.е.	$l=1$		5	3	4	3	2	3	3
	$l=2$		3	4	4	5	4	4	5
	$l=3$		3	3	2	3	4	5	3

На первом этапе работы алгоритма решения задачи (5) исключаются варианты на последовательных участках графа G_J с учетом правил доминирования и формируются агрегированные варианты. Затем исключаются варианты, заведомо не входящие в оптимальное решение, для этого на отдельных участках вычисляются максимально возможные вероятности их исправной работы. В (6) во всех сомножителях участков, кроме одного, вариативные части заменяются на вычисленные максимально возможные вероятности исправной работы участков. В незафиксированном участке последовательно анализируются и исключаются варианты, использование которых не удовлетворяет (6).

Далее осуществляется ветвление по сформированным вариантам построения отдельных участков системы. При этом число уровней ветвления равно суммарному числу отдельных участков системы. Оценки (5) в процессе ветвления находятся следующим образом: для зафиксированных вариантов выбираются соответствующие величины, а для незафиксированных переменных — минимально возможные затраты на соответствующих участках. Оценки для (6) вычисляются аналогично, однако используются максимально возможные вероятности исправной работы для незафиксированных участков. Граф возможных реализаций РТС для рассматриваемого модельного примера имеет вид, показанный на рисунке 5, где жирными линиями выделены пути, соответствующие оптимальному решению задачи (5) при ограничениях (6).

6. Заключение. Несмотря на видимую простоту модельного примера в виду использования лишь одной входной модальности, полученный результат наглядно демонстрирует возможности разработанного научно-методического инструментария синтеза топологической структуры РТС. При исследовании систем с большим числом входных и выходных модальностей и более сложными взаимосвязями между ее элементами не всегда представляется возможным выразить аналитически параметры системы через параметры элементов, поэтому в данном случае для получения оценок надежности могут использоваться методы «критических путей и сечений» [20].

После синтеза топологической структуры РТС на основе определенного множества УК и ВУК, обслуживающих пользователей различных категорий, и заданных функций коммуникаций для каждой из них решается задача оптимального (квазиоптимального) распределение задач (функций) по узлам системы [2]. Критериями оптимизации в данном случае выступают затраты на оснащение УК (ВУК, ЦУК) программно-аппаратными средствами и их эксплуатацию, оперативность, надежность, массогабаритные показатели и энергопотребление.

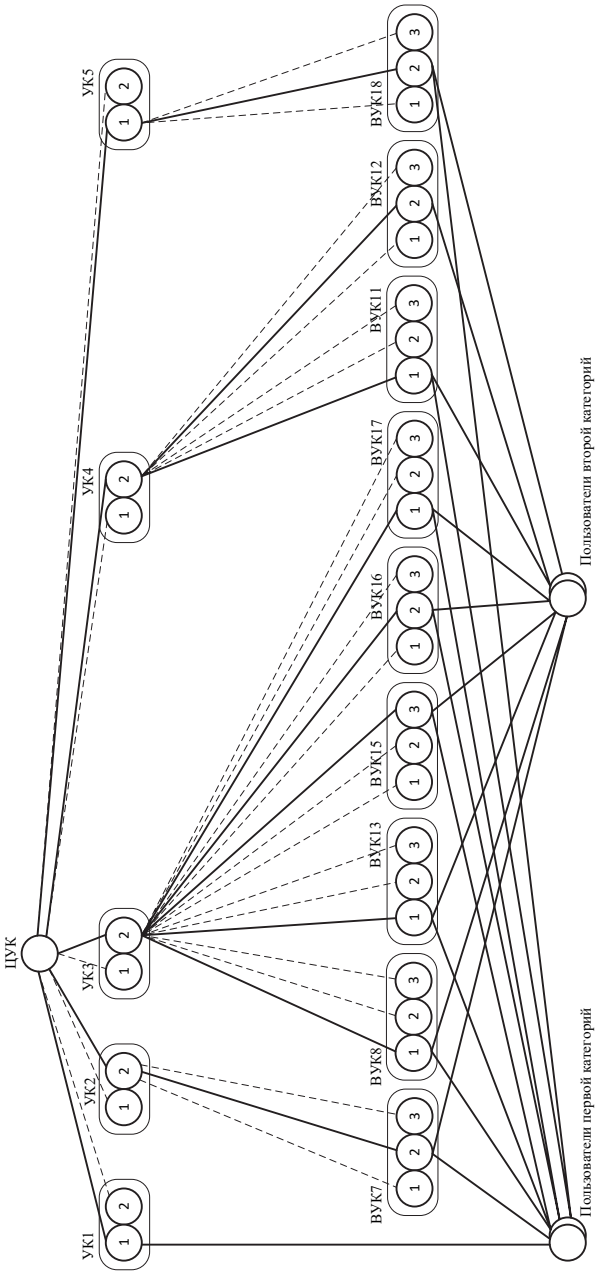


Рис. 5. Граф возможных реализаций РТС

Указанная задача синтеза функциональной структуры, равно как и рассмотренная в работе задача синтеза топологической структуры, является статической, поскольку номенклатура и распределение функций коммуникаций, а также состав реализующих их узлов и их расположение определяются для некоторого перспективного периода времени. Для РТС, функционирующей в течение достаточно длительного времени, приобретает актуальность совместная постановка задачи синтеза структуры и выбора плана (траектории) развития структуры системы.

Такая задача, являющаяся, по сути, динамической, в отличие от статической направлена на оптимизацию не только состава узлов, взаимосвязи между ними и распределение функций по УК (ВУК) для заданного интервала времени, но и на формирование (синтез) плана развития системы. Соответствующий научно-методический инструментарий подлежит разработке в процессе последующих исследований.

Литература

1. *Басов О.О., Карпов А.А., Саитов И.А.* Методологические основы синтеза полимодальных инфокоммуникационных систем государственного управления: монография // Орёл: Академия ФСО России. 2015. 272 с.
2. *Киселев Ю.В., Мотиенко А.И., Басов О.О., Саитов И.А.* Структурно-функциональная модель интеллектуальной инфокоммуникационной системы // Научно-технический вестник информационных технологий, механики и оптики. 2018. Т. 18. № 6. С. 1034–1046.
3. *Кузнецов Д.А., Офицеров А.И., Кузнецов А.В., Чистяков С.В., Басов О.О.* Предпосылки создания интеллектуального зала совещаний // Научный результат. Информационные технологии. 2018. Т.3. № 2.
4. *Ронжин Ал.Л., Ронжин Ан.Л.* Система аудиовизуального мониторинга участников совещания в интеллектуальном зале // Доклады ТУСУРа. 2011. № 1(22). Ч. 1. С. 153–157.
5. *Смирнов А.В. и др.* Подходы к выбору устройств ввода/вывода распределенных терминальных систем // Научные ведомости Белгородского государственного университета. Экономика. Информатика. 2019. Т. 46. № 2. С. 337–348.
6. *Alemdar H.Ö. et al.* Multi-modal fall detection within the WeCare framework // Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks. 2010. pp. 436–437.
7. *Hamm J., Money A.G., Atwal A., Paraskevopoulos I.* Fall prevention intervention technologies: A conceptual framework and survey of the state of the art // Journal of biomedical informatics. 2016. vol 59. pp. 319–345.
8. *Hamm J.J.* Technology-assisted healthcare: exploring the use of mobile 3D visualisation technology to augment home-based fall prevention assessments // Ph.D. Thesis. Brunel University London. 2018. 316 p.
9. *Papadopoulos A., Crump C., Wilson B.* Comprehensive home monitoring system for the elderly // Wireless Health 2010. 2010. pp. 214–215.
10. *Guan K., Shao M., Wu S.* A remote health monitoring system for the elderly based on smart home gateway // Journal of healthcare engineering. 2017. vol. 2017. 10 p.

11. *Шалковский А.Г., Купцов С. М., Берсенева Е.А.* Актуальные вопросы создания автоматизированной системы дистанционного мониторинга здоровья человека // Врач и информационные технологии. 2016. №. 1. С. 67–69.
12. *Giannoulis D., Benetos E., Stowell D., Plumbley M.D.* IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events – Training Dataset for Event Detection Task, subtasks 1 -OL and 2 – OS // Tech. Rep. 2012.
13. *Ширяев В.И.* Исследование операций и численные методы оптимизации // Ленанд. 2017. 224 с.
14. *Вентцель Е.С.* Исследование операций // М.: Советское радио. 1972. 552 с.
15. *Ficken F.A.* The simplex method of linear programming // Courier Dover Publications. 2015. 64 p.
16. *Цвиркун А.Д.* Основы синтеза структуры сложных систем // М.: Наука. 1982. 200 с.
17. *Баранов В.В., Цвиркун А.Д.* Управление развитием: структурный анализ, задачи, устойчивость // Автоматика и телемеханика. 2018. №. 10. С. 55–75.
18. *Corrado A.J.* Dynamics of complex systems // CRC Press. 2019.
19. *Brinsmead T.S., Hooker C.* Complex Systems Dynamics and Sustainability: Conception, Method and Policy // Philosophy of Complex Systems. 2011. vol. 10. pp. 809–838.
20. *Гадасин В.А., Ушаков И.А.* Надежность сложных информационно-управляющих систем // М.: Советское радио. 1975. 191 с.

Басов Олег Олегович — д-р техн. наук, старший научный сотрудник, факультет информационных технологий и программирования, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. (Университет ИТМО). Область научных интересов: инфокоммуникационная система, распределенная терминальная система, топологическая структура, доступность узлов, непрерываемость связи, проектирование полимодальных инфокоммуникационных систем. Число научных публикаций — 170. oobasov@mail.ru; Кронверкский пр., 49, 197101, Санкт-Петербург, Российская Федерация; р.т.: +7-919-201-18-97.

Саитов Игорь Акрамович — д-р техн. наук, профессор, сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: теория графов, теория массового обслуживания, теория вероятностей, теория множеств, применение методов математического моделирования в телекоммуникациях, анализ и синтез волоконно-оптических транспортных сетей. Число научных публикаций — 170. Akramovish@mail.ru; ул. Приборостроительная, 35, 302034, Орел, Российская Федерация; р.т.: +7(4862)549801.

Мотиенко Анна Игоревна — канд. техн. наук, старший научный сотрудник, отдел аспирантуры, информационно-образовательных технологий и услуг, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: инфокоммуникационная система, распределенная терминальная система, моделирование инфокоммуникационных систем медицинского назначения, сбор и анализ статистических медицинских данных, мониторинг здоровья населения. Число научных публикаций — 30. anna.gunchenko@gmail.com; 14-я линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7-812-328-03-82.

Астапов Сергей Сергеевич — Ph.D., доцент, кафедра программного обеспечения факультета информационных технологий, Таллинский Технический Университет. Область научных интересов: инфокоммуникационная система, распределенная терминальная система, топологическая структура, доступность узлов, непрерываемость связи. Число научных публикаций — 24. sergei.astapov@ttu.ee; Строительная дорога, 5, 19086, Таллин, Эстония; р.т.: +372 56 667 203.

Поддержка исследований. Работа выполнена при финансовой поддержке фонда РФФИ (проекты № 18-07-00380, № 19-07-00832).

DOI 10.15622/sp.2019.18.6.1357-1380

O.O. BASOV, I.A. SAITOV, A.I. MOTIENKO, S.S. ASTAPOV

SYNTHESIS OF THE TOPOLOGICAL STRUCTURE OF DISTRIBUTED TERMINAL SYSTEM FOR AUDIO MONITORING OF USERS OF LOCAL INFORMATION SPACES

Basov O.O., Saitov I.A., Motienko A.I., Astapov S.S, **Synthesis of the Topological Structure of Distributed Terminal System for Audio Monitoring of Users of Local Information Spaces.**

Abstract. A widespread use of multi-user interfaces, due to multimodality of traditional interpersonal communication, a transition to a polymerized presentation of information and systems, has allowed the creation of new approaches to their implementation based on distributed terminal systems. An approach to the synthesis of topological structures of such systems implemented in two stages is proposed in the article. The first stage determines a minimum set of communication nodes and their location based on the requirements for the availability of communication nodes for various categories of users and the globality of a distributed terminal system. The second stage determines options for constructing communication nodes and connections between them, which ensure the performance of audio monitoring functions of users of local information spaces while ensuring continuity of a bridge for different categories of users. A model example of the synthesis of a distributed terminal system for audio monitoring of two categories of users (adults and children) in the local information space (home), voice control subsystems of the "smart home" is presented. As a part of its solution, at each stage of the synthesis, the initial data are determined, a formal formulation of the synthesis problem is carried out, an algorithm for the solution and the results are presented. So the task of the first stage of the synthesis is a linear integer mathematical programming problem, solved in the model example by the simplex method, the solution of the second stage problem is based on the alternative graph formalization and the method of "branches and borders". The obtained results clearly demonstrate the capabilities of the proposed scientific and methodological tools for the synthesis of the topological structure of distributed terminal systems and the prospects of its use in the newly arising tasks of the technical implementation of new infocommunication technologies and services.

Keywords: Infocommunication System, Distributed Terminal System, Topological Structure, Node Availability, Communication Continuity.

Basov Oleg Olegovich — Senior Researcher, Information Technologies and Programming Faculty, ITMO University (Saint Petersburg National Research University of Information Technologies, Mechanics and Optics). Research interests: infocommunication system, distributed terminal system, topological structure, node availability, communication continuity, design of multimodal infocommunication systems. The number of publications — 170. oobasov@mail.ru; 49, Kronverksky pr., 197101, St. Petersburg, Russian Federation; office phone: +7-919-201-18-97.

Saitov Igor' Akramovich — Ph.D., Dr.Sci., Professor, Researcher, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: graph theory, queueing theory, probability theory, set theory, application of mathematical model approaches in telecommunications, analysis and synthesis fiber transport networks. The number of publications — 170. Akramovish@mail.ru; 35, Priborostroitelnaya str., 302034, Orel, Russian Federation; office phone: +7(4862)549801.

Motienko Anna Igorevna — Ph.D., Senior Researcher, Department of Post-Graduate Studies, Information and Educational Technologies and Services, St. Petersburg Institute for Informatics

and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: infocommunication system, distributed terminal system, Keywords modeling of information communication systems for medical purposes, collection and analysis of statistical medical data, public health monitoring. The number of publications — 30. anna.gunchenko@gmail.com; 39, 14-th Line V.O., 199178, St. Petersburg, Russian Federation; office phone: +7-812-328-03-82.

Astapov Sergei Sergeevich — Ph.D., Associate Professor, Department of Software Science of School of Information Technologies, Tallinn University of Technology. Research interests: infocommunication system, distributed terminal system, topological structure, node availability, communication continuity. The number of publications — 24. sergei.astapov@ttu.ee; 5, Ehitajate tee, 19086, Tallinn, Estonia; office phone: +372 56 667 203.

Acknowledgements. The research was supported by the Russian Foundation for Basic Research (projects No. 18-07-00380, 19-07-00832).

References

1. Basov O.O., Karpov A.A., Saitov I.A. *Metodologicheskie osnovy sinteza polimodal'nyh infokommunikacionnyh sistem gosudarstvennogo upravleniya: monografiya*. [Methodological Bases of Synthesis of Multimodal Communication Systems of Public Administration]. Orel: Akademiya FSO Rossii Publ. 2015. 277 p. (In Russ.).
2. Kiselev Yu.V., Motienko A.I., Basov O.O., Saitov I.A. [Structural-functional model of intelligent infocommunication system]. *Nauchno-tekhnicheskij vestnik informacionnyh tekhnologij, mekhaniki i optiki – Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 2018. Issue 18. vol. 6. pp. 1034–1046. (In Russ.).
3. Kuznetsov D.A. et all. [The Prerequisites of Creation of the Intellectual Hall of Meetings]. *Informacionnye tekhnologii – Information Technologies*. 2018. Issue 3. no. 2.(In Russ.).
4. Ronzhin Al.L., Ronzhin An.L. [The system of audiovisual monitoring of participants in the intelligent meeting room]. *Doklady TUSUR – Proceedings of TUSUR*. 2011. vol. 1(22). Chapter 1. pp. 153–157. (In Russ.).
5. Smirnov A.V. et al. [Approaches to Selection of Input and Output Devices of Distributed Terminal Systems]. *Nauchnye vedomosti Belgorodskogo gosudarstvennogo universiteta. Ekonomika. Informatika – Belgorod State University Scientific Bulletin. Series "Economics. Computer Science"*. 2019. Issue 46. vol. 2. pp. 337–348.
6. Alemdar H.Ö. et al. Multi-modal fall detection within the WeCare framework. Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks. 2010. pp. 436–437.
7. Hamm J., Money A.G., Atwal A., Paraskevopoulos I. Fall prevention intervention technologies: A conceptual framework and survey of the state of the art. *Journal of biomedical informatics*. 2016. vol. 59. pp. 319–345.
8. Hamm J.J. Technology-assisted healthcare: exploring the use of mobile 3D visualisation technology to augment home-based fall prevention assessments. Ph.D. Thesis. Brunel University London. 2018. 316 p.
9. Papadopoulos A., Crump C., Wilson B. Comprehensive home monitoring system for the elderly .Wireless Health 2010. 2010. pp. 214–215.
10. Guan K., Shao M., Wu S. A remote health monitoring system for the elderly based on smart home gateway. *Journal of healthcare engineering*. 2017. vol. 2017. 10 p.
11. Shalkovsky A.G., Kuptsov S.M., Berseneva E.A. [Topical issues of person remote monitoring automated system creation]. *Vrach i informacionnye tekhnologii – Information technologies for the Physician*. 2016. vol. 1. pp. 67–69. (In Russ.).
12. Giannoulis D., Benetos E., Stowell D., Plumbley M.D. IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events – Training Dataset for Event Detection Task, subtasks 1 -OL and 2 – OS . Tech. Rep. 2012.

13. Shiryaev V.I. *Issledovanie operacij i chislennye metody optimizacii* [Operations research and numerical optimization methods]. Lenand. 2017. 224 p. (In Russ.).
14. Ventcel' E.S. *Issledovanie operacij* [Operations research]. M.: Sovetskoe Radio. 1972. 552 c. (In Russ.).
15. Ficken F.A. The simplex method of linear programming. Courier Dover Publications. 2015. 64 p.
16. Tsvirkun A.D. *Osnovy sinteza struktury slozhnyh sistem* [Basics of the Structure Synthesis of Complex Systems]. M.: Sovetskoe Radio. 1982. 200 p. (In Russ.).
17. Baranov V.V., Tsvirkun A.D., [Development control: structural analysis, problems, stability]. *Avtomatika i telemehanika – Automation and Remote Control*. 2018. vol. 79(10). pp. 1780–1796.
18. Corrado A.J. Dynamics of complex systems. CRC Press. 2019.
19. Brinsmead T.S., Hooker C. Complex Systems Dynamics and Sustainability: Conception, Method and Policy. *Philosophy of Complex Systems*. 2011. vol. 10. pp. 809–838.
20. Gadasin V.A., Ushakov I.A. *Nadezhnost' slozhnyh informacionno-upravljajushhih system* [Robustness of a complex control information systems]. M.: Sovetskoe Radio. 1975. 191 p.

DOI 10.15622/sp.2019.18.6.1381-1406

L. BUREŠ, I. GRUBER, P. NEDUCHAL, M. HLAVÁČ, M. HRÚZ

SEMANTIC TEXT SEGMENTATION FROM SYNTHETIC IMAGES OF FULL-TEXT DOCUMENTS

Bureš L., Gruber I., Neduchal P., Hlaváč M., Hrúz M. **Semantic Text Segmentation from Synthetic Images of Full-Text Documents.**

Abstract. An algorithm (divided into multiple modules) for generating images of full-text documents is presented. These images can be used to train, test, and evaluate models for Optical Character Recognition (OCR). The algorithm is modular, individual parts can be changed and tweaked to generate desired images. A method for obtaining background images of paper from already digitized documents is described. For this, a novel approach based on Variational AutoEncoder (VAE) to train a generative model was used. These backgrounds enable the generation of similar background images as the training ones on the fly.

The module for printing the text uses large text corpora, a font, and suitable positional and brightness character noise to obtain believable results (for natural-looking aged documents).

A few types of layouts of the page are supported. The system generates a detailed, structured annotation of the synthesized image. Tesseract OCR to compare the real-world images to generated images is used. The recognition rate is very similar, indicating the proper appearance of the synthetic images. Moreover, the errors which were made by the OCR system in both cases are very similar. From the generated images, fully-convolutional encoder-decoder neural network architecture for semantic segmentation of individual characters was trained. With this architecture, the recognition accuracy of 99.28% on a test set of synthetic documents is reached.

Keywords: Generation of Synthetic Images, Semantic Text Segmentation, Variational Autoencoder, VAE, Optical Character Recognition, OCR, Aged-Looking Text Generation.

1. Introduction. Computer vision and machine learning fields are fast-growing, innovative, and well-researched fields nowadays. One of the research areas is Optical Character Recognition (OCR), which can be used for reading scanned documents. In this paper, we focus on digitized text documents and its corpus generation – we considered only the documents with typewritten text. The OCR methods are very often used as a follow-up step in the task of textual data digitization. E.g., full-text search, information retrieval, indexing, and document search can be achieved with the data digitalization for digital archiving purposes. Previously, most of OCR systems were rule-based systems, which were usually divided into a few steps – mainly detection and recognition phases. For instance, here are a few; OCRopus [4], GOCR, Ocrad, Tesseract OCR system [21], and many more. All noted systems have been published under the licenses which allow free use. Further, there exists multiple commercial systems, for instance: ABBYY FineReader or OmniPage. These days, methods of machine learning and artificial intelligence are used in the form of Deep Neural Networks (DNN). These approaches are mostly used in the way of end-to-end solutions, which cover all the necessary steps: detection,

localization, and reading [13-15, 24, 25]. The DNNs are successfully used in the area of text detection and recognition in everyday environment domain (e.g., a photo of a street – where text can be located on store signs, car plates, and other areas), so-called wild text reading. In other words, it is a task of detecting text in real-world environments, where text appears in a sparse form (not full-text documents like in our use case). A lot of noise and clutter are present in such data, and the text can be in any font/form/angle/deformation/style. A large amount of training data is required for the training of machine learning methods to be able to generalize the nature of the task. In the case of supervised methods also, the targets need to be available. Generally, one of the most time-consuming tasks in the process of training neural networks is to gather labeled data which are mostly labeled by humans and can be inaccurate. An algorithm for synthetic training data generation, which has been previously published in [10], can be used for the task of wild text reading. In their work, they estimate the depth of the image data. From this information, the geometry is estimated, and the synthetic text was transformed into this detected region to look natural. The algorithms for real-world text reading nowadays achieve good results, but they cannot be used for full-text document reading tasks. It is due to a significant difference in the task itself and in the image/scene appearance. The authors, in the paper [7], published a system for synthetic text generation, and it was for the aim of reading short texts in mobile environments. Our approach is different from theirs in this aspect. Our focus is mainly on full-text document generation, but they generated only short sentences. In this paper, we propose an automatic system for synthetic image data generation – for labeled training dataset creation. The dataset was further used for the training of multiple machine learning algorithms. The data source which we have used are collections of digitized full-text documents from the times after World War II from Czechoslovakia. This data we aim to emulate.

We followed and considerably extended our previously published work; [5] and [9]. The newly obtained results correspond to the ones presented and described by experiments in this paper.

Several algorithms were used in consecutive order to generate aged and natural-looking typewritten documents. In the first step, we automatically extract samples of document's backgrounds – aged-looking paper. We used our real document dataset for generating the empty documents (plain aged-looking paper backgrounds). Next, Variational AutoEncoder (VAE) [16] was trained and fine-tuned from obtained image dataset. Then, the VAE can generate backgrounds from random noise of given properties. Further, for the illusion of aged-looking text, we used brightness and per-character location noise to create even more authentic images. In the final phase, the corresponding text

annotation is automatically generated. Our process is capable of generating a huge amount of text documents – it needs only a text corpus as an input. The process of generation is language independent. Even though we used exact data in our experiments, the proposed system can generate images from any font and background. Only the proper dataset has to be provided. Next, we trained a DNN for semantic segmentation of the characters. During this segmentation process, a label is assigned to every image pixel. The assigning process is explained in [2], where it was used for road segmentation in urban environments. The method was updated and enhanced a few times, you can check [6] – e.g., network architecture, loss function and general approach to decoding were enhanced.

The main contributions of the paper and the described system are:

– an algorithm designed to remove text from the image of the full-text document. This part aims to keep the image as much natural-looking aged-document as possible;

– synthetic background generator based on VAE is trained by image data from the text removal algorithm (aged-looking documents). This output generates artificial/synthetic images – which are then used as background for generating our full-text typewritten documents;

– a system which is capable of generating trustworthy full-text typewritten documents. Moreover, the document layout can be edited during its generation. All layouts have to be defined before the generation starts;

– Full-text images are semantically segmented to provide the first step of an OCR system. Our character recognition accuracy is above 99%.

The paper is organized as follow: in the Introduction describes the topic of this paper and previous related works. As well as in the introduction, we discuss our contributions. Background Extraction section is focused on a text removal algorithm that is capable of obtaining document backgrounds.The way of generating synthetic artificial document backgrounds is described in the Background Generator section.In the next section, the Text Generator is described. With a synthetic background as an input, it is possible to generate an authentic-looking document. Moreover, it is possible to choose a layout of the text.Experiments and an evaluation of the proposed system are described in the next section. The main goal of this section is to verify whether the OCR results have similar accuracy for real and generated data. Semantic Segmentation section is focused on the description of a segmentation algorithm based on fully-convolutional encoder-decoder architecture. Moreover, the process is based on the segmentation of convex hulls. It is a different approach than the direct segmentation of characters.Summary and proposition of future work are discussed in the last section named Conclusion and Future Work.

2. Background Extraction. Our developed algorithm for image background extraction is described in this section. The background extraction algorithm needs digital scans of original typewritten image documents as input. Commonly the text is of a dark shade on a bright background. Our system expects it in this form. This assumption is the truth for the majority of tested free and commercial software (some can handle inverted colors too), but generally, all digitized documents usually meet this requirement. The developed algorithm is based on computing the mean color value, according to this assumption.

The algorithm loads image in RGB and grayscale color format, and the text is found with the use of Otsu's brightness thresholding method [19]. The thresholded binary image is used for localization of corresponding pixels where the text's pixels are stored as ones and non-text's pixels as zeros. The dilatation process with a square kernel of predefined suitable size is applied in order to suppress the effect of color values from the pixels between text and non-text in the mean color computation. The individual mean color of the background's pixels is calculated for every color component separately. The binary image, from the previous step, is used for masking, which removes all the text pixels from the computation. Each of the pixels which belong to a text label is replaced by mean values of the background.

Further steps are applied to our algorithm. The color difference between pixel values and calculated mean color can be large in some image areas – this can leads to brightness discontinuities which are unwanted. The re-computation process enhances the quality of the result by replacing the color in text pixels by a mean value from the text pixel's local neighborhood (NEIGHBOR_MEAN_ADAPT in the Algorithm 1. It creates a desired statistical bias towards the mean background value by using every pixel for the local mean computation, even the already replaced text pixels. The summary of the proposed method is shown in the Algorithm 1. An example of the results of the proposed algorithm is shown in Figure 1.

```
img_color, img_gray ← LOAD_DOCUMENT
binary ← OTSU(img_gray, BINARY_INVERSE)
binary ← DILATE(binary)
red, green, blue ← GET_CHANNELS(img_color)
red_mean ← MEAN(red[binary = 0])
green_mean ← MEAN(green[binary = 0])
blue_mean ← MEAN(blue[binary = 0])
red[binary = 1] ← red_mean
green[binary = 1] ← green_mean
```

```
blue[binary = 1] ← blue_mean
result ← [red, green, blue]
FOR NONZERO_PIXELS AS p IN binary:
   result[p] ← NEIGHBOR_MEAN_ADAPT(result, p, n_size=(k))
SAVE_IMG(result)
```

Algorithm 1. Background extraction algorithm

There are some flaws in the algorithm. When some dark artifacts are present in the input image, the algorithm is not able to remove them completely. Moreover, the procedure occasionally results in high contrast noise in the yielded image. However, these issues are well handled by the VAE method of generating synthetic backgrounds, and thus they do not need to be addressed by this algorithm.

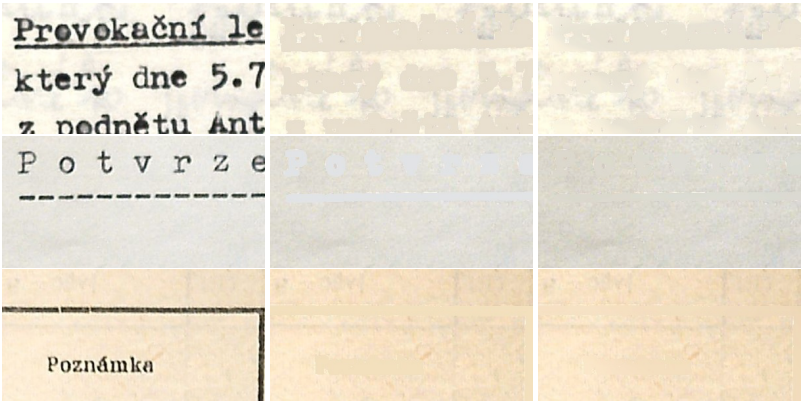


Fig. 1. Every step of the algorithm demonstrated on the tiles of input documents. The left column contains examples of input documents. The results after mean value replacement are presented in the middle column. The right column contains result after NEIGHBOR_MEAN_ADAPT – i.e., local adaptation of the mean color to eliminate "shadows" of erased characters

3. Background Generation. The task of synthetic image generation can be handled in different ways. This paper presents a method based on a neural network topology called autoencoder [3, 8]. This topology usually employs a bow-tie structure, where the first part encodes the input data into a latent space, and the second part then decodes the vector from latent space back to the shape of the input data. In the usual implementation of a plain

autoencoder, there are no restrictions applied to the latent space. Without any restrictions, the autoencoder acts as a memory with compression. To improve the idea of autoencoder into a generative model Variational Autoencoder was developed [12, 16, 23]. This approach forces the latent space to follow a Gaussian distribution. This typically leads to blurry images generated at the output of the VAE. We have discovered such behavior is to our advantage and developed the background generator for our synthetic data based on the VAE. The blur typical for the VAE introduced into the generated data helps us to improve the results of the Background Extraction algorithm (Section 2). The VAE eliminates the artifacts from the training set by compressing the data and then generating them back from the Gaussian latent space.

The training data for the VAE were obtained by Background Extraction algorithm in Section 2. The original scans of old documents were rid of the text by the background extraction algorithm, and 685 of them were used for training. The structure of the network is described in detail in Table 1. The latent space is created by two fully connected layers with size 250 neurons representing the mean and the variance of the Gaussian distribution with set values of zero mean and unit variance for each individual neuron. The training used the following parameters: learning rate of 0.001 with RMSprop optimizer over 1000 epochs. The input images were resized from the original resolution to the size of $128 \times 96 \times 3$. This size roughly represents the original ratio of the training data.

Only the decoder part of the VAE is used after training as a generator for synthetic backgrounds. The background is generated from an input vector of size 250. The values of the vector are sampled from the Gaussian distribution with zero mean and standard deviation of 0.15. This value was selected experimentally based on the analysis of the generated data. The images generated with a higher value of standard deviation produced various unrealistic backgrounds with different artifacts and colors which were not suitable for the task of synthesizing artificial images of old documents. Before the data can be processed by the decoder, there is a necessary step of resizing the single-dimensional vector of the input vector to a three-dimensional shape that can then be processed by deconvolutions. The needed shape is taken from the output of the last convolutional layer of the encoder. The output of the decoder is resized to the original resolution of the training data ($2480 \times 3504 \times 3$) using linear interpolation.

The main contribution of this approach is the ability to create a general background that is very similar to the real old paper. The algorithm is able to produce various colors from clean white paper, through gray paper, to yellow recycled paper typically used during older eras. The process removes typical

artifacts from the training data like water damage, coffee circle stains, and letters visible from the other side of the paper, which are typically present due to the nature of the typewriting. These may be added to the generated background if needed in our future research.

Table 1. Structure of VAE

Encoder	Decoder
64 conv(2×2), ReLU	deconv(3×3), ReLU
64 conv(2×2), BN, ReLU	deconv(3×3), ReLU
64 conv(3×3), ReLU	deconv(3×3), ReLU
64 conv(3×3), BN, ReLU	conv(2×2), sigmoid
500 fully-connected	

The encoder is composed of four convolutional layers with 64 kernels with the ReLU activation function. Every other convolutional layer is followed by batch normalization. The intermediate layer with 500 neurons has tanh activation function. The latent space is represented by two fully connected layers with 250 neurons and a linear activation function. The decoder mirrors the structure of the encoder.

4. Synthetic Image with Text Generation. The first part of creating a synthetic image is generating the background. The background image is created by our trained VAE model described in Section 3. The VAE neural network is fast enough to generate the images on the fly. The second part is adding synthetic text to the background. The font used to generate the synthetic text is Bohemian typewriter. A sample image with the font is depicted in Figure 2. We have used this font as it was corresponding to our real data from scanned typewritten documents. The font size was experimentally set to 45 pixels. This size is the closest to the real data, and it also depends on the resolution of the background. The font can be easily replaced by any other desirable font. This change does not impact the performance of our algorithm.

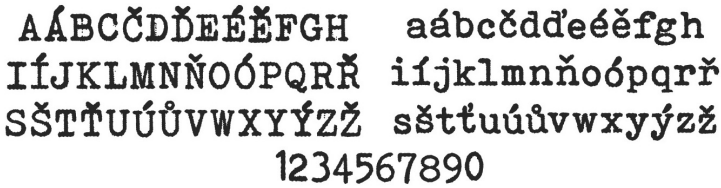


Fig. 2. Bohemian typewriter font used for generating synthetic documents

We have implemented an input from a text file to generate the text in the synthetic images. The text in the file can be composed of any UTF-8 characters. The contents of the text file are not limited by the output font. If the output font does not contain the character from the input file, it is replaced by a dummy placeholder.

The background is generated after the input file is processed. Our VAE neural network generator outputs an image with a resolution of 2480×3504. This resolution was chosen to correspond to the scanned documents in our dataset. The image has three color channels (RGB). Generated backgrounds are unique and unpredictable in the sense of color and noise. The general distribution of colors generated by our VAE network is given by the distribution in the training set. The pre-trained VAE model can be easily changed for another model if another type of background is needed.

The first instance of the text is generated on a white background with a resolution corresponding to the synthetic background (2480×3504). The white background is used to print out the characters with zero brightness. The text output is formatted into a predefined block. The block corresponds with a predefined area in the image where the text is supposed to be placed. These areas are easily added and/or modified to create a desired format of the output document, for example, two columns, two paragraphs, zig-zag, etc. The position of the area is given by the set of coordinates (x,y) that represents the top left corner. The second parameter of the area is its width and height of (w,h). The text is continuously printed into the predefined areas until all of them are filled, or there is no more text on the input. After that, if any text is remaining, it is omitted. A random offset is added to the coordinates (x,y) to add a more natural appearance to the synthetic data. Examples of the generated images are in Figures 3 and 4.

The position of every character is calculated with respect to its size. This position is further augmented with an offset generator. The offset can reach values of ± 3 pixels in both directions of x and y coordinates. This value is randomly added to each character. We have implemented this generator to simulate better the text produced by older types of typewriters where all the characters were not printed in a straight line. The offsets were selected experimentally to respect the size of the font and the resolution of the whole image. The value of the parameter can be easily changed if needed.

Various defects can be found in archived typewritten documents. One of them is a variance in the brightness of the old paper versus a new paper. We introduce a method for simulating this effect in our synthetic data. We generate a Gaussian noise $\mathcal{N}\left(\mu, \sigma^2\right)$ (where $\mu = 0$ and $\sigma^2 = 0.3$) and reshape it into the size width = $\frac{2480}{8}$ and height = $\frac{3504}{8}$. This noise matrix is then reshaped

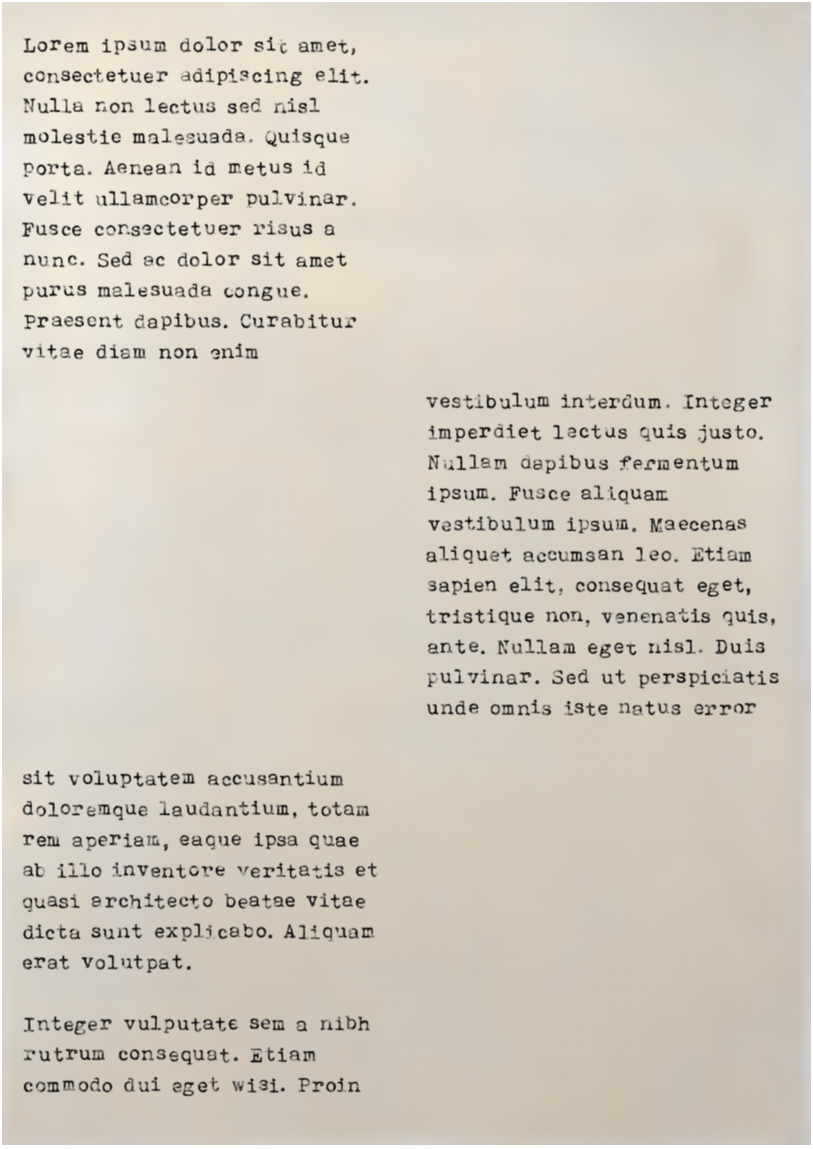


Fig. 3. Example of the layout areas: zig-zag text areas

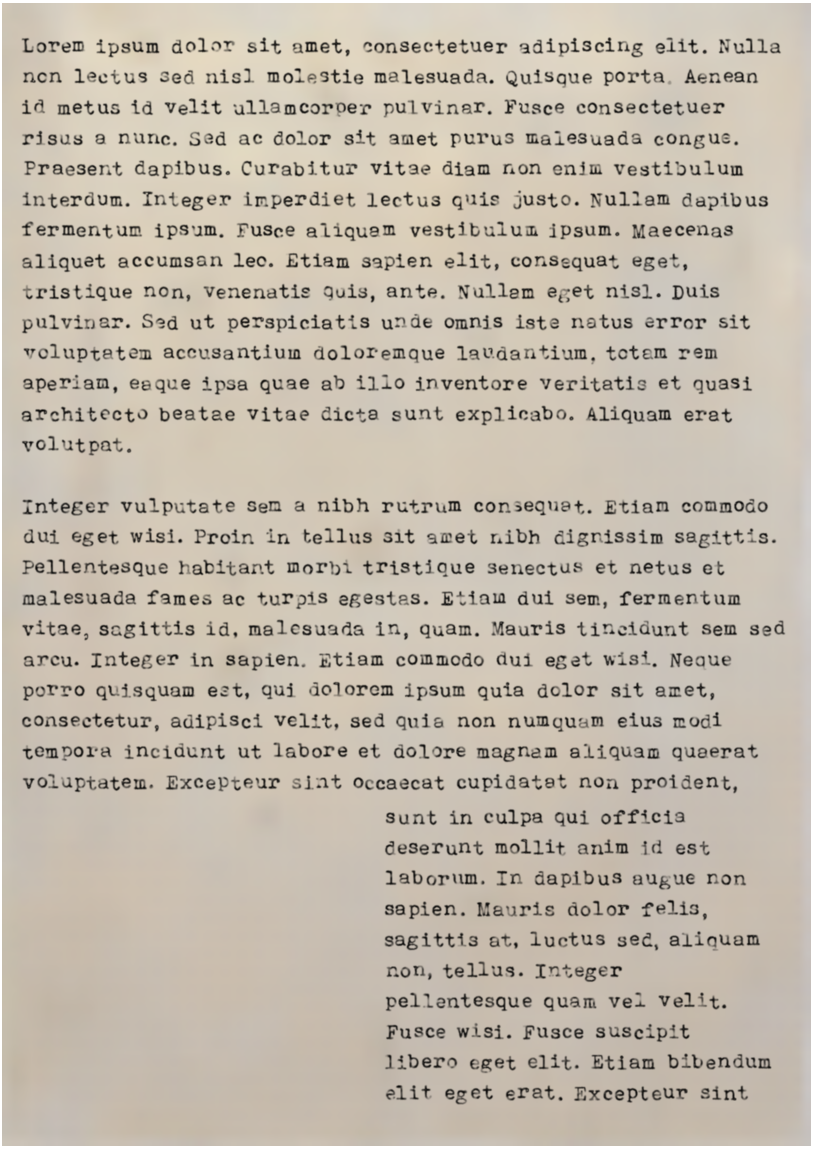


Fig. 4. Example of the layout areas: change of the width in the middle of the second paragraph

into the size of the synthetic image using linear interpolation. The noise is then added to the synthetic image (at the locations where the characters of the text are present). The final values are then clipped for the purpose of the consistency of the image data. The resulting image is further augmented by a blurring filter with a local averaging window of size 7×7. This process produces a synthetic image with blurred text areas. This image is then merged with the synthetic background generated by our VAE network. One last operation in the form of a blur filter is then applied to the whole image using a local averaging window of size 5×5.

We can create a bounding box from the knowledge of the parameters of each character (coordinates (x, y), character offset, and font size). The size of the bounding box is constant since the font width of the typewriter is not changing during writing. We have aimed for bounding boxes that are considered minimal. We have implemented a method to decrease the size of the bounding boxes. The synthetic image is transformed into grayscale at first. Then the contour of each character is extracted. Bounding boxes are then resized to touch the contours found in the previous step. The same has been done for page's, word's, and line's bounding boxes. Examples of character bounding boxes are shown in Figure 5.

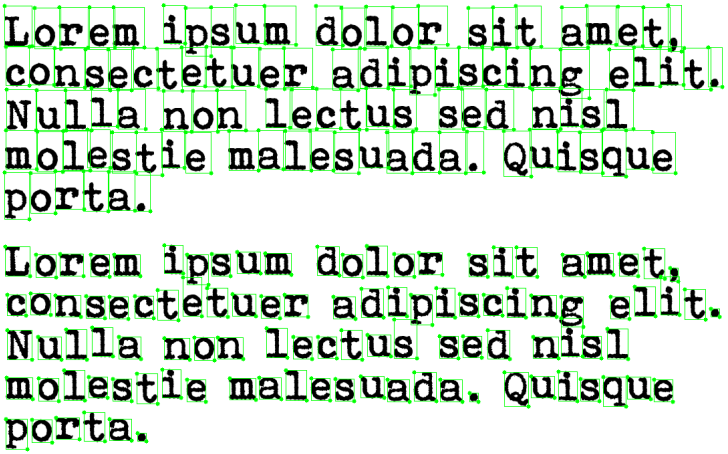


Fig. 5. Examples of character's bounding boxes: (top) the original bounding boxes, (bottom) fine–tuned character's bounding boxes

All collected data about images, words, characters, and lines are stored in files (in pickle data files and png images). This data can be utilized in further experiments and analysis.

The pickle data file for every page is structured in the following manner: the printing areas of a defined layout store the objects of lines. The line object contains a list of word objects, and each word object contains a list of character objects. The objects contain precise x and y global coordinates and the height and width of its bounding box. The file also contains information about the number of objects. We have also generated pixel-wise data with binary information about each character's pixel coordinates. This information was obtained by thresholding the white background image with the generated text before it was merged with the VAE background. The storage structure is depicted in Figure 6.

The data structure is designed to be universally usable in various tasks. We have used it for training several types of neural networks. It can also be used as ground truth data for benchmarking existing algorithms. A comparison of generated vs real text is depicted in Figure 7.

5. Experiments. To test the efficiency of the proposed method, we have designed an experiment to measure the accuracy of an OCR system on both the generated synthetic data and real scanned documents. We have chosen the Tesseract system for the OCR task. The hypothesis is that if the generated data have valid properties, Tesseract should perform similarly on the generated data as on the real data. For this purpose, we manually annotated 24 documents with a total of 25 292 valid characters. The formula to compute the accuracy was:

$$p = \left(1 - \frac{l}{m}\right) \cdot 100 \quad [\%], \tag{1}$$

where l is the Levenshtein distance between the two strings representing the OCR results of synthetic, respectively real data and m is the length of the longer string. The strings were rid of white characters, that do not bear any valuable information, such as spaces, newlines, dots, colons, etc.

Another metric we used was Word recall. In this case, we count the percentage of words from the ground truth text that we are able to recover with the OCR system. In the experiments, four datasets were tested. The first one is a dataset of real documents, and the other three are artificial datasets of documents generated by our proposed algorithms from the same text, as is in the real dataset. The artificial synthetic datasets were generated with various amount of character noise – i.e., damage in the structure of printed characters in the scan (see Section 4). The results are shown in Table 2. It can be seen that we are able to control the properties of the algorithm so that the results of the OCR system are close to the results from the real data. In Table 2 you can see Mean accuracy – i.e., character accuracy of OCR system, standard

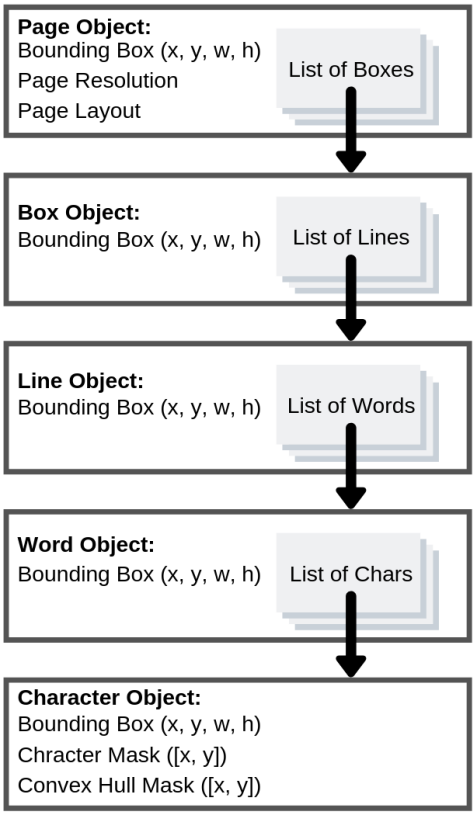


Fig. 6. Pickle data file structure hierarchy

deviation, and word recall. The word recall is important for measuring the performance of the OCR system for the purpose of information retrieval.

The results in Table 2 show that the dataset denoted Generated three is similarly challenging for the Tesseract system as the real data. Moreover, when observing the recognized texts, the OCR makes very similar mistakes as in the real scanned documents. The character noise used in this dataset can be used for generating benchmark datasets for comparing state of the art OCR systems. We can also control the amount of character noise to generate less challenging data. This attribute of the algorithm can be beneficial for training DNNs. First, a DNN can be trained on easier data, and then the difficulty of

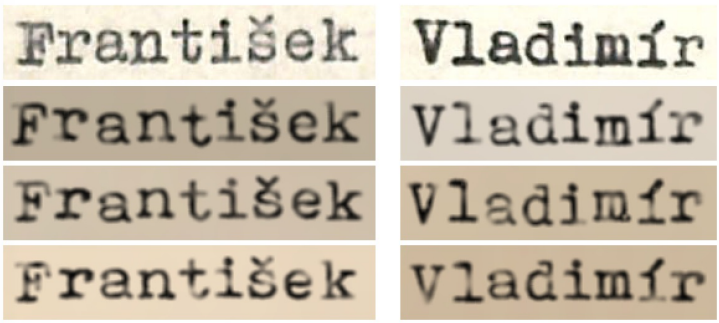


Fig. 7. Examples of generated text. (top row) original scanned text sample. (following rows) generated text

Table 2. OCR accuracy of real and generated scans. The generated datasets have different amount of noise applied: Generated 1 < Generated 2 < Generated 3

Dataset	Mean accuracy	STD	Word recall
Real data	0.73	0.14	0.79
Generated 1	0.94	0.08	0.90
Generated 2	0.92	0.10	0.83
Generated 3	0.88	0.12	0.69

the data can be progressively raised. One can also generate baseline datasets that should be recognized flawlessly.

6. Semantic Segmentation. Inspired by [2, 18, 20], we propose a method based on a Fully-Convolutional neural network with Encoder-Decoder structure, which proved itself to be perfectly suitable for semantic segmentation tasks. First, we designed a simple plain Encoder-Decoder network (our baseline architecture, for exact configuration, see Table 3), nevertheless, this setup did not provide satisfactory results. Therefore, we implemented a couple of improvements. To be more specific, our final neural network architecture is based on the work of Ronnenberger et al. [20] and their U-Net architecture (see Table 4). In comparison to standard Encoder-Decoder, this architecture utilizes skip connections between each layer i in the encoder and layer n-i in the decoder, where n is the total number of layers. The skip connections are utilized for element-wise addition of the activations from layer i and layer n-i. Experiments proved that usage of skip connections allowed better propagation of information through the network and, therefore, significantly improved

classification (or segmentation) results. In our future work, we would also like to test the concatenation of the activations instead of their summation.

A good practice before designing the architecture of a neural network is to understand the input data. In the case of visual processing, the receptive field is an important attribute of the individual layers. Since we want to read a text composed into lines, we need to figure out the average height of a text line so that the receptive field of the deepest layer is in compliance with this height. This approach is supported by the fact that text is highly dependent in one line and less dependent in the vertical direction. In the original resolution of 2480×3504 pixels, the average line-height is 36 pixels. Due to the memory limitations during training, the original resolution was reduced to 620×876, thus, the average height in the lower resolution is 9 pixels.

During the design of the neural network architecture, we utilize standard (de)convolution-batch normalization-ReLU sequence for all the convolutions and deconvolutions except the last one, which performs classification task and therefore uses classical Softmax activation function. All the convolutions and deconvolutions in this work are used with stride 1. We would also like to point out the total max-pooling operation omission, which stems from the optimal size of the convolutional receptive field considering the line average height.

Table 3. Structure of our baseline architecture

Encoder	**Decoder**
Conv(16, 3×3), BN, ReLU	Deconv(64, 5×5), BN, ReLU
Conv(32, 3×3), BN, ReLU	Deconv(32, 3×3), BN, ReLU
Conv(64, 5×5), BN, ReLU	Deconv(16, 3×3), BN, ReLU
	Conv(N, 1×1), Softmax

The encoder is composed of three convolutional layers with 16, 32, and 64 kernels, followed by batch normalization and ReLU activation function. The decoder mirrors this structure. N in the last convolutional layer of the decoder is the number of classes.

Table 4. Structure of our final architecture

Encoder	**Decoder**
Conv1(16, 3×3), BN, ReLU	Deconv1(64, 5×5), BN, ReLU
Conv2(32, 3×3), BN, ReLU	Deconv2(32, 3×3), BN, ReLU
Conv3(32, 3×3), BN, ReLU	Deconv3(32, 3×3), BN, ReLU
Conv4(64, 3×3), BN, ReLU	Deconv4(16, 3×3), BN, ReLU
	ConvF(N, 1×1), Softmax

The encoder is composed of four convolutional layers with 16, 32, 64, and 64 kernels, followed by batch normalization and ReLU activation function again. The decoder mirrors this structure. *N* in the last convolutional layer ConvF of the decoder is the number of classes. There is a skip connection between each layer *i* and layer *n-i*.

All tested neural network architectures were implemented in Python using Chainer deep learning framework [1, 22]. Also, in all experiments, we use the cross-entropy loss for the network training. In the following subsections, we describe the experimental setting and the results of two experiments. For both of them, we are providing quantitative results for testing synthetic data and qualitative results for both synthetic and real data.

6.1. Two-class classification. In the first experiment, we train baseline architecture to perform semantic segmentation of input image into two classes: text, and background, i.e., last convolutional layer has two kernels – one for each class and the softmax activation function is used to obtain a probability measure in each location. To further elaborate, for each class, the network produces one output map with segmentation for this specific class. The final segmentation map can be constructed by combining the results from all output maps in a simple manner.

Standard mini-batch SGD optimization with mini-batch size 2 – memory limitation – is used in our architecture. We used the learning rate with step decay. In particular, the learning rate started on value 0.01, and every 10 epochs, the step decay 0.1 was applied. The recognition accuracy of the network on the testing set is 99.28% after 30 epochs. An example of a result of a synthetic document segmentation is shown in Figure 8.

Our dataset contains 150 thousand synthetic documents divided into three sets – i.e., training, development, and testing set. The total number is split as follows. The training set contains 100 thousand documents, the development set 20 thousand documents, and finally testing set contains 30 thousand documents.

In Figure 9, an example of real document segmentation is shown. It can be noted that the network is capable of filtering out the document background. Moreover, it even removes paper inaccuracies and other defects. On the other hand, the network leaves parts of the text which are not wanted in our result. In the example, the underline under characters is unwanted because it can influence the accuracy of the OCR algorithm. The reason for this phenomenon is caused by the absence of this type of distraction in the synthetic data. Nevertheless, we argue that this approach can be used as a part of a pre-processing pipeline for the standard OCR algorithm and hopefully improves its final results.

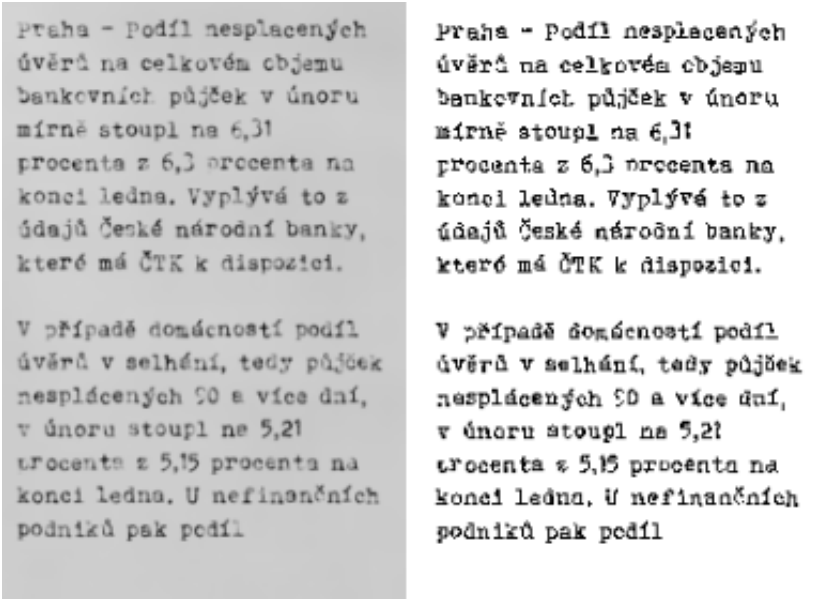


Fig. 8. A result (on the right) of the semantic segmentation for the class *text* of the synthetic document (on the left)

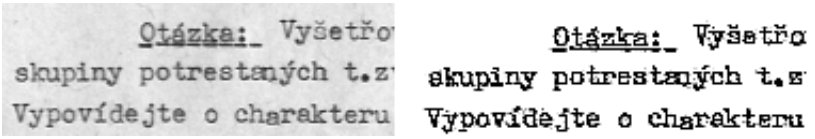


Fig. 9. A result (on the right) of the semantic segmentation for the class *text* of the real historical document (on the left)

Segmentation of convex hulls instead of direct character segmentation is used in our research. The idea was proposed in the paper Hajic et al. [11]. The paper contains a statement that the training of the network should be better with a convex hull approach. The idea is verified on the task of musical symbol recognition. In Figure 10, an example of convex hulls of letters a, b, and c are shown.

To try convex hulls on our segmentation problem, we had to regenerate synthetic data labels. We trained the network using the same training settings to predict hulls instead of characters. The recognition accuracy was improved by 0.27% on the test set using this approach. It represents an improvement of 37.50% w.r.t. the approach based on the direct segmentation of characters. On

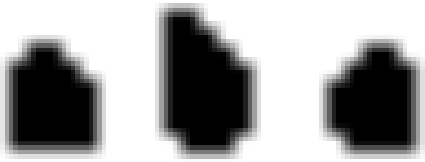


Fig. 10. Examples of convex hulls for letters *a*, *b*, and *c*

the other hand, it does not solve any of the systematic problems present in the first method.

Table 5. Baseline architecture results of per-pixel recognition accuracy of two-class segmentation on the synthetic documents

NN architecture	Development set	Test set
Direct-char	99.32%	99.28%
Convex-hull	99.60%	99.55%

Generation of the problematic unwanted parts of the text is one of the tasks that we want to address in our future research. Their presence in the training sets should improve the results of the neural network.

6.2. Single-character classification. The second experiment is focused on the semantic segmentation into 107 classes – i.e., classes of Czech language text characters plus one for the background. This task is called a single-character classification. We assume that it can avoid problems of two-class segmentation. During the experiment, we train the network using both direct-character and convex hull segmentation. The recognition accuracy of the network reached 97.75%, 97.58%, respectively.

Despite the good quality of results, we find out that the network provides almost perfect prediction for frequent characters – e.g. `a` and `e` – but significantly worse results for the rare ones – e.g. `F` and `G`. The reason for this fact is the nature of the Czech language because some of the characters are rare. Consistently, there are a small number of occurrences of these characters in the training data.

There are basically two possible solutions to overcome this flaw. First, we can regenerate the training set with a uniform distribution of the characters. This approach will definitely solve the problem for the synthetic data, but it will likely provide worse results for the real documents, which meet the unbalanced frequency of characters of the Czech language.

Table 6. Comparison per-pixel recognition accuracy of single-character segmentation on the synthetic document test set for the baseline and the final architecture

NN architecture	Direct-char	Convex-hull
Baseline	97.75%	97.58%
Baseline_weighted	97.83%	97.75%
Final	99.14%	99.17%
Final_weighted	99.53%	99.52%

The second option is that we preserve the same training set, but we motivate our network to stop ignoring the rare classes. Therefore, we weight the loss from individual classes w.r.t. their frequency, i.e., the loss from the less frequent classes introduces a larger penalty than from the frequent ones. The frequencies of individual characters were obtained from 10 thousand real internet news documents. Weighted categorical cross-entropy slightly improves results. However, the obtained results still were not satisfying. We tested a few modifications to our baseline neural network architecture. In Table 6, the results of the baseline and our final architecture can be found. The results while using standard categorical cross-entropy are listed too. We would like to point out the fact that convex-hull prediction reached only comparable results despite that in theory, it should significantly improve them. In our opinion, it is caused by the relative similarity of the characters' convex hull and the characters themselves as opposed to the findings in the work [11] with convex hulls of musical symbols. We are also providing qualitative results (see Figure 11 and Figure 12) of the final architecture using weighted cross-entropy loss.

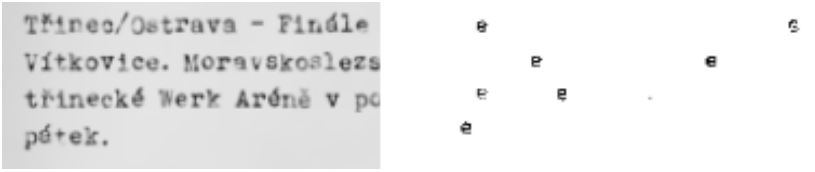


Fig. 11. A result (on the right) of the semantic segmentation for letter *e* of a synthetic document (on the left) using the final architecture using weighted cross-entropy loss

The final network architecture and training process reached good accuracy. Therefore, we want to use it as a pre-processing part for a novel OCR algorithm for scanned historical documents in our future research. On the other hand, there are still problems with the segmentation of rare characters that should be addressed. There is also some noise caused by the appearance of

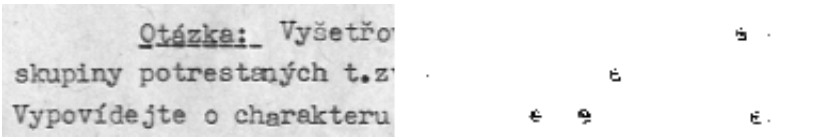


Fig. 12. A result (on the right) of the semantic segmentation for letter *e* of the real historical document (on the left) using the final architecture using weighted cross-entropy loss

scanned documents. In the future, we want to try more complex architectures together with an extended training set to solve these problems. Moreover, the parallelization of the training process can increase the batch size and lead to results of better quality.

7. Conclusion and Future Work. We have presented an algorithm for generating synthetic images resembling real-world scanned typewritten documents. The algorithm is highly modifiable in terms of generating different classes of documents. The algorithm is composed of modules handling smaller tasks. They can be easily switched and modified, which also enables appearance tweaks for any given class of documents. The modules include a dynamic generation of backgrounds using VAE, different standard fonts, and layout generation. The attributes of the printed text can be easily interchanged. There is the possibility of location noise addition, blurring of the printed text, or the whole generated image. We test the properties of the generated documents, and we observe that through the lens of an existing OCR system, they are very similar to real-world scanned documents. The generated data can be used to train, test, and evaluate new or existing OCR algorithms. In this work, we utilize the algorithm to generate vast amounts of training data for our semantic segmentation approach of text detection and character classification. We use an approach called semantic segmentation, which involves a fully-convolutional neural network.

Two different tests were performed: two-class classification and character classification. With the best training settings and final fully-convolutional encoder-decoder architecture inspired by U-Net, we reached 99.28% recognition accuracy, 99.52% respectively, for the synthetic data. We also proposed segmentation of convex hulls instead of direct segmentation of characters, which, unfortunately, did not provide any legible improvement. We also provide qualitative results for real data. For our future research, we have several plans. First, we would like to extend our generator with options to generate new text features like text underlining or line margins. Second, we plan to use two-class classification results as a part of a pre-processing pipeline for the standard

OCR algorithm. Third, we will employ more complex architecture in the character classification task. Last, but not least, we will develop a text decoder to be able to compare reached results with other OCR algorithms.

References

1. Akiba T., Fukuda K., Suzuki S. ChainerMN: Scalable distributed deep learning framework. arXiv preprint arXiv:1710.11351. 2017.
2. Badrinarayanan V., Kendall A., Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. vol. 39(12). pp. 2481–2495.
3. Bengio Y. et al. Learning deep architectures for AI. Foundations and trends® in Machine Learning. 2009. vol. 2(1). pp. 1–127.
4. Breuel T. Recent progress on the OCRopus OCR system. Proceedings of the International Workshop on Multilingual OCR. 2009. pp. 2.
5. Bureš L., Neduchal P., Hlavác M., Hrúz M. Generation of synthetic images of full-text documents. International Conference on Speech and Computer. 2018. pp. 68–75.
6. Chen L.C., Papandreou G., Schroff F., Adam H. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587. 2017.
7. Chernyshova Y.S., Gayer A.V., Sheshkus A.V. Generation method of synthetic training data for mobile OCR system. Tenth International Conference on Machine Vision (ICMV 2017). 2018. vol. 10696. pp. 106962G.
8. Dumas T., Roumy A., Guillemot C. Autoencoder based image compression: can the learning be quantization independent? 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018. pp. 1188–1192.
9. Gruber I., Hlavác M., Hrúz M., Železný M. Semantic segmentation of historical documents via fully-convolutional neural network. International Conference on Speech and Computer. 2019. pp. 142–149.
10. Gupta A., Vedaldi A., Zisserman A. Synthetic data for text localization in natural images. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2016. pp. 2315–2324.
11. Hajic J., Dorfer M., Widmer G., Pecina P. Towards full-pipeline handwritten OMR with musical symbol detection by u-nets. ISMIR. 2018. pp. 225–232.
12. Huang H., He R., Sun Z., Tan T. Introvae: Introspective variational autoencoders for photographic image synthesis. Advances in Neural Information Processing Systems. 2018. pp. 52–63.
13. Huang W., Qiao Y., Tang X. Robust scene text detection with convolution neural network induced mser trees. European Conference on Computer Vision. 2014. pp. 497–511.
14. Jaderberg M., Vedaldi A., Zisserman A. Deep features for text spotting. European Conference on Computer Vision. 2014. pp. 512–528.
15. Jaderberg M., Simonyan K., Vedaldi A., Zisserman A. Reading text in the wild with convolutional neural networks. International Journal of Computer Vision. 2016. vol. 116(1). pp. 1–20.
16. Kingma D.P., Welling M. Auto-encoding variational bayes. International Conference on Learning Representations. 2014. 21 p.
17. Lin G. et al. Refinenet: Multi-path refinement networks for dense prediction. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019.
18. Noh H., Hong S., Han B. Learning deconvolution network for semantic segmentation. Proceedings of the IEEE International conference on computer vision. 2015. pp. 1520–1528.

19. Otsu N. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics. 1979. vol. 9(1). pp. 62–66.
20. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. 2015. pp. 234–241.
21. Smith R. An overview of the tesseract OCR engine. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007). 2007. vol. 2. pp. 629–633.
22. Tokui S., Oono K., Hido S., Clayton J. Chainer: a next-generation open source framework for deep learning. Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Twenty-ninth Annual Conference on Neural Information Processing Systems (NIPS). 2015. vol. 5. pp. 1–6.
23. Wen S. et al. Variational autoencoder based image compression with pyramidal features and context entropy model. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2019. pp. 0–0.
24. Zhao H. et al. Pyramid scene parsing network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 2881–2890.
25. Zhou X. et al. EAST: An efficient and accurate scene text detector. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 5551–5560.

Lukáš Bureš — Ph.D. Student, New Technologies for the Information Society (Research centre of Faculty of Applied Sciences), University of West Bohemia. Research interests: machine learning, computer vision, visual keypoint detection and description. The number of publications — 1. lbures@ntis.zcu.cz; 8, Technická, 306 14, Plzen, Czechia; office phone: +420 377 632 145.

Ivan Gruber — Ph.D. Student, NTIS - New Technologies for the Information Society (Research Centre of Faculty of Applied Sciences), University of West Bohemia. Research interests: machine learning, face recognition, image recognition. The number of publications — 14. grubiv@ntis.zcu.cz; 8, Technická, 306 14, Plzen, Czech Republic; office phone: +420 377 632 137.

Petr Neduchal — Ph.D. Student, NTIS - New Technologies for the Information Society (Research Centre of Faculty of Applied Sciences), University of West Bohemia. Research interests: mobile robotics, simultaneous localization and mapping, robot exploration, computer vision, machine learning. The number of publications — 7. neduchal@kky.zcu.cz; 8, Technická, 306 14, Plzen, Czech Republic; office phone: +420 377 632 145.

Miroslav Hlaváč — Ph.D. Student, NTIS - New Technologies for the Information Society (Research Centre of Faculty of Applied Sciences), University of West Bohemia. Research interests: machine learning, image recognition, lip reading, audiovisual speech recognition. The number of publications — 1. mhlavac@ntis.zcu.cz; 8, Technická, 306 14, Plzen, Czech Republic; office phone: +420 377 632 136.

Marek Hrúz — Ph.D., Senior Researcher, NTIS - New Technologies for the Information Society (Research Centre of Faculty of Applied Sciences), University of West Bohemia. Research interests: computer vision, machine learning, deep learning, data mining, multi-modal data processing. The number of publications — 42. mhruz@ntis.zcu.cz; 8, Technická, 306 14, Plzen, Czech Republic; office phone: +420 377 632 555.

Acknowledgements. This work was supported by the Ministry of Education of the Czech Republic, project No. LTARF18017 and Ministry of Education, Youth and Sports of the Czech Republic project No. LO1506. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the

programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated. The work has been supported by the grant of the University of West Bohemia, project No. SGS-2019-027.

УДК 004.9 **DOI** 10.15622/sp.2019.18.6.1381-1406

Л. Буреш, И. Грубер, П. Недухал, М. Главач, М. Груз

СЕГМЕНТАЦИЯ СЕМАНТИЧЕСКОГО ТЕКСТА ПО ИСКУССТВЕННОМУ ИЗОБРАЖЕНИЮ ПОЛНОТЕКСТОВЫХ ДОКУМЕНТОВ

Буреш Л., Грубер И., Недухал П., Главач М., Груз М. **Сегментация семантического текста по искусственному изображению полнотекстовых документов.**

Аннотация. Предлагается разделенный на несколько модулей алгоритм для создания изображений полнотекстовых документов. Эти изображения можно использовать для обучения, тестирования и оценки моделей оптического распознавания символов (ОПР). Алгоритм является модульным, отдельные части могут быть изменены и настроены для создания желаемых изображений. Описывается метод получения фоновых изображений бумаги из уже оцифрованных документов. Для этого используется новый, основанный на вариационном автоэнкодере подход к обучению генеративной модели. Эти фоны позволяют сразу же сгенерировать такие же фоновые изображения, как те, на которых производилось обучение.

Для получения правдоподобного эффекта старения в модуле печати текста используются большие текстовые блоки, типы шрифтов и вариативность изменения яркости символов.

Поддерживаются несколько типов макетов страницы. Система генерирует подробную структурированную аннотацию искусственного изображения. Для сравнения реальных изображений с искусственно созданными используется программа Тессеракт ОПР. Точность распознавания приблизительно схожа, что указывает на правильность сгенерированных искусственных изображений. Более того, допущенные системой ОПР ошибки в обоих случаях очень похожи. На основе сгенерированных изображений была обучена архитектура сверточная кодер-декодер нейронная сеть полностью для семантической сегментации отдельных символов. Благодаря этой архитектуре достигнута точность распознавания 99,28% в тестовом наборе синтетических документов.

Ключевые слова: генерация искусственных изображений, сегментация семантического текста, вариационный автоэнкодер, OCR, оптическое распознавание символов, распознавание текста, генерация искусственно состаренного текста.

Буреш Лукаш — аспирант, НТИО - Новые технологии для информационного общества (исследовательский центр факультета прикладных наук), Западно-чешский университет. Область научных интересов: машинное обучение, компьютерное зрение, визуальное обнаружение и описание ключевых точек. Число научных публикаций — 1. lbures@ntis.zcu.cz; ул. Техническая, 8, 306 14, Пльзень, Чехия; р.т.: +420 377 632 145.

Грубер Иван — аспирант, НТИО - Новые технологии для информационного общества (исследовательский центр факультета прикладных наук), Западно-чешский университет. Область научных интересов: машинное обучение, распознавание лиц, распознавание изображений. Число научных публикаций — 14. grubiv@ntis.zcu.cz; ул. Техническая, 8, 306 14, Пльзень, Чехия; р.т.: +420 377 632 137.

Недухал Петр — аспирант, НТИО - Новые технологии для информационного общества (исследовательский центр факультета прикладных наук), Западно-чешский университет. Область научных интересов: мобильная робототехника, одновременная локализация и картирование, исследование робота, компьютерное зрение, машинное обучение. Число

научных публикаций — 7. neduchal@kky.zcu.cz; ул. Техническая, 8, 306 14, Пльзень, Чехия; р.т.: +420 377 632 145.

Главач Мирослав — аспирант, НТИО - Новые технологии для информационного общества (исследовательский центр факультета прикладных наук), Западно-чешский университет. Область научных интересов: машинное обучение, компьютерное зрение, лабиомансия, аудиовизуальное распознавание речи. Число научных публикаций — 1. mhlavac@ntis.zcu.cz; ул. Техническая, 8, 306 14, Пльзень, Чехия; р.т.: +420 377 632 136.

Груз Марек — Ph.D., старший научный сотрудник, НТИО - Новые технологии для информационного общества (исследовательский центр факультета прикладных наук), Западно-чешский университет. Область научных интересов: компьютерное зрение, машинное обучение, глубинное обучение, интеллектуальный анализ данных, мультимодальная обработка данных. Число научных публикаций — 42. mhruz@ntis.zcu.cz; ул. Техническая, 8, 306 14, Пльзень, Чехия; р.т.: +420 377 632 555.

Поддержка исследований. Данное исследование проведено при поддержке Министерства образования Чешской Республики (проект № LTARF18017) и Министерства образования, молодежи и спорта Чешской Республики (проект № LO1506). Также благодарим за предоставление доступа к хранилищам, принадлежащим участникам создания метацентра национальной грид-инфраструктуры в рамках программы «Проекты крупных инфраструктур для исследований, разработок и инноваций» (CESNET LM2015042). Работа выполнена при поддержке гранта Западно-чешского университета (проект № SGS-2019-027).

Литература

1. *Akiba T., Fukuda K., Suzuki S.* ChainerMN: Scalable distributed deep learning framework // arXiv preprint arXiv:1710.11351. 2017.
2. *Badrinarayanan V., Kendall A., Cipolla R.* Segnet: A deep convolutional encoder-decoder architecture for image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. vol. 39(12). pp. 2481–2495.
3. *Bengio Y. et al.* Learning deep architectures for AI // Foundations and trends® in Machine Learning. 2009. vol. 2(1). pp. 1–127.
4. *Breuel T.* Recent progress on the OCRopus OCR system // Proceedings of the Inter-national Workshop on Multilingual OCR. 2009. pp. 2.
5. *Bureš L., Neduchal P., Hlavác M., Hrúz M.* Generation of synthetic images of full-text documents // International Conference on Speech and Computer. 2018. pp. 68–75.
6. *Chen L.C., Papandreou G., Schroff F., Adam H.* Rethinking atrous convolution for semantic image segmentation // arXiv preprint arXiv:1706.05587. 2017.
7. *Chernyshova Y.S., Gayer A.V., Sheshkus A.V.* Generation method of synthetic training data for mobile OCR system // Tenth International Conference on Machine Vision (ICMV 2017). 2018. vol. 10696. pp. 106962G.
8. *Dumas T., Roumy A., Guillemot C.* Autoencoder based image compression: can the learning be quantization independent? // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018. pp. 1188–1192.
9. *Gruber I., Hlavác M., Hrúz M., Železný M.* Semantic segmentation of historical documents via fully-convolutional neural network // International Conference on Speech and Computer. 2019. pp. 142–149.
10. *Gupta A., Vedaldi A., Zisserman A.* Synthetic data for text localization in natural images // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2016. pp. 2315–2324.
11. *Hajic J., Dorfer M., Widmer G., Pecina P.* Towards full-pipeline handwritten OMR with musical symbol detection by u-nets // ISMIR. 2018. pp. 225–232.

12. *Huang H., He R., Sun Z., Tan T.* Introvae: Introspective variational autoencoders for photographic image synthesis // Advances in Neural Information Processing Systems. 2018. pp. 52–63.
13. *Huang W., Qiao Y., Tang X.* Robust scene text detection with convolution neural network induced mser trees // European Conference on Computer Vision. 2014. pp. 497–511.
14. *Jaderberg M., Vedaldi A., Zisserman A.* Deep features for text spotting // European Conference on Computer Vision. 2014. pp. 512–528.
15. *Jaderberg M., Simonyan K., Vedaldi A., Zisserman A.* Reading text in the wild with convolutional neural networks // International Journal of Computer Vision. 2016. vol. 116(1). pp. 1–20.
16. *Kingma D.P., Welling M.* Auto-encoding variational bayes // International Conference on Learning Representations. 2014. 21 p.
17. *Lin G. et al.* Refinenet: Multi-path refinement networks for dense prediction // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019.
18. *Noh H., Hong S., Han B.* Learning deconvolution network for semantic segmentation // Proceedings of the IEEE International conference on computer vision. 2015. pp. 1520–1528.
19. *Otsu N.* A threshold selection method from gray-level histograms // IEEE transactions on systems, man, and cybernetics. 1979. vol. 9(1). pp. 62–66.
20. *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical image computing and computer-assisted intervention. 2015. pp. 234–241.
21. *Smith R.* An overview of the tesseract OCR engine // Ninth International Conference on Document Analysis and Recognition (ICDAR 2007). 2007. vol. 2. pp. 629–633.
22. *Tokui S., Oono K., Hido S., Clayton J.* Chainer: a next-generation open source framework for deep learning // Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Twenty-ninth Annual Conference on Neural Information Processing Systems (NIPS). 2015. vol. 5. pp. 1–6.
23. *Wen S. et al.* Variational autoencoder based image compression with pyramidal features and context entropy model // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2019. pp. 0–0.
24. *Zhao H. et al.* Pyramid scene parsing network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 2881–2890.
25. *Zhou X. et al.* EAST: An efficient and accurate scene text detector // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. pp. 5551–5560.

DOI 10.15622/sp.2019.18.6.1407-1433

V.F. OCHKOV, I.E. VASILEVA

APPLICATION OF DIFFERENCE SCHEMES TO DECISION THE PURSUIT PROBLEM

Ochkov V.F., Vasileva I.E. **Application of Difference Schemes to Decision the Pursuit Problem.**

Abstract. The problem of the pursuit curve construction in the case when the tangent to pursuer's motion trajectory passes at any time through the point representing the pursued is considered. A new approach to construct the pursuit curves using difference schemes is proposed. The proposed technique eliminates the need to derive the differential equations for the description of the pursuit curves, which is quite difficult task in the general case. In addition, the application of difference methods is justified in a situation where it is complicated to find the analytical solution of an existing differential equation and it is possible to obtain the pursuit curve only numerically. Various modifications of difference schemes respectively equivalent to the Euler, to the Adams – Bashforth and to the Milne methods are constructed. Their software implementation is realized by using the mathematical package Mathcad. We consider the case of a uniform rectilinear motion of the pursued whose differential equation describing the path of the pursuer and its analytical solution are known. We compare the numerical solutions obtained by the different methods with the well-known analytical solution. The error of the obtained numerical solutions is examined. Moreover, an application is considered illustrating the construction of the difference schemes for the case of an arbitrary trajectory of the pursued. Also, we extend the proposed method to the case of cyclic pursuit with several participants in the three-dimensional space. In particular, we construct a difference scheme equivalent to the Euler method for a three-dimensional analogue of the "bugs problem". The results obtained are demonstrated by means of animated examples for either two-dimensional or three-dimensional cases.

Keywords: Differential Games, Pursuit Problem, Pursuit Curve, Numerical Methods, Difference Methods, the Euler Method, "Three Bugs" or "Three Mice" Problem, Mathcad.

1. Introduction. The pursuit problem belongs to a class of long and widely studied problems. The leading role in the formulation of this problem belongs to research conducted in the field of the differential games, where it is necessary to choose the optimal pursuit strategy [1-6]. A common strategy is the pursuit method [7] which determines the motion of the pursuer in such way that the tangent to the trajectory of its motion at any time passes through the position of the point associated with the pursued. The problem of constructing the trajectory of the pursuer — the pursuit curve — is relevant in various areas and it has a wide practical value, in particular, in mechanics, military affairs, control systems [8-25].

Despite the variety of applied problems, only particular cases of pursuit are sufficiently studied, for example, the case of a simple motion when the pursued entity moves uniformly along a straight line. For this case, it is possible to explicitly write a second-order nonlinear differential equation describing the pursuit curve and to find its analytical solution [9, 10]. It is assumed, for

definiteness, that the pursued entity begins a uniform motion with speed $\bar{v}$ along the axis Oy from the origin of coordinates, while the pursuer starts with speed $\bar{V}$ from the point with coordinates (1,0) (Figure 1).

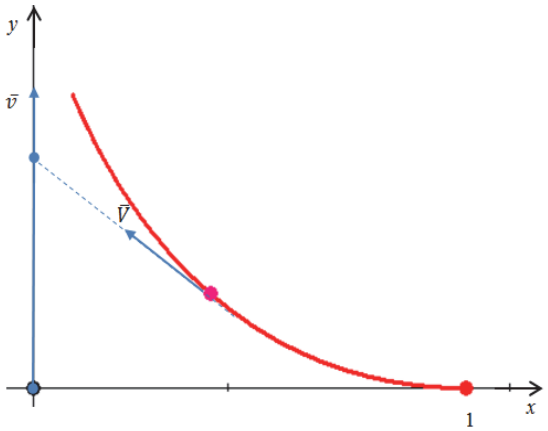


Fig. 1. Scheme of a simple motion

Then the sought differential equation is given by [9, 10]:

$$y''(x) = \frac{1}{kx} \cdot \sqrt{1 + (y'(x))^2}, \tag{1}$$

where $k = v / V$ is the ratio of the speeds of the pursued and of the pursuer.

Taking into account the zero initial conditions $y(1) = 0, y'(1) = 0,$ which emerge from the statement of the problem, the solution of equation (1) is described as follows [10]:

$$\begin{gathered} y(x) = \frac{1}{2}\left(\frac{x^{1+k}}{1+k} - \frac{x^{1-k}}{1-k} \right) + \frac{k}{1-k^2}, \text{if } k \neq 1; \\ y(x) = \frac{1}{4}(x^2 - \ln x^2 - 1), \text{if } k = 1. \end{gathered} \tag{2}$$

To find the pursuit curve in particular cases, either kinematics methods [11-14] or parametrization methods [8] are also used.

However, the question of describing the curve of the pursuer in the general case remains open due to the complexity of deriving the differential equation itself, as well as finding its analytical solution.

This article proposes a numerical approach for constructing the pursuit curve based on difference schemes. Numerical methods are often used in

differential games problems, also in relation to the pursuit curve problem. In particular, the grid method and its variations are widely known [22-24], they are designed to derive the value function in a time-optimal game and the optimal trajectory [22]. This method is also successfully applied to solve the pursuit problem in distributed control systems [25]. The difference schemes investigated in the mentioned works are constructed for an available differential equation describing the pursuit process.

The proposed approach instead makes it possible to abandon the need to derive the differential equation that describes the pursuit trajectory. In addition, the application of this approach is justified in a situation where it is difficult to find the analytical solution of the existing differential equation and it is possible to obtain the pursuit curve only by numerical methods. The possibility of this approach is described, in particular, in reference [26].

2. Construction of a difference scheme equivalent to the Euler method. Initially, it is constructed a difference scheme for the case of a simple motion (Figure 1) in order to show the approximation of a well-known analytical solution (2).

Then we deal with the parametrization with respect to time t and denote at each time point the known coordinates of the pursued with $x(t), y(t)$ and the unknown coordinates of the pursuer with $X(t), Y(t)$ [8].

According to the law of the uniform rectilinear motion, the coordinates $x(t), y(t)$ of the pursued along the axis *Oy*, are determined by the formulas:

$$\begin{aligned} x(t) &= 0, \\ y(t) &= v \cdot t. \end{aligned} \tag{3}$$

The pursuit time is divided into n time intervals and the time step is denoted Δt, and the approximate coordinates of the pursued and of the pursuer at each step: x_i, y_i and $X_i, Y_i, i = \overline{0,n}$, respectively. At the initial time, the coordinates have the following values:

$$\begin{aligned} x_0 &= 0, \quad y_0 = 0, \\ X_0 &= 0, \quad Y_0 = 0. \end{aligned} \tag{4}$$

From (3) it follows that the coordinates of the pursued $x_{i+1}, y_{i+1}, i = \overline{0,n-1}$, are defined as follows:

$$\begin{aligned} x_{i+1} &= 0, \\ y_{i+1} &= y_i + v \cdot \Delta t, \end{aligned} \quad i = \overline{0,n-1}. \tag{5}$$

In order to determine the coordinates of the pursuer at step $i+1$, geometric constructions are used (Figure 2).

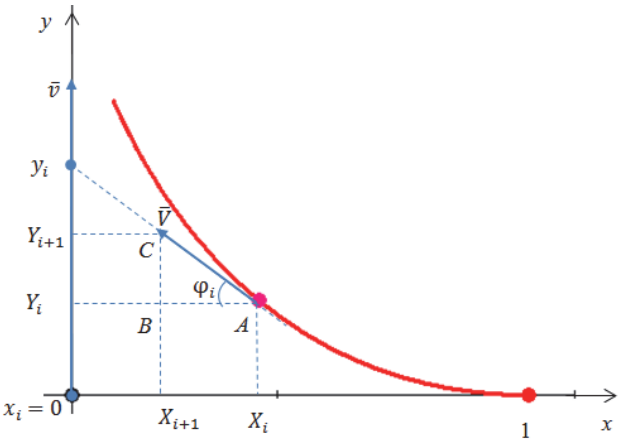


Fig. 2. Construction of the difference scheme

The acute angle between the tangent to the pursuit curve and the Ox axis at step i is denoted by φ_i. Obviously, at the initial time $\varphi_0 = 0$. The tangent of the right angle (Figure 2) is given by:

$$\operatorname{tg}(\varphi_i) = \frac{y_i - Y_i}{|x_i - X_i|}, \; i = \overline{0, n-1}. \tag{6}$$

From the right triangle ABC we obtain:

$$\begin{aligned} X_{i+1} - X_i &= -AC \cdot \cos\varphi_i, \\ Y_{i+1} - Y_i &= AC \cdot \sin\varphi_i, \end{aligned} \quad i = \overline{0, n-1}, \tag{7}$$

or

$$\begin{aligned} X_{i+1} &= X_i + AC \cdot \cos(\pi + \varphi_i), \\ Y_{i+1} &= Y_i + AC \cdot \sin(\pi + \varphi_i), \end{aligned} \quad i = \overline{0, n-1}. \tag{8}$$

Since the velocity vector $\bar{V}$ is directed tangentially, the following relation applies:

$$AC = \left|\overline{V}\right| \cdot \Delta t. \tag{9}$$

Taking into account (6) and generalizing formula (8) for the cases where the angle φ_i is located in any coordinates quarter, the following recurrence formulas are derived for determining the coordinates of the pursuer:

$$\begin{aligned} X_{i+1} &= X_i + \left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_i), \\ Y_{i+1} &= Y_i + \left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_i), \end{aligned} \quad i = \overline{0, n-1}, \tag{10}$$

where

$$\overline{\varphi}_i = \begin{cases} \operatorname{arctg}\left(\dfrac{y_i - Y_i}{x_i - X_i}\right) + \pi, \text{if } X_i > x_i, \\ \operatorname{arctg}\left(\dfrac{y_i - Y_i}{x_i - X_i}\right), \text{if } X_i < x_i \\ \dfrac{\pi}{2} \cdot \operatorname{sgn}(y_i - Y_i), \text{if } X_i = x_i, \end{cases} \tag{11}$$

$$i = \overline{0, n}.$$

It should be noted that the calculation by the difference scheme (10) should be stopped if at any step $k \le n$ the equality $x_k = X_k, y_k = Y_k$ is verified. In fact, this means that the pursuer caught up with the pursued, i.e. the pursuit is complete.

Thus, the central element of the proposed method is the calculation at each step i of the angle $\overline{\varphi}_i$ between the tangent to the pursuit curve and the axis Ox. The expressions $\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_i)$ and $\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_i)$ are respectively approximate increments of functions $X(t)$ and $Y(t)$ per step Δt (Figure 2), and, therefore, at a small step, these expressions serve as analogues of the differentials of the functions $X(t)$ and $Y(t)$ at the point t_i. Therefore, it can be said that the difference scheme (10) is equivalent to the classical difference scheme, which is constructed by the Euler method for finding the solution $u(t)$ of a Cauchy problem of the type:

$$\begin{aligned} u'(t) &= f(t), \\ u(t_0) &= u_0, \end{aligned} \tag{12}$$

and it has the form [27]:

$$u_{i+1} = u_i + \mathrm{d}u(t_i), \quad i = \overline{0, n-1}, \tag{13}$$

where t_i is a partition point, u_i are the values of the approximate solution at the partition points, Δt is the step.

The difference scheme (10) is implemented using the mathematical package Mathcad [28, 29] for given values of the number of steps n, step Δt and speeds v and $V (v \neq V)$. A numerical solution of the problem is obtained in terms of the vectors $(X_i, Y_i), i = \overline{1, n}$, that is, we obtain an approximation of the pursuit curve. A comparison with the analytical solution (2) is made. The results are shown in Figure 3a, b (the animation for Figure 3a is posted on an electronic resource https://community.ptc.com/t5/PTCMathcad/OchkovArticleAnimaions/td-p/583048).

From the graphs shown in Figure 3, it is clear that the obtained numerical solution actually implements the pursuit method (the tangent at the selected point of the pursuer's trajectory passes through the point of the pursued entity — Figure 3b) and approximates the analytical solution quite well Figure 3a.

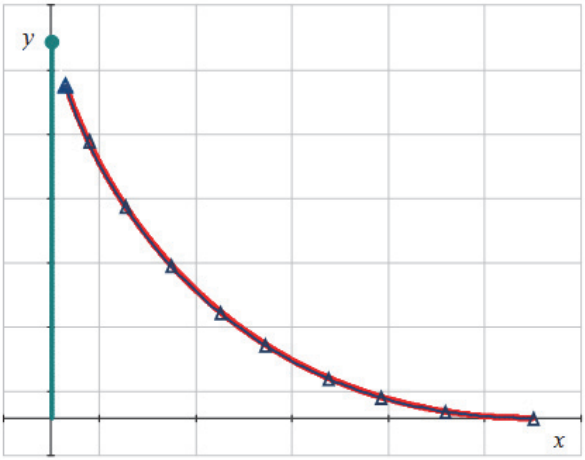


Fig. 3(a). Comparison of numerical and analytical pursuit curves: general view

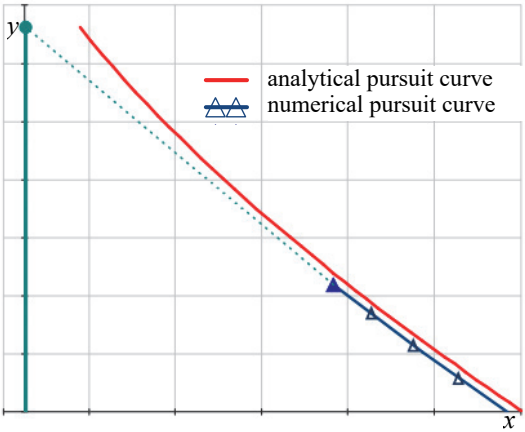


Fig. 3(b). Comparison of numerical and analytical pursuit curves: scaled view of part of the curve

In Figure 4 it is reported a plot showing the error ε of the numerical solution, defined according to the formula:

$$\varepsilon = \left| y(X_i) - Y_i(X_i) \right|, \tag{14}$$

where $y(X_i)$ is defined according to (2).

As can be seen from Figure 4, because of accumulation, the error becomes equal to the step $\Delta t = 10^{-3}$, that is, the difference scheme (10) has an accuracy of the first order corresponding to the Euler method [27].

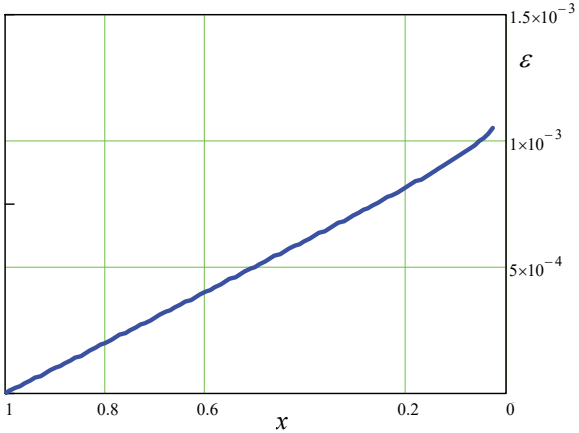


Fig. 4. Numerical solution error

Труды СПИИРАН. 2019. Том 18 № 6. ISSN 2078-9181 (печ.)
ISSN 2078-9599 (онлайн) www.proceedings.spiiras.nw.ru

3. Construction of difference schemes using linear multi-step methods. The Euler method, that is used to construct the difference scheme described above, is a rather "rough" method of approximate calculation. To increase accuracy, linear multi-step methods are usually applied. These methods use not only one, but several previously calculated values of the sought function [27]. Let us consider the construction of a difference scheme equivalent to the two-step Adams – Bashforth method, which for a Cauchy problem of the form (12) is given by the formula [27]:

$$\begin{aligned} u_1 &= u_0 + \mathrm{d}u(t_0), \\ u_{i+1} &= u_i + \frac{3}{2}\mathrm{d}u(t_i) - \frac{1}{2}\mathrm{d}u(t_{i-1}), i = \overline{1, n-1}. \end{aligned} \tag{15}$$

This method has, in contrast to the Euler method, an accuracy of the second order [27].

The difference scheme (10) is modified in accordance with (15). Taking into account the fact that the expressions $\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_i)$ and $\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_i)$ have the meaning of differentials of the functions $X(t)$ and $Y(t)$ at the point t_i, the modified difference method will be:

$$\begin{aligned} X_1 &= X_0 + \left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_0), Y_1 = Y_0 + \left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_0), \\ X_{i+1} &= X_i + \frac{3}{2}\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_i) - \frac{1}{2}\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{i-1}), \\ Y_{i+1} &= Y_i + \frac{3}{2}\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_i) - \frac{1}{2}\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{i-1}), \\ & i = \overline{1, n-1}, \end{aligned} \tag{16}$$

where $\overline{\varphi}_i$ is defined by the formula (11).

The difference scheme (16) is implemented using the Mathcad mathematical package for given values of the number of steps n, step Δt and speeds $v, V (v \neq V)$. The obtained numerical solution is compared with the analytical solution (2) and with the solution obtained using the difference method (10). The results are shown in Figure 5a, b (the animation for Figure 5a is posted on an electronic resource https://community.ptc.com/t5/PTCMathcad/OchkovArticleAnimaions/td-p/583048).

From Figure 5b it is clear that the numerical solution obtained by the difference scheme (16) (equivalent to the Adams – Bashforth method)

approximates the analytical solution better than the numerical solution obtained by the difference scheme (10) (equivalent to the Euler method): the plot of this solution coincides visually with the analytical solution.

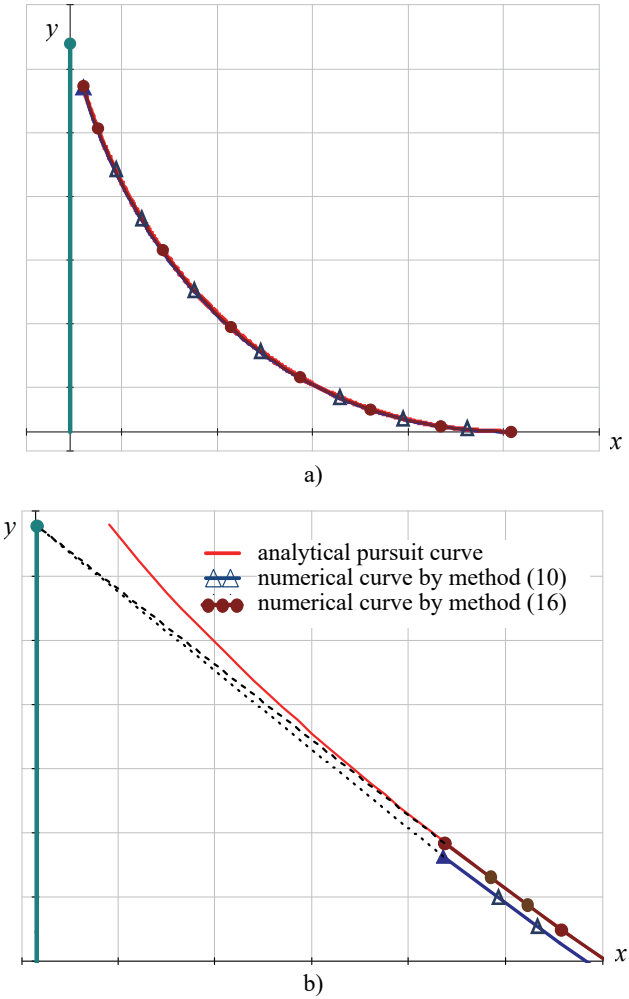


Fig. 5. Comparison of numerical and analytical pursuit curves:
a) general view, b) scaled view of part of the curve

The plot in Figure 6 shows the errors of the numerical solutions, determined by formula (14). The error of the numerical solution obtained by this method (16), for a given step size $\Delta t = 10^{-3}$ is of the order

$O(\Delta t^2) \sim 10^{-6}$, which is far less than the error of the numerical solution obtained by method (10).

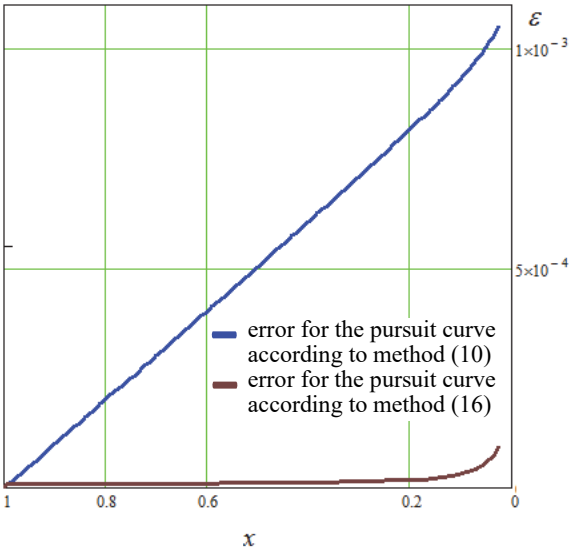


Fig. 6. Numerical error of the solutions obtained by the two methods

If it is necessary to construct the pursuit trajectory with a much smaller error, numerical methods of a higher order should be used, for example, the predictor-corrector methods which are widely used in applications [30-36]. Let us construct the difference scheme equivalent to the fourth-order Milne method [37, 38]. This method requires four initial steps and uses a couple of finite-difference formulas (a predictor and a corrector).

For a given initial value u_0, additional initial values u_1, u_2, u_3 are calculated by some other methods, for example, by the Runge-Kutta method, which has fourth-order accuracy [27, 39, 40].

Let us consider the predictor-corrector formulas by the Milne method starting with the fourth step.

The predictor formula for the Cauchy problem (12) is defined as follows [37]:

$$u_{i+1}^{pred} = u_{i-3} + \frac{4}{3}\left(2\mathrm{d}u\left(t_{i-2}\right) - \mathrm{d}u\left(t_{i-1}\right) + 2\mathrm{d}u\left(t_i\right)\right) \qquad (17)$$

$$i = \overline{3, n-1}.$$

The corrector formula is defined as follows [37]:

$$u_{i+1}^{corr} = u_{i-1} + \frac{1}{3}\left(\mathrm{d}u\left(t_{i-1}\right) + 4\mathrm{d}u\left(t_i\right) + \mathrm{d}u^{pred}\left(t_{i+1}\right)\right), \quad i = \overline{3, n-1}, \tag{18}$$

where expression $\mathrm{d}u^{pred}\left(t_{i+1}\right)$ defines a predicted differential of function $u(t)$ at the point t_{i+1} based on (17).

Let us consider the difference scheme according to (17)-(18).

As noted above, the key element of the proposed method is the calculation at each step i of the angle $\overline{\varphi}_i$ between the tangent to the pursuit curve and the axis Ox. The expressions $\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_i)$ and $\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_i)$ are respectively analogues of the differentials of the functions $X(t)$ and $Y(t)$ at the point t_i.

Accordingly, the essential link of the difference method is the calculation of the predicted angle $\overline{\varphi}_{i+1}^{pred}$.

Thus, the difference method is defined as follows:

$$\begin{aligned}
X_{i+1}^{pred} &= X_{i-3} + \frac{4}{3}(2\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}-2}) - \left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}-1}) + \\
&\quad + 2\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}})), \\
Y_{i+1}^{pred} &= Y_{i-3} + \frac{4}{3}(2\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{\mathrm{i}-2}) - \left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{\mathrm{i}-1}) + \\
&\quad + 2\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{\mathrm{i}})), \\
X_{i+1}^{corr} &= X_{i-1} + \frac{1}{3}(\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}-1}) + 4\left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}}) + \\
&\quad + \left|\overline{V}\right| \cdot \Delta t \cdot \cos(\overline{\varphi}_{\mathrm{i}+1}^{pred})), \\
Y_{i+1}^{corr} &= Y_{i-1} + \frac{1}{3}(\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{\mathrm{i}-1}) + 4\left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{\mathrm{i}}) + \\
&\quad + \left|\overline{V}\right| \cdot \Delta t \cdot \sin(\overline{\varphi}_{i+1}^{pred})), \\
&\quad i = \overline{3, n-1},
\end{aligned} \tag{19}$$

where

$$\overline{\varphi}_{i+1}^{pred} = \begin{cases} \operatorname{arctg}\left(\dfrac{y_{i+1} - Y_{i+1}^{pred}}{x_{i+1} - X_{i+1}^{pred}}\right) + \pi, \text{if } X_{i+1}^{pred} > x_{i+1}, \\ \operatorname{arctg}\left(\dfrac{y_{i+1} - Y_{i+1}^{pred}}{x_{i+1} - X_{i+1}^{pred}}\right), \text{if } X_{i+1}^{pred} < x_{i+1}, \\ \dfrac{\pi}{2} \cdot \operatorname{sgn}\left(y_{i+1} - Y_{i+1}^{pred}\right), \text{if } X_{i+1}^{pred} = x_{i+1}, \end{cases} \quad (20)$$

$$i = \overline{3, n-1}.$$

The difference scheme (19) is implemented using the Mathcad mathematical package for given values of the number of steps n, step Δt and speeds $v, V (v \neq V)$. The obtained numerical solution is compared with the analytical solution (2) and with the solutions obtained using the difference methods (10), (16). The general view of the plot of the analytical solution and of the numerical solution obtained using the difference methods (19) is identical to the ones in Figures 3, 5.

The plot in Figure 7 shows the errors of the numerical solutions determined according to formula (14) for each of the three solutions obtained by the difference scheme (10), (16), (19).

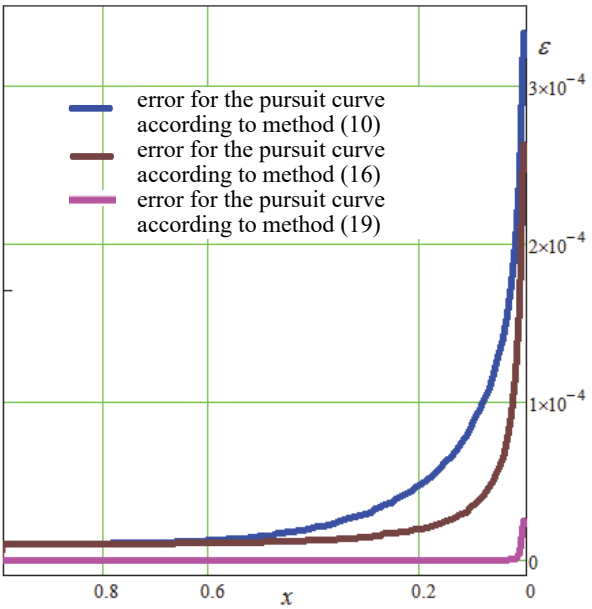


Fig. 7. Numerical error of the solutions obtained by the three methods

The error of the numerical solution obtained by method (19) (equivalent to the Milne method), for a given step size $\Delta t = 10^{-3}$, is of the order $O(\Delta t^4) \sim 10^{-12}$, which is significantly less than the errors of the numerical solutions obtained by methods (10) and (16) (analogue to the Euler method and the Adams-Bashforth method respectively).

A high accuracy of a numerical solution is achieved by implementing of the proposed method.

Therefore, it can be concluded that the proposed difference schemes (10), (16) and (19) allow to construct numerical solutions that approximate the analytical pursuit curve with varying accuracy according to the specific problem.

4. Application of difference schemes for constructing the pursuit curve in case of an arbitrary trajectory of the pursued. The derived difference schemes (10), (16), (19) can be used to obtain a numerical solution also if the pursued entity moves not along a straight line, as it was considered previously, but along an arbitrary trajectory.

Let us consider as an example the situation when the pursued moves uniformly along an elliptical path, i.e. the coordinates $x(t), y(t)$ are determined by the formulas:

$$\begin{aligned} x(t) &= A \cdot \sin(v \cdot t), \\ y(t) &= B \cdot \cos(v \cdot t), \end{aligned} \tag{21}$$

where v is the speed of the pursued, while A, B are constant coefficients.

Let the pursued start moving from the point with coordinates:

$$x_0 = 0,\ y_0 = 1. \tag{22}$$

Then the difference relations (5) take the form:

$$\begin{aligned} x_{i+1} &= A \cdot \sin\left(v \cdot \Delta t \cdot (i+1)\right), \\ y_{i+1} &= A \cdot \cos\left(v \cdot \Delta t \cdot (i+1)\right), \end{aligned} \quad i = \overline{0, n-1}. \tag{23}$$

The initial values of the sought coordinates of the pursuer $X_i, Y_i, i = \overline{0,n}$ are assumed as follows:

$$X_0 = 0, Y_0 = -1. \tag{24}$$

Then, by performing the calculation using formulas (10), (16), (19) for given values of the number of steps n, step Δt and speeds $v, V (v \neq V)$. three numerical versions of the pursuit curve are obtained.

The implementation of the described algorithm using the Mathcad mathematical package is shown in Figures 8a, b (the animation for Figure 8a is posted on an electronic resource https://community.ptc.com/t5/PTCMathcad/OchkovArticleAnimaions/td-p/583048).

From Figure 8b we notice that the numerical pursuit curves constructed according to the methods (16) and (19) (analogue to the Adams – Bashforth and to the Milne methods respectively) provide a better prediction for the next step of the pursued than the pursuit curve constructed according method (10) (equivalent to the Euler method): the $n+1$ steps of the pursuer trajectories are directed not at the point of the pursued location (as in method (10)), but with anticipation towards the $(n+1)-$ th step of the pursued.

It is obvious that method (19) (equivalent to the Milne method) provides a better approximation to the $(n+1)-$ th step of the pursued than method (16) (equivalent to the Adams – Bashforth method).

Thus, the possibility of using the difference schemes is demonstrated also in the case when the pursued entity moves along an arbitrary trajectory.

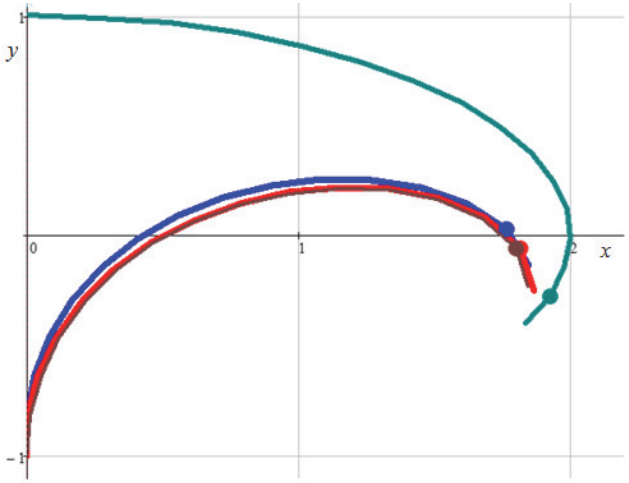


Fig. 8(a). Pursuit curves obtained by different methods for the case of an elliptical path of the pursued: general view

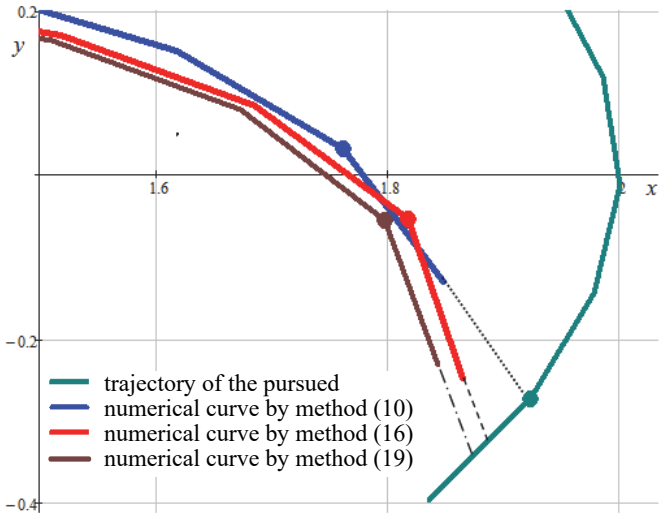


Fig. 8(b). Pursuit curves obtained by different methods for the case of an elliptical path of the pursued: view of part of the curve

The advantage of the proposed method is that it allows to construct numerical interpretations of pursuit curves for given strategies of the pursued, which are not, generally speaking, solutions of any differential equations, as it is usually done in the theory of the differential games (see, for example, [41]).

5. Application of difference schemes for constructing a pursuit curve in the three-dimensional space. The proposed method for constructing difference schemes can be extended to the case of the three-dimensional space. The problem of constructing a pursuit curve in three-dimensional space is most often considered in the context of control theory problems with cyclic pursuit, where several enumerated objects start moving from different points of the space, with each object catching up with the next [42-44]. This case has been studied quite well on the plane (see, for example, [45-47]), and it reduces to the so-called "bugs problem" (this problem is also known as "mice problem"). [48, 49], when the objects begin to move uniformly at the same speed from the vertices of a regular polygon, with each object moving in the direction of its nearest neighbour. In this case, the trajectory of motion of each object is a logarithmic spiral [50].

We construct a difference scheme equivalent to the Euler method for the "bugs problem" in the three-dimensional space, i.e. for the problem of finding the trajectories of objects that begin a uniform motion with equal speed from the vertices of a regular polyhedron.

As a polyhedron, it was considered a tetrahedron with vertices at the points $(0,0,0),\left(\frac{\sqrt{3}}{2},\frac{1}{2},0\right),(0,1,0),\left(\frac{\sqrt{3}}{6},\frac{1}{2},\sqrt{\frac{2}{3}}\right)$. Let four objects begin their uniform motion with the same speed from the vertices of the tetrahedron, with the first object following the second, the second following the third, and so on (Figure 9).

The speeds of the four objects are denoted by $\overline{v_1},...,\overline{v_4}$ (the modules of the speeds are set to be equal). The coordinates of each of the objects in a time-parametric form will be denoted by $\left(x^1(t),y^1(t),z^1(t)\right),\ldots,\left(x^4(t),y^4(t),z^4(t)\right)$.

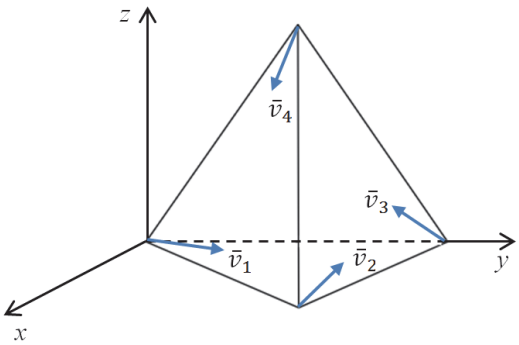


Fig. 9. Cyclic pursuit in a tetrahedron

To construct a difference scheme which is similar to the two-dimensional case, we will not explore the trajectories of the objects of the pursuit in the space itself, but their projections on the $0xy$ and $0xz$ planes.

A partition of the pursuit time into n time intervals is introduced and the time step is denoted by Δt. Then for the coordinates $x^k(t),y^k(t),k=\overline{1,4}$, the difference relation (10) is written for the two-dimensional case in the $0xy$ plane, and for the coordinate $z^k(t),k=\overline{1,4}$ in the $0xz$ plane.

To do this, firstly, it is necessary to calculate, respectively, in the $0xy$ and $0xz$ planes, the angles $\overline{\varphi}_i^{oxy}$, $\overline{\varphi}_i^{oxz}$, $i=\overline{1,n}$, between the tangent to the pursuit curve and the axis Ox by the formula (11).

Secondly, one should use not the modules of objects velocities $\left|\overline{v_k}\right|, k=\overline{1,4},$ but the modules of their projections $\left|\overline{v_k}^{oxy}\right|, \left|\overline{v_k}^{oxz}\right|, k=\overline{1,4}$ on the corresponding planes.

The difference relations are derived for finding the coordinates $\left(x^1(t), y^1(t), z^1(t)\right)$ of the first object, starting from the origin and following the second object with coordinates $\left(x^2(t), y^2(t), z^2(t)\right),$ starting the motion from the vertex $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right).$ Then the initial values of the coordinates are:

$$\begin{gathered} x_0^1 = y_0^1 = z_0^1 = 0; \\ x_0^2 = \frac{\sqrt{3}}{2}, y_0^2 = \frac{1}{2}, z_0^2 = 0. \end{gathered} \tag{25}$$

According to (10) and taking into account the above remarks, the difference relations are converted to the form:

$$\begin{gathered} x_{i+1}^1 = x_i^1 + \left|\overline{v}_1^{oxy}\right| \times \Delta t \times \cos(\overline{\varphi}_i^{oxy}), \\ y_{i+1}^1 = y_i^1 + \left|\overline{v}_1^{oxy}\right| \times \Delta t \times \sin(\overline{\varphi}_i^{oxy}), \\ z_{i+1}^1 = z_i^1 + \left|\overline{v}_1^{oxz}\right| \times \Delta t \times \sin(\overline{\varphi}_i^{oxz}), \\ i = \overline{0, n-1}, \end{gathered} \tag{26}$$

where

$$\overline{\varphi}_i^{oxy} = \begin{cases} \operatorname{arctg}\left(\dfrac{y_i^2 - y_i^1}{x_i^2 - x_i^1}\right) + \pi, \text{for } x_i^1 > x_i^2, \\ \operatorname{arctg}\left(\dfrac{y_i^2 - y_i^1}{x_i^2 - x_i^1}\right), \text{ for } x_i^1 < x_i^2, \\ \dfrac{\pi}{2} \cdot \operatorname{sgn}\left(y_i^2 - y_i^1\right), \text{for } x_i^1 = x_i^2, \end{cases} \tag{27}$$

$$\overline{\varphi}_i^{oxz} = \begin{cases} \operatorname{arctg}\left(\dfrac{z_i^2 - z_i^1}{x_i^2 - x_i^1}\right) + \pi, \text{for } x_i^1 > x_i^2, \\ \operatorname{arctg}\left(\dfrac{z_i^2 - z_i^1}{x_i^2 - x_i^1}\right), \text{for } x_i^1 < x_i^2, \\ \dfrac{\pi}{2} \cdot \operatorname{sgn}\left(z_i^2 - z_i^1\right), \text{for } x_i^1 = x_i^2, \end{cases}$$

$$i = \overline{0, n}.$$

The projections $\left|\overline{v}_1^{oxy}\right|$, $\left|\overline{v}_1^{oxz}\right|$ are found using the geometric constructions in Figure 10.

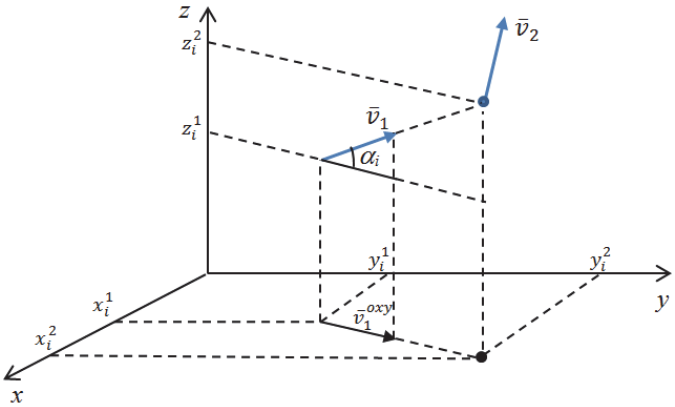


Fig. 10. Determination of the projection of the velocity of an object on the plane

Figure 10 shows that the projection is defined as follows:

$$\left|\overline{v}_1^{oxy}\right| = \left|\overline{v_1}\right| \cdot \cos\left(\alpha_i\right), \tag{28}$$

where a_i is calculated from the relation:

$$tg\left(\alpha_i\right) = \frac{\left|z_i^2 - z_i^1\right|}{\sqrt{\left|x_i^2 - x_i^1\right|^2 + \left|y_i^2 - y_i^1\right|^2}}. \tag{29}$$

Similarly, the projection $\left|\overline{v}_1^{oxz}\right|$ is defined:

$$\left|\overline{v_1}^{oxz}\right| = \left|\overline{v_1}\right| \cdot \cos(\beta_i), \tag{30}$$

where β_i is calculated from the relationship:

$$\operatorname{tg}(\beta_i) = \frac{\left|y_i^2 - y_i^1\right|}{\sqrt{\left|x_i^2 - x_i^1\right|^2 + \left|z_i^2 - z_i^1\right|^2}}. \tag{31}$$

Using (28)-(31), the difference method (26) is obtained in the final form:

$$\begin{gathered} x_{i+1}^1 = x_i^1 + \left|\overline{v_1}\right| \cdot \cos(\alpha_i) \cdot \Delta t \cdot \cos(\overline{\varphi}_i^{oxy}), \\ y_{i+1}^1 = y_i^1 + \left|\overline{v_1}\right| \cdot \cos(\alpha_i) \cdot \Delta t \cdot \sin(\overline{\varphi}_i^{oxy}), \\ z_{i+1}^1 = z_i^1 + \left|\overline{v_1}\right| \cdot \cos(\beta_i) \cdot \Delta t \cdot \sin(\overline{\varphi}_i^{oxz}), \\ i = \overline{0, n-1}, \end{gathered} \tag{32}$$

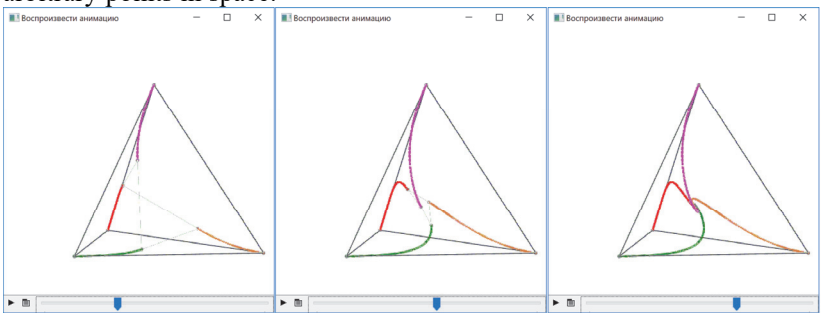
where $\overline{\varphi}_i^{oxy}$, $\overline{\varphi}_i^{oxz}$ are determined according to (27), and α_i and β_i are determined by the relations:

$$\begin{aligned} \alpha_i &= \operatorname{arctg}\left(\frac{\left|z_i^2 - z_i^1\right|}{\sqrt{\left|x_i^2 - x_i^1\right|^2 + \left|y_i^2 - y_i^1\right|^2}}\right), \\ \beta_i &= \operatorname{arctg}\left(\frac{\left|y_i^2 - y_i^1\right|}{\sqrt{\left|x_i^2 - x_i^1\right|^2 + \left|z_i^2 - z_i^1\right|^2}}\right). \end{aligned} \tag{33}$$

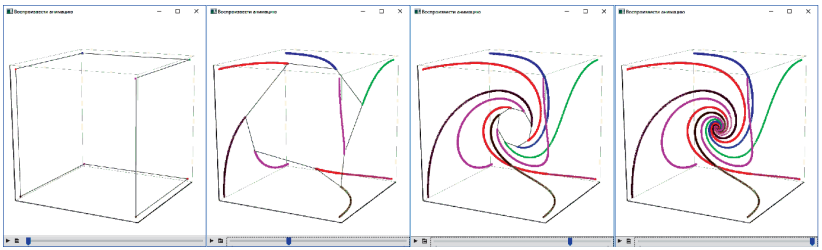
By establishing similar difference relations (32) for the second, third and fourth objects and combining them into a system we obtain a complete difference method for constructing the pursuit curves in the tetrahedron. Figure 11a (animations for Figure 11 are posted on electronic resources https://community.ptc.com/t5/PTC-Mathcad/Is-it-my-own-or-Mathcad-15-error/m-p/576164, https://community.ptc.com/t5/PTC-Mathcad/Bats-

problem/m-p/576137) shows the results of the implementation of the difference method in the Mathcad package for the tetrahedron, while Figures 11b, c show the results of implementing similarly constructed pursuit curves for cases when objects start their motion from the vertices of a cube and of a dodecahedron respectively.

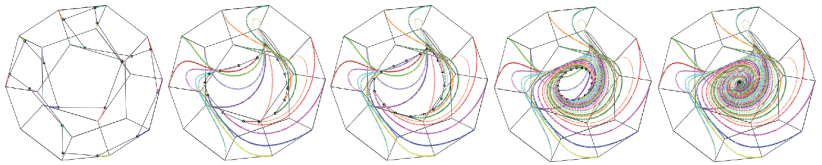
It could be noted that the proposed approach allows to construct difference schemes not only for objects that start their motion with identical speeds from the vertices of Platonic solids, but also for objects that move uniformly with different speeds and begin their motion at arbitrary points in space.



a)



b)



c)

Fig. 11. Constructing the pursuit curves in the three-dimensional space

6. Conclusions. This article proposes a new approach to the construction of pursuit curves through the use of difference schemes. The

advantage of the proposed approach is the possibility of describing pursuit curves in a numerical way without deriving the differential equation. The constructed modifications of difference schemes equivalent respectively to the Euler, to the Adams – Bashforth and to the Milne methods approximate the analytical solution with high accuracy. The proposed approach can be applied to the numerical construction of pursuit trajectories with arbitrary pursued strategy either in two-dimensional plans or in three-dimensional spaces.

References

1. Petrosyan L.A., Mazalov V.V. Recent Advances in Game Theory and Applications. Springer International Publishing. 2016. 284 p.
2. Mazalov V., Chirkova J.V. Networking Games. Network Forming Games and Games on Networks. Academic Press. 2019. 322 p.
3. Petrosyan L.A. Differential Games of Pursuit. World Scientific. 1993. 332 p.
4. Isaacs R. Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization. Courier Corporation. 1999. 416 p.
5. Kumkov S.S., Le Menec S., Patsko V.S. Zero-sum pursuit-evasion differential games with many objects: survey of publications. *Dynamic Games and Applications*. 2017. vol. 7. no. 4. pp. 609–633.
6. Ramana M.V., Kothari M. Pursuit-evasion games of high speed evade. *Journal of Intelligent & Robotic Systems*. 2017. vol. 85. no. 2. pp 293–306.
7. Pankov S.Ya., Zaburaev E.Yu., Matveev A.M. *Teoriya i metodika upravleniya aviatsiei* [Theory and Methods of Aviation Control]. Ulyanovsk: UVAU GA. 2006. 190 p. (In Russ.).
8. Barton J.C., Eliezer C.J. On pursuit curves. *The ANZIAM Journal.* 2000. vol. 41. no. 3. pp. 358–371.
9. Samoyavcheva M.V., Fedorov L.I. [The pursuit problem]. *Vestnik moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Fizika-matematika – Bulletin of the Moscow Region State University. Series: Physics-Mathematics.* 2011. no. 1. pp. 65–69. (In Russ.).
10. Pták P., Tkadlec J. The Dog-and-Rabbit Chase Revisited. *Acta Polytechnica.* 1996. vol. 36. pp. 5–10.
11. Mungan C.A. A Classic Chase Problem Solved from a Physics Perspective. *European Journal of Physics.* 2005. vol. 26. no. 6. pp. 985–990.
12. Sigaladze Z.K., Chashina O.I. [The problem of pursuing a hare by a wolf as an exercise of elementary kinematics]. *Vestnik NGU. Seriya "Fizika" – Siberian Journal of Physics.* 2010. Issue 5. vol. 2. pp. 111–115. (In Russ.).
13. Pogrebskaya T.N., Soltakhanov Sh.Kh. [The control of chasing a target by the pursuit method as a nonholonomic problem in mechanics]. *Vestnik Sankt-Peterburgskogo universiteta. Seriya 1. Matematika. Mekha-nika. Astronomiya – Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy.* 2007. vol. 1. pp. 117–126. (In Russ.).
14. Kuzmina L.I., Osipov Yu.V. [Calculation of the length of the trajectory for the problem of pursuit]. *Vestnik MGSU – Vestnik MGSU.* 2013. vol 12. pp. 20–26. (In Russ.).
15. Azamov A.A., Kuchkarov A.Sh., Samatov B.O. [The relation between problems of pursuit, controllability and stability in the large in linear systems with different types of constraints]. *Prikladnaya matematika i mehanika – Journal of Applied Mathematics and Mechanics.* 2007. Issue 71. vol. 2. pp. 259–263. (In Russ.).

16. Ivanov A.A., Shmakov O.A. [Algorithm for defining the inner geometry of a snakelike manipulator in case of leading link movements along the incremental trajectory]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2016. vol. 6(49). pp. 190–207. (In Russ.).
17. Lazarev V.S., Agadjanov D.E. [Using graphic-analytical methods for robots group movement trajectories formation in the two-dimensional environment]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2016. vol. 2(45). pp. 45–57. (In Russ.).
18. Shan Y. et al. CF-Pursuit: A Pursuit Method with a Clothoid Fitting and a Fuzzy Controller for Autonomous Vehicles. *International Journal of Advanced Robotic Systems.* 2015. vol. 12. no. 9. pp. 134.
19. Amer N.H., Zamzuri H., Hudha K., Kadir Z.A. Modelling and Control Strategies in Path Tracking Control for Autonomous Ground Vehicles: A Review of State of the Art and Challenges. *Journal of Intelligent & Robotic Systems.* 2017. vol. 86. no. 2. pp. 225–254.
20. Spakov O. et al. PursuitAdjuster: an exploration into the design space of smooth pursuit-based widgets. Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. 2016. pp. 287–290.
21. Khamis M. et al. Eyescout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays. Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST 2017). 2017. pp. 155–166.
22. Bardi M., Falcone M., Soravia P. Numerical Methods for Pursuit-Evasion Games via Viscosity Solutions. Stochastic and Differential Game. 1999. pp. 105–175.
23. Kumkov S., Ménec S., Patsko V. Zero-Sum Pursuit-Evasion Differential Games with Many Objects: Survey of Publications. *Dynamic Games and Applications*. 2017. vol. 7. no. 4. pp. 609–633.
24. Munts N., Kumkov S. [On the coincidence of the minimax solution and the value function in a time-optimal game with a lifeline]. *Trudy Instituta Matematiki i Mekhaniki UrO RAN – Proceedings of Krasovskii Institute of Mathematics and Mechanics*. 2018. vol. 24(2). pp. 200–214. (In Russ.).
25. Alimov K.N., Mamatov M.S. Solving a Pursuit Problem in High-order Controlled Distributed Systems. *Siberian Advances in Mathematics.* 2014. vol. 24(4). pp. 229–239.
26. Kazakov Yu.V. Zadacha o presledovanii na vrashchayushchemsya diske [Chase problem on a turntable]. Available at: http://old.exponenta.ru/EDUCAT/systemat/kasakov/pursuit/index.asp (accessed: 16.03.2019). (In Russ.).
27. Bakhvalov N.S., Zhidkov N.P., Kobelkov G.M. *Chislennyie metodyi* [Numerical Methods]. BINOM. Laboratoriya znaniy. 2008. 636 p. (In Russ.).
28. Ochkov V.F. *Mathcad 14 dlya studentov, inzhenerov u construktorov* [Mathcad 14 for Students, Engineers and Designers]. BHV-Peterburg. 2007. 368 p. (In Russ.).
29. Ochkov V.F., Bogomolova E.P., Ivanov D.A. *Fiziko-matematicheskie etyudy s Mathcad i Internet: Uchebnoe posobie* [Physics and Mathematics Studies with Mathcad and Internet: Tutorial]. Lan Publ. 2018. 560 p. (In Russ.).
30. Ponomarev A.A. [Suboptimal control construction for the model predictive controller]. *Vestnik Sankt-Peterburgskogo universiteta. Seriya 10. Prikladnaya matematika. Informatika. Processy upravleniya – Vestnik of Saint Peterburg University Applied Mathematics. Computer Science. Control Processes.* 2017. vol. 2. pp. 193–208. (In Russ.).
31. Ma Y.J. Reconstruction of a Robin coefficient by a predictor-corrector method. *Mathematical Problems in Engineering.* 2015. vol. 11. pp 1–7.
32. Binder A.J., Luskin M., Ortner C. Analysis of a predictor-corrector method for computationally efficient modeling of surface effects in 1D. 2016. Available at: https://arxiv.org/pdf/1605.05750v1.pdf (accessed: 16.03.2019).

33. Onwubuoya C., Akinyemi S.T., Odabi O.I., Odachi G.N. Numerical simulation of a computer virus transmission model using Euler predictor-corrector method. *IDOSR Journal of Applied Sciences.* 2018. vol. 3(1). pp. 16–28.
34. Abdullahi Y.A., Omar Z., Kuboye J.O. Derivation of block predictor – block corrector method for direct solution of third order ordinary differential equations. *Global Journal of Pure and Applied Mathematics.* 2016. vol. 12(1). pp. 343–350.
35. Daftardar-Gejji V., Sukale Y., Bhalekar S. A new predictor-corrector method for fractional differential equations. *Applied Mathematics and Computation*. 2014. vol. 244. pp. 158–182.
36. Ndanusa A., Tafida F.U. Predictor-corrector methods of high order for numerical integration of initial value problems. *International Journal of Scientific and Innovative Mathematical Research (IJSIMR).* 2016. vol. 4. no. 2. pp. 47–55.
37. Oghonyon J., Okunuga S.A., Lyase S.A. Milne's implementation on block predictor-corrector methods. *Journal of Applied Sciences.* 2016. vol. 16. no. 5. pp. 236–241.
38. Söderlind G. Multistep Methods. Encyclopedia of Applied and Computational Mathematics. 2015. 10 p.
39. Islam M.A. A comparative study on numerical solutions of initial value problems for ordinary differential equations with Euler and Runge Kutta methods. *American Journal of Computational Mathematics.* 2015. vol. 5. no. 3. pp. 393–404.
40. Fathoni M.F., Wuryandari A.I. Comparison between Euler, Heun, Runge-Kutta and Adams-Bashforth-Moulton integration methods in the particle dynamic simulation. 2015 4th International Conference on Interactive Digital Media (ICIDM). 2015. pp. 1–7.
41. Kuchkarov A.S. Solution of simple pursuit-evasion problem when evader moves on a given curve. *International Game Theory Review*. 2010. vol. 12. no. 3. pp. 223–238.
42. Galloway K.S., Justh E.W., Krishnaprasad P.S. Cyclic pursuit in three dimensions. 49th IEEE Conference on Decision and Control (CDC). 2010. pp. 7141–7146.
43. Galloway K.S., Justh E.W., Krishnaprasad P.S. Geometry of cyclic pursuit. Proceedings of the 48th IEEE Conference on Decision and Control (CDC). 2009. pp. 7485–7490.
44. Mukherjee D., Kumar S. Finite-time heterogeneous cyclic pursuit with application to target interception. 2018. Available at: https://arxiv.org/pdf/1811.10827.pdf (accessed: 16.03.2019).
45. Arnold M., Baryshnikov Y., Liberzon D. Cyclic pursuit without coordinates: convergence to regular polygon formations. 53rd IEEE Conference on Decision and Control. 2014. pp. 6191–6196.
46. Marshall J.A., Brouck M.E., Francis B.A. Pursuit formations of unicycles. *Automatica.* 2006. vol. 42. no. 1. pp. 3–12.
47. Sharma B., Ramakrishnan S., Kumar M. Cyclic pursuit in a multi-agent robotic system with double-integrator dynamics under linear interactions. *Robotica.* 2013. vol. 31. no. 7. pp. 1037–1050.
48. Arnold M., Zharnitsky V. Cyclic evasion in the three bug problem. *The American Mathematical Monthly*. 2015. vol. 122. no. 4. pp. 377–380.
49. Chapman S.J., Lottes J., Trefethen L.N. Four bugs on a rectangle. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2010. vol. 467. no. 2127. pp. 881–896.
50. Ding W., Yan G., Lin Z. Formations on Two-Layer Pursuit Systems. 2009 IEEE International Conference on Robotics and Automation. 2009. pp. 3496–3501.

Ochkov Valery Fedorovich — Ph.D., Dr.Sci., Professor, Professor, Department of Theoretical Bases of Heat Engineering, National Research University Moscow Power Engineering Institute. Research interests: information technologies in applications. The number of publications — 330. ochkov@twt.mpei.ac.ru; 14, Krasnokazarmennay str., 111250, Moscow, Russian Federation; office phone: +7(495) 362-71-71.

Vasileva Inna Evgen'evna — Ph.D., Associate Professor, Department of Mathematics, Military Educational and Scientific Center of the Air Force "N.E. Zhukovsky and Y.A. Gagarin Air Force Academy". Research interests: mathematical modelling, methods of mathematics and informatics in applications. The number of publications — 60. ivasad@mail.ru; 54 a, Old Bolsheviks str., 394064, Voronezh, Russian Federation; office phone: +7 (473) 244-76-13.

УДК 519.62 **DOI** 10.15622/sp.2019.18.6.1407-1433

В.Ф. Очков, И.Е. Васильева

ПРИМЕНЕНИЕ РАЗНОСТНЫХ СХЕМ К РЕШЕНИЮ ЗАДАЧИ О ПОГОНЕ

Очков В.Ф., Васильева И.Е. **Применение разностных схем к решению задачи о погоне.**

Аннотация. В работе рассматривается один из аспектов задачи о преследовании: построение траекторий движения преследователя для случая, когда преследование осуществляется по методу погони, то есть касательная, проведенная к траектории движения преследователя в любой момент времени, проходит через положение точки, которая ассоциируется с преследуемым. Предлагается новый подход построения кривых погони путем использования разностных схем. Данная методика позволяет отказаться от необходимости составлять дифференциальные уравнения для описания кривых погони, что бывает достаточно сложно сделать в общем случае. Кроме того, применение разностных схем обосновано в ситуации, когда нахождение аналитического решения уже имеющегося дифференциального уравнения затруднительно, и дает возможность получить кривую погони численным способом. Построены различные модификации разностных схем, являющиеся аналогами схем на основе методов Эйлера, Адамса — Башфорта и Милна. Осуществлена их программная реализация с помощью математического пакета Mathcad. Рассмотрен случай равномерного прямолинейного движения преследуемого, для которого известно дифференциальное уравнение, описывающее траекторию преследователя, и его аналитическое решение. Проведен сравнительный анализ полученных разными методами численных решений и известного аналитического решения. Найдена погрешность полученных численных реализаций. Рассмотрено применение построенных разностных схем для более общего случая произвольной траектории преследуемого. Также описан алгоритм распространения предложенного метода для случая циклического преследования с несколькими участниками в трехмерном пространстве. В частности, построена разностная схема, аналогичная методу Эйлера, для трехмерного аналога «задачи о жуках». Полученные результаты продемонстрированы на анимационных примерах как для двумерного, так и трехмерного случаев.

Ключевые слова: дифференциальные игры, задача о преследовании, метод погони, кривая погони, численные методы, разностные схемы, метод Эйлера, «задача о трех жуках», Mathcad.

Очков Валерий Федорович — д-р техн. наук, профессор, профессор, кафедра теоретических основ теплотехники, Федеральное государственное бюджетное образовательное учреждение высшего образования «Национальный исследовательский университет МЭИ». Область научных интересов: информационные технологии в прикладных задачах. Число научных публикаций — 330. ochkov@twt.mpei.ac.ru; ул. Красноказарменная, 14, 111250, Москва, Российская Федерация; р.т.: +7(495) 362-71-71.

Васильева Инна Евгеньевна — канд. физ.-мат. наук, доцент, кафедра математики, Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина». Область научных интересов: математическое моделирование, методы математики и информатики в прикладных задачах. Число научных публикаций — 60. ivasad@mail.ru; ул. Старых Большевиков, 54 а, 394064, Воронеж, Российская Федерация; р.т.: +7 (473) 244-76-13.

Литература

1. *Petrosyan L.A., Mazalov V.V.* Recent Advances in Game Theory and Applications // Springer International Publishing. 2016. 284 p.
2. *Mazalov V., Chirkova J.V.* Networking Games. Network Forming Games and Games on Networks // Academic Press. 2019. 322 p.
3. *Petrosyan L.A.* Differential Games of Pursuit // World Scientific. 1993. 332 p.
4. *Isaacs R.* Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization // Courier Corporation. 1999. 416 p.
5. *Kumkov S.S., Le Menec S., Patsko V.S.* Zero-sum pursuit-evasion differential games with many objects: survey of publications // Dynamic Games and Applications. 2017. vol. 7. no. 4. pp. 609–633.
6. *Ramana M.V., Kothari M.* Pursuit-evasion games of high speed evade // Journal of Intelligent & Robotic Systems. 2017. vol. 85. no. 2. pp. 293–306.
7. *Паньков С.Я., Забураев Е.Ю., Матвеев А.М.* Теория и методика управления авиацией // Ульяновск: УВАУ ГА. 2006. 190 с.
8. *Barton J.C., Eliezer C.J.* On pursuit curves // The ANZIAM Journal. 2000. vol. 41. no. 3. pp. 358–371.
9. *Самоявчева М.В., Федоров Л.И.* Задача о погоне // Вестник Московского государственного университета. Серия: Физика-Математика. 2011. № 1. С. 65–69.
10. *Pták P., Tkadlec J.* The Dog-and-Rabbit Chase Revisited // Acta Polytechnica. 1996. vol. 36. pp. 5–10.
11. *Mungan C.A.* A Classic Chase Problem Solved from a Physics Perspective // European Journal of Physics. 2005. vol. 26. no. 6. pp. 985–990.
12. *Сигаладзе З.К., Чацина О.И.* Задача преследования зайца волком как упражнение элементарной кинематики // Вестник НГУ. Серия: Физика. 2010. Т. 5. № 2. С. 111–115.
13. *Погребская Т.Н., Солтаханов Ш.Х.* Управление преследованием цели по методу погони как неголономная задача механики // Вестник Санкт-Петербургского университета. Серия 1. Математика. Механика. Астрономия. 2007. Вып. 1. С. 117–126.
14. *Кузьмина Л.И., Осипов Ю.В.* Расчет длины траектории для задачи преследования // Вестник МГСУ. 2013. № 12. С. 20–26.
15. *Азамов А.А., Кучкаров А.Ш., Саматов Б.О.* О связи между задачами преследования, управляемости и устойчивости в целом в линейных системах с разнотипными ограничениями // Прикладная математика и механика. 2007. Т. 71. Вып. 2. С. 259–263.
16. *Иванов А.А., Шмаков О.А.* Алгоритм определения внутренней геометрии манипулятора змеевидного типа при движении лидирующего звена по наращиваемой траектории // Труды СПИИРАН. 2016. Вып. 6(49). С. 190–207.
17. *Лазарев В.С., Агаджанов Д.Э.* Использование графоаналитических методов для формирования траектории группы подвижных объектов в двумерной среде // Труды СПИИРАН. 2016. Вып. 2(45). С. 45–57.
18. *Shan Y. et al.* CF-Pursuit: A Pursuit Method with a Clothoid Fitting and a Fuzzy Controller for Autonomous Vehicles // International Journal of Advanced Robotic Systems. 2015. vol. 12. no. 9. pp. 134.
19. *Amer N.H., Zamzuri H., Hudha K., Kadir Z.A.* Modelling and Control Strategies in Path Tracking Control for Autonomous Ground Vehicles: A Review of State of the Art and Challenges // Journal of Intelligent & Robotic Systems. 2017. vol. 86. no. 2. pp. 225–254.
20. *Spakov O. et al.* PursuitAdjuster: an exploration into the design space of smooth pursuit-based widgets // Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. 2016. pp. 287–290.

21. *Khamis M. et al.* Eyescout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays // Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST 2017). 2017. pp. 155–166.
22. *Bardi M.*, *Falcone M.*, *Soravia P.* Numerical Methods for Pursuit-Evasion Games via Viscosity Solutions // Stochastic and Differential Game. 1999. pp. 105–175.
23. *Kumkov S., Ménec S., Patsko V.* Zero-Sum Pursuit-Evasion Differential Games with Many Objects: Survey of Publications // Dynamic Games and Applications. 2017. vol. 7. no. 4. pp. 609–633.
24. *Мунц Н.В., Кумков С.С.* О совпадении минимаксного решения и функции цены игры быстродействия с линией жизни // Труды Института математики и механики УрО РАН. 2018. Вып. 24(2). С. 200–214.
25. *Alimov K.N., Mamatov M.S.* Solving a Pursuit Problem in High-order Controlled Distributed Systems // Siberian Advances in Mathematics. 2014. vol. 24(4). pp. 229–239.
26. *Казаков Ю.В.* Задача о преследовании на вращающемся диске. URL: http://old.exponenta.ru/EDUCAT/ systemat/kasakov/pursuit/index.asp (дата обращения: 16.03.2019).
27. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы // БИНОМ. Лаборатория знаний. 2008. 636 с.
28. *Очков В.Ф.* Mathcad 14 для студентов, инженеров и конструкторов // БХВ-Петербург. 2007. 368 с.
29. *Очков В.Ф., Богомолова Е.П., Иванов Д.А.* Физико-математические этюды с Mathcad и Интернет: учебное пособие // Лань. 2018. 560 с.
30. *Пономарев А.А.* Аппроксимация обратной связи в регуляторе «Предиктор-корректор» явной функцией // Вестник Санкт-Петербургского университета. Серия 10. Прикладная математика. Информатика. Процессы управления. 2017. № 2. С. 193–208.
31. *Ma Y.J.* Reconstruction of a Robin coefficient by a predictor-corrector method // Mathematical Problems in Engineering. 2015. vol. 11. pp 1–7.
32. *Binder A.J., Luskin M., Ortner C.* Analysis of a predictor-corrector method for computationally efficient modeling of surface effects in 1D. 2016. URL: https://arxiv.org/pdf/1605.05750v1.pdf (дата обращения: 16.03.2019).
33. *Onwubuoya C., Akinyemi S.T., Odabi O.I., Odachi G.N.* Numerical simulation of a computer virus transmission model using Euler predictor-corrector method // IDOSR Journal of Applied Sciences. 2018. vol. 3(1). pp. 16–28.
34. *Abdullahi Y.A., Omar Z., Kuboye J.O.* Derivation of block predictor – block corrector method for direct solution of third order ordinary differential equations // Global Journal of Pure and Applied Mathematics. 2016. vol. 12(1). pp. 343–350.
35. *Daftardar-Gejji V., Sukale Y., Bhalekar S.* A new predictor-corrector method for fractional differential equations // Applied Mathematics and Computation. 2014. vol. 244. pp. 158–182.
36. *Ndanusa A., Tafida F.U.* Predictor-corrector methods of high order for numerical integration of initial value problems // International Journal of Scientific and Innovative Mathematical Research (IJSIMR). 2016. vol. 4. no. 2. pp. 47–55.
37. *Oghonyon J.,Okunuga S.A., Lyase S.A.* Milne's implementation on block predictor-corrector methods // Journal of Applied Sciences. 2016. vol. 16. no. 5. pp. 236–241.
38. *Söderlind G.* Multistep Methods // Encyclopedia of Applied and Computational Mathematics. 2015. 10 p.
39. *Islam M.A.* A comparative study on numerical solutions of initial value problems for ordinary differential equations with Euler and Runge Kutta methods // American Journal of Computational Mathematics. 2015. vol. 5. no. 3. pp. 393–404.

40. *Fathoni M.F., Wuryandari A.I.* Comparison between Euler, Heun, Runge-Kutta and Adams-Bashforth-Moulton integration methods in the particle dynamic simulation // 2015 4th International Conference on Interactive Digital Media (ICIDM). 2015. pp. 1–7.
41. *Kuchkarov A.S.* Solution of simple pursuit-evasion problem when evader moves on a given curve // International Game Theory Review. 2010. vol. 12. no. 3. pp. 223–238.
42. *Galloway K.S., Justh E.W., Krishnaprasad P.S.* Cyclic pursuit in three dimensions // 49th IEEE Conference on Decision and Control (CDC). 2010. pp. 7141–7146.
43. *Galloway K.S., Justh E.W., Krishnaprasad P.S.* Geometry of cyclic pursuit // Proceedings of the 48th IEEE Conference on Decision and Control (CDC). 2009. pp. 7485–7490.
44. *Mukherjee D., Kumar S.* Finite-time heterogeneous cyclic pursuit with application to target interception. 2018. URL: https://arxiv.org/pdf/1811.10827.pdf (дата обращения: 16.03.2019).
45. *Arnold M., Baryshnikov Y., Liberzon D.* Cyclic pursuit without coordinates: convergence to regular polygon formations // 53rd IEEE Conference on Decision and Control. 2014. pp. 6191–6196.
46. *Marshall J.A., Brouck M.E., Francis B.A.* Pursuit formations of unicycles // Automatica. 2006. vol. 42. no. 1. pp. 3–12.
47. *Sharma B., Ramakrishnan S., Kumar M.* Cyclic pursuit in a multi-agent robotic system with double-integrator dynamics under linear interactions // Robotica. 2013. vol. 31. no. 7. pp. 1037–1050.
48. *Arnold M., Zharnitsky V.* Cyclic evasion in the three bug problem // The American Mathematical Monthly. 2015. vol. 122. no. 4. pp. 377–380.
49. *Chapman S.J., Lottes J., Trefethen L.N.* Four bugs on a rectangle // Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2010. vol. 467. no. 2127. pp. 881–896.
50. *Ding W., Yan G., Lin Z.* Formations on Two-Layer Pursuit Systems // 2009 IEEE International Conference on Robotics and Automation. 2009. pp. 3496–3501.

УДК 621.398(075.8) **DOI** 10.15622/sp.2019.18.6.1434-1461

Е.В. Копкин, И.М. Кобзарев

ИСПОЛЬЗОВАНИЕ МЕРЫ ЦЕННОСТИ ИНФОРМАЦИИ СТРАТОНОВИЧА ДЛЯ ОПТИМИЗАЦИИ ГИБКИХ ПРОГРАММ ДИАГНОСТИРОВАНИЯ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Копкин Е.В., Кобзарев И.М. **Использование меры ценности информации Стратоновича для оптимизации гибких программ диагностирования технических объектов.**

Аннотация. Существующие методы вычисления ценности диагностической информации, циркулирующей в автоматизированных системах мониторинга технического состояния объектов, не учитывают потери (выигрыши), связанные с принятием неправильных решений при идентификации этого состояния.

Цель работы — разработка алгоритма, позволяющего решить задачу распознавания технического состояния, в котором находится анализируемый объект, методом динамического программирования, используя в качестве оптимизируемого показателя ценность диагностической информации. Решение задачи оптимизации диагностической процедуры основано на использовании меры ценности информации Р. Л. Стратоновича, модифицированной применительно к предметной области технического диагностирования и при использовании диагностических признаков, представленных в виде интервалов на вещественной числовой оси. Максимальная ценность диагностической информации достигается при минимизации средних потерь (максимизации средних выигрышей), получаемых при выполнении проверок диагностических признаков, в процессе распознавания технического состояния объекта.

Для решения задачи предложено обладающее научной новизной рекуррентное выражение, позволяющее вычислять ценность информации, получаемой при выполнении проверок диагностических признаков в каждом из анализируемых информационных состояний процесса диагностирования. В процессе реализации программы диагностирования при распознавании технического состояния объекта возможны как потери, так и выигрыши. Разность их априорных и апостериорных средних значений численно характеризует ценность диагностической информации. Величина показателя ценности информации зависит от вероятностей исходов проверок диагностических признаков и пропорциональна разности апостериорных и априорных вероятностей достижения цели диагностирования.

Использование предложенного решения позволяет синтезировать оптимальную по критерию максимума ценности диагностической информации гибкую программу диагностирования в виде ориентированного графа или упорядоченных по очередности их выполнения наборов проверок, необходимых для распознавания конкретного технического состояния, в котором находится объект.

Реализация разработанного алгоритма возможна в программно-алгоритмическом обеспечении автоматизированных систем мониторинга состояния сложных технических объектов.

Ключевые слова: техническое состояние, гибкая программа диагностирования, ценность информации, диагностический признак.

1. Введение. Вопросами ценности информации научное сообщество начало заниматься во второй половине XX века. Академик А. А. Харкевич сформулировал понятие ценности информации одним из первых [1]. Он предположил, что если получение информации увеличивает вероятность достижения цели, то эта информация будет

обладать некоторой положительной ценностью. Предложенная А. А. Харкевичем мера ценности информации определяется как двоичный логарифм отношения вероятности P_1 достижения цели после получения информации к вероятности P_0 достижения цели до получения этой информации, то есть в соответствии с выражением:

$$V = \log_2 \frac{P_1}{P_0}. \quad (1)$$

В работе М. М. Бонгарда [2] рассматривались ситуации, в которых получаемая информация может вообще не обладать никакой ценностью или даже быть отрицательной, если осуществляется передача ложной информации, что приводит к увеличению неопределенности.

Очевидно, что существует логическая связь между количеством информации по Шеннону, то есть величиной снятой неопределенности, и показателем (1). Ценность информации измеряется в битах, а ее числовые значения лежат в диапазоне от -∞ до +∞.

Предложенный В. И. Корогодиным в работе [3] показатель ценности информации определяется по формуле:

$$V = \frac{P_1 - P_0}{1 - P_0}. \quad (2)$$

Этот показатель численно изменяется от 0 до 1 и по свойствам не отличается от меры ценности А. А. Харкевича (1).

Предложенный Р. Л. Стратоновичем метод определения ценности информации [4] предусматривает, что достижение цели за счет получения некоторого количества информации возможно различными путями. При этом более ценной будет та информация, получение которой снижает условные штрафы, которые приходится платить для достижении цели, или, наоборот, увеличивает получаемый в этом процессе выигрыш.

Свою теорию ценности информации предложил также Р. Ховард [5] в 1966 году. В отличие от теории информации Шеннона, которая измеряет ее количество в битах, Р. Ховард количественно определяет ценность информации как функцию от ее потенциала, помогающего в принятии решения. Центральная идея этой теории представляется в виде уравнения:

$$V(I) = U(d^*) - U(d), \quad (3)$$

где I — информация; $V(I)$ — ее ценность; d — решение, принимаемое до получения информации; d^* — измененное решение, принимаемое

после получения информации; $U(d)$ и $U(d^*)$ — функции ценности соответствующих решений.

Ключевым компонентом в процессе принятия решения является функция ценности U, которая, используя численное значение $V(I)$ ценности информации, определяет выигрыш (доход) от принятого решения. Если принятое решение позволяет получить предпочтительный результат, то ему назначается более высокая ценность по сравнению с решением, способствующим достижению менее предпочтительного результата. В соответствии с выражением (3) ценность информации представляет собой разность функций ценности решений, принимаемых до и после получения информации. Таким образом, различная информация обладает различными значениями ценности.

Теория ценности информации Ховарда была применена ко множеству исследований в различных предметных областях, таких как инвестиционный анализ [6-8], клинические испытания [9-11], инфраструктурные проекты [12-15], антикризисный менеджмент [16], география, геофизика и геология [17-20]. В этих исследованиях для вычисления ценности информации использовалась денежная стоимость.

В технических приложениях понятие «ценность информации» особенно актуально при обеспечении защиты информации. Аксиоматика понятия ценности информации и математические модели, с помощью которых эту ценность можно количественно оценить, предложены в монографии Г. П. Шанкина [21], развивающей идеи М. М. Бонгарда и Р. Л. Стратоновича. Величина возможного выигрыша, который можно получить, обладая некоторым количеством информации, и определяет ее ценность для пользователя при решении конкретных задач.

Процессы передачи информации по каналам связи имеют различные цели, достижение которых можно оценить соответствующими частными показателями, имеющими различную степень важности. Свертка этих частных показателей позволяет сформировать обобщенный показатель ценности информации, основанный на использовании информации для повышения эффективности процесса ее защиты [22, 23]. Численное значение обобщенного показателя ценности информации определяется на основе логико-лингвистической модели оценки свойств, характеризующих этот показатель, и метода попарных сравнений, используемого для определения значений коэффициентов важности этих свойств.

Показатель ценности информации, предложенный в [24], представляет собой отношение величины смыслового значения общего объема информации к количеству единиц (объемов), необходимых для ее передачи. Другими словами, ценность определяется глубиной

смысла и обратно пропорциональна количеству символов для его передачи. При этом, однако, не уточняются понятия «смысловое значение» и «единицы информации», которые зависят от конкретной предметной области используемой информации.

Современные информационно-измерительные системы, применяемые при управлении сложными техническими объектами и технологическими процессами, обрабатывают огромные потоки разнородной измерительной информации и, как правило, имеют в своем составе автоматизированные системы мониторинга (АСМ) состояния этих объектов и процессов. В контурах АСМ обрабатывается и анализируется диагностическая информация, используемая для принятия решений о техническом состоянии объектов и для управления этими состояниями. Оперативность принятия решений можно существенно повысить при задействовании только такой диагностической информации, которая будет считаться ценной для определения технического состояния. В связи с этим приложение теории ценности информации к задачам контроля и диагностирования представляет значительный научный и практический интерес.

В частности, в [25] предложена информационная мера диагностического признака (показатель ценности диагностической информации), используемого для определения состояния объекта, как количество информации, вносимое рассматриваемым признаком для определения этого состояния.

В [26, 27] представлены оптимальные алгоритмы построения гибких программ диагностирования (ГПД) объектов на основе использования меры ценности информации А. А. Харкевича и В. И. Корогодина, соответственно. Использование меры ценности Р. Л. Стратоновича для построения ГПД предложено в [28], однако только для случая, когда диагностические признаки имеют дискретную форму представления.

Следует отметить, что системы диагностирования, использующие дискретные диагностические признаки, имеют сравнительно невысокую помехоустойчивость. Если создаваемая АСМ предназначена для контроля и диагностирования бортовой аппаратуры образцов ракетно-космической техники, то радиотелеметрические каналы передачи диагностической информации в наибольшей степени подвержены воздействию естественных и искусственных возмущающих воздействий. В этом случае способ представления информации в виде дискретных признаков оказывается малоэффективным.

В этом отношении более надежными и эффективными являются непрерывные признаки, представленные в виде абстрактных

вещественных чисел, содержащих в себе сведения об измеренных на борту параметрах, причем характер зависимости каждого из чисел от того или иного параметра не раскрывается. В процессе диагностирования предусматривается интервальное оценивание непрерывных признаков (а не точечное, как дискретных), благодаря чему возмущающие воздействия на передаваемые с борта сообщения оказываются значительно ослабленными.

В силу указанных обстоятельств разработка алгоритма построения ГПД, оптимальной в смысле ценности (по Стратоновичу) анализируемой диагностической информации и устойчивой к воздействию помех благодаря использованию непрерывных признаков, является актуальной задачей, имеющей важное научно-прикладное значение. Разумеется, переход к непрерывным признакам существенно осложняет процесс построения алгоритма, что является платой за вновь приобретенные его полезные качества.

2. Постановка задачи. Для формальной постановки задачи воспользуемся известными моделями объекта и процесса диагностирования, представленными в работе [26]. При этом будем считать заданными: множество $S = \left\{ S_i \mid i = \overline{1,m} \right\}$ возможных технических состояний (ТС) объекта; множество $\Pi = \left\{ \pi_j \mid j = \overline{1,n} \right\}$ диагностических признаков, обеспечивающих попарную различимость этих ТС; множество $L = \left\{ \ell_{ij} \mid i = \overline{1,m}; j = \overline{1,n} \right\}$ интервалов на вещественной числовой оси, характеризующих допустимые значения каждого признака в каждом ТС. Измеренные значения y_j соответствующих признаков $\pi_j \left(j = \overline{1,n} \right)$, которые являются вещественными числами, имеющими равномерный закон распределения, представляют собой поступающую с объекта измерительную информацию.

Процесс измерения текущего значения y_j признака π_j и выявления попадания этого значения в интервал $\ell_{ij} \in L$ (или в несколько интервалов в случае их пересечения) называется проверкой и обозначается символом $\hat{\pi}_j$.

Основные элементы процесса диагностирования — информационные состояния (ИС) $R_k \subseteq S$ и выполняемые в них проверки $\hat{\pi}_j$, образующие множества $\Omega = \left\{ R_k \middle| k = \overline{1,M} \right\}$ и $\hat{\Pi} = \left\{ \hat{\pi}_j \mid j = \overline{1,n} \right\}$ соответственно. ИС процесса диагностирования является виртуальной

структурой — подмножеством множества S, в каждом из элементов которого с некоторой вероятностью может находиться объект в момент анализа его состояния. Будем различать конечные ИС $R_i = \{S_i\}$, $i = \overline{1, m}$, являющиеся распознанными ТС объекта, от всех остальных ИС $R_k \subseteq S(k = \overline{m+1, M})$, для которых $2 \le \text{card}\{R_k\} \le m$.

Выбирая для проверки в каждом из ИС R_k только те признаки π_j, по которым хотя бы одно ТС отличается от других, можно сделать процесс диагностирования более эффективным. Сформируем для каждого ИС R_k подмножество Π_k признаков, допустимых для проверки в этом ИС, определяемое согласно условию:

$$\pi_j \in \Pi_k, \text{если} \left(\exists S_i, S_f \in R_k\right) : \left(\ell_{ij} \bigcap \ell_{fj} = \varnothing\right). \quad (4)$$

Для рассматриваемых признаков $\pi_j \in \Pi_k$ на вещественной числовой оси можно выделить некоторое количество подынтервалов Δ_{kj}, определяющих диапазоны, в котором интервалы ℓ_{ij}, характеризующие ТС S_i, пересекаются:

$$\Delta_{kj} = \bigcap_{\{i\}} \ell_{ij}.$$

Событие ($y_j \in \Delta_{kj}$), заключающееся в попадании измеренного значения y_j признака π_j в один из подынтервалов Δ_{kj}, называется исходом проверки $\hat{\pi}_j$, выполненной в ИС R_k. Обозначим через Δ_{kj}^v подынтервал Δ_{kj}, имеющий порядковый номер v из их общего возможного числа ω_{kj}. Тогда событие $\left(y_j \in \Delta_{kj}^v\right), v = \overline{1, \omega_{kj}}$, представляющее собой v-й исход проверки $\hat{\pi}_j$, выполненной в ИС R_k, обозначим через $\hat{\pi}_j^v$.

При этом проверка $\hat{\pi}_j$ формально описывается отображением:

$$\hat{\pi}_j : R_k \to R_{kj}^v, \text{если } y_j \in \Delta_{kj}^v \ \left(v = \overline{1, \omega_{kj}}\right), \quad (5)$$

где

$$R_{kj}^v = \left\{ S_i \in R_k \mid i : \bigcap_{\{i\}} \ell_{ij} \neq \varnothing \right\}.$$

Новое ИС R_{kj}^{v}, которое получается из ИС R_k при *v*-м исходе проверки $\hat{\pi}_j$, имеет в своем составе меньшее количество ТС, то есть $\text{card}\left\{R_{kj}^{v}\right\} < \text{card}\{R_k\}$.

Вероятность реализации события $\left(y_j \in \Delta_{kj}^{v}\right)$ назовем вероятностью *v*-го исхода проверки признака π_j в ИС $R_k \subseteq S$, обозначим через $P_k\left(\hat{\pi}_j^{v}\right)$ и будем вычислять по формуле:

$$P_k\left(\hat{\pi}_j^{v}\right) = P\left(R_{kj}^{v} / R_k\right) = P\left(y_j \in \Delta_{kj}^{v}\right) = \frac{|\Delta_{kj}^{v}|}{|\nabla_{kj}|}, \qquad (6)$$

где $\left|\Delta_{kj}^{v}\right| = \left|\bigcap_{\left\{i:S_i \in R_{kj}^{v}\right\}} \ell_{ij}\right|$ и $\left|\nabla_{kj}\right| = \left|\bigcup_{\{i:S_i \in R_k\}} \ell_{ij}\right|$ — длины пересечения и объединения соответствующих интервалов.

При указанных условиях синтезируем ориентированный граф *G*, вершины которого представляют собой ИС $R_k \subset S$ процесса диагностирования, а дуги — исходы $\hat{\pi}_j^{v}$ проверок в этих ИС. Граф *G*, состоящий из ветвей $G_r \in U$ (*r* — порядковый номер ветви, *U* — множество всех ветвей), имеет одну начальную и *m* конечных вершин, соответствующих опознанным ТС объекта $S_i\left(i = \overline{1, m}\right)$. Ветви G_r графа образуют полную группу несовместных событий, то есть $\sum_{G_r \in U} P\left(G_r\right) = 1$.

Состав проверок, входящих в каждую из ветвей G_r графа, а также последовательность их выполнения, образуют гибкую программу диагностирования, в ходе реализации которой обеспечивается распознавание каждого из заданных ТС $S_i \in S\left(i = \overline{1, m}\right)$ объекта с максимальной в среднем ценностью *V* диагностической информации, получаемой при выполнении проверок.

Построенный граф *G* должен удовлетворять условию:

$$G = \arg\max_{x=1,2,\ldots} \left\{V\left(G_x\right)\right\},$$

где $V(G_x)$ — усредненный по всем ветвям показатель ценности диагностической информации, получаемой с помощью некоторого варианта искомой ГПД; x — номер варианта.

В процессе синтеза ГПД формируются упорядоченные подмножества $\Pi_r \subseteq \Pi$ признаков, последовательная проверка которых обеспечивает распознавание i-го ТС объекта (r — номер ветви, ведущей в i-е ТС). Поскольку эти подмножества являются структурно взаимосвязанными в рамках составляемой программы, то они не могут быть сформированы по отдельности. Кроме того, при использовании диагностических признаков, имеющих непрерывную форму представления, в одно и то же ТС S_i могут приводить несколько ветвей G_r, поэтому для каждого ТС может существовать несколько подмножества Π_r.

3. Вычисление ценности диагностической информации на основе меры Р. Л. Стратоновича. Любое ИС $R_k \in \Omega$ (кроме конечных $R_i = \{S_i\}$, $i = \overline{1, m}$) обладает некоторой энтропией (неопределенностью) и может рассматриваться как отдельный источник информации. Тогда выполняемую в этом ИС проверку $\hat{\pi}_j$ можно рассматривать как средство получения диагностической информации, снижающей исходную неопределенность и способствующей достижению цели.

Проверка $\hat{\pi}_j$, выполненная в ИС R_k, имеет произвольное конечное число ω_{kj} случайных исходов $\hat{\pi}_j^v \left(v = \overline{1, \omega_{kj}}\right)$. В соответствии с отображением (5) в результате каждого исхода $\hat{\pi}_j^v$ получается новое ИС $R_{kj}^v \left(v = \overline{1, \omega_{kj}}\right)$, для которого $\text{card}\left\{R_{kj}^v\right\} < \text{card}\{R_k\}$.

Обозначим через R_{kj}^u и $R_{kj}^v \left(u, v = \overline{1, \omega_{kj}}\right)$ конкурирующие гипотезы о возможных исходах проверки $\hat{\pi}_j$ до и после ее выполнения. Соответственно, вероятности $P_k\left(\hat{\pi}_j^u\right)$ и $P_k\left(\hat{\pi}_j^v\right)$ исходов, в результате которых получаются гипотетические ИС R_{kj}^u и R_{kj}^v, будем рассматривать как априорные и апостериорные вероятности достижения цели.

Р. Л. Стратонович предложил количественную оценку ценности информации в виде разности априорных и апостериорных средних потерь (штрафов), либо разности апостериорных и априорных выигрышей, получаемых в процессе достижения цели [4].

Пусть исходом проверки $\hat{\pi}_j$, выполненной в ИС R_k, является ИС $R_{kj}^v = R_{kj}^u$, то есть гипотеза R_{kj}^u подтверждается. Тогда получаемые при этом потери оценим величиной $\alpha \geq 0$. Если же полученное ИС $R_{kj}^v \neq R_{kj}^u$, то будем оценивать потери величиной $\beta_i > \alpha \left(i = \overline{1, \omega_{kj} - 1} \right)$.

Если использовать стратегию максимизации «выигрышей», то величиной $\alpha \geq 0$. будет оцениваться выигрыш, получаемый в случае, когда $R_{kj}^v = R_{kj}^u$, а величиной $\beta_i < \alpha$ — выигрыш, получаемый в ином случае.

Априорные средние потери, получаемые до выполнения проверки $\hat{\pi}_j$ в ИС R_k, можно оценить как:

$$L\left(R_{kj}^u\right) = P_k\left(\hat{\pi}_j^1\right)\beta_1 + P_k\left(\hat{\pi}_j^2\right)\beta_2 + \ldots + P_k\left(\hat{\pi}_j^{u-1}\right)\beta_{u-1} + P_k\left(\hat{\pi}_j^u\right)\alpha + \\ + P_k\left(\hat{\pi}_j^{u+1}\right)\beta_{u+1} + P_k\left(\hat{\pi}_j^{u+2}\right)\beta_{u+2} + \ldots + P_k\left(\hat{\pi}_j^{\omega_{kj}}\right)\beta_{\omega_{kj}} = \\ = P_k\left(\hat{\pi}_j^u\right)\alpha + \sum_{R_{kj}^w \in R_k \setminus \{R_{kj}^u\}} P_k\left(\hat{\pi}_j^w\right)\beta_w.$$

Апостериорные средние потери, получаемые после выполнения проверки $\hat{\pi}_j$, можно оценить следующим образом:

$$L\left(R_{kj}^v / R_{kj}^u\right) = P_k\left(\hat{\pi}_j^1\right)\beta_1 + P_k\left(\hat{\pi}_j^2\right)\beta_2 + \ldots + P_k\left(\hat{\pi}_j^{v-1}\right)\beta_{v-1} + P_k\left(\hat{\pi}_j^v\right)\alpha + \\ + P_k\left(\hat{\pi}_j^{v+1}\right)\beta_{v+1} + P_k\left(\hat{\pi}_j^{v+2}\right)\beta_{v+2} + \ldots + P_k\left(\hat{\pi}_j^{\omega_{kj}}\right)\beta_{\omega_{kj}} = \\ = P_k\left(\hat{\pi}_j^v\right)\alpha + \sum_{R_{kj}^w \in R_k \setminus \{R_{kj}^v\}} P_k\left(\hat{\pi}_j^w\right)\beta_w.$$

Аналогично, априорный и апостериорный средние выигрыши оцениваются величинами:

$$G\left(R_{kj}^u\right) = P_k\left(\hat{\pi}_j^u\right)\alpha + \sum_{R_{kj}^w \in R_k \setminus \{R_{kj}^u\}} P_k\left(\hat{\pi}_j^w\right)\beta_w;$$

$$G\left(R_{kj}^v / R_{kj}^u\right) = P_k\left(\hat{\pi}_j^v\right)\alpha + \sum_{R_{kj}^w \in R_k \setminus \{R_{kj}^v\}} P_k\left(\hat{\pi}_j^w\right)\beta_w.$$

Если в качестве основной выбрана гипотеза R_{kj}^{u}, а исходом проверки является $\hat{\pi}_j^{v}$, то при использовании стратегии минимизации потерь ценность $V_k\left(\hat{\pi}_j^{v,u}\right)$ получаемой информации определяется по формуле:

$$V_k\left(\hat{\pi}_j^{v,u}\right) = L\left(R_{kj}^{u}\right) - L\left(R_{kj}^{v} \middle/ R_{kj}^{u}\right) = \alpha\left[P_k\left(\hat{\pi}_j^{u}\right) - P_k\left(\hat{\pi}_j^{v}\right)\right] + \\ + \sum_{R_{kj}^{w} \in R_k \setminus \{R_{kj}^{u}\}} P_k\left(\hat{\pi}_j^{w}\right)\beta_w - \sum_{R_{kj}^{w} \in R_k \setminus \{R_{kj}^{v}\}} P_k\left(\hat{\pi}_j^{w}\right)\beta_w;\ v,u = \overline{1,\omega_{kj}}. \qquad (7)$$

При использовании стратегии максимизации выигрышей ценность получаемой информации определяется по формуле

$$V_k\left(\hat{\pi}_j^{v,u}\right) = G\left(R_{kj}^{v} \middle/ R_{kj}^{u}\right) - G\left(R_{kj}^{u}\right) = \alpha\left[P_k\left(\hat{\pi}_j^{v}\right) - P_k\left(\hat{\pi}_j^{u}\right)\right] + \\ + \sum_{R_{kj}^{w} \in R_k \setminus \{R_{kj}^{v}\}} P_k\left(\hat{\pi}_j^{w}\right)\beta_w - \sum_{R_{kj}^{w} \in R_k \setminus \{R_{kj}^{u}\}} P_k\left(\hat{\pi}_j^{w}\right)\beta_w;\ v,u = \overline{1,\omega_{kj}}. \qquad (8)$$

После выполнения проверки $\hat{\pi}_j$ основная гипотеза R_{kj}^{u} может подтвердиться ($v = u$) или не подтвердиться ($v \neq u$) с вероятностью $P_k\left(\hat{\pi}_j^{v}\right)$, определяемой по формуле (6).

В зависимости от той гипотезы R_{kj}^{u}, которая принята в качестве основной, а также от исхода $\hat{\pi}_j^{v}$ выполняемой проверки, будет меняться и величина показателя $V_k\left(\hat{\pi}_j^{v,u}\right)$, оцениваемого формулами (7) и (8).

Двойное суммирование (по числу ω_{kj} выдвинутых гипотез R_{kj}^{u} и по числу возможных исходов $\hat{\pi}_j^{v}$ проверки $\hat{\pi}_j$) взвешенных с помощью вероятностей $P_k\left(\hat{\pi}_j^{v}\right)$ значений показателя $V_k\left(\hat{\pi}_j^{v,u}\right)$ позволяет оценить среднюю ценность $V_k\left(\hat{\pi}_j\right)$ информации, получаемой в результате выполнения проверки $\hat{\pi}_j$ в ИС R_k,:

$$V_k\left(\hat{\pi}_j\right) = \sum_{v=1}^{\omega_{kj}} \sum_{u=1}^{\omega_{kj}} P_k\left(\hat{\pi}_j^{v}\right) V_k\left(\hat{\pi}_j^{v,u}\right). \qquad (9)$$

При использовании показателя (7) формулу (9) можно записать в виде:

$$V_k\left(\hat{\pi}_j\right)=\sum_{v=1}^{\omega_{kj}}P_k\left(\hat{\pi}_j^v\right)\sum_{u=1}^{\omega_{kj}}\left\{\alpha\left[P_k\left(\hat{\pi}_j^u\right)-P_k\left(\hat{\pi}_j^v\right)\right]+\right.$$
$$\left.+\sum_{R_{kj}^w\in R_k\setminus\{R_{kj}^u\}}P_k\left(\hat{\pi}_j^w\right)\beta_w-\sum_{R_{kj}^w\in R_k\setminus\{R_{kj}^v\}}P_k\left(\hat{\pi}_j^w\right)\beta_w\right\}, \quad (10)$$

а при использовании показателя (8) — в виде:

$$V_k\left(\hat{\pi}_j\right)=\sum_{v=1}^{\omega_{kj}}P_k\left(\hat{\pi}_j^v\right)\sum_{u=1}^{\omega_{kj}}\left\{\alpha\left[P_k\left(\hat{\pi}_j^v\right)-P_k\left(\hat{\pi}_j^u\right)\right]+\right.$$
$$\left.+\sum_{R_{kj}^w\in R_k\setminus\{R_{kj}^v\}}P_k\left(\hat{\pi}_j^w\right)\beta_w-\sum_{R_{kj}^w\in R_k\setminus\{R_{kj}^u\}}P_k\left(\hat{\pi}_j^w\right)\beta_w\right\}. \quad (11)$$

Несложные математические преобразования позволяют привести формулы (10) и (11) к следующему виду (в зависимости от используемой стратегии):

$$V_k\left(\hat{\pi}_j\right)=\sum_{v=1}^{\omega_{kj}}\left(\beta_v-\alpha\right)P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right)-1\right] \quad (12)$$

при использовании стратегии минимизации «потерь»;

$$V_k\left(\hat{\pi}_j\right)=\sum_{v=1}^{\omega_{kj}}\left(\alpha-\beta_v\right)P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right)-1\right] \quad (13)$$

при использовании стратегии максимизации выигрышей.

Чтобы вычислить оценку $V(G)$ ценности информации для уже построенной ГПД, необходимо усреднить величины $V_k\left(\hat{\pi}_j\right)$ по множеству $\Omega_k\subset\Omega$ всех неконечных ИС $R_k\subseteq S$, входящих в состав ориентированного графа G:

$$V\left(G\right)=M\left\{V_k\left(\hat{\pi}_j\right)\right\}=\sum_{R_k\in\Omega_k}P\left(R_k\right)V_k\left(\hat{\pi}_j\right), \quad (14)$$

где $P(R_k)$ — вероятность соответствующего ИС; M — символ математического ожидания.

Чтобы вычислить вероятности $P(R_k)$, обозначим ветвь G_r графа G, по которой процесс распознавания ТС последовательно переходит из начального состояния $R_k = S$ в промежуточное ИС $R_k \subset S$ через G_r^k. Тогда вероятность $P\left(G_r^k\right)$ реализации этой ветви можно вычислить по формуле:

$$P\left(G_r^k\right) = \prod_{\pi_j \in \Pi_r^k} P_k\left(\hat{\pi}_j^v\right), v = \overline{1, \omega_{kj}}, \tag{15}$$

где Π_r^k — упорядоченное подмножество признаков, проверяемых при реализации ветви G_r^k.

Соответственно, вероятность $P(R_k)$ определяется как сумма вероятностей ветвей, приводящих в ИС R_k, то есть

$$P\left(R_k\right) = \sum_r P\left(G_r^k\right). \tag{16}$$

При использовании показателя (12) формула (14) преобразуется к виду:

$$V\left(G\right) = \sum_{R_k \in \Omega_k} P\left(R_k\right) \sum_{v=1}^{\omega_{kj}} \left(\beta_v - \alpha\right) P_k\left(\hat{\pi}_j^v\right) \left[\omega_{kj} P_k\left(\hat{\pi}_j^v\right) - 1\right], \tag{17}$$

а при использовании показателя (13) — к виду:

$$V\left(G\right) = \sum_{R_k \in \Omega_k} P\left(R_k\right) \sum_{v=1}^{\omega_{kj}} \left(\alpha - \beta_v\right) P_k\left(\hat{\pi}_j^v\right) \left[\omega_{kj} P_k\left(\hat{\pi}_j^v\right) - 1\right]. \tag{18}$$

Допустим, что когда основная гипотеза подтверждается, то потери отсутствуют ($\alpha = 0$), а когда отклоняется — потери равны единице ($\beta_v = 1$). Также допустим, что выигрыш при подтверждении основной гипотезы равен единице, то есть $\alpha = 1$, а при ее отклонении этот выигрыш равен нулю, то есть $\beta_v = 0$.

Тогда формулы (12) и (13) преобразуются к виду:

$$V_k\left(\hat{\pi}_j\right)=\sum_{v=1}^{\omega_{kj}}P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right)-1\right], \tag{19}$$

а формулы (17) и (18) — к виду:

$$V(G)=\sum_{R_k\in\Omega_k}P(R_k)\sum_{v=1}^{\omega_{kj}}P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right)-1\right]. \tag{20}$$

Для того чтобы установить непосредственную зависимость показателя (20) от выбираемых признаков, запишем это выражение в насколько измененном виде, то есть

$$V_k\left(\hat{\pi}_j\right)=\sum_{R_k\in\Omega_k^*}P(R_k)\sum_{v=1}^{\omega_{kj}}P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right)-1\right], \tag{21}$$

где $\Omega_k^*\subseteq\Omega_k$ — подмножество неконечных ИС, содержащихся в R_k-подпрограмме, начинающейся с проверки признака $\pi_j\in\Pi_k$. Под R_k-подпрограммой понимается отдельная часть графа G, начинающаяся в ИС R_k, вместе с выходящими из него дугами, которые приводят к конечным ИС $R_i=\{S_i\}$.

Используя метод динамического программирования, с помощью выражения (21) можно последовательно вычислять значения показателей $V_k\left(\hat{\pi}_j\right)$, начиная с ИС R_k, для которых $\mathrm{card}\{R_k\}=2$, и заканчивая $R_k=S$, для которого $\mathrm{card}\{R_k\}=m$, и осуществлять выбор наиболее ценных признаков.

Однако при последовательном переходе от ИС с меньшей мощностью к ИС с большей мощностью все вычисления по формуле (21) придется выполнять заново, поэтому необходимо преобразовать ее в рекуррентное соотношение.

Обозначим проверяемый в ИС R_k признак через π_s, а относящиеся к нему члены выделим в виде отдельного слагаемого. В результате из (21) получим выражение:

$$\begin{aligned}V_k\left(\hat{\pi}_s\right)&=\sum_{v=1}^{\omega_{ks}}P_k\left(\hat{\pi}_s^v\right)\left[\omega_{ks}P_k\left(\hat{\pi}_s^v\right)-1\right]+\\&+\sum_{v=1}^{\omega_{ks}}P_k\left(\hat{\pi}_s^v\right)\sum_{u=1}^{\omega_{ksj}}P_{ks}^v\left(\hat{\pi}_j^u\right)\left[\omega_{ksj}P_{ks}^v\left(\hat{\pi}_j^u\right)-1\right],\end{aligned} \tag{22}$$

где ω_{ksj} — число исходов проверки $\hat{\pi}_j$, выполненной в ИС R_{ks}^v согласно отображению:

$$\hat{\pi}_j : R_{ks}^v \to \left(R_{ks}^v\right)_j^u, u = \overline{1, \omega_{ksj}};$$

$P_{ks}^v\left(\hat{\pi}_j^u\right)$ — вероятность u-го исхода $\left(R_{ks}^v\right)_j^u$ этой проверки, которая вычисляется по формуле, аналогичной (6), то есть

$$P_{ks}^v\left(\hat{\pi}_j^u\right) = P\left[\left(R_{ks}^v\right)_j^u \Big/ R_{ks}^v\right] = \frac{|\left(\Delta_{ks}^v\right)_j^u|}{|\left(\nabla_{ks}\right)_j|}.$$

Введем обозначение

$$V_{ks}^v\left(\hat{\pi}_j\right) = \sum_{u=1}^{\omega_{ksj}} P_{ks}^v\left(\hat{\pi}_j^u\right)\left[\omega_{ksj} P_{ks}^v\left(\hat{\pi}_j^u\right) - 1\right] \qquad (23)$$

и сравним его с формулой (19). Очевидно, что $V_{ks}^v\left(\hat{\pi}_j\right)$ представляет собой значение ценности информации, получаемой при выполнении проверки признака $\pi_j \in \Pi_{ks}^v$ в ИС R_{ks}^v. Π_{ks}^v — определяемое по условию (4) подмножество признаков, допустимых для проверки в ИС R_{ks}^v. Другими словами, величина $V_{ks}^v\left(\hat{\pi}_j\right)$ определяет ценность информации, получаемой при выполнении дальнейших проверок.

Рекуррентная формула, с помощью которой вычисляется средняя ценность диагностической информации, получаемой при выполнении проверки $\hat{\pi}_s$ в ИС R_k, образуется путем подстановки в формулу (22) выражения (23), то есть

$$V_k\left(\hat{\pi}_s\right) = \sum_{v=1}^{\omega_{ks}} P_k\left(\hat{\pi}_s^v\right)\left[\omega_{ks} P_k\left(\hat{\pi}_s^v\right) - 1 + V_{ks}^v\left(\hat{\pi}_j\right)\right]. \qquad (24)$$

Если какой-либо из исходов выполняемой проверки является конечным, то есть $R_{ks}^v = R_i = \{S_i\}$, и ТС объекта считается распознанным, то соответствующее ему значение $V_{ks}^v\left(\hat{\pi}_j\right) = 0$.

Следует также отметить, что при вычислении показателя $V_k(\hat{\pi}_s)$ для ИС R_k, у которых $\operatorname{card}\{R_k\} = 2$, вместо формулы (24) можно воспользоваться более простым выражением:

$$V_k(\hat{\pi}_s) = \left[P_k(\hat{\pi}_s^1) - P_k(\hat{\pi}_s^2) \right]^2. \quad (25)$$

Вычислив значение показателя $V_k(\hat{\pi}_s)$ для всех допустимых признаков $\pi_s \in \Pi_k$, выберем в качестве оптимального признак π_j, удовлетворяющий условию

$$\pi_j = \arg \max_{\pi_s \in \Pi_k} \{V_k(\hat{\pi}_s)\}. \quad (26)$$

Вычислив по формуле (24) значения показателей ценности информации для всех ИС $R_k = \Omega_k$ и выбрав для каждого из них оптимальные признаки по условию (26), составим оптимальную ГПД в виде ориентированного графа.

Для проверки правильности выполненных расчетов, используя формулу (20), рассчитаем среднюю ценность информации, получаемой при реализации синтезированной диагностической процедуры. Если вычисленное по этой формуле значение $V(G) = V_k(\hat{\pi}_j)$, k: $R_k = S$, то ГПД составлена правильно.

4. Алгоритм построения оптимальной гибкой программы диагностирования. Алгоритм построения оптимальной ГПД на основе метода динамического программирования состоит из трех этапов, каждый из которых включает выполнение ряда последовательных шагов.

Этап 1. Определение множества $\Omega = \{R_k \mid R_k \subseteq S\}$ информационных состояний процесса диагностирования.

Шаг 1.1. Включение в множество Ω начального ИС $R_k = S$, в котором для проверки допустимы все признаки, то есть $\Pi_k = \Pi$.

Шаг 1.2. Выполнение проверок всех признаков $\pi_j \in \Pi_k$ в ИС $R_k = S$ согласно отображению (5) и получение их исходов $R_{kj}^v \left(v = \overline{1, \omega_{kj}} \right)$.

Шаг 1.3. Выбор ИС R_{kj}^v, для которых $2 \le \operatorname{card}\{R_{kj}^v\} < m$, присвоение каждому из них нового порядкового номера k, то есть

$R_{kj}^{v} \to R_k \subset S$, определение для этих ИС R_k в соответствии с условием (4) множества Π_k допустимых для проверки признаков и включение в множество Ω. Повторяющиеся ИС $R_k \subset S$ в множество Ω не включаются.

Шаг 1.4. Выполнение шагов 1.2 и 1.3 для каждого из ИС $R_k \subset S$, включенных в множество Ω на шаге 1.3, за исключением ИС, у которых $\text{card}\{R_k\} = 2$, до тех пор, пока все полученные ИС R_{kj}^{v} не будут удовлетворять условию $\text{card}\{ R_{kj}^{v} \} \le 2$.

Шаг 1.5. Упорядочение включенных в множество Ω ИС R_k по их мощности, то есть по возрастанию числа содержащихся в элементов.

Этап 2. Выбор оптимальных признаков для каждого ИС $R_k \subseteq S$.

Шаг 2.1. Извлечение из множества Ω ИС R_k, для которых $\text{card}\{R_k\} = 2$, вычисление по формуле (25) показателей $V_k\left(\hat{\pi}_s\right)$ ценности информации для каждого из признаков $\pi_s \in \Pi_k$ и выбор оптимальных признаков в соответствии с условием (26).

Шаг 2.2. Последовательное извлечение из множества Ω ИС R_k, для которых $2 < \text{card}\{R_k\} \le m$, вычисление по формуле (24) показателей $V_k\left(\hat{\pi}_s\right)$ ценности информации для каждого из признаков $\pi_s \in \Pi_k$ и выбор оптимальных в соответствии с условием (26). В качестве рекуррентных добавок $V_{ks}^{v}\left(\hat{\pi}_j\right)$ в формуле (24) следует использовать соответствующие значения $V_k\left(\hat{\pi}_j\right)$, полученные на предыдущих шагах. Выбор оптимального признака для ИС $R_k = S$ завершает выполнение этапа 2.

Этап 3. Построение оптимальной ГПД по результатам расчетов.

Шаг 3.1. Применение оптимальной проверки $\hat{\pi}_j$, выбранной на последнем шаге этапа 2, к ИС $R_k = S$ и получение ее исходов $R_{kj}^{v}\left(v = \overline{1, \omega_{kj}}\right)$.

Шаг 3.2. Применение оптимальных проверок $\hat{\pi}_j$, выбранных на этапе 2, к каждому из ИС R_{kj}^{v}, полученных на предыдущем шаге, получение исходов этих проверок.

Шаг 3.3. Выполнение операций, предусмотренных шагом 3.2, до получения всех конечных состояний $R_i = \{S_i\}$ $\left(i = \overline{1, m}\right)$.

В результате выполнения этапа 3 определим все ветви $G_r \in U$ графа G, представляющие собой наборы $\Pi_r \subseteq \Pi$ признаков $\pi_j \in \Pi$,

упорядоченных по очередности их проверки, необходимых для распознавания каждого из заданных ТС $S_i \in S$ объекта.

5. Пример реализации алгоритма. Синтезируем, используя метод динамического программирования, оптимальную по критерию максимума ценности диагностической информации ГПД объекта. В качестве исходных данных будем использовать представленные в виде таблицы 1 множество $S = \left\{ S_i \mid i = \overline{1,5} \right\}$ ТС объекта, множество $\Pi = \left\{ \pi_j \mid j = \overline{1,5} \right\}$ диагностических признаков, обеспечивающих попарную различимость этих ТС, а также множество $L = \left\{ \ell_{ij} \mid i = \overline{1,5}; j = \overline{1,5} \right\}$ интервалов, характеризующих возможный диапазон значений признаков в каждом ТС.

Таблица 1. Таблица состояний объекта диагностирования

ТС S_i	Диагностические признаки π_j				
	π_1	π_2	π_3	π_4	π_5
S_1	(0,0; 0,4)	(0,2; 0,5)	(0,1; 0,3)	(0,0; 0,5)	(0,5; 1,0)
S_2	(0,2; 0,6)	(0,7; 1,0)	(0,3; 0,8)	(0,2; 0,6)	(0,0; 0,3)
S_3	(0,5; 0,8)	(0,0; 0,4)	(0,6; 1,0)	(0,4; 0,6)	(0,6; 0,8)
S_4	(0,6; 1,0)	(0,2; 0,7)	(0,4; 0,8)	(0,7; 1,0)	(0,3; 0,5)
S_5	(0,3; 0,5)	(0,6; 0,8)	(0,0; 0,3)	(0,5; 0,7)	(0,3; 0,7)

Решение. Этап 1. Сформируем для заданных исходных данных множество Ω всех возможных ИС $R_k \subseteq S$ процесса диагностирования и множества Π_k допустимых для проверки признаков в каждом из этих состояний.

Полученное множество Ω упорядочим по числу содержащихся в ИС R_k $\left(k = \overline{1,22} \right)$ элементов S_i $\left(i = \overline{1,5} \right)$ и сведем в таблицу 2. В первый столбец занесем ИС R_k (кроме конечных ИС $R_i = \{S_i\}$). Во второй столбец таблицы 2 занесем допустимые для проверки в каждом ИС признаки $\pi_j \in \Pi_k$, определяемые из условия (4).

Этап 2. Рассчитаем значения показателей ценности информации для каждого из полученных ИС $R_k \subseteq S$ и выберем наиболее ценные признаки.

Рассмотрим сначала ИС $R_6 - R_{15}$, для которых $\text{card}\{R_k\} = 2$ $\left(k = \overline{6,15} \right)$. Поскольку любая проверка (из числа допустимых), выполненная в этих состояниях, приведет к распознаванию конкретного ТС, то расчеты будем проводить по формуле (25).

Таблица 2. Информационные состояния и выбираемые в них признаки

ИС $R_k \subseteq S$	Π_k	Оптимальный признак π_j	$V_k\left(\hat{\pi}_j\right)$
$R_6 = \{S_1, S_2\}$	π_2, π_3, π_5	π_3	0,1837
$R_7 = \{S_1, S_3\}$	π_1, π_3	π_3	0,1111
$R_8 = \{S_1, S_4\}$	$\pi_1, \pi_3, \pi_4, \pi_5$	π_5	0,1837
$R_9 = \{S_1, S_5\}$	π_2, π_4	π_4	0,1837
$R_{10} = \{S_2, S_3\}$	π_2, π_5	π_5	0,04
$R_{11} = \{S_2, S_4\}$	$\pi_1, \pi_2, \pi_4, \pi_5$	π_2	0,0625
$R_{12} = \{S_2, S_5\}$	π_3, π_5	π_3	0,0625
$R_{13} = \{S_3, S_4\}$	π_4, π_5	π_4	0,04
$R_{14} = \{S_3, S_5\}$	π_1, π_2, π_3	π_2	0,1111
$R_{15} = \{S_4, S_5\}$	π_1, π_3, π_4	π_1	0,1111
$R_{16} = \{S_1, S_2, S_3\}$	$\pi_1, \pi_2, \pi_3, \pi_5$	π_2	0,1528
$R_{17} = \{S_1, S_2, S_5\}$	$\pi_2, \pi_3, \pi_4, \pi_5$	π_3	0,4522
$R_{18} = \{S_1, S_3, S_4\}$	$\pi_1, \pi_3, \pi_4, \pi_5$	π_4	0,3457
$R_{19} = \{S_1, S_3, S_5\}$	$\pi_1, \pi_2, \pi_3, \pi_4$	π_4	0,5828
$R_{20} = \{S_2, S_3, S_4\}$	$\pi_1, \pi_2, \pi_4, \pi_5$	π_1	0,14
$R_{21} = \{S_2, S_3, S_5\}$	$\pi_1, \pi_2, \pi_3, \pi_5$	π_2	0,3828
$R_{22} = \{S_1, S_2, S_3, S_4, S_5\}$	$\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$	π_5	0,31

Так, согласно условию (4), в ИС $R_6 = \{S_1, S_2\}$ допустимыми для проверки являются признаки π_2, π_3 и π_5. Проверка $\hat{\pi}_5$, например, согласно отображению (5) имеет следующие конечные исходы:

$$\hat{\pi}_5 : R_6 \to \begin{cases} R^1_{6;5} = R_2 = \{S_2\}, \text{ если } y_5 \in (0,0;\, 0,3) = \Delta^1_{6;5}, \\ R^2_{6;5} = R_1 = \{S_1\}, \text{ если } y_5 \in (0,5;\, 1,0) = \Delta^2_{6;5}. \end{cases}$$

Вероятности этих исходов вычислим по формуле (6):

$$\left|\nabla_{6;5}\right| = \left|\ell_{15} \cup \ell_{25}\right| = 0,8;$$

$$P_6\left(\hat{\pi}^v_5\right) = \frac{\left|\Delta^v_{6;5}\right|}{\left|\nabla_{6;5}\right|} = \begin{cases} 0,375 \ \ (v=1); \\ 0,625 \ \ (v=2). \end{cases}$$

Полученные значения вероятностей исходов подставим в формулу (25) и вычислим:

$$V_6\left(\hat{\pi}_5\right) = \left[P_6\left(\hat{\pi}^1_5\right) - P_6\left(\hat{\pi}^2_5\right)\right]^2 = (0,375 - 0,625)^2 = 0,0625.$$

Аналогичные вычисления для проверок $\hat{\pi}_2$ и $\hat{\pi}_3$ дают следующие результаты:

$$V_6\left(\hat{\pi}_2\right)=0\ ;\ V_6\left(\hat{\pi}_3\right)=0{,}1837.$$

По условию (26) выберем признак π_3 в качестве оптимального для проверки в ИС R_6. Занесем этот признак, а также значение показателя $V_6\left(\hat{\pi}_3\right)$ в соответствующие ячейки таблицы 2.

Аналогичным образом выберем наиболее ценные для проверки признаки в ИС R_k $\left(k=\overline{7,15}\right)$ и заполним соответствующие ячейки таблицы 2.

Рассмотрим теперь ИС R_k $\left(k=\overline{16,21}\right)$, для которых $\mathrm{card}\{R_k\}=3$. Например, для ИС $R_{16}=\{S_1, S_2, S_3\}$ допустимые для проверки признаки составляют множество $\Pi_{16}=\{\pi_1, \pi_2, \pi_3, \pi_5\}$. Проверка $\hat{\pi}_1$, выполненная в этом ИС, имеет пять исходов:

$$\hat{\pi}_1 : R_{16} \rightarrow \begin{cases} R^1_{16;1}=R_1=\{S_1\}, \text{ если } y_1\in(0,0;\,0,2)=\Delta^1_{16;1};\\ R^2_{16;1}=R_6=\{S_1, S_2\}, \text{ если } y_1\in(0,2;\,0,4)=\Delta^2_{16;1};\\ R^3_{16;1}=R_2=\{S_2\}, \text{ если } y_1\in(0,4;\,0,5)=\Delta^3_{16;1};\\ R^4_{16;1}=R_{10}=\{S_2, S_3\}, \text{ если } y_1\in(0,5;\,0,6)=\Delta^4_{16;1};\\ R^5_{16;1}=R_3=\{S_3\}, \text{ если } y_1\in(0,6;\,0,8)=\Delta^5_{16;1}. \end{cases}$$

Используя формулу (6), рассчитаем вероятности этих исходов:

$$\left|\nabla_{16;1}\right|=\left|\ell_{11}\cup\ell_{21}\cup\ell_{31}\right|=0{,}8\ ;$$

$$P_{16}\left(\hat{\pi}^v_1\right)=\frac{\left|\Delta^v_{16;1}\right|}{\left|\nabla_{16;1}\right|}=\begin{cases}0{,}25\ \ (v=1;\,2;\,5);\\ 0{,}125\ \ (v=3;\,4).\end{cases}$$

Исходы $R^v_{16;1}$ $\left(v=\{1;\,3;\,5\}\right)$ являются конечными, поэтому для них в формуле (24) значения $V^v_{16;1}\left(\hat{\pi}_j\right)\Big|_{v=\{1;\,3;\,5\}}=0$. Для исходов $R^2_{16;1}=R_6$ и $R^4_{16;1}=R_{10}$ на предыдущих шагах были уже определены значения $V_6\left(\hat{\pi}_3\right)=0{,}1837$ и $V_{10}\left(\hat{\pi}_5\right)=0{,}04$. Подставим эти значения в формулу (24) и получим:

$$V_{16}\left(\hat{\pi}_1\right)=\sum_{v=1}^{5}P_{16}\left(\hat{\pi}_1^v\right)\left[5P_{16}\left(\hat{\pi}_1^v\right)-1+V_{16;1}^v\left(\hat{\pi}_j\right)\right]=$$
$$=0,25\left[5\cdot 0,25-1+0\right]+0,25\left[5\cdot 0,25-1+0,1837\right]+$$
$$+0,125\left[5\cdot 0,125-1+0\right]+0,125\left[5\cdot 0,125-1+0,04\right]+$$
$$+0,25\left[5\cdot 0,25-1+0\right]=0,1447.$$

Аналогичные вычисления, проведенные для оставшихся нерассмотренными признаков $\{\pi_2, \pi_3, \pi_5\}$, дают следующие результаты:

$$V_{16}\left(\hat{\pi}_2\right)=0,1528\ ;\ V_{16}\left(\hat{\pi}_3\right)=0,0459\ ;\ V_{16}\left(\hat{\pi}_5\right)=0,059\ .$$

По условию (26) для проверки в ИС R_{16} выберем признак π_2. Занесем его и соответствующее значение показателя $V_{16}\left(\hat{\pi}_2\right)$ в третий и четвертый столбцы таблицы 2.

Действуя таким же образом, выберем оптимальные для проверки признаки π_j для ИС R_{17} — R_{21} и заполним соответствующие ячейки таблицы 2.

Завершая синтез ГПД, выберем оптимальный признак для проверки в начальном ИС $R_{22} = S$, учитывая, что $\Pi_{22} = \Pi$. В частности, проверка $\hat{\pi}_5$ имеет шесть возможных исходов с соответствующими вероятностями:

$$\hat{\pi}_5:\ R_{22}\ \rightarrow\begin{cases}R_{22;5}^1=R_2=\{S_2\},\ \text{если}\ y_5\in(0,0;\,0,3)=\Delta_{22;5}^1;\\ R_{22;5}^2=R_{15}=\{S_4,S_5\},\ \text{если}\ y_5\in(0,3;\,0,5)=\Delta_{22;5}^2;\\ R_{22;5}^3=R_9=\{S_1,S_5\},\ \text{если}\ y_5\in(0,5;\,0,6)=\Delta_{22;5}^3;\\ R_{22;5}^4=R_{19}=\{S_1,S_3,S_5\},\ \text{если}\ y_5\in(0,6;\,0,7)=\Delta_{22;5}^4;\\ R_{22;5}^5=R_7=\{S_1,S_3\},\ \text{если}\ y_5\in(0,7;\,0,8)=\Delta_{22;5}^5;\\ R_{22;5}^6=R_1=\{S_1\},\ \text{если}\ y_5\in(0,8;\,1,0)=\Delta_{22;5}^6;\end{cases}$$
$$\left|\nabla_{22;5}\right|=\left|\ell_{15}\cup\ell_{25}\cup\ell_{35}\cup\ell_{45}\cup\ell_{55}\right|=1,0\,;$$

$$P_{22}\left(\hat{\pi}_5^v\right)=\frac{\left|\Delta_{22;5}^v\right|}{\left|\nabla_{22;5}\right|}=\begin{cases}0,3\,(v=1);\\ 0,2\,(v=2;\,6);\\ 0,1\,(v=3;\,4;\,5).\end{cases}$$

Заметим, что исходы $R^{v}_{22;5}\left(v=\{1;6\}\right)$ являются конечными, то есть никакой дополнительной информации для распознавания ТС объекта уже не требуется, поэтому при использовании формулы (24) величины $V^{v}_{22;5}\left(\hat{\pi}_j\right)\Big|_{v=\{1;6\}}$ примем равными нулю. А для исходов $R^{v}_{22;5}\left(v=\overline{2,5}\right)$ соответствующие значения $V^{v}_{22;5}\left(\hat{\pi}_j\right)$ были получены на предыдущих шагах расчета и зафиксированы в таблице 2. Подставим имеющиеся данные в формулу (24) и получим:

$$V_{22}\left(\hat{\pi}_5\right)=\sum_{v=1}^{6}P_{22}\left(\hat{\pi}^{v}_5\right)\left[6P_{22}\left(\hat{\pi}^{v}_5\right)-1+V^{v}_{22;5}\left(\hat{\pi}_j\right)\right]=$$

$$=0,3\left[6\cdot 0,3-1+0\right]+0,2\left[6\cdot 0,2-1+0,1111\right]+$$

$$0,1\left[6\cdot 0,1-1+0,1837\right]+0,1\left[6\cdot 0,1-1+0,5828\right]+$$

$$0,1\left[6\cdot 0,1-1+0,1111\right]+0,2\left[6\cdot 0,2-1+0\right]=0,31.$$

Путем аналогичных вычислений определим среднее значение ценности диагностической информации, получаемой при выполнении других проверок:

$$V_{22}\left(\hat{\pi}_1\right)=0,2018;\ \ V_{22}\left(\hat{\pi}_2\right)=0,2249;$$
$$V_{22}\left(\hat{\pi}_3\right)=0,1572;\ \ V_{22}\left(\hat{\pi}_4\right)=0,2903.$$

Очевидно, что по условию (26) оптимальным является признак π_5. Заполним таблицу 2 окончательно и выполним третий этап построения ГПД.

Этап 3. В начальном ИС $R_{22} = \{S_1, S_2, S_3, S_4, S_5\}$ выполняется проверка $\hat{\pi}_5$. Неконечными исходами этой проверки являются ИС R_7, R_9, R_{15} и R_{19}, для которых в качестве оптимальных выбраны проверки $\hat{\pi}_3$, $\hat{\pi}_4$, $\hat{\pi}_1$ и $\hat{\pi}_4$, соответственно. Одним из неконечных исходов проверки $\hat{\pi}_4$, выполняемой в ИС R_{19}, является R_{14}, для которого в качестве оптимальной выбирается проверка $\hat{\pi}_2$.

Таким образом, все упорядоченные подмножества признаков $\Pi_r \subseteq \Pi$ ($r=\overline{1,14}$), которые задают состав и очередность выполнения проверок $\pi_j \in \Pi$ ($j=\overline{1,5}$), необходимых для распознавания каждого из

заданных технических состояний $S_i \in S$ ($i = \overline{1,5}$) объекта определены. Представим эти подмножества в виде таблицы 3.

Таблица 3. Упорядоченные наборы признаков, необходимых для распознавания ТС объекта

ТС S_i	Подмножества признаков $\Pi_r \subseteq \Pi$
S_1	$\Pi_{10} = \{\pi_5, \pi_4, \pi_3\}$; $\Pi_{11} = \{\pi_5, \pi_3\}$; $\Pi_{12} = \Pi_{13} = \{\pi_5, \pi_4\}$; $\Pi_{14} = \{\pi_5\}$
S_2	$\Pi_1 = \{\pi_5\}$
S_3	$\Pi_7 = \{\pi_5, \pi_4, \pi_2\}$; $\Pi_8 = \{\pi_5, \pi_4, \pi_3\}$; $\Pi_9 = \{\pi_5, \pi_3\}$
S_4	$\Pi_2 = \{\pi_5, \pi_1\}$
S_5	$\Pi_3 = \{\pi_5, \pi_1\}$; $\Pi_4 = \Pi_5 = \{\pi_5, \pi_4\}$; $\Pi_6 = \{\pi_5, \pi_4, \pi_2\}$

Оптимальная по критерию максимума ценности диагностической информации ГПД изображена в виде ориентированного графа на рисунке 1.

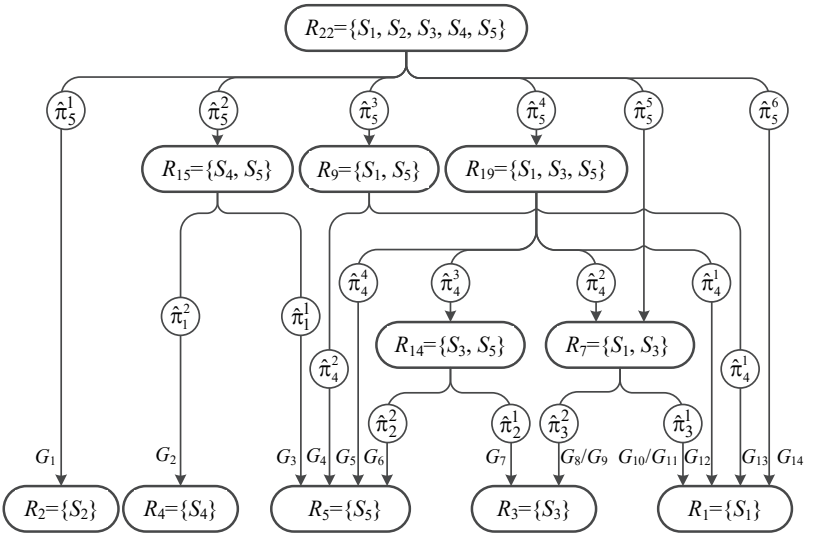


Рис. 1. Гибкая программа диагностирования, оптимальная по критерию максимума ценности диагностической информации

Для проверки оптимальности синтезированной ГПД рассчитаем среднюю ценность информации, получаемой при ее реализации, с помощью формулы (20).

Сначала по формулам (15) и (16) определим вероятности неконечных ИС R_k (k = 7; 9; 14; 15; 19; 22), имеющихся в построенной программе:

$$P(R_7) = P_{22}\left(\hat{\pi}_5^4\right)P_{19}\left(\hat{\pi}_4^2\right) + P_{22}\left(\hat{\pi}_5^5\right) = 0,1\cdot 0,143 + 0,1 = 0,1143;$$

$$P(R_9) = P_{22}\left(\hat{\pi}_5^3\right) = 0,1;$$

$$P(R_{14}) = P_{22}\left(\hat{\pi}_5^4\right)P_{19}\left(\hat{\pi}_4^3\right) = 0,1\cdot 0,143 = 0,0143;$$

$$P(R_{15}) = P_{22}\left(\hat{\pi}_5^2\right) = 0,2;\ P(R_{19}) = P_{22}\left(\hat{\pi}_5^4\right) = 0,1;\ P(R_{22}) = 1.$$

Подставим полученные данные в формулу (20) и рассчитаем:

$$V(G) = \sum_{k=\{7;9;14;15;19;22\}} P(R_k)\sum_{v=1}^{\omega_{kj}} P_k\left(\hat{\pi}_j^v\right)\left[\omega_{kj}P_k\left(\hat{\pi}_j^v\right) - 1\right] =$$

$$= 0,1143\left(0,333\left[2\cdot 0,333 - 1\right] + 0,667\left[2\cdot 0,667 - 1\right]\right) +$$

$$+0,1\left(0,714\left[2\cdot 0,714 - 1\right] + 0,286\left[2\cdot 0,286 - 1\right]\right) +$$

$$+0,0143\left(0,333\left[2\cdot 0,333 - 1\right] + 0,667\left[2\cdot 0,667 - 1\right]\right) +$$

$$+0,2\left(0,333\left[2\cdot 0,333 - 1\right] + 0,667\left[2\cdot 0,667 - 1\right]\right) +$$

$$+0,1\left(0,571\left[4\cdot 0,571 - 1\right] + 3\cdot 0,143\left[4\cdot 0,143 - 1\right]\right) +$$

$$+1\left(0,3\left[6\cdot 0,3 - 1\right] + 3\cdot 0,1\left[6\cdot 0,1 - 1\right] + 2\cdot 0,2\left[6\cdot 0,2 - 1\right]\right) = 0,31.$$

Поскольку $V(G) = V_{22}\left(\hat{\pi}_5\right)$, значит ГПД составлена правильно и является оптимальной по выбранному критерию.

6. Заключение. Математический аппарат, предложенный для формализации положений теории ценности информации Р. Л. Стратоновича применительно к задачам технической диагностики позволяет сделать вывод о том, что диагностическая информация обладает ценностью, которую можно вычислять. При этом наибольшей ценностью обладает та диагностическая информация, при получении которой в наибольшей степени различаются между собой априорные и апостериорные вероятности распознавания технического состояния анализируемого объекта. Этот вывод согласуется также с концепцией ценности информации, предложенной А. А. Харкевичем. Об этом свидетельствует тот факт, что ГПД, которая изображена на рисунке 1, получилась аналогичной программе, построенной таким же методом для тех же исходных данных при использовании показателя ценности информации Харкевича.

Следует отметить, что предложенный алгоритм будет корректным и при использовании дискретных диагностических признаков. При этом необходимо задать значения $P(S_i)$ вероятностей анализируемых технических состояний объекта, а некоторые формулы претерпят незначительные изменения.

Рассмотренный алгоритм может быть положен в основу разработки прикладного программного обеспечения программно-аппаратных комплексов, предназначенных для автоматизированного мониторинга состояния сложных технических объектов, функционирующих в масштабе времени, близком к реальному.

Литература

1. *Харкевич А.А.* О ценности информации // Проблемы кибернетики. 1960. Вып. 4. С. 53–57.
2. *Бонгард М.М.* Проблема узнавания // М.: Наука. 1967. 320 с.
3. *Корогодин В.И., Корогодина В.Л.* Информация как основа жизни // Дубна: Феникс. 2000. 208 с.
4. *Стратонович Р.Л.* Теория информации // М.: Сов. радио. 1975. 424 с.
5. *Howard R.A.* Information Value Theory // IEEE Transactions on Systems Science and Cybernetics. 1966. vol. 2. no. 1. pp. 22–26.
6. *Sulganik E., Zilcha I.* The Value of Information in the Presence of Futures Markets // The Journal of Futures Markets. 1996. vol. 16. no. 2. pp. 227–240.
7. *Wilson E.C.* A Practical Guide to Value of Information Analysis // Pharmacoeconomics. 2015. vol. 33. no. 2. pp. 105–121.
8. *Bhalla D.* Weight of Evidence (WOE) and Information Value Explained. URL: https://www.listendata.com/2015/03/weight-of-evidence-woe-and-information.html (дата обращения: 15.08.2018).
9. *Welton N.J, Thom H.H.Z.* Value of Information: We've Got Speed, What More Do We Need? // International Journal of the Society for Medical Decision Making. 2015. vol. 35. no. 5. pp. 564–566.
10. *Mohseninejad L. et al.* Value of Information Analysis from a Societal Perspective: a Case Study in Prevention of Major Depression // Value in Health. 2013. vol. 16. no. 4. pp. 490–497.
11. *Bartha E. et al.* Value of Information: Interim Analysis of a Randomized, Controlled Trial of Goal-Directed Hemodynamic Treatment for Aged Patients // Trials. 2013. vol. 14. no. 1. pp. 205. URL: http://www.trialsjournal.com/content/14/1/205 (дата обращения: 15.08.2018).
12. *Malings C., Pozzi M.* Value of Information for Spatially Distributed Systems: Application to Sensor Placement // Reliability Engineering & System Safety. 2016. vol. 154. pp. 219–233.
13. *Malings C., Pozzi M.* Conditional Entropy and Value of Information Metrics for Optimal Sensing in Infrastructure Systems // Structural Safety. 2016. vol. 60. pp. 77-90.
14. *Malings C., Pozzi M.* Value of Information in Spatio-Temporal Systems: Sensor Placement and Scheduling // Reliability Engineering & System Safety. 2018. vol. 172. pp. 45–57.
15. *Memarzadeh M., Pozzi M.* Value of Information in Sequential Decision Making: Component Inspection, Permanent Monitoring and System-Level Scheduling // Reliability Engineering & System Safety. 2016. vol. 154. pp. 137–151.
16. The Value of Timely Information During Disasters (Measured in Hours). URL: https://irevolutions.org/2016/04/06/the-value-of-timely-information-during-disasters-measured-in-hours/ (дата обращения: 15.08.2018).

17. *Eidsvik J., Mukerji T., Bhattacharjya D.* Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis // Cambridge University Press. 2015. 400 p.
18. *Hu Y., Janowicz K., Chen Y.* Task-Oriented Information Value Measurement Based on Space-Time Prisms // International Journal of Geographical Information Science. 2015. vol. 30. no. 6. pp. 1228–1249.
19. *Hu Y., Janowicz K., Couclelis H.* Prioritizing Disaster Mapping Tasks for Online Volunteers Based on Information Value Theory // Geographical Analysis. 2017. vol. 49. no. 2. pp. 175–198.
20. *Минниахметов Д.Ф., Николаева М.А.* Обзор подходов к оценке ценности информации // Всероссийская конференция «Информационные технологии интеллектуальной поддержки принятия решений». 2016. С. 115–116.
21. *Шанкин Г.П.* Ценность информации. Вопросы теории и приложений // М.: Филоматис. 2004. 128 с.
22. *Авсентьев О.С., Авсентьев А.О.* Формирование обобщенного показателя ценности информации в каналах связи // Вестник Воронежского института МВД России. 2015. № 3. С. 55–63.
23. *Авсентьев А.О.* Определение ценности информации // Доклады Томского государственного университета систем управления и радиоэлектроники. 2016. Т. 19. № 1. С. 21–24.
24. *Голубев В.И.* Новый подход к определению ценности информации // Радиоэлектронные устройства и системы для инфокоммуникационных технологий: Сборник трудов 69-й международной конференции. 2014. Т. 4. С. 132–136.
25. *Машошин О.Ф.* Оценка диагностической ценности информации при решении задач в области эксплуатации авиационной техники // Научный вестник Московского государственного технического университета гражданской авиации. 2015. № 219(9). С. 53–56.
26. *Мышко В.В., Кравцов А.Н., Копкин Е.В., Чикуров В.А*. Теоретические основы и методы оптимизации анализа технического состояния сложных систем: монография // СПб.: ВКА имени А.Ф.Можайского. 2013. 303 с.
27. *Копкин Е.В., Кравцов А.Н., Чикуров В.А.* Выбор диагностических признаков с учетом их ценности для распознавания технического состояния объекта // Труды Военно-космической академии имени А.Ф. Можайского. 2015. Вып. 647. С. 22–29.
28. *Копкин Е.В., Кобзарев И.М.* Оптимальный алгоритм анализа технического состояния объекта на основе меры ценности диагностической информации // Труды Военно-космической академии имени А.Ф. Можайского. 2018. Вып. 661. С. 15–31.

Копкин Евгений Вениаминович — д-р техн. наук, доцент, профессор, кафедра технологий и средств автоматизации обработки и анализа информации космических средств, Военно-космическая академия имени (ВКА им. А.Ф. Можайского). Область научных интересов: автоматизированная обработка и анализ информации космических средств. Число научных публикаций — 50. kopkins@mail.ru; ул. Ждановская, 13, 197198, Санкт-Петербург, Российская Федерация; р.т.: +7(921)961-1338.

Кобзарев Игорь Михайлович — адъюнкт, кафедра технологий и средств автоматизации обработки и анализа информации космических средств, Военно-космическая академия имени А.Ф.Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: автоматизированная обработка и анализ информации космических средств. Число научных публикаций — 4. kobzaryan@mail.ru; ул. Ждановская, 13, 197198, Санкт-Петербург, Российская Федерация; р.т.: +7(911)235-8494.

DOI 10.15622/sp.2019.18.6.1434-1461

E.V. KOPKIN, I.M. KOBZAREV

INFORMATION VALUE MEASURE FOR OPTIMIZATION OF FLEXIBLE DIAGNOSIS PROGRAMS OF TECHNICAL OBJECTS

Kopkin E.V., Kobzarev I.M. **Information Value Measure for Optimization of Flexible Diagnosis Programs of Technical Objects.**

Abstract. Existing methods of calculating of the value of diagnostic information circulating in the automated systems of monitoring of technical condition of objects do not take into account "losses" ("gains") resulting from making "wrong" decisions when identifying this state. The purpose of the work is to develop an algorithm that allows to solve the problem of recognizing the technical state of the object being analyzed by means of dynamic programming, the value of the diagnostic information as an optimized indicator being used. The solution to the optimization problem of a diagnostic procedure is based on the use of a measure of the information value proposed by R. L. Stratonovich. It is modified according to the subject area of the technical diagnostics and in the case when the diagnostic features presented in the form of intervals on the real numerical axis are used. The maximum value of the diagnostic information is achieved by minimizing the average "losses" (maximizing the average "gains") obtained when performing tests of diagnostic signs in the process of recognizing the technical condition of an object. To solve the problem, a recurrent expression possessing a scientific novelty has been proposed. It allows to calculate the value of the information obtained when performing tests of diagnostic signs in each of the analyzed information states of the diagnostic process. In the process of the diagnostics program implementation when recognizing the technical condition of the object both "losses" and "winnings" are possible. The difference between their a priori and a posteriori means values characterizes the value of the diagnostic information numerically. The magnitude of the information value indication depends on the probabilities of the results of the diagnostic signs checks and is proportional to the difference between the a posteriori and a priori probabilities of achieving the diagnostic goal. By using the proposed solution, it is possible to synthesize the flexible diagnostics program that is optimal according to the maximum value of diagnostic information in the form of a oriented graph or sets of tests in proper sequence of their execution. This is necessary in order to recognize the specific technical state in which the object is located. The implementation of the algorithm developed is possible in the software and algorithmic support of the automated systems for monitoring the state of complex technical objects.

Keywords: Technical State, Flexible Diagnosis Program, Information Value, Diagnostic Sign.

Kopkin Evgeniy Veniaminovich — Ph.D., Dr.Sci., Associate Professor, Professor, Department of Technologies and Automation Equipment for Processing and Analysis of Space Objects Information, Mozhaisky Military Space Academy. Research interests: analysis of technical state of the complex systems, data acquisition. The number of publications — 50. kopkins@mail.ru; 13, Zhdanovskaya str., 197198, St. Petersburg, Russian Federation; office phone: +7(921)961-1338.

Kobzarev Igor Mihaylovich — Ph.D. Student, Department of Technologies and Automation Equipment for Processing and Analysis of Space Objects Information, Mozhaysky Military Space Academy. Research interests: analysis of technical state of the complex systems, data acquisition. The number of publications — 4. kobzaryan@mail.ru; 13, Zhdanovskaya str., 197198, St. Petersburg, Russian Federation; office phone: +7(911)235-8494.

References

1. Kharkevich A.A. [About the Value of Information]. *Problemy kibernetiki – Problems of Cybernetics*. 1960. vol. 4. pp. 53–57. (In Russ.).

2. Bongard M.M. *Problema uznavaniya* [Recognition Problem]. M.: Nauka Publ. 1967. 320 p. (In Russ.).
3. Korogodin V.I., Korogodina V.L. *Informaciya kak osnova zhizni* [Information as the Basis of Life]. Dubna: Feniks. 2000. 208 p. (In Russ.).
4. Stratonovich R.L. *Teoriya informacii* [The Theory of Information]. M.: Sov. radio. 1975. 424 p. (In Russ.).
5. Howard R.A. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics*. 1966. vol. 2. no. 1. pp. 22–26.
6. Sulganik E., Zilcha I. The Value of Information in the Presence of Futures Markets. *The Journal of Futures Markets*. 1996. vol. 16. no. 2. pp. 227–240.
7. Wilson E.C. A Practical Guide to Value of Information Analysis. *Pharmacoeconomics*. 2015. vol. 33. no. 2. pp. 105–121.
8. Bhalla D. Weight of Evidence (WOE) and Information Value Explained. Available at: https://www.listendata.com/2015/03/weight-of-evidence-woe-and-information.html (accessed: 15.08.2018).
9. Welton N.J., Thom H.H. Value of Information: We've Got Speed, What More Do We Need? Medical Decision making. *An International Journal of the Society for Medical Decision Making*. 2015. vol. 35. no. 5. pp. 564–566.
10. Mohseninejad L. et al. Value of Information Analysis from a Societal Perspective: a Case Study in Prevention of Major Depression. *Value in Health.* 2013. vol. 16. no. 4. pp. 490–497.
11. Bartha E. et al. Value of Information: Interim Analysis of a Randomized, Controlled Trial of Goal-Directed Hemodynamic Treatment for Aged Patients. *Trials.* 2013. vol. 14. no. 1. pp. 205. Available at: http://www.trialsjournal.com/content/14/1/205 (accessed: 15.08.2018).
12. Malings C., Pozzi M. Value of Information for Spatially Distributed Systems: Application to Sensor Placement *Reliability Engineering & System Safety*. 2016. vol. 154. pp. 219–233.
13. Malings C., Pozzi M. Conditional Entropy and Value of Information Metrics for Optimal Sensing in Infrastructure Systems. *Structural Safety*. 2016. vol. 60. pp. 77–90.
14. Malings C., Pozzi M. Value of Information in Spatio-Temporal Systems: Sensor Placement and Scheduling. *Reliability Engineering & System Safety*. 2018. vol. 172. pp. 45–57.
15. Memarzadeh M., Pozzi M. Value of Information in Sequential Decision Making: Component Inspection, Permanent Monitoring and System-Level Scheduling *Reliability Engineering & System Safety*. 2016. vol. 154. pp. 137–151.
16. The Value of Timely Information During Disasters (Measured in Hours). Available at: https://irevolutions.org/2016/04/06/the-value-of-timely-information-during-disasters-measured-in-hours/ (accessed: 15.08.2018).
17. Eidsvik J., Mukerji T., Bhattacharjya D. Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis. Cambridge University Press. 2015. 400 p.
18. Hu Y., Janowicz K., Chen Y. Task-Oriented Information Value Measurement Based on Space-Time Prisms. *International Journal of Geographical Information Science*. 2015. vol. 30. no. 6. pp. 1228–1249.
19. Hu Y., Janowicz K., Couclelis H. Prioritizing Disaster Mapping Tasks for Online Volunteers Based on Information Value Theory. *Geographical Analysis*. 2017. vol. 49. no. 2. pp. 175–198.
20. Minniakhmetov D.F., Nikolaeva M.A. [A review of approaches to value of information estimation]. *Vserossijskaya konferenciya «Informacionnye tekhnologii intellektual'noj podderzhki prinyatiya reshenij»* [Proceedings of the All-Russian Conference on Information Technologies for Intelligent Decision Making Support (ITIDS'2016)]. 2016. pp. 115–116. (In Russ.).

21. Shankin G.P. *Cennost' informatcii. Voprosy teorii i prilozheniy* [The Value of Information. Theory and Applications]. M.: Filomatis. 2004. 128 p. (In Russ.).
22. Avsent'ev O.S., Avsent'ev A.O. [The Formation of the Information Value Generalized Index in the Communication Channels]. *Vestnik Voronezhskogo instituta ministerstva vnutrennih del Rossii – The Bulletin of Voronezh Institute of the Ministry of Internal Affairs of Russia*. 2015. vol. 3. pp. 55–63. (In Russ.).
23. Avsent'ev A.O. [Estimation of the Information Value]. *Doklady Tomskogo gosudarstvennogo universiteta system upravleniya i radioelektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics.* 2016. Issue 19. vol. 1. pp. 21–24. (In Russ.).
24. Golubev V.I. [A new approach to value of information determination]. *Radioelektronnye ustroystva i sistemy dlya infokommunikatsionnyh tekhnologiy (REDS-2014)*: *Sbornik trudov 69-y mezhdunarodnoj konferentsii* [Radio-Electronic Devices and Systems for the Infocommunication Technologies: Collected papers]. 2014. Issue 4. pp. 131-136. (In Russ.).
25. Mashoshin O.F. [Evaluation of Diagnostic Information in Solving the Task of Aircraft Operation]. *Nauchnyi vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta grazhdanskoy aviatsii – Civil Aviation High Technologies*. 2015. vol. 219(9). pp. 53–56. (In Russ.).
26. Myshko V.V., Kravtsov A.N., Kopkin E.V., Chikurov V.A. *Teoreticheskie osnovy i metody optimizatsii analiza tekhnicheskogo sostoianiia slozhnykh system. Monografija* [Theoretical Bases and Methods for Optimizing the Analysis of the Technical State of Complex Systems. Monography]. SPb: VKA imeni A.F.Mozhajskogo. 2013. 303 p. (In Russ.).
27. Kopkin E.V., Kravtsov A.N., Chikurov V.A. [Diagnostic Sign Selection by their Value for Recognition of Object Technical Condition]. *Trudy voenno-kosmicheskoy akademii imeni A.F.Mozhayskogo – Proceedings of the Military Space academy named after A.F.Mozhaisky*. 2015. vol. 647. pp. 22–29. (In Russ.).
28. Kopkin E.V., Kobzarev I.M. [The Optimal Algorithm for Analysis of Object Technical State on the Base of Value Measure of Diagnostic Information]. *Trudy voenno-kosmicheskoy akademii imeni A.F.Mozhayskogo – Proceedings of the Military Space academy named after A.F.Mozhaisky*. 2018. vol. 661. pp. 15–31 (In Russ.).

УДК 629.7 **DOI** 10.15622/sp.2019.18.6.1462-1490

Е.П. Минаков, Б.В. Соколов, С.Е. Шалдаев, М.А. Александров

РАСЧЕТ И ИССЛЕДОВАНИЕ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК РУБЕЖЕЙ АТАКИ АСТЕРОИДОВ ОРБИТАЛЬНЫМИ СРЕДСТВАМИ

Минаков Е.П., Соколов Б.В., Шалдаев С.Е., Александров М.А. **Расчет и исследование пространственно-временных характеристик рубежей атаки астероидов орбитальными средствами.**

Аннотация. Развитие работ по борьбе с астероидной опасностью требует построения и исследования областей космического пространства, перемещаясь в которых космические аппараты-перехватчики могут воздействовать на астероиды. В предлагаемой статье такие области названы рубежами атаки, пространственно-временные характеристики которых зависят от параметров орбит астероидов, а также фазовых координат узловых точек. В указанных точках происходит пересечение траекторией движения астероидов и плоскостей орбит космических аппаратов-перехватчиков. В случае воздействия космических аппаратов-перехватчиков на астероиды в узловых точках особую значимость приобретает исследование пространственно-временных характеристик рубежей атаки с учетом ограничений на относительные скорости сближения астероидов и космических аппаратов-перехватчиков. Для этого предлагается построить и проанализировать соответствующие зоны обратной досягаемости.

В состав разработанного комплекса моделей включена имитационная модель, с использованием которой генерируются случайные углы между проекциями векторов скоростей астероидов на плоскости орбит космических аппаратов-перехватчиков и текущими направлениями на годографы векторов их скоростей в узловых точках, а также аналитическая модель оценивания пространственно-временных характеристик рубежей атаки астероидов, задаваемых радиусами их наружных и внутренних границ и используемых для определенных значений соответствующих аргументов широт и времени прибытия космических аппаратов-перехватчиков в узловые точки.

Апробация разработанных моделей и исследование соответствующих характеристик рубежей атаки проведена в ходе вычислительных экспериментов по двухцикловому моделированию величин углов между проекциями векторов скоростей астероидов на плоскости орбит космических аппаратов-перехватчиков и текущими направлениями на годографы векторов их скоростей в узловых точках. Полученные результаты позволили провести верификацию и валидацию разработанных моделей, на основе чего был сделан вывод о требуемой степени их адекватности. Также предложена процедура оценивания параметров рубежей атаки, зависящих как от значений аргументов широт космических аппаратов-перехватчиков, так и высот их полета над поверхностью Земли. При этом обоснован подход к оцениванию пространственно-временных характеристик рубежей атаки астероидов космическими аппаратами-перехватчиками для любых внутриплоскостных параметров их орбит.

Ключевые слова: астероидная опасность, космический аппарат-перехватчик, отделяемая часть, узловая точка, рубеж атаки, пространственно-временные характеристики.

1. Введение. Одной из главных причин возможного исчезновения жизни на Земле является возможное столкновение планеты с астероидами [1-14]. Многочисленные исследования

зарубежных и отечественных ученых показывают, что возможность (вероятность) появления такого события ничтожна мала, но все-таки она не нулевая [4, 9, 12-19]. Падение Тунгусского метеорита в начале XX века, недавнее падение метеорита в районе города Челябинска говорит о том, что опасность возникновения планетарной катастрофы существует. Одно из направлений развития работ по борьбе с астероидной опасностью, проводимых в настоящее время в нашей стране и за рубежом и посвященных вопросам создания соответствующих систем автономного автоматического управления космическими аппаратами-перехватчиками (КАП) [14-19, 21-26], связано с построением в околоземном или окололунном космическом пространстве таких областей, перемещаясь в пределах которых предназначаемые для уничтожения движущихся к Земле астероидов КАП могут гарантированно (с баллистической точки зрения) достигать и поражать их [27-30]. Другими словами, для любого момента времени в случае нахождения КАП в рассматриваемой области всегда может быть найдена допустимая баллистическая траектория перехвата астероида указанным аппаратом (либо аппаратами) в соответствующей узловой точке. Быстротечность событий, связанных с падением на Землю астероидов, объективно существующее запаздывание в процессах их обнаружения и тому подобное существенно затрудняет своевременное выведение КАП на орбиты, компланарные с орбитами движения астероидов. В этой связи возможные сценарии уничтожение астероидов должны заранее быть спланированы и ориентированы на так называемые узловые точки (УТ) траекторий движения КАП, которые образуются при пересечении астероидом (целью) плоскости движения КАП и являются наиболее предпочтительными с точки зрения их энергетических затрат при выполнении перехвата. Данные точки задаются дальностью от центра Земли или Луны, а также углом между направлением на УТ и линией пересечения плоскостей орбит астероида и КАП. Еще один важный параметр УТ — время пересечения астероидом плоскости движения КАП [28-30].

На эффективность уничтожения астероидов влияет множество факторов и в первую очередь тактико-технические характеристики (ТТХ) КАП [23, 28-29]. Очевидно, что эти характеристики имеют первостепенное значение при формировании пространственного положения областей, в пределах которых в определенное время должны начинать движение КАП для уничтожения астероидов. Определение пространственно-временных характеристик указанных областей (рубежей атаки астероидов) основывается на учете таких параметров КАП, как их предельные

относительные скорости перемещения относительно перехватываемых ими астероидов. При этом *под рубежом атаки* (РА) понимается такая пространственно-временная область (ПРВО), двигаясь в которой КАП может поразить астероид в УТ. В этой связи разработка математических моделей и исследование пространственно-временных характеристик (параметров) РА с учетом соответствующих баллистических и технических ограничений на условия применения КАП является новой и актуальной научно-технической задачей, требующей своего решения.

2. Основные понятия и постановка задачи исследования. Решение широкого круга задач, связанных с организацией целенаправленных воздействий на потенциально опасные астероиды, должно основываться, как уже было сказано ранее, на формировании (расчете) в космическом пространстве пространственно-временных областей (ПРВО), находясь в которых КАП гарантированно (с баллистической точки зрения) перехватывает соответствующий конкретный астероид. К таким ПРВО относятся зоны атаки астероидов (ЗА), то есть ПРВО, в пределах которых должно быть осуществлено требуемое воздействие по ним, и рубежи атаки (РА) астероидов, представляющие собой, как уже говорилось ранее, ПРВО, находясь в которых КАП может осуществить движение по соответствующей баллистической траектории для совершения конкретных воздействий на астероид в зонах атаки. Предварительный расчет границ указанных зон и их пространственно-временных характеристик позволяет существенно сократить затраты времени на подготовку исходных данных, формирование и реализацию соответствующих управляющих воздействий на борту КАП при выполнении им маневра-перехвата астероида в конкретных условиях обстановки. Более того, как показывает анализ, предварительный расчет РА и их характеристик при решении задач динамического целераспределения и целенаведения группировки КАП на несколько одновременно летящих в сторону Земли астероидов или их осколков позволит существенно повысить оперативность синтеза и реализации программ перехвата указанных целей [29, 31, 32].

В соответствии с введенными понятиями зона атаки астероида Q определяется на декартовом произведении двух базисных множеств — множества точек W, из которых может быть осуществлено требуемое воздействие по ним, и множества времен R, когда это воздействие может произойти: $Q = W \times R$. В свою очередь, РА, представляющий собой множество точек (область) A, которое также может быть задано на декартовом произведение двух множеств: множества B —

множества точек пространственного положения РА и множества D — времен его существования, то есть $A = B \times D$.

С учетом вышеизложенного особую актуальность приобретает задача расчета и оценивания элементов множества A на основе элементов множеств Q и G (G — множество тактико-технических и баллистических характеристик КАП), а также поиска (конструктивного описания) отображения $F : \{Q, G\} \to A$.

В состав ТТХ КАП могут входить относительные расстояния между ними и астероидами, их относительные линейные и угловые скорости в момент воздействия КАП на астероиды и некоторые другие характеристики. В качестве баллистических характеристик могут выступать параметры орбит КАП и астероидов, фазовые координаты областей воздействия на астероиды и тому подобное. Помимо того, должны быть заданы способы применения КАП. Из множества указанных характеристик в статье исследуются только пространственно-временные характеристики РА, формируемые в околоземном космическом пространстве с учетом ограничений на относительные скорости между ними и КАП.

Исходя из вышеизложенного содержательного описания исследуемых задач, будем считать, что нам заданы:

– параметры орбиты КАП — $\{a_{ki}, e_{ki}, i_{ki}\} \subset G$;

– допустимая относительная скорость движения астероида и КАП в момент применения КАП — $\Delta V_{OTH} \in G$;

– параметры орбиты *j*-го астероида — $\{a_j, e_j, i_j, \omega_j\}$.

Для *i*-й УТ по $\{a_j, e_j, i_j, \omega_j\}$ могут быть определены параметры движения *j*-го астероида — $\{r_{ji}, \varphi_{ji}, t_{ji}, V_{ji}, \Theta_{ji}\}$, где r_{ji} — расстояние УТ от центра Земли или Луны; φ_{ji}, — угол между направлением на УТ и линией пересечения плоскостей орбит астероида и КАП; t_{ji} — время пролета астероидом плоскости движения КАП; V_{Aji} — скорость астероида в УТ; Θ_{Aji} — угол бросания астероида в УТ [2,7];

В ходе исследований пространственно-временных характеристик РА были приняты следующие допущения:

1) движение астероида и КАП моделируется по законам Кеплера;

2) КАП способен выполнять только компланарное маневрирование;

3) воздействие КАП на астероид осуществляется в УТ;

4) геометрические размеры Земли (Луны) не учитываются.

Новым объектом в исследовании характеристик РА астероидов в УТ является «зона обратной досягаемости» при ограничении на относительную скорость движения КАП и астероида в момент применения КАП, представляющая собой ПРВО, получаемую на основе обратного отображения множества допустимых значений начальных условий (НУ) движения КАП — $\{< r_{Kji}, V_{Kji}, \theta_{K1ji} >\}$ в ПРВО возможных движений КАП, обеспечивающих требуемое воздействия на астероид с учетом заданных ограничений на относительную скорость их сближения (рисунок 1).

Основа проводимого исследования заключается в установлении закономерностей изменения максимального и минимального радиусов РА астероида и соответствующих времен в зависимости от ориентации его вектора допустимой скорости движения — $\beta \in [0°,360°)$, обеспечивающего удовлетворение ограничения на относительную скорость движения КАП и астероида в момент применения КАП — ΔV_{OTH}, для различных значений аргумента широты КАП — $u \in [0°,360°)$ на основе построения «зон обратных досягаемостей».

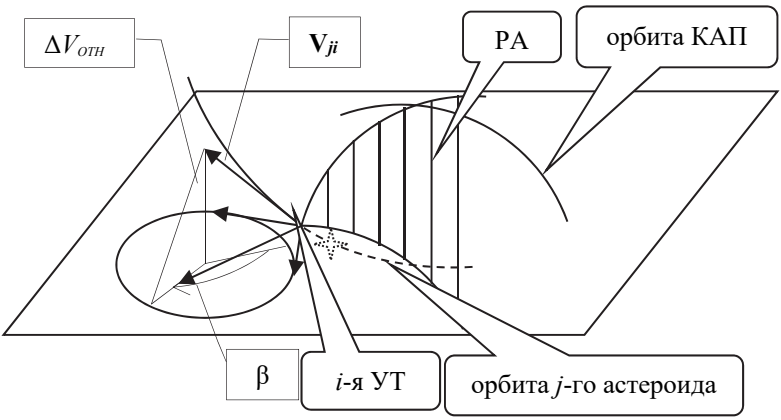


Рис. 1. Рубежи атаки астероида

Для проведения вычислительных экспериментов помимо указанных исходных данных должны быть также заданы шаг варьирования положения допустимого вектора скорости КАП в момент применения по астероиду — $\Delta\beta$ и шаг изменения аргумента широты для оценивания пространственных характеристик РА.

3. Модель оценивания характеристик РА при ограничениях на относительные скорости применения КАП по астероиду методом Монте-Карло. Основными математическими зависимостями в указанной модели являются:

– определение трансверсальной составляющей вектора скорости *j*-го астероида в *i*-й УТ [29, 30]:

$$V_{A\tau ji} = V_{Aji} \cos \Theta_{Aji}, \tag{1}$$

– вычисление отрезка, проектирующего годограф вектора скорости j-го астероида в i-й УТ на плоскость орбиты КАП:

$$V_{Ahji} = V_{A\tau ji} \sin \gamma_j, \tag{2}$$

где $\gamma_j = i_K - i_{Aj}$ — угол некомпланарности между орбитами КАП и *j*-го астероида; i_K — наклонение орбиты КАП; i_{Aj} — наклонение орбиты *j*-го астероида;

– расчет радиуса зоны годографов векторов скоростей движения КАП в *i*-й УТ, пригодных для воздействия по *j*-му астероиду:

$$\Delta V_{rji} = \sqrt{\Delta V_{OTH}^2 - V_{Ahji}^2}. \tag{3}$$

Таким образом, по приведенным зависимостям может быть сформировано множество значений начальных условий движения [30] КАП в УТ, обеспечивающих формирование «зон обратной досягаемости» при ограничении на относительную скорость движения КАП и астероида в момент применения КАП:

1) время — t_{ji};

2) радиус — r_{ji};

3) аргумент широты — φ_{ji};

4) множество пригодных для применения по *j*-му астероиду в *i*-й УТ значений скоростей КАП — $V_{Kji} \in [0, \Delta V_{rji}]$;

5) множество пригодных для применения по *j*-му астероиду в *i*-й УТ значений углов бросания КАП — $\Theta_{Kji} \in [0°, 360°)$.

В качестве одного из способов оценивания пространственно-временных характеристик РА астероидов в УТ был выбран способ,

базирующийся на реализациия метода Монте-Карло, используя который удалось рассчитать и проанализировать такие их параметры, как допустимые большие полуоси — $a_{PA} = a_{PA}(V_{Kji}, \Theta_{Kj}; t_{ji}, r_{ji})$, относительные эксцентриситеты — $e_{PA} = e_{PA}(V_{Kji}, \Theta_{Kj}; t_{ji}, r_{ji})$, аргументы широт перигеев — $\omega_{PA} = \omega_{PA}(V_{Kji}, \Theta_{Kj}; t_{ji}, r_{ji})$ и времена прохождения перигеев — $\tau_{ПРА} = \tau_{ПРА}(V_{Kji}, \Theta_{Kji}; t_{ji}, r_{ji})$. Для проведения указанных расчетов осуществлялось стохастическое моделирование возможных реализаций допустимых значений величин V_{Kji} и Θ_{Kji} при заданных t_{ji} и r_{ji}:

$$a_{PA} = r_{ji} / (2 - k), \tag{4}$$

где $k = r_{ji} \cdot V_{Kji}^2 / K_3$,

$$e_{PA} = \sqrt{1 - k(2-k)\cos\Theta_{Kji}}; \tag{5}$$

$$\omega_{PA} = \varphi_{ji} - \arccos[(a_{PA}(1 - e_{PA}^2) / r_{ji} - 1) / e_{PA}]; \tag{6}$$

$$\tau_{ПРА} = t_{ji} + \sqrt{a_{PA}^3 / K_3}(E_{ji} - e_{PA}\sin E_{ji}); \tag{7}$$

где $\begin{cases} \sin E_{ji} = (\sqrt{1 - e_{PA}^2 \sin\varphi_{ji}}) / (1 + e_{PA}\cos(\varphi_{ji} - \omega_{PA})) \\ \cos E_{ji} = (e_{PA} + \cos\varphi_{ji}) / (1 + e_{PA}\cos(\varphi_{ji} - \omega_{PA})) \end{cases}$.

По значениям a_{PA}, e_{PA}, ω_{PA} и $\tau_{ПРА}$ для различных значений аргументов широт $u \in [\varphi_{ji}, \varphi_{ji} - 360°)$ были оценены радиусы r и соответствующие моменты времен t. Соответствующие зависимости имеют вид:

$$r(u) = a_{PA}(1 - e_{PA}^2) / (1 + e_{PA}\cos(u - \omega_{PA})); \tag{8}$$

$$t(u) = \tau_{ПРА} + \sqrt{a_{PA}^3 / K_3}(E - e_{PA}\sin E), \tag{9}$$

где $K_3 = 398600,44\,км^3/c^2$ — гравитационная постоянная Земли; $E(u)$ — эксцентрическая аномалия:

$$\begin{cases} \sin E(u) = (\sqrt{1-e_{PA}^2}\sin u)/(1+e_{PA}\cos(u-\omega_{PA})) \\ \cos E(u) = (e_{PA}+\cos u)/(1+e_{PA}\cos(u-\omega_{PA})) \end{cases}. \quad (10)$$

Граничное время существования РА определяется временем t_{ji}. В любой момент времени $t(u) < t_{ji}$ пространственное положение РА в плоскости движения КАП (ОЧ) может быть оценено в соответствии с зависимостью:

$$\begin{cases} r_B(u) = \max\limits_N \{r_N(u)\}; \\ r_H(u) = \min\limits_N \{r_N(u)\}. \end{cases} \quad (11)$$

Время *t* определяется по формуле:

$$t = t_{ji} - \Delta t, \quad (12)$$

где $\Delta t = \sqrt{a_{PA}^3 / K_3}\cdot[(E_{ji} - E(u)) - e_{PA}^2(\sin E_{ji} - \sin E(u))]$.

В то же время задание случайным образом значений угла Θ_{Kji} приводит к неоднозначности в определении модуля вектора скорости V_{Kji}. В этой связи при проведении исследований характеристик РА при ограничениях на относительные скорости целесообразно перейти к моделированию угла β между проекции вектора скорости *j*-го астероида в *i*-й УТ на плоскость орбиты КАП и текущим направлением на годограф вектора скорости КАП (ОЧ), детерминировано изменяющегося от 0^0 до 360^0 (рисунок 2).

Тогда для моделирования V_{Kji} и Θ_{Kji} необходимо:

1) вычисление проекции вектора скорости *j*-го астероида в *i*-й УТ на плоскость орбиты КАП:

$$V_{Afji} = V_{A\tau ji}\cdot\cos\gamma_j; \quad (13)$$

2) при заданном угле $\beta = \beta_0$ моделирование по равномерному закону распределения случайных величин, принадлежащих отрезку $s \in [0, \Delta V_{rji}]$;

3) определение V_{Kji}:

$$V_{Kji} = \sqrt{s^2 + V_{Afji}^2 - 2 \cdot s \cdot V_{Afji} \cdot \cos\beta}; \quad (14)$$

4) оценивание угла бросания Θ_{Kji}:

$$\Theta_{Kji} = \alpha + \Theta_{dji}; \quad (15)$$

где $\sin\alpha = s \cdot \sin\beta / V_{Kji}$; $tg\Theta_{dji} = V_{Adji} / V_{Ahji}$; $V_{Adji} = V_{Aji} \cdot \sin\Theta_{Aji}$.

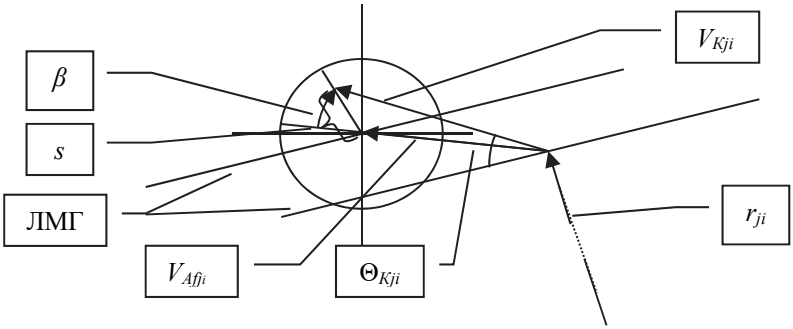


Рис. 2. Геометрическое описание взаимосвязи V_{Kji} и Θ_{Kji}

4. Алгоритм оценивания характеристик РА при ограничениях на относительные скорости применения КАП по астероиду, а также результаты вычислительных экспериментов.

Основными этапами реализации алгоритма являются:

1. Ввод исходных данных.
2. Расчет $V_{A\tau ji}$, V_{Ahji}, ΔV_{rji}, V_{Afji} по формулам (1)-(3).
3. Задание начального значения аргумента широты КАП — $u = \varphi_{ji}$.
4. Задание начального значения угла $\beta = 0°$.
5. Задание начального значения счетчика числа испытаний в методе Монте-Карло $N := 1$.
6. Моделирование случайной величины s по равномерному закону распределения на интервале $[0, \Delta V_{rji}]$.

7. Расчет величин V_{Kji}, α, Θ_{dji}, V_{Adji}, Θ_{Kji} по формулам (14)-(15).

8. Расчет величин k_N, a_{PAN}, e_{PAN}, ω_{PAN}, E_{jiN}, τ_{PAN} по формулам (4)-(10).

9. Расчет $r_N(u)$, $E_N(u)$, $t_N(u)$, Δt, t по формулам (11)-(12) и вывод значений $r_N(u)$, $t_N(u)$.

10. Задание следующего значения счетчика числа испытаний в методе Монте-Карло $N := N + 1$.

11. Оценивание величины N: если $N \le N_3$, то переход на п.6; если $N > N_3$, то переход на п.12 (в ходе вычислительных экспериментов величина N_3 принималась равной 200).

12. Изменение значения угла β на величину $\Delta\beta$.

13. Оценивание величины β: если $\beta \le 360°$, то переход на пункт 5; если $\beta > 360°$, то переход на пункт 14.

14. Определение величин $r_B(u)$, $r_H(t)$, t по формулам (11), (12).

15. Изменение значения аргумента широты КАП: $u := u - \Delta u$ (шаг изменения Δu в ходе вычислительных экспериментов принимался равным 10^0).

16. Оценивание величины u: если $u \ge 0°$, то переход на пункт 4; если $u < 0°$, то переход на пункт 17.

17. Вывод $\{r_N(u), t_N(u)\}$.

Исходные данные для проведения вычислительных экспериментов приведены в таблице 1

Таблица 1. Исходные данные для проведения экспериментов

r_{ji}, км	φ_{ji}, град	t_{ji}, с	V_{Aji}, км/с	Θ_{Aji}, гр	i_K, град	i_{Aj}, гр	$\Delta V_{отн}$, км/с	Δt, с	$\Delta\beta$, гр	K_3, км^3/с^2
7000	45	1000	10	1	63	63,5	0,5	60	10	398600,44

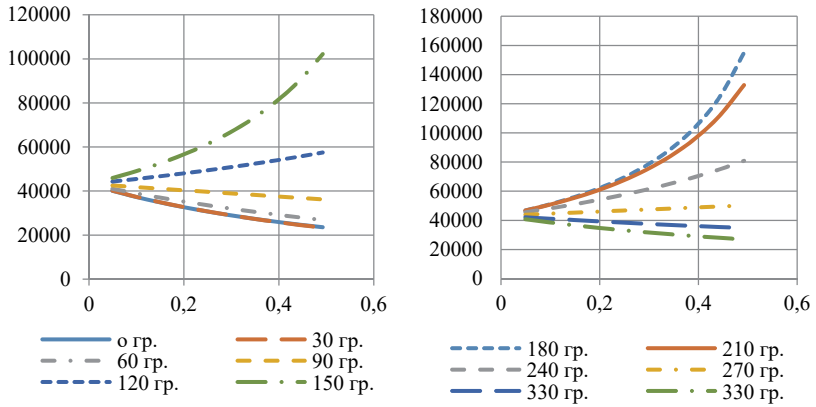
В ходе вычислительных экспериментов была исследована зависимость изменения величины модуля радиуса-вектора $r(u)$ от изменения величины радиуса точки годографа вектора скорости КАП — s при различных значениях угла β. Результаты исследования указанных зависимостей иллюстрируются при $u = 45°$ для $\beta = 0°(30°)150°$ и десяти значений s от 0 км до $\Delta V_{rji}c$ шагом $\Delta V_{rji}/10$ по каждому β данными таблице 2.

Таблица 2. Результаты исследования зависимости величины модуля $r(u)$ от s

β , гр.	0	30	60	90	120	150
s, км/с	r(u),км	r(u),км	r(u),км	r(u),км	r(u),км	r(u),км
0,0492328	40111,787	40183,953	41043,96	42538,038	44335,636	45945,492
0,0984656	37365,749	37471,54	39006,345	41833,652	45487,299	48994,108
0,1476985	34937,957	35050,181	37113,167	41129,947	46702,639	52459,779
0,1969313	32776,195	32875,631	35349,807	40427,118	47986,919	56434,385
0,2461641	30839,077	30912,181	33703,553	39725,362	49345,991	61038,932
0,2953969	29093,378	29130,694	32163,301	39024,881	50786,39	66436,097
0,3446298	27512,12	27507,163	30719,307	38325,882	52315,429	72849,842
0,3938626	26073,166	26021,623	29362,986	37628,572	53941,323	80597,284
0,4430954	24758,183	24657,342	28086,744	36933,163	55673,331	90142,746
0,4923282	23551,857	23400,186	26883,839	36239,868	57521,925	102194,39

Соответствующие графики приведены на рисунке 3(а).

Аналогичные результаты для $\beta = 180°(30°)330°$ приведены в таблице 3. Соответствующие графики приведены на рисунке 3(б).



а) б)

Рис. 3. Графики зависимостей $r(u)$ от s для: а) $\beta = 0°(30°)150°$; б) для $\beta = 180°(30°)330°$

Значения $r_B(u)$, $r_H(u)$ получаемые по формулам (11) для тех же значений β и соответствующие времена $t_B(u)$, $t_H(u)$ сведены в таблице 4.

Таблица 3. Результаты исследования зависимости величины модуля $r(u)$ от s

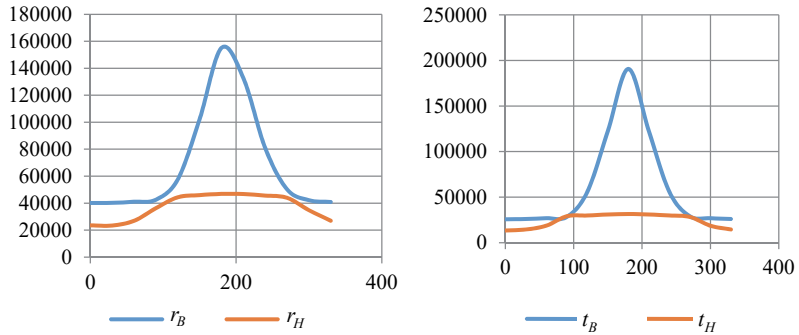
β , гр.	180	210	240	270	300	330
s, км/с	r(u),км	r(u),км	r(u),км	r(u),км	r(u),км	r(u),км
0,0492328	46846,097	46718,983	45622,845	43948,092	42204,845	40830,612
0,0984656	51036,632	50703,28	48206,722	44653,392	41217,555	38664,199
0,1476985	55970,684	55315,543	51021,567	45358,637	40277,52	36707,898
0,1969313	61865,168	60716,341	54099,393	46063,654	39381,538	34932,552
0,2461641	69030,468	67126,082	57478,419	46768,276	38526,687	33314,161
0,2953969	77926,863	74855,769	61204,643	47472,339	37710,296	31832,789
0,3446298	89267,434	84358,866	65333,93	48175,686	36929,923	30471,74
0,3938626	104219,25	96323,42	69934,808	48878,164	36183,325	29216,926
0,4430954	124833,05	111846,29	75092,271	49579,627	35468,443	28056,376
0,4923282	155074,9	132789,74	80913,041	50279,931	34783,38	26979,851

Таблица 4. Результаты исследования зависимостей величин $r_B(u)$, $r_H(u)$, $t_B(u)$, $t_H(u)$ от β

β , гр.	r_B , км	r_H , км	t_B , с	t_H , с
0	40111,787	23551,857	25697,535	13457,179
30	40183,953	23400,186	26031,271	14592,66
60	41043,96	26883,839	26977,724	18736,992
90	42538,038	36239,868	28368,744	29137,213
120	57521,925	44335,636	55956,634	29887,121
150	102194,39	45945,492	122787,64	31088,553
180	155074,9	46846,097	190672,19	31547,223
210	132789,74	46718,983	122456,82	31082,969
240	80913,041	45622,845	55748,887	29877,945
270	50279,931	43948,092	29035,722	28358,862
300	42204,845	34783,38	26969,722	18688,367
330	40830,612	26979,851	26026,867	14572,674

Соответствующие графики приведены на рисунках 4(а) и 4(б).

В ходе вычислительных экспериментов были получены значения $r_B(u)$, $r_H(u)$ для различных значений $u = 0°,15°,30°$ и $45°$ (таблица 5).



а) б)

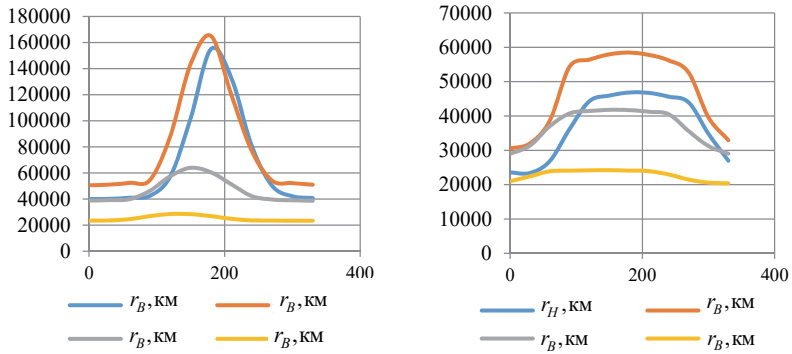
Рис. 4. Графики зависимостей от β : а) $r_B(u)$, $r_H(u)$; б) $t_B(u)$, $t_H(u)$ от β

Таблица 5. Результаты исследования зависимостей величин $r_B(u)$, $r_H(u)$ от β

β, гр.	r_B, км	r_B, км	r_B, км	r_B, км	r_H, км	r_H, км	r_H, км	r_H, км
u, гр.	0	15	30	45	0	15	30	45
0	23512,331	38818,7	50613,533	40111,787	20979,733	28964,212	30534,493	23551,857
30	23684,04	39250,309	51113,799	40183,953	22434,147	31386,242	32016,972	23400,186
60	24635,098	39959,209	52465,156	41043,96	23892,948	37028,253	38687,279	26883,839
90	27004,564	46418,893	54692,866	42538,038	24082,176	40768,862	54396,421	36239,868
120	28513,22	57639,238	88365,275	57521,925	24197,803	41453,692	56430,487	44335,636
150	28415,386	64007,521	143733,22	102194,39	24206,61	41807,048	57962,284	45945,492
180	26941,345	60533,183	165038,83	155074,9	24107,31	41720,792	58475,11	46846,097
210	24901,26	51366,242	119914,61	132789,74	23929,742	41228,837	57796,383	46718,983
240	23723,559	42334,601	77007,483	80913,041	22966,903	40486,093	56184,127	45622,845
270	23542,918	39702,877	54169,847	50279,931	21480,9	35588,436	52619,65	43948,092
300	23433,152	39078,403	52313,168	42204,845	20590,766	31235,944	39595,338	34783,38
330	23421,617	38758,956	51042,845	40830,612	20390,046	29034,013	32941,763	26979,851

Соответствующие графики приведены на рисунках 5(а) и 5(б).

Анализ результатов, полученных в ходе вычислительных экспериментов, показывает, что годограф вектора скорости КАП, обеспечивающий $r_B(u)$, $r_H(u)$, должен принадлежать окружности скоростей движения КАП в i-й УТ, которые пригодны для воздействия по j-му астероиду.



а) б)

Рис. 5. Графики зависимостей от β : а) $r_B(u)$; б) $r_H(u)$

5. Аналитическая модель оценивания пространственно-временных характеристик РА астероидов. Оценивания пространственно-временных характеристик РА включает в себя:

1) определение величины проекции вектора скорости j-го астероида в i-й УТ на плоскость орбиты КАП — V_{Afji} по формуле (13);

2) определение начальных условий движения КАП по орбите, соответствующей «зоне обратной досягаемости» при $\beta = \beta_0 = 0° - V_{ji}(0°)$, $r_{ji}, \theta_1(0°)$:

$$V_{ji}(0°) = V_{Afji} + \Delta V_{rji} \text{ж}; \tag{16}$$

$$\theta_1(0°) = \arccos(V_{A\tau ji} \cdot \cos\gamma_j / \Delta V_{rji}); \tag{17}$$

3) вычисление по известным начальным условиям кеплеровских элементов орбиты КАП, соответствующей $\beta = \beta_0 = 0°$ по формулам (4)-(6);

4) определение начальных условий движения КАП по орбите, соответствующей «зоне обратной досягаемости» при $\beta = 180° : V_{ji}(180°)$, $r_{ji}, \theta_1(180°)$:

$$V_{ji}(180°) = V_{Afji} - \Delta V_{rji}; \tag{18}$$

$$\theta_1(0°) = \arccos(V_{A\tau ji} \cdot \cos\gamma_j / \Delta V_{rji}); \tag{19}$$

5) вычисление по известным начальным условиям кеплеровских элементов орбиты КАП, соответствующей $\beta = 180°$ по формулам (4)–(6);

6) задание начального значения угла $\beta := \beta_0 = 0°$;

7) вычисление модуля предельного пригодного вектора скорости КАП в момент применения по *j*-му астероиду для заданного значения угла β:

$$V_{ji}(\beta) = \sqrt{V_{Afji}^2 + \Delta V_{rji}^2 + 2V_{Afji} \cdot \Delta V_{rji} \cdot \cos\beta};$$

8) определение угла бросания проекции вектора скорости *j*-го астероида в *i*-й УТ на плоскость орбиты КАП:

$$\cos\theta_1(\beta) = V_{Afji} / V_{ji};$$

9) расчет дополнительного угла бросания вектора предельного пригодного вектора скорости КАП:

$$\delta(\beta) = \arcsin(\Delta V_{rji} / V_{ji} \sin\beta);$$

10) определение угла бросания вектора предельного пригодного вектора скорости КАП:

$$\theta(\beta) = \theta_1(\beta) - \delta(\beta);$$

11) вычисление по известным начальным условиям $< r_{ji}, V_{ji}(\beta), \theta(\beta) >$ плоскостных кеплеровских элементов орбиты КАП — $< a_{ji}(\beta), e_{ji}(\beta), \omega_{ji}(\beta) >$, где $a_{ji}(\beta)$ определяется по формуле (4), $e_{ji}(\beta)$ — по формуле (5), $\omega_{ji}(\beta)$ — по формуле (6);

12) задание начального значения угла аргумента широты $u := u_0 = 0°$;

13) определение модуля радиуса-вектораКАП на орбите, соответствующей $\beta = \beta_0 = 0°$: $V_{ji}(0°)$, r_{ji}, $\theta_1(0°)$ по формуле (8);

14) вычисление модуля радиуса-вектора КАП на орбите, соответствующей $\beta = 180°$: $V_{ji}(180°)$, r_{ji}, $\theta_1(180°)$ по формуле (8);

15) расчет модуля радиуса-вектора КАП — $r_{KAII}(u;\beta)$ на орбите, соответствующей углам uи β по формуле (8);

16) оценивание эксцентрической аномалии УТ по формулам (10);

17) определение эксцентрической аномалии, соответствующей углу u по формулам (10);

18) вычисление времени движения КАП от точки с аргументом широты u до УТ по формуле (9);

19) расчет момента времени пролета КАП точек орбиты с аргументом широты u

$$t(u;\beta) = t_{ji} + \Delta t(u);$$

20) перерасчет значения аргумента широты $u := u + \Delta u$; если $u \in [0°,360°)$, переход на пункты 16 и 15 и накопление значений $r_{СЧ}(u;\beta)$, $t(u;\beta)$; если — нет, переход на пункт 21;

21) перерасчет значения угла $\beta := \beta + \Delta\beta$; если $\beta \in [0°,360°)$, переход на пункт 6 и накопление значений $r_{СЧ}(u;\beta)$; если — нет, то переход на пункт 22;

22) определение радиусов наружной границы РА для определенных значений u:

$$r_H(u) = \min r_{КАП}(u;\beta)$$

и соответствующих времен $t_H(u)$;

23) определение радиусов внутренней границы РА для определенных значений u:

$$r_B(u) = \max r_{КАП}(u;\beta)$$

и соответствующих времен $t_B(u)$.

6. Результаты оценивания пространственно-временных характеристик РА астероидов. Апробация предлагаемой модели была проведена для исходных данных, представленных в таблицах 6 и 7.

Таблица 6. Исходные данные для проведения экспериментов

$\Delta V_{отн}$, км/с	a_j, км	e_j	$i_{КАП}$, гр.
4	7000	0	80

Таблица 7. Исходные данные для проведения экспериментов

r_{ji}, км	φ_{ji}, гр	t_{ji}, с	V_{ji}, км/с	θ_{ji}, гр.	Δu, гр.	$\Delta\beta$, гр
7000	10	200	7,54605327	0	30	30

Порядок проведения вычислительных экспериментов состоял в однократном получении указанных в пунктах 1-8 величин и вычислении величин указанных в пунктах 10-27 при двух цикловом варьировании величинами углов β и u в указанных выше диапазонах с

заданными шагами (внешний цикло по углу β, внутренний по углу u). С учетом того, что β и u могут изменяться непрерывно в указанных диапазонах для зависимостей $r_H(u)$ и $r_B(u)$, $t_H(u)$ и $t_B(u)$ соответствующие графики были построены как огибающие линии, проходящие через полученные в расчетах и приведенные в таблицах точки значений $r_{бпр}(u;\beta)$.

Результаты расчетов величин модулей радиусов-векторов КАП при $\beta = 0°$ и $\beta = 180°$ приведены в таблице 8 во 2 и 8 столбцах. В этой же таблице для различных значений аргумента широты — $u = 0°(30°)360°$ представлены значения $r_{КАП}(u;\beta)$ при $\beta = 0°(30°)150°$ (столбцы 3-7).

Полученные результаты сведены в графики на рисунках 6, 7.

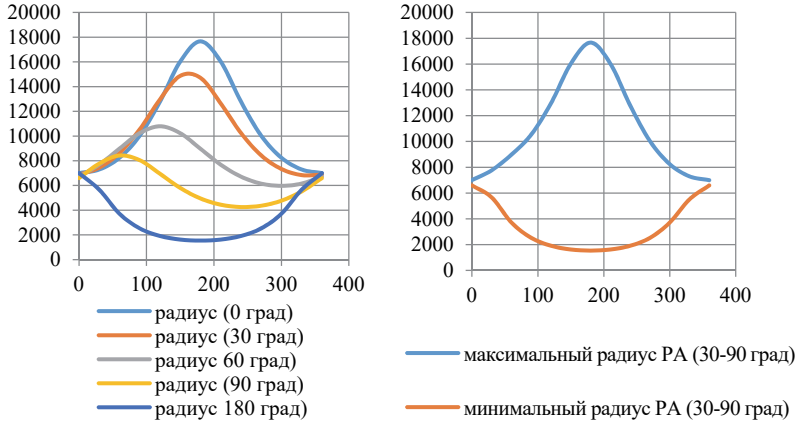
На рисунке 6а приведены графики для $\beta = 0°(30°)90°$ и для $\beta = 0°$, $\beta = 180°$, а на рисунке 6б — для $r_B(u)$ и $r_H(u)$.

На рисунке 7а приведены графики для $\beta = 120°(30°)180°$ и для $\beta = 0°$, $\beta = 180°$, а на рисунке 7б — для $r_B(u)$ и $r_H(u)$.

Таблица 8. Результаты исследования зависимости величины $r_{КАП}(u;\beta)$ от u

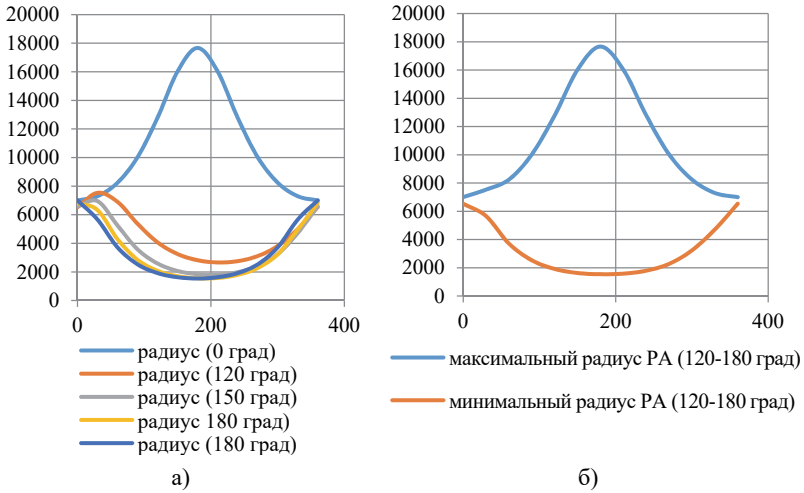
, гр	$r_{КАП}(u;0°)$ км	$r_{КАП}(u;30°)$ км	$r_{КАП}(u;60°)$ км	$r_{КАП}(u;90°)$ км	$r_{КАП}(u;120°)$ км	$r_{КАП}(u;150°)$ км	$r_{КАП}(u;180°)$ км
0	7000	6872,683	6721,607	6596,361	6532,261	6590,084	7000
30	7295,011	7441,268	7689,113	7769,653	7533,058	6950,923	5657,871
60	8244,263	8631,117	8980,556	8423,844	6899,894	5244,177	3712,944
90	10026,49	10538,51	10237,69	8045,619	5365,757	3575,733	2526,535
120	12791,79	12971	10787,78	6946,743	4065,754	2584,309	1914,719
150	16027,79	14871,4	10230,81	5799,596	3249,579	2066,956	1626,405
180	17663,33	14711,5	8971,381	4948,483	2815,787	1837,326	1541,448
210	16027,79	12643,52	7681,347	4444,938	2663,268	1811,114	1626,405
240	12791,79	10244,04	6716,465	4255,858	2752,568	1978,926	1914,719
270	10026,49	8432,586	6151,527	4359,395	3106,943	2401,829	2526,535
300	8244,263	7332,317	5968,652	4768,068	3812,865	3235,601	3712,944
330	7295,011	6838,333	6154,016	5517,087	4987,665	4712,335	5657,871
360	7000	6872,683	6721,607	6596,361	6532,261	6590,084	7000

Результаты расчетов величин моделей радиусов-векторов КАП при $u = 0°(30°)360°$ и $\beta = 210°(30°)330°$ приведены в таблице 9 (столбцы 2-6). В этой же таблице для различных значений аргумента широты — $u = 0°(30°)360°$ представлены значения $r_H(u)$ и $r_B(u)$ (столбцы 7, 8).



а) б)

Рис. 6. Графики зависимостей от u : а) $r_{КАП}(u;\beta)$; б) $r_B(u)$ и $r_H(u)$



а) б)

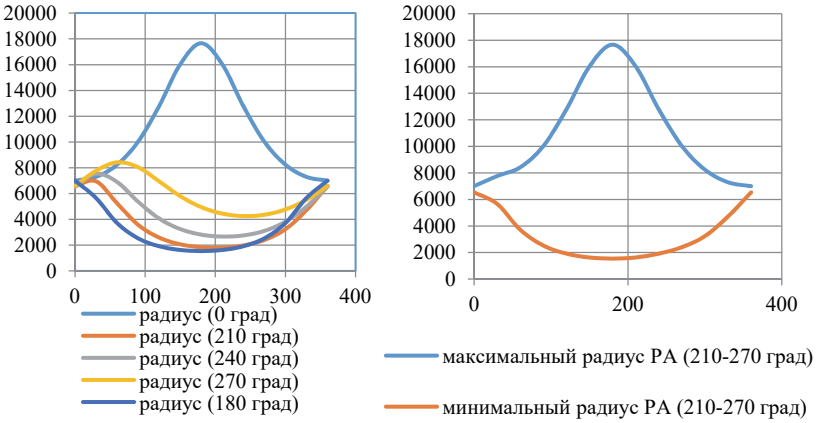
Рис. 7. Графики зависимостей от u : а) $r_{КАП}(u;\beta)$; б) $r_B(u)$ и $r_H(u)$

На рисунке 8а приведены графики для $\beta = 210°(30°)270°$ и для $\beta = 0°$, $\beta = 180°$, а на рисунке 8б — для $\max r_{КАП}(u;\beta)$ и $\min r_{КАП}(u;\beta)$.

Таблица 9. Результаты исследования зависимостей величин $r_{КАП}(u;\beta)$, $r_H(u)$, $r_B(u)$ от u

u, град	$r_{КАП}(u;210°)$, км	$r_{КАП}(u;240°)$, км	$r_{КАП}(u;270°)$, км	$r_{КАП}(u;300°)$, км	$r_{КАП}(u;330°)$, км	$r_B(u)$, км	$r_H(u)$, км
0	6590,084	6532,261	6596,361	6721,607	6872,683	7026,915	6532,261
30	6950,923	7533,058	7769,653	7689,113	7441,268	7769,653	5657,871
60	5244,177	6899,894	8423,844	8980,556	8631,117	8980,556	3712,944
90	3575,733	5365,757	8045,619	10237,69	10538,51	10538,51	2526,535
120	2584,309	4065,754	6946,743	10787,78	12971	12971	1914,719
150	2066,956	3249,579	5799,596	10230,81	14871,4	16027,79	1626,405
180	1837,326	2815,787	4948,483	8971,381	14711,5	17663,33	1541,448
210	1811,114	2663,268	4444,938	7681,347	12643,52	16823,12	1626,405
240	1978,926	2752,568	4255,858	6716,465	10244,04	13791,97	1914,719
270	2401,829	3106,943	4359,395	6151,527	8432,586	10767,95	2401,829
300	3235,601	3812,865	4768,068	5968,652	7332,317	8690,169	3235,601
330	4712,335	4987,665	5517,087	6154,016	6838,333	7509,226	4712,335
360	6590,084	6532,261	6596,361	6721,607	6872,683	7026,915	6532,261

Полученные результаты сведены в графики на рисунках 8-10.

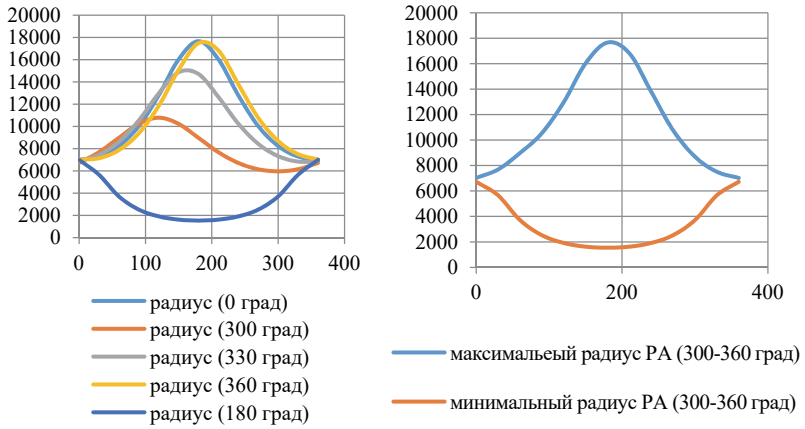


а) б)

Рис. 8. Графики зависимостей от u : а) $r_{КАП}(u;\beta)$; б) $r_B(u)$ и $r_H(u)$

На рисунке 9(а) приведены графики для $\beta = 300°(30°)360°$ и для $\beta = 0°$, $\beta = 180°$, а на рисунке 9(б) — для $\max r_{КАП}(u;\beta)$ и $\min r_{КАП}(u;\beta)$.

На рисунке 10 представлены графики зависимостей максимального радиуса РА и минимального радиуса РА в функции от аргумента широты.



а) б)

Рис. 9. Графики зависимостей от u : а) $r_{KAП}(u;\beta)$; $r_B(u)$ и $r_H(u)$

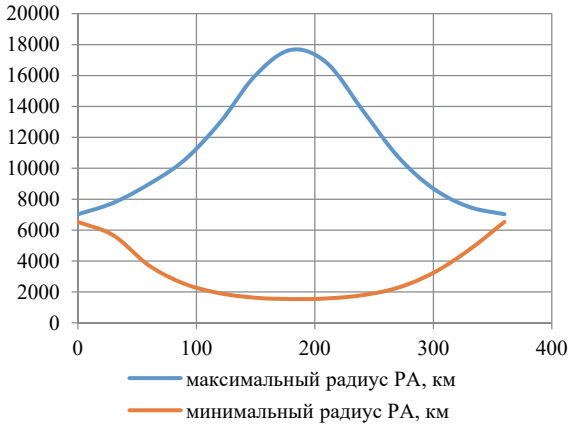


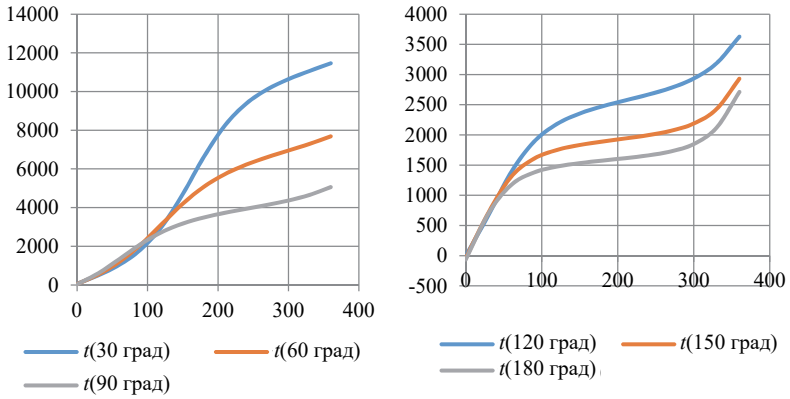
Рис. 10. Графики зависимостей $\max r_{KAП}(u;\beta)$ и $\min r_{KAП}(u;\beta)$ от $\max r_{KAП}(u;\beta)$ и u

Результаты расчетов времен, необходимых для пролета границ РА при $u = 0°(30°)360°$ и $\beta = 30°(30°)180°$, приведены в таблице 10.

Таблица 10. Результаты исследования зависимости величины $t(u;\beta)$ от u

u, гр.	$t(u;30°)$, с	$t(u;60°)$, с	$t(u;90°)$, с	$t(u;120°)$, с	$t(u;150°)$, с	$t(u;180°)$, с
0	62,77918	51,70948	30,07894	-2,24299	-39,5597	-64,7661
30	496,1395	539,1359	602,385	672,9495	716,975	703,8212
60	1042,695	1194,83	1339,062	1387,911	1311,159	1174,099
90	1822,6	2074,764	2102,84	1890,05	1608,096	1378,654
120	3007,393	3137,876	2728,629	2180,921	1751,395	1475,661
150	4698,252	4200,25	3175,464	2354,935	1833,271	1532,984
180	6625,065	5078,65	3491,168	2474,595	1891,147	1575,659
210	8253,396	5732,978	3731,811	2572,337	1941,759	1615,335
240	9374,242	6219,521	3938,198	2667,851	1996,332	1661,255
270	10114,02	6607,392	4140,543	2779,559	2069,034	1727,811
300	10639,8	6951,314	4367,707	2935,234	2189,061	1849,491
330	11064,53	7295,378	4656,542	3187,861	2428,828	2120,676
360	11462,2	7683,711	5059,117	3627,715	2930,957	2712,792

На рисунке 11а приведены графики зависимостей t(u;β) для $\beta = 30°(30°)90°$ и, а на рисунке 11б — для $\beta = 120°(30°)180°$.



а) б)

Рис. 11. Графики зависимостей $t(u;\beta)$ от u : а) для $\beta = 30°(30°)90°$; б) для $\beta = 120°(30°)180°$

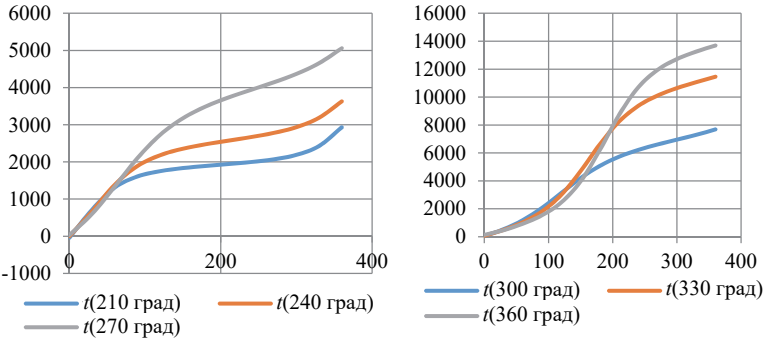
Труды СПИИРАН. 2019. Том 18 № 6. ISSN 2078-9181 (печ.)
ISSN 2078-9599 (онлайн) www.proceedings.spiiras.nw.ru

Результаты расчетов времен, необходимых для пролета границ РА при $u = 0°(30°)360°$ и $\beta = 210°(30°)270°$, $\beta = 300°(30°)360°$ приведены в таблице 11.

Таблица 11. Результаты исследования зависимости величины $t(u;\beta)$ от u

u, гр.	$t(u;210°)$, с	$t(u;240°)$, с	$t(u;270°)$, с	$t(u;300°)$, с	$t(u;330°)$, с	$t(u;360°)$, с
0	-39,5597	-2,24299	30,07894	51,70948	62,77918	135,5967
30	716,975	672,9495	602,385	539,1359	496,1395	474,3378
60	1311,159	1387,911	1339,062	1194,83	1042,695	934,6346
90	1608,096	1890,05	2102,84	2074,764	1822,6	1542,287
120	1751,395	2180,921	2728,629	3137,876	3007,393	2464,613
150	1833,271	2354,935	3175,464	4200,25	4698,252	3967,229
180	1891,147	2474,595	3491,168	5078,65	6625,065	6216,622
210	1941,759	2572,337	3731,811	5732,978	8253,396	8737,53
240	1996,332	2667,851	3938,198	6219,521	9374,242	10706,83
270	2069,034	2779,559	4140,543	6607,392	10114,02	11949,14
300	2189,061	2935,234	4367,707	6951,314	10639,8	12722,79
330	2428,828	3187,861	4656,542	7295,378	11064,53	13258,37
360	2930,957	3627,715	5059,117	7683,711	11462,2	13689,82

На рисунке 12а приведены графики зависимостей t(u;β) для $\beta = 210°(30°)270°$, а на рисунке 12б — для $\beta = 300°(30°)360°$.



а) б)

Рис. 12. Графики зависимостей $t(u;\beta)$ от u : а) для $\beta = 210°(30°)270°$; б) для $\beta = 300°(30°)360°$

В таблице 12 сведены данные по РА астероидов в УТ, полученные при принятых исходных данных, как пространственно-временных областях для различных значений $u = 0°(30°)360°$.

Таблица 12. Результаты исследования зависимостей пространственно-временных характеристик РА астероидов от u

u, гр.	радиус, км	радиус, км	время, с	время, с
	максимум	минимум	максимум	минимум
0	7026,915	6532,261	135,5967	-64,7661
30	7769,653	5657,871	716,975	496,1395
60	8980,556	3712,944	1387,911	1042,695
90	10538,51	2526,535	2102,84	1378,654
120	12971	1914,719	3137,876	1475,661
150	16027,79	1626,405	4698,252	1532,984
180	17663,33	1541,448	6625,065	1575,659
210	16823,12	1578,566	8737,53	1615,335
240	13791,97	1790,873	10706,83	1661,255
270	10767,95	2274,165	11949,14	1727,811
300	8690,169	3233,235	12722,79	1849,491
330	7509,226	4712,335	13258,37	2120,676
360	7026,915	6532,261	13689,82	2712,792

Полученные результаты позволили разработать схему оценивания параметров РА от u и высоты над поверхностью Земли — $H_{кр}$, представленную на рисунке 13.

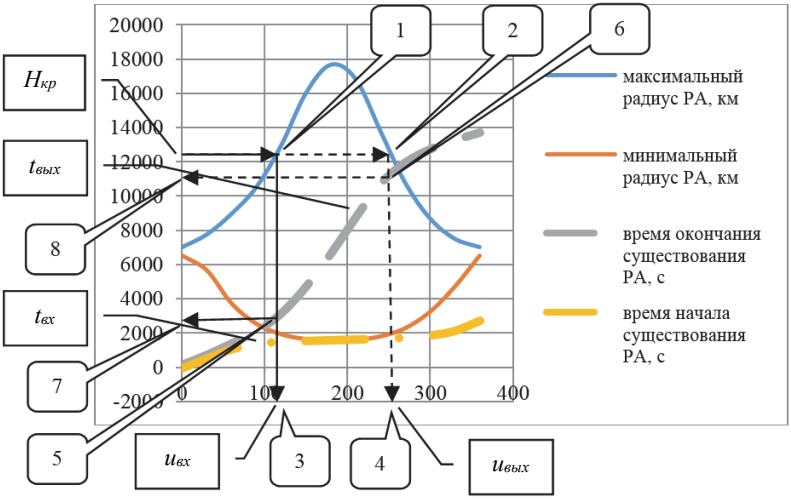


Рис. 13. Схема оценивания параметров РА от u и высоты над поверхностью Земли — $H_{кр}$

Труды СПИИРАН. 2019. Том 18 № 6. ISSN 2078-9181 (печ.)
ISSN 2078-9599 (онлайн) www.proceedings.spiiras.nw.ru

Особенностью графиков на рисунке 13 является то, что диапазон изменения времен и диапазон изменения времен по оси ординат совпадают.

Представленные на рисунке 13 графики позволяют оценить параметры РА не только в зависимости от различных значений аргумента широты, но и от радиуса (высоты над поверхностью Земли) по следующей схеме:

1) по заданной высоте круговой орбиты КАП – $H_{кр}$ определяются точки (как правило две) пересечения соответствующей горизонтальной линии с графиком зависимости максимального радиуса РА от аргумента широты — точки «1» и «2»;

2) проектирование этих точек на ось абцисс и определение соответствующих значений аргументов широт — $u_{вх}$ и $u_{вых}$ — точки «3» и «4»;

3) определение точек пересечения проектирующих на ось абцисс с графиками зависимостей времени существования РА — точки «5» и «6»;

4) определение абцисс этих точек — соответствующих времен $t_{вх}$ и $t_{вых}$ — точки «7» и «8».

4. Заключение. Приведенные в статье результаты позволяют сделать вывод о том, что размеры РА колеблются в значительных диапазонах для околоземных орбит КАП даже при незначительных диапазонах относительных скоростей КАП и астероидов, что позволяет прогнозировать достаточно широкий диапазон параметров рубежей применения этих КА. При этом предлагаемый в статье подход к оцениванию пространственно-временных характеристик РА астероидов позволяет получать их независимо от параметров движения КАП, что существенно облегчает последующий поиск параметров рубежей применения этих КА. Кроме того, разработанная аналитическая модель оценивания пространственно-временных характеристик РА астероидов позволяет определять аналогичные характеристик РА при ограничениях на относительное расстояние между астероидом и КАП в момент его применения.

Литература

1. *Ross I.M., Park S.Y., Porter S.E.* Gravitational Effects of Earth in Optimizing Delta-V for Deflecting Earth-Crossing Asteroids // Journal of Spacecraft and Rockets. 2001. vol. 38. no. 5. pp. 759–764.
2. Asteroid Impact & Deflection Assessment (AIDA) collaboration. URL: https://www.esa.int/Safety_Security/Hera/Asteroid_Impact_Deflection_Assessment_AIDA_collaboration (дата обращения: 15.08.2019).
3. *Dwayne A.* Giant bombs on giant rockets: Project Icarus // The SpaceReview. 2004.
4. Asteroid deflection mission seeks smashing ideas. URL: http://www.esa.int/Enabling_Support/Operations/Asteroid_deflection_mission_seeks_smashing_ideas (дата обращения: 15.08.2019).

5. *Powell C.S.* Developing Early Warning Systems for Killer Asteroids // Discover. 2013. pp. 60–61.
6. *Harper P.* Earth will be hit by asteroid with 100% certainty – space experts warn // Daily Star. 2018.
7. *Homer A.* Earth Will Be Hit By An Asteroid With 100 Percent Certainty, Says Space-Watching Group B612 // Inquisitr. 2018.
8. *Stanley-Becker I.* Stephen Hawking feared race of 'superhumans' able to manipulate their own DNA. URL: https://www.washingtonpost.com/news/morning-mix/wp/2018/10/15/stephen-hawking-feared-race-of-superhumans-able-to-manipulate-their-own-dna/ (дата обращения: 15.08.2019).
9. *Константиновская Л.В.* Прогноз и предотвращение кометно-астероидной опасности. URL: http://www.astronom2000.info/different/pipao/ (дата обращения: 11.09.2013).
10. Угрозы из космоса // Популярная Механика. 2010. № 9(95). С. 31–46.
11. *Назиров Р.Р., Эйсмонт Н.А.* Гравитационные маневры как способ направить малые астероиды на траекторию встречи с опасными околоземными объектами // Космические исследования. 2010. Т. 48. № 5. С. 1–6.
12. *Седых Р.* Угрозы из космоса растут // Военное обозрение. 2013. URL: http://vpknews.ru/ (дата обращения: 21.06.2017).
13. *Колин К.К.* Глобальные угрозы развитию цивилизации в XXI веке // Стратегические приоритеты. 2014. № 1. С. 6–30.
14. *Бакланов О. и др.* От астероидно-космической опасности Землю защитит «Цитадель» // Воздушно-космическая сфера. 2017. № 1. С. 90–95.
15. *Billings L.* Words matter: A call for responsible communication about asteroid impact hazards and plans for planetary defense // Space Policy. 2015. vol. 33. pp. 8–12.
16. *Wall M.* A Killer Asteroid Is Coming – We Don't Know When (So Let's Be Ready), Bill Nye Says. URL: https://www.space.com/killer-asteroids-warning-bill-nye.html (дата обращения: 15.08.2019).
17. *Johns H.U.* Asteroids are stronger, harder to destroy than previously thought. URL: https://www.sciencedaily.com/releases/2019/03/190304095926.htm (дата обращения: 15.08.2019).
18. *Charles E.M., Ramesh K.T., Richardson D.C.* A new hybrid framework for simulating hypervelocity asteroid impacts and gravitational reaccumulation // Icarus. 2019. vol. 321. pp. 1013–1025.
19. *Dillow C.* How it Would Work: Destroying an Incoming Killer Asteroid With a Nuclear Blast // Bonnier. 2012. URL: https://www.flightglobal.com/news/articles/ nasa-plans-armageddon-spacecraft-to-blast-asteroid-215924 (дата обращения: 15.08.2019).
20. *Ekpo S.C., George D.* A system engineering analysis of highly adaptive small satellites // IEEE Systems Journal. 2012. vol. 7. no. 4. pp. 642–648.
21. *Zhao L., Jia Y.* Decentralized adaptive attitude synchronization control for spacecraft formation using nonsingular fast terminal sliding mode // Nonlineat Dyan. 2014. vol. 78. no. 4. pp. 2779–2794.
22. *Du H., Li S.* Attitude synchronization control for a group of flexible spacecraft // Automatica. 2014. vol. 50. no. 2. pp. 646–651.
23. *Lan Q., Yang J., Guo L.* Finitic-time soft landing on asteroids using nonsingular terminal sliding mode control // Transactions of the Institute of Measurement and Control. 2014. vol. 36. no. 2. pp. 216–223.
24. *Zhan L., Lia Q., Chen G., Sun H.* Pre-impact trajectory planning for minimizing base attitude disturbance in spave manipulator systems for a capture task // Chinese Journal of Aeronautics. 2015. vol. 28. no. 4. pp. 1199–1208.
25. *Zou A.M.* Finite-time output feedback attitude tracking control for rigid spacecraft // IEEE Transactions on Control Systems Technology. 2014. vol. 22. no. 1. pp. 338–345.

26. *Zhao I., Jia Y.* Finite-time attitude stabilization for a class of stochastic spacecraft systems // IET Control Theory & Applications. 2015. vol. 9. no. 8. pp. 1320–1327.
27. *Шустов Б.М., Рыхлова Л.В.* Астероидно-кометная опасность: вчера, сегодня, завтра // М.: Физматлит. 2010. 384 с.
28. *Минаков Е.П., Соколов Б.В.* Исследование характеристик размещения и вариантов применения моноблочных стационарных наземных средств поражения астероидов // Труды СПИИРАН. 2016. Вып. 5(48). С. 181–196.
29. *Минаков Е.П., Соколов Б.В., Шалдаев С.Е.* Исследование характеристик и вариантов применения окололунной системы поражения астероидов // Труды СПИИРАН. 2017. Вып. 5(54). С. 106–129.
30. *Баринов К.Н., Бурдаев М.Н., Мамон П.А.* Динамика и принципы построения орбитальных систем космических аппаратов // М.: Машиностроение. 1975. 232 с.
31. *Ivanov D., Dolgui A., Sokolov B., Werner F.* Schedule robustness analysis with the help of attainable sets in continuous flow problem under capacity disruptions // International Journal of Production Research. 2016. vol. 54. no. 1. pp. 3397–3413.
32. *Ivanov D., Dolgui A., Sokolov B.* Robust dynamic schedule coordination control in the supply chain // Computers & Industrial Engineering. 2016. vol. 94. pp. 18–31.

Минаков Евгений Петрович — д-р техн. наук, профессор, профессор, кафедра оценивания эффективности, Военно-космическая академия им. А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: системный анализ, баллистическое обеспечение полетов космических аппаратов, эффективность применения космических комплексов и систем. Число научных публикаций — 150. ep.minakov12345@mail.ru; Ждановская, 13, 197082, Санкт-Петербург, Российская Федерация; р.т.: +7 (812) 552-6341.

Соколов Борис Владимирович — д-р техн. наук, профессор, Заслуженный деятель науки РФ, главный научный сотрудиник, лаборатория информационных технологий в системном анализе и моделировании, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук. Область научных интересов: разработка научных основ теории управления структурной динамикой сложных организационно-технических систем. Число научных публикаций — 650. sokol@iias.spb.su; 14 линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7(812) 328–01–03; факс: +7(812)328-44-50.

Шалдаев Сергей Евгеньевич — д-р техн. наук, доцент, начальник управления, Военно-космическая академия им. А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: радиолокация, радионавигация, построение и испытание сложных технических систем. Число научных публикаций — 100. 100.shs99@yandex.ru; ул. Ждановская, 13, 197082, Санкт-Петербург, Российская Федерация; р.т.: +7(812) 347-95-59.

Александров Максим Андреевич — канд. техн. наук, доцент, преподаватель, Военно-космическая академия им. А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: моделирование применения космических комплексов и систем. Число научных публикаций — 15. maks.aleksandrov.vka@mail.ru; ул. Ждановская, 13, 197082, Санкт-Петербург, Российская Федерация; р.т.: +7 (812) 552-6341.

Поддержка исследований. Исследование выполнено в рамках бюджетной темы №№0073–2019–0004 и Госзадания Министерства образования и науки РФ №2.3135.2017/4.6.

DOI 10.15622/sp.2019.18.6.1462-1490

E.P. MINAKOV, B.V. SOKOLOV, S.E. SHALDAEV, M.A. ALEKSANDROV

CALCULATION AND RESEARCH OF SPACE-TEMPORAL CHARACTERISTICS OF ATTACK ABRASES OF ASTEROIDS BY ORBITAL MEANS

Minakov E.P., Sokolov B.V., Shaldaev S.E., Aleksandrov M.A. **Calculation and Research of Space-Temporal Characteristics of Attack Abrases of Asteroids by Orbital Means.**

Abstract. A development of work to combat the asteroid hazard requires construction and study of areas of outer space in which moving spacecraft-interceptors can affect asteroids. In this paper, such areas are called attack lines, the spatio-temporal characteristics of which depend on the parameters of the asteroid's orbits and the phase coordinates of the nodal points. At these points the trajectory intersects the asteroids and the orbital planes of the spacecraft-interceptors. In the case of the impact of spacecraft-interceptors on asteroids at nodal points, the study of the spatio-temporal characteristics of the lines of attack, taking into account restrictions on the relative speeds between asteroids and spacecraft-interceptors, is of particular importance. Building and analyzing the corresponding zones of reverse reach are suggested.

In the article, the developed models include a simulation model, using which random angles between the projections of the velocity vectors of asteroids on a plane of the orbits of spacecraft-interceptors and the current directions on the hodographs of their velocity vectors at nodal points, as well as an analytical model for estimating the spatio-temporal characteristics of boundaries are simulated attacks of asteroids, including: the radii of their external and internal boundaries for certain values of the latitude arguments and arrival time of spacecraft-interceptors at modal points.

Testing these models and the corresponding characteristics of the attack lines were carries out during computational experiments on two cyclic modeling of the angles between the projections of the velocity vectors of asteroids on the plane of the orbits of interceptor spacecraft and the current directions on the hodographs of their velocity vectors at nodal points. The results obtained made it possible to verify and validate the developed models, on the basis of which a conclusion was drawn about the required degree of their applicability. In the paper also a procedure for estimating the parameters of attack lines, depending on the values of the arguments of the latitudes of interceptor spacecraft and their altitudes above the Earth's surface is proposed. At the same time, an approach is substantiated for estimating the spatio-temporal characteristics of the boundaries of attack of asteroids by spacecraft-interceptors for any inside the planar parameters of their orbits.

Keywords: Asteroid Danger, Spacecraft Interceptor, Detachable Part, Anchor Point, Line of Attack, Space-Time Characteristics.

Minakov Evgeniy Petrovich — Ph.D., Dr.Sci., Professor, Professor, Department of Effectiveness Evaluation, Mozhaisky Military Space Academy. Research interests: system analysis, provision of ballistic spacecraft, effectiveness of space complexes and systems. The number of publications — 150. ep.minakov12345@mail.ru; 13, Gdanovskaya, 197082, St. Petersburg, Russian Federation; office phone: +7 (812) 552-6341.

Sokolov Boris Vladimirovich — Ph.D., Dr.Sci., Professor, Honored scientist of Russian Federation, Chief Reseatrcxher, Laboratory of Information Technologies in Systems Analysis and Modeling, St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences. Research interests: development of research fundamentals for the control theory by structural dynamics of complex organizational-technical systems. The number of publications — 650. sokol@iias.spb.su; 39, 14th line V.O., 199178, St.Petersburg, Russian Federation; office phone: +7(812) 328–01–03; fax: +7(812)328-44-50.

Shaldaev Sergey Evgenjevich — Ph.D., Dr.Sci., Associate Professor, Head of Department, Mozhaisky Military Space Academy. Research interests: radiolocation, radio navigation, construction and testing of complex technical systems. The number of publications — 100. 100.shs99@yandex.ru; 13, Zhdanovskaya str., 197082, St. Petersburg, Russian Federation; office phone: +7(812) 347-95-59.

Aleksandrov Maksim Andreevich — Ph.D., Associate Professor, Lecturer, Mozhaisky Military Space Academy. Research interests: modeling the use of space systems and systems. The number of publications — 15. maks.aleksandrov.vka@mail.ru; 13, Zhdanovskaya str., 197082, St. Petersburg, Russian Federation; office phone: +7 (812) 552-6341.

Acknowledgements. This research is supported by state research 0073–2019–0004 and state order of the Ministry of Education and Science of the Russian Federation №2.3135.2017/4.6.

References

1. Ross I.M., Park S.Y., Porter S.E. Gravitational Effects of Earth in Optimizing Delta-V for Deflecting Earth-Crossing Asteroids. *Journal of Spacecraft and Rockets.* 2001. vol. 38. no. 5. pp. 759–764.
2. Asteroid Impact & Deflection Assessment (AIDA) collaboration. Available at: https://www.esa.int/Safety_Security/Hera/Asteroid_Impact_Deflection_Assessment_AIDA_collaboration (accessed: 15.08.2019).
3. Dwayne A. Giant bombs on giant rockets: Project Icarus. The Space Review. 2004.
4. Asteroid deflection mission seeks smashing ideas. Available at: http://www.esa.int/Enabling_Support/Operations/Asteroid_deflection_mission_seeks_smashing_ideas (accessed: 15.08.2019).
5. Powell C.S. Developing Early Warning Systems for Killer Asteroids. *Discover.* 2013. pp. 60–61.
6. Harper P. Earth will be hit by asteroid with 100% certainty – space experts warn. Daily Star. 2018.
7. Homer A. Earth Will Be Hit By An Asteroid With 100 Percent Certainty, Says Space-Watching Group B612. Inquisitr. 2018.
8. Stanley-Becker I. Stephen Hawking feared race of 'superhumans' able to manipulate their own DNA. Available at: https://www.washingtonpost.com/news/morning-mix/wp/2018/10/15/stephen-hawking-feared-race-of-superhumans-able-to-manipulate-their-own-dna/ (accessed: 15.08.2019).
9. Konstantinovskaya L.V. Prognoz I predotvrashchenie kometno-asteroidnoj opasnosti [Forecast and prevention of comet-asteroid hazard]. Available at: http://www.astronom2000.info/different/pipao/ (accessed: 11.09.2013). (In Russ.).
10. [Threats from spase]. *Populyarnaya Mekhanika — Popular Mechanics.* 2010. vol. 9(95). pp. 31–46. (In Russ.).
11. Nazirov R.R., E'jsmont N.A. [Gravitational maneuvers as a way to direct small asteroids to a meetings trajectory with dangerous undegraund objects]. *Kosmicheskie issledovaniya – Cosmic Research.* 2010. vol. 48. no. 5. pp. 1–6. (In Russ.).
12. Sedyh R. [Treats from space are growing]. *Voennoe obozrenie – Military Watch*. 2013. Available at: http://vpk-news.ru/ (accessed: 21.06.2017). (In Russ.).
13. Kolin K.K. [Global threats of civilization development in the XXI century]. *Strategicheskie prioritety – Strategic Priorities*. 2014. vol. 1. pp. 6–30. (In Russ.).
14. Baklanov O. et al. [From asteroid and space danger The Earth will be protected by the Citadel]. *Vozdushno-kosmicheskaya sfera – Aerospace Sphere Journal.* 2017. vol. 1. pp. 90–95. (In Russ.).
15. Billings L. Words matter: A call for responsible communication about asteroid impact hazards and plans for planetary defense. *Space Policy*. 2015. vol. 33. pp. 8–12.

16. Wall M. A Killer Asteroid Is Coming – We Don't Know When (So Let's Be Ready), Bill Nye Says. Available at: https://www.space.com/killer-asteroids-warning-bill-nye.html (accessed: 15.08.2019).
17. Johns H.U. Asteroids are stronger, harder to destroy than previously thought. Available at: https://www.sciencedaily.com/releases/2019/03/190304095926.htm (accessed: 15.08.2019).
18. Charles E.M., Ramesh K.T., Richardson D.C. A new hybrid framework for simulating hypervelocity asteroid impacts and gravitational reaccumulation. *Icarus*. 2019. vol. 321. pp. 1013–1025.
19. Dillow C. How it Would Work: Destroying an Incoming Killer Asteroid With a Nuclear Blast // Bonnier. 2012. Available at: https://www.flightglobal.com/news/articles/ nasa-plans-armageddon-spacecraft-to-blast-asteroid-215924 (accessed: 15.08.2019).
20. Ekpo S.C., George D. A system engineering analysis of highly adaptive small satellites. *IEEE Systems Journal.* 2012. vol. 7. no. 4. pp. 642–648.
21. Zhao L., Jia Y. Decentralized adaptive attitude synchronization control for spacecraft formation using nonsingular fast terminal sliding mode. *Nonlinear Dynamics.* 2014. vol. 78. no. 4. pp. 2779–2794.
22. Du H., Li S. Attitude synchronization control for a group of flexible spacecraft. *Automatica.* 2014. vol. 50. no. 2. pp. 646–651.
23. Lan Q., Yang J., Guo L. Finitic-time soft landing on asteroids using nonsingular terminal sliding mode control. *Transactions of the Institute of Measurement and Control.* 2014. vol. 36. no. 2. pp. 216–223.
24. Zhan L., Lia Q., Chen G., Sun H. Pre-impact trajectory planning for minimizing base attitude disturbance in spave manipulator systems for a capture task. *Chinese Journal of Aeronautics.* 2015. vol. 28. no. 4. pp. 1199–1208.
25. Zou A.M. Finite-time output feedback attitude tracking control for rigid spacecraft. *IEEE Transactions on Control Systems Technology.* 2014. vol. 22. no. 1. pp. 338–345.
26. Zhao I., Jia Y. Finite-time attitude stabilization for a class of stochastic spacecraft systems. *IET Control Theory & Applications*. 2015. vol. 9. no. 8. pp. 1320–1327.
27. Shustov B.M., Rikhlovoy L.V. *Asteroidno-kometnajaopasnost': vchera, segodnja, zavtra* [Asteroid-cometdanger: yesterday, today, tomorrow]. M.: Phyznatlit. 2010. 384 p. (In Russ.).
28. Minakov E.P., Sokolov B.V. [Investigation of the location and application of monoblock stationary ground-based asteroid weapons]. *Trudy SPIIRAN — SPIIRAS Proceedings*. 2016. vol. 5(48). pp. 181–196. (In Russ.).
29. Minakov E.P., Sokolov B.V., Shaldaev S.E. [Investigation of the characteristics and applications of the circumlunar system of asteroid damage]. *Trudy SPIIRAN —SPIIRAS Proceedings.* 2017. vol. 5(54). pp. 106–129. (In Russ.).
30. Barinov K.N., Burdaev M.N., Mamon P.A. *Dinamika i principy postroeniya orbital'nyh sistem kosmicheskih apparatov* [Dynamics and principles of construction of spacecraft orbital systems]. M.: Mashinostroenie. 2013. 90 p. (In Russ.).
31. Ivanov D., Dolgui A., Sokolov B., Werner F. Schedule robustness analysis with the help of attainable sets in continuous flow problem under capacity disruptions. *International Journal of Production Research*. 2016. vol. 54. no. 1. pp. 3397–3413.
32. Ivanov D., Dolgui A., Sokolov B. Robust dynamic schedule coordination control in the supply chain. Computers & Industrial Engineering. 2016. vol. 94. pp. 18–31.

УДК 519.6 **DOI** 10.15622/sp.2019.18.6.1491-1514

Ю.А. Бычков, Е.Б. Соловьева, С.В. Щербаков

АНАЛИТИЧЕСКИ-ЧИСЛЕННЫЙ АЛГОРИТМ ВЫЧИСЛЕНИЯ КОРНЕЙ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ С ЗАДАННЫМИ ПРЕДЕЛЬНЫМИ ПОГРЕШНОСТЯМИ

Бычков Ю.А., Соловьева Е.Б., Щербаков С.В. **Аналитически-численный алгоритм вычисления корней алгебраических уравнений с заданными предельными погрешностями.**

Аннотация. Предложен алгоритм вычисления приближенных значений корней алгебраических уравнений с заданными предельными абсолютными погрешностями. Математическую основу алгоритма составляет аналитически-численный метод решения нелинейных интегрально-дифференциальных уравнений с нестационарными коэффициентами. Аналитически-численный метод относится к классу одношаговых непрерывных методов переменного порядка с адаптивной процедурой выбора шага расчета, формализованной оценкой погрешности производимых вычислений на каждом шаге и погрешности, накапливаемой в ходе расчета. Предлагаемый алгоритм вычисления приближенных значений корней алгебраического уравнения с заданными предельными абсолютными погрешностями состоит из двух этапов. Результатом выполнения первого этапа служат числовые интервалы, содержащие неизвестные точные значения корней алгебраического уравнения. На втором этапе вычисляем приближенные значения этих корней с заданными предельными абсолютными погрешностями. В качестве примера использования предложенного алгоритма приведено нахождение корней алгебраического уравнения пятого порядка с тремя различными значениями предельной абсолютной погрешности.

На основе полученных результатов сделаны следующие выводы. Предложенный алгоритм позволяет выделить числовые интервалы, содержащие неизвестные точные значения корней. Знание этих интервалов дает возможность вычисления приближенных значений корней с любой заданной предельной абсолютной погрешностью. Результативность алгоритма, то есть гарантия достижения поставленной цели не зависит от выбора начальных условий. Алгоритм не итерационный, поэтому число шагов расчета, которое необходимо для выделения числового интервала, содержащего неизвестное точное значение какого-либо из корней алгебраического уравнения, всегда ограничено. Алгоритм при поиске определенного корня алгебраического уравнения вычислительно полностью автономен.

Ключевые слова: алгебраическое уравнение, нелинейное дифференциальное уравнение, аналитически-численный метод, приближенное значение корня.

1. Введение. Задача решения алгебраических уравнений обладает самостоятельным научным интересом и широким кругом применения. Она мотивирует разработку новых и совершенствование уже имеющихся алгоритмов ее решения, например итерационных методов высокого порядка сходимости [1-7] (в частности для нахождения кратных корней [8-16]), эволюционных [17] и адаптивных [18] алгоритмов, методов нелинейной регрессии [19, 20] и экстраполяции [21, 22]. Сравнительный анализ используемых численных методов решения алгебраических уравнений выявляет следующие их недостатки.

Во-первых, отсутствуют указания к выбору начального приближенного значения искомого корня алгебраического уравнения, успех

которого во многом определяет не только объем последующих вычислений, но и сходимость всей процедуры расчета.

Во-вторых, расчетные схемы методов, как правило, итерационные и не содержат критерия самоостанова, связанного с заданной предельной абсолютной погрешностью расчета приближенного значения корня алгебраического уравнения.

В-третьих, отсутствуют результативные процедуры оценок, возникающих в ходе расчета погрешностей. Это не позволяет контролировать степень удаления вычисляемых приближенных значений корня от его неизвестного точного значения. В результате о их близости можно судить лишь по уровню невязки алгебраического уравнения относительно вычисленного приближенного значения корня.

Предлагается свободный от указанных недостатков алгоритм вычисления с заданными предельными абсолютными погрешностями приближенных значений корней алгебраических уравнений.

Решаемое алгебраическое уравнение имеет следующий вид:

$$A(p) = p^N - \sum_{n=1}^{N} A^*_{N-n} p^{N-n} = \sum_{n=1}^{N_m} (p - \lambda_n)^{M_n} = 0, \qquad (1)$$

где N — порядок уравнения; A^*_{N-n} — коэффициенты уравнения; N_m — число корней с учетом их кратности; M_n — кратность корня λ_n.

В качестве математической основы предлагаемого алгоритма используется аналитически-численный метод решения обыкновенных нелинейных интегрально-дифференциальных уравнений с нестационарными коэффициентами [23, 24]. Существенное отличие предлагаемого метода от всех широко используемых методов в том, что он контролирует и оценивает абсолютную погрешность найденного приближенного значения корня алгебраического уравнения. Преимущества алгоритма заключаются в следующем. Статическая задача решения алгебраического уравнения эквивалентно преобразуется в динамическую задачу решения дифференциального уравнения. Эту динамическую задачу решает аналитически-численный метод, который в силу одного из своих основных достоинств позволяет найти не только приближенное значение корня, но и указать границы области, в которой находится его точное недосягаемое значение. Границами области можно управлять, что в результате позволяет найти значение корня с любой заданной точностью. Предлагаемый вычислительный алгоритм может быть реализован программно в виде некоторого универсального вычислительного инструмента. Однако это самостоятельная задача, ее постановка и решение находятся за рамками данной работы.

Решение нелинейных алгебраических уравнений с высокой точностью на основе аналитически-численный метода важно в прецизионной технике, где главной целью является точность обработки сигналов (информации), достигаемая, например, в пространстве траекторий. В качестве примера укажем синтез блоков управления для робототехнических систем, которые обеспечивают точность движения манипуляторов, согласованное перемещение конечностей роботов и так далее.

У аналитически-численного метода есть ограничения, например организация перехода через точки ветвления при поиске близких по числовым значениям корней может привести к некоторым алгоритмическим проблемам, определяющим необходимость реализации итерационных процедур расчета.

Решение поставленной задачи аналитически-численным методом связано с предварительным формированием для уравнения (1) его динамического аналога в виде известным образом соответствующего ему дифференциального уравнения. Далее для такого дифференциального уравнения формируем сопряженное ему дифференциальное уравнение, значения решений которого при определенном значении независимой переменной совпадают с корнями λ_n, n=1, 2, …, N уравнения (1). Этим предложениям отвечает последовательность преобразований, определяющих содержание предлагаемого алгоритма решения поставленной задачи.

Для примера рассмотрим наиболее простой случай, когда все корни уравнения (1) некратные и вещественные, то есть M_n=1, N_m=N и n=1, 2, …, N. Задачу поиска корней уравнения (1) в этом случае сопровождает рассмотрение однозначно определенной аналитической функции $x(t)$, которая эквивалентно описывает решаемую задачу:

$$x(t) = \sum_{n=0}^{N} A_n^* t^n,\ A_N^* = 1, t \in R; \tag{2}$$

$$x(t)\big|_{t=\lambda_n} = x\left(\lambda_n\right) = 0, \quad n = 1, 2, \ldots, N. \tag{3}$$

Для однозначной аналитической функции $x(t)$, описываемой выражением (2), существует многозначная обратная функция $t(x)$. Переход к многозначной функции $t(x)$ сводит задачу поиска корней λ_n, n=1, 2, …, N уравнения (1) к вычислению N значений многозначной функции $t(x)$ при значении независимой переменно x=0. Обратная для аналитической функции $x(t)$ функция $t(x)$ определяет по отношению к корням λ_n, n=1, 2, …, N уравненияа (1) следующие равенства, эквивалентные равенствам (3):

$$t(x)\big|_{x=0} = t(0) = \lambda_n, \quad n = 1, 2, \ldots, N. \tag{4}$$

Процедура построения функции $t(x)$, обратной для известной аналитической функции $x(t)$, в общем случае пока не доступна формализации и алгоритмизации. Возможен, однако, следующий прием.

Доступно определению обыкновенное дифференциальное уравнение, решением которого служит описываемая выражением (2) однозначная аналитическая функция $x(t)$. Для этого дифференциального уравнения можно сформировать сопряженное ему обыкновенное дифференциальное уравнение, решением которого служит функция $t(x)$, обратная функции $x(t)$ [25, 26]. Взаимосвязь между функциями $x(t)$ и $t(x)$ в отношении корней λ_n, n=1, 2, …, N уравнения (1), описываемая равенствами (3) и (4), определяет логику и математическое содержание предлагаемого алгоритма вычисления приближенных значений этих корней. Основу алгоритма составляет направленная замена исходной алгебраической задачи новой дифференциальной задачей, решение $t(x)$ которой удовлетворяет соотношениям (4). Несмотря на кажущееся усложнение задачи, предлагаемая замена позволяет достичь качественно новых результатов в сравнении с широко используемыми численными методами решения алгебраических уравнений.

2. Краткая характеристика аналитически-численного метода. Аналитически-численный метод относится к классу одношаговых непрерывных методов переменного порядка с адаптивной процедурой выбора шага расчета, формализованной оценкой погрешности производимых вычислений на каждом шаге и погрешности, накапливаемой в ходе расчета [23, 24]. Метод предназначен для поиска решений следующих обыкновенных интегрально-дифференциальных уравнений с нестационарными коэффициентами:

$$\mathbf{A}(D, D^{-1})\mathbf{x}(t) = \mathbf{G}(D, D^{-1})\mathbf{f}(t) + \mathbf{H}(\mathbf{x}, \mathbf{f}, t), \qquad (5)$$

где D — оператор обобщенного дифференцирования по независимой переменной t, D^{-1} — оператор интегрирования с переменным верхним пределом t и нижним пределом, который определяет левая граница текущего интервала интегрирования; $\mathbf{A}(D, D^{-1})$ и $\mathbf{G}(D, D^{-1})$ — квадратная, размером $L_x \times L_x$, и прямоугольная, размером $L_x \times L_f$, матрицы с полиномиальными от D и D^{-1} элементами; $\mathbf{x}(t)$ и $\mathbf{f}(t)$ — матрицы-столбцы искомых решений и внешних воздействий (заданных аналитических функций); $\mathbf{H}(x, f, t)$ — матрица-столбец со строками в виде сумм произведений, сомножители которых есть нестационарные коэффициенты, классические производные любого порядка и интегралы любой кратности, начиная с нулевого (нулевой), от искомых решений и (или) внешних воздействий, в произвольных дробно-рациональных степенях.

Расчетная схема аналитически-численного метода поиска в заданном ограниченном интервале времени $[t_0; T]$ решений $x_r(t)$, $r=1, 2, \ldots, L_x$ уравнения (5) при заданных предначальных условиях, основываясь на неравномерной дискретизации независимой переменной t, в каждом интервале расчета состоит из двух частей: аналитической и численной.

Математическую основу аналитической части метода составляют аппараты обобщенных функций, интегрального преобразования Лапласа и степенных рядов [27-32]. Использование аппарата обобщенных функций обусловлено возможным наличием в искомых решениях уравнения (5) дифференцируемых разрывов первого рода. Из этого вытекает необходимость выделения и последующего описания в решении сингулярной и регулярной составляющих. Сингулярная составляющая искомого решения по определению имеет замкнутую форму описания в виде суммы импульсных функций соответствующих порядков и в каких-либо иных представлениях не нуждается. Регулярная составляющая решения, имея произвольную форму описания, нуждается в унификации своего представления с учетом корректного отражения в таком представлении характерных свойств и качественных особенностей искомых решений уравнения (5).

Для унификации математического базиса расчетной схемы метода описание регулярной составляющей должно отвечать следующим требованиям:

– приводить описание нелинейной части $\mathbf{H}(x, f, t)$ уравнения (5) к единому и пригодному для последующих преобразований виду;

– выявлять существующие в решении особые точки;

– описывать получаемое приближенное решение в математической и информационной взаимосвязи с существующим и неизвестным точным решением уравнения (5).

Желательно также, чтобы унифицированное представление регулярной составляющей решения было не только формальным, но и информационно содержательным.

Перечисленным требованиям отвечает представление регулярной составляющей искомого решения $x_l(t)$, $l \in [1; L_x]$ в форме функционально-степенного ряда [23, 24]. Во-первых, подстановка в уравнение (5) вместо искомых решений их регулярных составляющих, описываемых соответствующими степенными рядами, тождественно изменяет произвольное описание элементов матрицы нелинейной части $\mathbf{H}(x, f, t)$ этого уравнения на определенное и унифицированное, на описание по целым степеням независимой переменной t. Во-вторых, знание радиуса сходимости степенного ряда для регулярной составляющей искомого решения обусловливает выявление особой точки этого решения,

включая однозначную идентификацию ее местоположения по независимой переменной *t*. В-третьих, в пределах радиуса сходимости степенной ряд превращается в ряд Тейлора и сходится абсолютно и равномерно к разложенной в него аналитической функции, то есть к регулярной составляющей решения. Заменяя при практических вычислениях ряд Тейлора его частичной суммой, можно сформировать описание возникающей при этом вычислительной погрешности и затем организовать процедуру ее оценки. Наконец, описание регулярной составляющей решения рядом Тейлора не только формально, но и информационно содержательно ввиду известного математического содержания его коэффициентов.

Отметим, что корректное формирование сингулярной составляющей решения и правильное вычисление начальных значений регулярной составляющей решения в точках ее разрывов первого рода по известным предначальным значениям возможно с помощью обобщенного интегрального преобразования Лапласа [23, 24]. Использование интегрального преобразования Лапласа для поиска решения уравнения (5) становится допустимым после описания элементов матрицы $\mathbf{H}(x, f, t)$ нелинейной части уравнения (5) степенными рядами относительно независимой переменной t.

Процедура аналитической части аналитически-численного метода на основе указанного математического аппарата сводится к следующему. При известных для текущего интервала расчета $[t_{k-1};\ t_k]$ предначальных условиях входящие в матрицу $\mathbf{H}(x, f, t)$ нелинейной части уравнения (5) искомые решения замещаем регулярными составляющими, которые формально описываем соответствующими степенными рядами. Коэффициенты этих степенных рядов неизвестны. Описанные таким образом искомые решения, а также при необходимости предварительно разложенные в степенные ряды нестационарные коэффициенты и внешние воздействия подставляем в соответствующие описания элементов матрицы $\mathbf{H}(x, f, t)$. Подстановку выполняем так, чтобы в матрице $\mathbf{A}(D, D^{-1})$ линейной части уравнении (5) были выделены по крайней мере старшие операторы дифференцирования по всем искомым решениям. В результате описание уравнения (5) относительно точек разложения искомых решений, нестационарных коэффициентов и внешних воздействий в соответствующие степенные ряды будет преобразовано к новой форме. Эта новая форма описания сопряжена с изменением только характера описания элементов матрицы $\mathbf{H}(x, f, t)$. Результатом такого изменения служит замена описаний всевозможных нелинейных и нестационарных зависимостей унифицированным описанием в виде степенных рядов относительно времени t. Вследствие этого

уравнение (5) будет приведено к виду, пригодному для применения обобщенного интегрального преобразования Лапласа. Преобразовав сформированное уравнение по Лапласу, получим линейную относительно изображений искомых решений систему алгебраических уравнений. Решения этой системы уравнений определяем по формуле Крамера и раскладываем их в соответствующие ряды Лорана в окрестности бесконечно удаленной точки. Коэффициенты главных и правильных частей этих рядов Лорана вычисляем с помощью предложенных в [23, 24] формул. Переведя сформированные ряды Лорана во временную область, для искомых решений $x_r(t)$, r=1, 2, …, L_x, уравнения (5) получим описания в виде обобщенных функций. Коэффициенты сингулярных и регулярных составляющих этих решений теперь известны. Регулярные составляющие решений в правой полуокрестности точек с абсциссой, совпадающей с началом текущего интервала расчета $[t_{k-1}; t_k]$, будут описаны соответствующими степенными рядами.

Итак, результатом выполнения в текущем интервале расчета $[t_{k-1;}\, t_k]$ аналитической части метода для искомого решения $x_l(t)$, $l \in [1; L_x]$ уравнения (5) служит следующее его описание [23, 24]:

$$x_l(t) = x_l^-(t) + x_l^+(t) = \sum_{j=0}^{-J_l} S_{l.j}\ \delta_j(t) + \sum_{i=0}^{\infty} R_{l.i}\ t^i / i!, \qquad (6)$$

где $x_l^-(t)$ и $x_l^+(t)$ — сингулярная и регулярная составляющие искомого решения $x_l(t)$; $\delta_j(t)$ — определенные в точке, начальной для текущего интервала расчета, импульсные функции от нулевого до J_l-го порядка включительно; $S_{l.j}$ — весовые коэффициенты импульсных функций; $R_{l.j}$ — коэффициенты разложения регулярной составляющей решения $x_l^+(t)$ в степенной ряд в правой полуокрестности начальной для текущего интервала расчета точке.

Сингулярную составляющую $x_l^-(t)$ искомого решения $x_l(t)$, $l \in [1; L_x]$, если она существует, определяем в аналитической части метода. Погрешность вычисления ее коэффициентов $S_{l.j}$ определяет погрешность расчета предначальных условий для текущего шага расчета [23, 24]. Регулярную составляющую решения $x_l^+(t)$ после выполнения аналитической части метода описывает следующий степенной ряд:

$$x_l^+(t) = \sum_{i=0}^{\infty} R_{l.i} t^i / i!,\ l \in [1, L_x]. \qquad (7)$$

Для вычисления в дискретные моменты времени $t=t_k$, $t_k=t_{k-1}+h_k$, k=1, 2, …, K из заданного интервала исследования $[t_0; T]$ приближенных

значений $x_l^+(t_k; I_l)$ регулярной составляющей x_l^+ искомого решения $x_l(t)$, $l \in [1; L_x]$ предназначена численная часть метода. Процедуру численной части метода в текущем интервале расчета $[t_{k-1}; t_k]$ начинаем с выбора величины шага расчета $h=h_k$, которая не должна превышать радиусов сходимости сформированных для описания регулярных составляющих решений $x_r^+(t)$ степенных рядов (7) при $l=r$.

Выбор величины текущего шага расчета сопровождает ряд взаимосвязанных условий. Во-первых, величина шага расчета должна обеспечивать численную устойчивость всей вычислительной процедуры, обусловливая устойчивое накопление погрешности в ходе расчета. Во-вторых, непосредственно влияя на уровень погрешности производимых вычислений, длина шага должна обеспечивать соответствие заданному предельному показателю возникающей при расчетах погрешности. В-третьих, регламентируя уровень неравномерной дискретизации независимой переменной t, величина шага должна соответствовать скорости фактического изменения динамических показателей регулярных составляющих $x_r^+(t)$ искомых решений $x_r(t)$, $r=1, 2, \ldots, L_x$, обеспечивая получение приближенных значений этих составляющих решений на оптимальной сетке по времени. В качестве таких динамических показателей регулярных составляющих решений $x_r^+(t)$ логично рассматривать коэффициенты сформированных для их описания степенных рядов (7), при $l=r$.

Процедура выбора величины шага расчета, удовлетворяющей приведенным условиям, описана в [23, 24]. В основе этой процедуры лежит исследование сходимости числовых мажорант степенных рядов (7), $l=r$, сформированных для регулярных составляющих решений $x_r^+(t)$. Условия сходимости этих числовых мажорант регламентируют ограничения величины шага расчета, которые обеспечивают непревышение интервала сходимости указанных степенных рядов. Выбранная таким образом величина шага, гарантируя сходимость степенных рядов (7), $l=r$ к разложенным в них аналитическим функциям $x_r^+(t)$, определяет протяженность интервала времени, в пределах которого эти составляющие решений существуют и могут быть описаны рядами Тейлора. При этом все имеющиеся в регулярных составляющих искомых решений разрывы второго рода будут гарантированно выявлены, и обеспечены условия их последующего корректного выделения [23, 24].

Дополняя сказанное, отметим, что выбранная величина шага расчета h обеспечивает необходимое согласование показателей естественной динамики регулярных составляющих $x_r^+(t)$ искомых решений $x_r(t)$,

r=1, 2, …, L_x. Точное числовое значение регулярной составляющей решения $x_l^+(t)$ при выбранной величине шага расчета h определяет следующий, полученный из степенного ряда (7) при t=h, числовой ряд:

$$x_l^+(h) = \sum_{i=0}^{\infty} R_{l.i} h^i / i!,\ l \in [1, L_x]. \quad (8)$$

При практических вычислениях числовой ряд (8) замещаем частичной суммой его первых I_l членов. Такая частичная сумма, определяя в результате выполнения текущего шага расчета h приближенное значение $x_l^+(h;\ I_l)$ регулярной составляющей $x_l^+(t)$ искомого решения $x_l(t)$, $l \in [1; L_x]$ имеет следующее описание:

$$x_l^+(h; I_l) = \sum_{i=0}^{I_l} R_{l.i} h^i / i!. \quad (9)$$

Заметим, что выбором необходимого значения порядка I_l частичной суммы, оценивая модуль остаточного члена числового ряда (7), t=h можно контролировать погрешность расчета.

Остаточный член числового ряда (7), t=h при замене на текущем шаге расчета h=h_k этого числового ряда частичной суммой (9) определяет локальную погрешность расчета. Эта погрешность явным образом зависит от величины текущего шага расчета h=h_k и значения I_l. Выбранная величина h=h_k позволяет составить формулы для вычисления верхней оценки $|\Delta x_l^+(h; I_l)|$ абсолютной локальной погрешности расчета. Эти формулы приведены в работах [23, 24]. Знание верхней оценки $|\Delta x_l^+(h; I_l)|$ позволяет после выполнения текущего шага расчета h=h_k выделить границы числового интервала, содержащего неизвестное точное значение $x_l^+(h)$ суммы $x_l^+(t)$ числового ряда (7) при t=h. Этот числовой интервал описывает следующее двойное неравенство:

$$x_l^+(h; I_l) - \left|\Delta x_l^+(h; I_l)\right| \le x_l^+(h) \le x_l^+(h; I_l) + \left|\Delta x_l^+(h; I_l)\right|,\ h \in [t_0; T]. \quad (10)$$

Ограниченность числовых показателей двойного неравенства (8), отражая устойчивость процесса накопления погрешности в пределах шага, позволяет организовать, начиная со второго шага, процедуру определения оценки полной (накопленной) погрешности расчета. Рекуррентная формула для вычисления верхней оценки $|\Delta x_l^+(t_k; I_l)|$ абсолютной полной погрешности расчета приближенного значения $x_l^+(t_k; I_l)$ регулярной составляющей $x_l^+(t)$ искомого решения $x_l(t)$, $l \in [1;\ L_x]$ приведена в работах [23, 24]. Знание верхней

оценки $|\Delta x_l^+(t_k;\ I_l)|$ позволяет определить границы числового интервала, содержащего неизвестное точное значение $x_l^+(t_k)$ регулярной составляющей решения $x_l^+(t)$ в дискретный момент времени $t=t_k$. Этот числовой интервал описывает следующее двойное неравенство:

$$x_l^+\left(t_k;I_l\right)-\left|\Delta x_l^+\left(t_k;I_l\right)\right|\le x_l^+\left(t_k\right)\le x_l^+\left(t_k;I_l\right)+\left|\Delta x_l^+\left(t_k;I_l\right)\right|. \qquad (11)$$

Выполнив текущий шаг расчета $h=h_k$ ось ординат смещаем вправо на его длину. После этого описанную вычислительную процедуру аналитически-численного метода повторяем, начиная с его аналитической части. В качестве предначальных условий для следующего шага расчета принимаем вычисленные на предыдущем шаге приближенные значения регулярных составляющих решений и их производных.

Вычисленные с оптимальным шагом дискретизации независимой переменной t приближенные значения $x_l^+(t_k;\ I_l)$, $l=r$ регулярных составляющих $x_r^+(t)$ искомых решений $x_r(t)$, $r=1,\ 2,\ \ldots,\ L_x$ и сформированные двойные неравенства (11) $l=r$ образуют итоговые числовые показатели расчета аналитически-численным методом. Двойные неравенства (11) $l=r$ описывают качественное отличие аналитически-численного метода от известных численных методов решения обыкновенных нелинейных неавтономных дифференциальных уравнений и обосновывают выбор этого метода для решения поставленной задачи.

3. Алгоритм вычисления приближенных значений корней алгебраического уравнения с заданной предельной абсолютной погрешностью. Предлагаемый алгоритм решения поставленной задачи состоит из двух этапов. Результатом выполнения первого этапа служат числовые интервалы, содержащие неизвестные точные значения λ_n корней алгебраического уравнения. На втором этапе вычисляем приближенные значения этих корней с заданными предельными абсолютными погрешностями $\varepsilon(\lambda)=[\varepsilon_1(\lambda)\ \varepsilon_2(\lambda)\ \ldots\ \varepsilon_N(\lambda)]$.

Первый этап. Выделение числовых интервалов, содержащих точные значения корней алгебраического уравнения.

Дифференциальное уравнение, решением которого служит функция $x(t)$, описываемая выражением (2) и удовлетворяющая соотношениям (3), таково:

$$\begin{aligned}&\dot{x}(t)=\sum_{n=1}^{N}a_n\,t^{n-1};\\&\dot{t}=1,\end{aligned} \qquad (12)$$

где $a_n=nA_n^*$; $a_N=N$; $x(t)\big|_{t=t_0}=x(t_0)=t_0$; $\dot{x}(t)$ — производная по переменной времени t.

Нелинейное дифференциальное уравнение, сопряженное дифференциальному уравнению (12), решением которого служит функция $t(x)$, обратная для функции $x(t)$, таково [25, 26]:

$$\frac{dt(x)}{dx} = \left(\sum_{n=1}^{N} a_n t^{n-1}(x) \right)^{-1}; \qquad (13)$$
$$x' = \mp 1.$$

где x' — производная по переменной x.

Решением системы нелинейных дифференциальных уравнений (13) при заданном начальном условии $t(x_0)=t_0$ служит многозначная функция $t(x)$, обратная для известной функции $x(t)$. Отметим, что в зависимости от выбора начальных условий $t(x_0)=t_0$ искомым решением $t(x)$ является одна из N возможных ветвей многозначной функции $t(x)$.

Знак начального значения независимой переменной $x=x_0$, $x_0=x(t_0)$ устанавливает знак шага расчета h_x при поиске решения системы нелинейных дифференциальных уравнений (13) аналитически-численным методом. Расчет осуществляем от начального значения независимой переменной $x=x_0$ до конечного $x=0$, поэтому правая часть второго из уравнений этой системы содержит знаки «$\mp$», указывая на возможность положительного либо отрицательного шага расчета h_x. Выделение числовых интервалов, каждый из которых содержит неизвестное точное значение λ_n, $n=1, 2, ..., N$ одного из корней уравнения (1), сопряжено с упорядочением задания начальных условий $x=x_0$, $x_0=x(t_0)$ для направленного перебора N ветвей неоднозначного решения $t(x)$ системы нелинейных дифференциальных уравнений (13). Для решения такой задачи необходимы: знание точек ветвления многозначной функции $t(x)$; организация корректного перехода через эти точки в ходе пошагового расчета [23, 24]. Используя известные формулы Коши, Монтеня, Кармайкла и Мейсона [6], можно вычислить верхнюю $|\lambda_{su}|$ и нижнюю $|\lambda_{in}|$ оценки абсолютных значений λ_n, $n=1, 2, ..., N$ корней уравнения (1). Эти оценки определяют границы числового интервала, содержащего неизвестные точные значения λ_n, $n=1, 2, ..., N$ всех корней уравнения (1). Чтобы сократить объем вычислительной работы, решение поставленной задачи целесообразно начать с максимального или минимального по модулю корня. Для этого из двух возможных начальных условий $t_0=\mp|\lambda_{su}|$ или $t_0=\mp|\lambda_{in}|$ выбираем то, которое определяет минимальное по модулю начальное значение независимой переменной $x=x_0$, $x_0=x(t_0)$ вычисленное по формуле (2).

Выбрав начальное условие $t_0=t(x_0)$, $x_0=x(t_0)$, в зависимости от знака этого условия, с положительным или отрицательным шагом расчета h_x приступаем к поиску решения $t(x)$ системы нелинейных дифференциальных уравнений (13) аналитически-численным методом. В результате для искомого решения $t(x)$ при $x=0$ выделяем числовой интервал, содержащий согласно равенствам (4) неизвестное точное значение максимального или минимального по модулю корня уравнения (1).

Выделенный числовой интервал с учетом принятых обозначений описывает аналогичное (11) следующее двойное неравенство:

$$t\left(0;I_t\right)-\left|\Delta t\left(0;I_t\right)\right|\le t(0)\le t\left(0;I_t\right)+\left|\Delta t\left(0;I_t\right)\right|, \tag{14}$$

где I_t — порядок полинома Тейлора, определяющего на шаге расчета h_x приближенное значение $t(x;\, I_t)$, $x=0$ искомого решения $t(x)$ системы нелинейных дифференциальных уравнений (13).

Двойное неравенство (14) содержит первый из корней уравнения (1), максимальный или минимальный по модулю. Его сопровождают следующие пояснения: $t(0;\, I_t)=\lambda_1(0;\, I_1)$ — приближенное значение первого, максимального или минимального по модулю корня уравнения (1); $|\Delta t(0;\, I_t)|=|\Delta\lambda_1(0;\, I_1)|$ — верхняя оценка абсолютной полной погрешности расчета приближенного значения $t(0;\, I_t)=\lambda_1(0;\, I_1)$ этого корня уравнения (1). Во вновь введенном обозначении двойное неравенство (14) приобретает следующий вид:

$$\lambda_1\left(0;I_1\right)-\left|\Delta\lambda_1\left(0;I_1\right)\right|\le\lambda_1\le\lambda_1\left(0;I_1\right)+\left|\Delta\lambda_1\left(0;I_1\right)\right|. \tag{15}$$

Получение двойного неравенства, аналогичного (15), для следующего второго корня $\lambda_2(0;\, I_2)$ уравнения (1) начинаем с выбора нового начального условия $t_0=t(x_0)$, $x_0=x(t_0)$. Оно должно принадлежать другой ветви многозначной функции $t(x)$. Для этого необходимо следующим образом определить точку ветвления функции $t(x)$.

Задав начальное условие $t_0=\lambda_1(0;\, I_1)$, $x_0=x(t_0)$, приступаем к поиску решения $t(x)$ уравнения (13) аналитически-численным методом, сохраняя тот знак шага расчета h_x, который был взят для составления двойного неравенства (14). Численный расчет продолжаем до значения независимой переменной $x=x^*$, при достижении которой выполняются следующие предельные соотношения:

$$\lim_{x\to x^*}\left|h_x\right|=0;\quad \lim_{x\to x^*}\frac{dt\left(x;I_t\right)}{dx}=\infty. \tag{16}$$

Символы «0» и «∞» в (16) определяются используемым вычислительным комплексом. Так как в точке ветвления претерпевает разрыв второго рода первая производная от многозначной функции $t(x)$, а величина шага расчета h_x не превосходит радиуса сходимости степенного ряда для искомого решения $t(x)$ уравнения (13), то выполнение предельных соотношений (16) информирует о приближении по x к точке ветвления многозначной функции $t(x)$ [23, 24]. По аналогии с двойным неравенством (14) в точке с абсциссой $x=x^*$, находящейся в непосредственной близости от абсциссы точки ветвления искомого решения $t(x)$, числовой интервал, который содержит неизвестное точное значение этого решения, описывает с учетом принятых обозначений следующее двойное неравенство:

$$t\left(x^*;I_t\right)-\left|\Delta t\left(x^*;I_t\right)\right|\le t\left(x^*\right)\le t\left(x^*;I_t\right)+\left|\Delta t\left(x^*;I_t\right)\right|. \qquad (17)$$

Новое начальное условие, принадлежащее другой ветви многозначной функции $t(x)$ и необходимое для выделения числового интервала, содержащего неизвестное точное значение следующего второго корня $\lambda_2(0; I_2)$ уравнения (1), вычисляем так:

$$t_0=t\left(x^*;I_t\right)\pm\Delta t;\quad x_0=x\left(t_0\right), \qquad (18)$$

где Δt — некоторая произвольная малая величина.

Выбор знака «+» или «–» в первом неравенстве (16) регламентирует условие сохранения того же характера изменения функции $t(x)$, который был при формировании двойного неравенства (17).

Задав в соответствии с равенствами (18) новое начальное условие и определив в зависимости от знака начального значения независимой переменной $x=x_0$ знак шага расчета h_x, приступаем к поиску решения $t(x)$ уравнения (11) в интервале $[x_0; 0]$ аналитически-численным методом. В результате при значении независимой переменной $x=0$ выделяем числовой интервал, содержащий согласно равенствам (4) неизвестное точное значение следующего второго корня $\lambda_2(0; I_2)$ уравнения (1). Выделенный числовой интервал описывает двойное неравенство (15) при соответствующей замене индекса «1» на индекс «2».

Далее описанный вычислительный алгоритм повторяем, последовательно выделяя числовые интервалы, которые содержат неизвестные точные значения λ_n, $n=1, 2, \ldots, N$ остальных корней уравнения (1).

Завершая изложение вычислительного алгоритма первого этапа, заметим: если при переходе через какую-либо точку ветвления многозначной функции $t(x)$ величина Δt окажется неконтролируемо больше,

чем интервал по t между соседними корнями уравнения (1), то общее число выделенных числовых интервалов окажется меньше N. Тогда, задавая меньшие значения Δt в равенствах (18), надо повторить расчеты с целью выделения остальных числовых интервалов.

Второй этап. Вычисление приближенных значений корней λ_n, n=1, 2, …, N *алгебраического уравнения* (1) *с заданной предельной абсолютной погрешностью* $\varepsilon(\lambda)=[\varepsilon_1(\lambda)\ \varepsilon_2(\lambda)\ \dots\ \varepsilon_N(\lambda)]$.

Для первого корня известен числовой интервал, описываемый двойным неравенством (15) и содержащий его неизвестное точное значение. Выбрав принадлежащее двойному неравенству (15) приближенное значение $\lambda_1^{[1]}(0;\ I_1)$ первого корня $\lambda_1(0;\ I_1)$ и приняв его в качестве начального условия $t_0=\lambda_1^{[1]}(0;\ I_1)$, $x_0=x(t_0)$, в интервале независимой переменной $[x_0;\ 0]$ приступаем к поиску решения $t(x)$ уравнения (13) аналитически-численным методом. Знак шага расчета h_x определяет знак начального значения $x_0=x(t_0)$ независимой переменной x, обеспечивая достижение этой переменной ее конечного значения x=0. Результатом расчета при x=0 служит числовой интервал, описываемый двойным неравенством (14) и содержащий неизвестное точное значение первого корня $\lambda_1(0;\ I_1)$ уравнения (1). Протяженность вновь выделенного числового интервала вследствие уменьшения числа шагов, необходимых для расчета в интервале $[x_0;\ 0]$, будет меньше размера числового интервала, выделенного после выполнения первого этапа алгоритма. Выбрав затем новое приближенное значение $\lambda_1^{[2]}(0;\ I_1)$ первого корня $\lambda_1(0;\ I_1)$, который принадлежит новому выделенному числовому интервалу, определяем новое начальное условие $t_0=\lambda_1^{[2]}(0;\ I_1)$, $x_0=x(t_0)$ и повторяем расчет в интервале $[x_0;\ 0]$. Расчет повторяем до тех пор, пока протяженность требуемого интервала расчета $[x_0;\ 0]$ по x не станет меньше, чем длина первого и единственного шага h_x. Как следует из двойного неравенства (10), при выполнении единственного шага расчета h_x выбором порядка I_t полинома Тейлора, используемого для вычисления приближенного значения $t(x;\ I_t)$, x=0, искомого решения $t(x)$ всегда можно добиться выполнения следующего двойного неравенства:

$$\left|t(0)-t\left(0;I_t\right)\right|\le\left|\Delta t\left(h_x;I_t\right)\right|\le\varepsilon_1(\lambda), \tag{19}$$

где $|\Delta t(h_x;\ I_t)|$ — верхняя оценка абсолютной локальной погрешности расчета, вычисленная с учетом принятых обозначений по формулам, приведенным в работах [23, 24]; $\varepsilon_1(\lambda)$ — заданная предельная абсолютная погрешность вычисления приближенного значения $t(0;\ I_1)=\lambda_1(0;\ I_1)$ первого корня λ_1 алгебраического уравнения (1).

Из двойного неравенства (19) с учетом обозначений, сопровождающих переход от двойного неравенства (14) к двойному неравенству (15), следует:

$$\left|\lambda_1 - \lambda_1\left(0; I_1\right)\right| \le \varepsilon_1(\lambda). \tag{20}$$

Неравенство (20) описывает результат вычисления приближенного значения $t(0; I_t)=\lambda_1(0; I_1)$ первого корня λ_1 алгебраического уравнения (1) с заданной предельной абсолютной погрешностью $\varepsilon_1(\lambda)$. Для остальных корней уравнения (1) вычислительный алгоритм решения поставленной задачи аналогичен.

Дополняя сказанное, отметим особенности решения поставленной задачи в тех случаях, когда среди корней уравнения (1) есть кратные и (или) комплексные сопряженные.

Если среди корней уравнения (1) есть кратные, то исходное комбинированное описание (2), (3) задачи решения уравнения (1) предварительно преобразуем к следующему необходимому виду:

$$y(t) = \frac{x(t)}{\dot{x}(t)} = \frac{\sum\limits_{n=0}^{N} A_n^* t^n}{\sum\limits_{n=1}^{N} a_n t^{n-1}}, \quad t \in R; \tag{21}$$

$$y(t)\big|_{t=\lambda_n} = y\left(\lambda_n\right) = 0, \quad n = 1, 2, \ldots, N_{mu}. \tag{22}$$

Функция $y(t)$, описываемая двойным равенством (21), такова, что сформированное на ее основе алгебраическое уравнение $y(t)=0$ имеет те же корни, что и исходное уравнение (1), но кратность этих корней равна единице [23, 24]. Используя описанный выше вычислительный алгоритм для (21), (22) решаем задачу вычисления с заданной предельной абсолютной погрешностью приближенных значений корней уравнения (1).

Если среди искомых корней уравнения (1) есть комплексные сопряженные, то подстановкой в выражение (2) вместо вещественной переменной t комплексной переменной $s=u+jv$ взамен исходной однозначной аналитической функции $x(t)$ рассматриваем функцию $z(s)=b(u,v) + jg(u,v)$, где $b(u,v)$, $g(u,v)$ — функции, образующие вещественную и мнимую части переменной $z(s)$ соответственно. Решение поставленной задачи в этом случае сводится к решению следующей системы алгебраических уравнений:

$$b(u,v) = 0; \quad g(u,v) = 0. \tag{23}$$

Вектор-функция $\mathbf{s}(z)=[u(b,\mathrm{g})\ v(b,\mathrm{g})]^T$, обратная для исходной вектор-функции, двумя компонентами которой служат вещественная и мнимая части функции $z(s)=b(u,v)+jg(u,v)$, является решением двух систем обыкновенных нелинейных дифференциальных уравнений, которые формируем на основе системы алгебраических уравнений (23) в соответствии с алгоритмом, изложенным в работах [23, 24]. Для поиска решений этих двух систем дифференциальных уравнений, которые при $b=0$ и $g=0$ совпадают с решениями системы алгебраических уравнений (23), используем описанный выше вычислительный алгоритм.

4. Пример расчета. Вычисление с заданной предельной абсолютной погрешностью $\varepsilon(\lambda)=1\cdot10^{-3}$; $1\cdot10^{-6}$; $1\cdot10^{-12}$ приближенных значений корней алгебраического уравнения пятой степени.

Задача решения заданного алгебраического уравнения в форме (2), (3) имеет следующее описание:

$$x(t)=t^5+15t^4+83t^3+225t^2+274t+120; \tag{24}$$

$$x(t)\big|_{t=\lambda_n}=x\left(\lambda_n\right)=0,\quad n=\overline{1,5}. \tag{25}$$

Первый этап. Нахождение числовых интервалов, содержащих точные значения корней алгебраического уравнения (24).

В форме (11) система обыкновенных нелинейных дифференциальных уравнений, решением которых служит функция $t(x)$, обратная известной функции $x(t)$ и описываемой выражением (24), имеет следующий вид:

$$\frac{dt(x)}{dx}=\left(5t^4(x)+60t^3(x)+249t^2(x)+450t(x)+274\right)^{-1};\qquad x'=\mp1. \tag{26}$$

Используя известные результаты алгебры, установили, что искомые корни уравнения (25) вещественные, отрицательные и некратные [27-32]. Вычисленные по формуле Коши верхняя $|\lambda_{su}|$ и нижняя $|\lambda_{in}|$ оценки модулей этих неизвестных корней λ_n таковы:

$$\left|\lambda_{su}\right|=1+\max\{120;274;225;83;15\}=275;$$

$$\left|\lambda_{in}\right|=\frac{120}{120+\max\{1;15;83;225;274\}}=0.30456.$$

Все корни λ_n, $n=\overline{1,5}$ лежат внутри числового по t интервала $[-|\lambda_{su}|; -|\lambda_{in}|]$. Подставив известные значения оценок λ_{su}, λ_{in} в выражение (24), получим:

$$x_0 = x\left(-\left|\lambda_{su}\right|\right) = -1.4486858\cdot 10^{12};$$
$$x_0 = x\left(-\left|\lambda_{in}\right|\right) = 55.2123.$$

Поскольку $|x(-|\lambda_{su}|)| \gg |x(-|\lambda_{in}|)|$, то выделение числовых интервалов, содержащих неизвестные точные значения искомых корней λ_n, $n=\overline{1,5}$, целесообразно начать с минимального по модулю корня алгебраического уравнения (25). Задали начальное условие $t_0=-|\lambda_{in}|=-0.30456$, $x_0=x(t_0)=55.2123$ и, учитывая знак x_0, выбрали $\mathrm{sgn}(h_x)=-1$. В интервале $[x_0; 0]$ аналитически-численным методом решили систему дифференциальных уравнений (26). В результате для первого минимального по модулю корня λ_1 выделили числовой интервал, содержащий его неизвестное точное значение. Выделенный интервал описывает двойное неравенство (15), которое в рассматриваемом случае приобрело следующие числовые показатели:

$$-1.0087450066 - 0.066844865 \le \lambda_1 \le -1.0087450066 + 0.066844865. \quad (27)$$

Используя двойное неравенство (27), выбрали начальное условие $t_0=\lambda_1(0; I_1)=-1.0087450066$, $x_0=x(t_0)$ и продолжили поиск решения $t(x)$ уравнения (26). В результате нашли абсциссу точки ветвления $x=x^*$, соответствующую предельным соотношениям (16). Совершив в соответствии с равенством (16) переход на очередную ветвь многозначной функции $t(x)$ выделили числовой интервал, который содержит неизвестное точное значение следующего второго корня λ_2 алгебраического уравнения (23). Описанный вычислительный алгоритм повторили для остальных корней.

Результаты расчета, проведенного на первом этапе с заданным предельным уровнем абсолютной локальной погрешности $\varepsilon(\lambda)=1\cdot 10^{-5}$, приведены в таблице 1.

В таблице 1 использованы следующие обозначения: t_0, $x_0=x(t_0)$ — начальные условия для поиска решения $t(x)$; $\lambda_n(0; I_n)$ — приближенные значения корней; $|\Delta\lambda_n(0; I_n)|$ — верхние оценки абсолютных полных погрешностей расчета приближенных значений корней; $t(x^*; I_t)$ — ординаты точек в непосредственной близости от точек ветвления многозначной функции $t(x)$; Δt — произвольная малая величина, необходимая для перехода на очередную ветвь многозначной функции $t(x)$; $\mathrm{sgn}(h_x)$ — знак шага расчета h_x в интервале $[x_0; 0]$.

Таблица 1. Результаты выделения одномерных интервалов, содержащих неизвестные точные значения корней алгебраического уравнения (24)

n	t_0	$x_0 = x(t_0)$	$\lambda_n(0; I_n)$	$\lvert\Delta\lambda_n(0; I_n)\rvert$	$\dfrac{t(x^*; I_t)}{\Delta t}$	$\operatorname{sgn} h_x$
1	–0.3046	55.2123	–1.00874501	0.06684865	–1.3546 –0.0054	–1
2	–1.36	3.631045	–1.99993460	0.05395789	–2.4557 –0.0443	+1
3	–2.5	1.406251	–2.98898512	0.02754328	–3.5424 –0.0076	–1
4	–3.55	–1.418453	–4.00207918	0.03875414	–4.6443 –0.0057	+1
5	–4.65	3.63081	–4.99865317	0.05047823	–	–1

Выделенные и описываемые при соответствующей замене индексов двойными неравенствами вида (27) числовые интервалы содержат точные значения $\lambda_n = -n$ корней алгебраического уравнения (24).

Второй этап. Вычисление с заданной предельной абсолютной погрешностью $\varepsilon(\lambda)=[\varepsilon_1(\lambda)\ \varepsilon_2(\lambda)\ \dots\ \varepsilon_N(\lambda)]^T$ приближенных значений корней λ_n алгебраического уравнения (24).

Выбрав принадлежащее двойному неравенству (27) приближенное значение первого корня, например $\lambda_1^{[1]}(0; I_1) = -1.00874501$, и задав начальное условие $t_0=\lambda_1^{[1]}(0; I_1)$, $x_0=x(t_0)$, в интервале $[x_0; 0]$ аналитически-численным методом решили уравнения (26). Результатом стал новый более узкий числовой интервал, содержащий неизвестное точное значение первого корня λ_1 и описываемый двойным неравенством (15) с новыми числовыми показателями. Выбрав принадлежащее выделенному числовому интервалу новое приближенное значение первого корня $\lambda_1^{[2]}(0; I_1)$ и задав новое начальное условие $t_0=\lambda_1^{[2]}(0; I_1)$, $x_0=x(t_0)$ в интервале $[x_0; 0]$, повторили расчет. Затем расчет повторяли до тех пор, пока он не был сведен к выполнению единственного шага. При одношаговом расчете, выбирая соответствующий порядок метода I_n, обеспечили выполнение неравенства (20), вычислив в итоге приближенное значение $\lambda_n(0; I_n)$, $n=1$ первого корня λ_1 алгебраического уравнения (24) с предельной абсолютной погрешностью $\varepsilon_1(\lambda)$. Таким же образом вычислили приближенные значения всех остальных корней λ_n, $n=\overline{2, 5}$.

Результаты расчета на втором этапе сведены в таблице 2. В таблице 2 использованы обозначения из таблицы 1. Корни уравнения (24) представлены в седьмом столбце таблицы 2 соответственно заданным предельным абсолютным погрешностям $\varepsilon(\lambda)$, которые указаны в начале пункте 4.

Таблица 2. Вычисление с заданной предельной абсолютной погрешностью корней алгебраического уравнения (24)

n	t_0	$x_0,\times10^{-3}$	$h_x,\times10^{-3}$	$\varepsilon_n(\lambda)$	I_n	$\lambda_n(0;I_n)$	$\lvert\Delta\lambda_n(h_x;I_n)\rvert$
1	–0.99991545	0.204003	–0.204003	$1\cdot10^{-3}$	6	–1.00004531	0.826272×10^{-3}
				$1\cdot10^{-6}$	9	–1.00000012	0.302883×10^{-6}
				$1\cdot10^{-12}$	14	–1.00000000	0.814897×10^{-12}
2	–1.99993466	–0.396021	0.396021	$1\cdot10^{-3}$	6	–2.00005138	0.952531×10^{-3}
				$1\cdot10^{-6}$	9	–2.00000021	0.605752×10^{-6}
				$1\cdot10^{-12}$	15	–2.00000000	0.100361×10^{-12}
3	–2.99996518	0.140225	–0.140225	$1\cdot10^{-3}$	6	–3.00004121	0.678810×10^{-3}
				$1\cdot10^{-6}$	9	–3.00000017	0.908165×10^{-6}
				$1\cdot10^{-12}$	15	–3.00000000	0.150545×10^{-12}
4	–4.00207918	12.495	–12.495	$1\cdot10^{-3}$	6	–3.99999857	0.452618×10^{-3}
				$1\cdot10^{-6}$	10	–3.99999997	0.705432×10^{-6}
				$1\cdot10^{-12}$	15	–4.00000000	0.392155×10^{-12}
5	–4.99865317	–31.122	–31.122	$1\cdot10^{-3}$	7	–4.99995678	0.805343×10^{-3}
				$1\cdot10^{-6}$	10	–4.99999937	0.918324×10^{-6}
				$1\cdot10^{-12}$	14	–5.00000000	0.103112×10^{-12}

Как видно из таблицы 2, при уменьшении уровня предельной абсолютной погрешности расчета $\varepsilon_n(\lambda)$, $n\in[1; 5]$ на девять порядков порядок метода увеличивается всего лишь в 2–2.5 раза.

5. Заключение. Отметим следующие особенности предложенного алгоритма, отличающие его от широко используемых.

Во-первых, алгоритм позволяет выделять числовые интервалы, содержащие неизвестные точные значения корней. Знание таких интервалов определяет возможность вычисления приближенных значений корней с любой заданной предельной абсолютной погрешностью.

Во-вторых, результативность алгоритма, то есть гарантия достижения поставленной цели не зависит от выбора «удачных» или «близких» начальных условий — достаточно вычислить лишь верхнюю $|\lambda_{su}|$ и нижнюю $|\lambda_{in}|$ оценки неизвестных точных значений корней алгебраического уравнения.

В-третьих, алгоритм на первом этапе не итерационный, поэтому число шагов расчета, необходимое для выделения числового интервала, который содержит неизвестное точное значение какого-либо из корней λ_n, n=1, 2, …, N алгебраического уравнения, всегда ограничено.

В-четвёртых, алгоритм при поиске определенного корня алгебраического уравнения вычислительно полностью автономен, что исключает неконтролируемый «дрейф» погрешности и взаимовлияние погрешностей, сопровождающих поиск различных корней.

Литература

1. *Ugboh J.A., Esuabana I.M.* Marching Method: A New Numerical Method for Finding Roots of Algebraic and Transcendental Equations // American Journal of Computational and Applied Mathematics. 2019. vol. 9(1). pp. 6–11.
2. *Petkovica M.S., Petkovic L.D., Dzunic J.* On an Efficient Simultaneous Method for Finding Polynomial zeros // Applied Mathematics Letters. 2014. vol. 28. pp. 60–65.
3. *Nedzhibov G.H.* A Derivative-Free Iterative Method for Simultaneously Computing an Arbitrary Number of Zeros of Nonlinear Equations // Computers and Mathematics with Applications. 2012. vol. 63. no. 7. pp. 1185–1191.
4. *Khamisov O.V.* Finding Roots of Nonlinear Equations Using the Method of Concave Support Functions // Mathematical Notes. 2015. vol. 98. no. 3-4. pp. 484–491.
5. *Амосов А.А., Дубинский Ю.А., Копченова Н.В.* Вычислительные методы: учеб. пособие // СПб.: Лань. 2014.
6. *Hamming R.W.* Numerical Method for Scientists and Engineers // New York: Dover Publications Inc. 1987. 752 p.
7. *Воронова М.Е., Симакова М.Н., Симаков Е.Е.* Методы решения нелинейных уравнений // Юный ученый. 2016. № 3. С. 102–105.
8. *Alharbi A.R. et al.* Higher Order Numerical Approaches for Nonlinear Equations by Decomposition Technique // IEEE Access. 2019. vol. 7. pp. 44329–44337.
9. *Zafar F., Cordero A., Quratulain R., Torregrosa J.R.* Optimal Iterative Methods for Finding Multiple Roots of Nonlinear Equations Using Free Parameters // Journal of Mathematical Chemistry. 2017. vol. 56. pp. 1884–1891.
10. *Akram S., Zafar F., Yasmin N.* An Optimal Eighth-Order Family of Iterative Methods for Multiple Roots // Mathematics. 2019. vol. 7. no. 8. pp. 672.
11. *Behl R., Cordero A., Motsa S.S., Torregrosa J.R.* An Eighth-order Family of Optimal Multiple Root Finders and its Dynamics // Numerical Algorithms. 2018. vol. 77. no. 4. pp. 1249–1272.
12. *Soleymani F., Babajee D.K.R., Lotfi T.* On a Numerical Technique for Finding Multiple Zeros and its Dynamic // Journal of the Egyptian Mathematical Society. 2013. no. 21. no. 3. pp. 346–353.
13. *Winkler J.R., Lao X., Hasan M.* The Computation of Multiple Roots of a Polynomial // Journal of Computational and Applied Mathematics. 2012. vol. 236. no. 14. pp. 3478–3497.
14. *Yun B.I.* A Derivative Free Iterative Method for Finding Multiple Roots of Nonlinear Equations // Applied Mathematics Letters. 2009. vol. 22. no. 12. pp. 1859–1863.
15. *Jaiswal J.P.* An Optimal Order Method for Multiple Roots in Case of Unknown Multiplicity // Algorithms. 2016. vol. 9. no. 1. pp. 10.
16. *Тубольцев М.Ф., Маторин С.И., Тубольцева О.М.* Эвристический компьютерный алгоритм вычисления кратных корней нелинейного уравнения // Научные ведомости БелГУ. Серия: Экономика. Информатика. 2015. Вып. 34. № 7(204). С. 78–83.
17. *Liao Z. et al.* Solving Nonlinear Equations System With Dynamic Repulsion-Based Evolutionary Algorithms // IEEE Transactions on Systems, Man and Cybernetics: Systems. 2018. vol. 7. pp. 1–12.
18. *Gong W., Wang Y., Cai Z., Wang L.* Finding Multiple Roots of Nonlinear Equation Systems via a Repulsion-Based Adaptive Differential Evolution // IEEE Transactions on Systems, Man and Cybernetics: Systems. 2018. vol. 5. pp. 1–15.

19. *Neamvonk J., Phuenaree B., Neamvonk A.* A New Method for Finding Root of Nonlinear Equations by Using Nonlinear Regression // Asian Journal of Applied Sciences. 2015. vol. 3. no. 6. pp. 818–822.
20. *Yang X.J., Machado J.A.T., Srivastava H.M.* A New Numerical Technique for Solving the Local Fractional Diffusion Equation: Two-dimensional Extended Differential Transform Approach // Applied Mathematics and Computation. 2016. vol. 274. pp. 143–151.
21. *Kalitkin N.N.*, *Kuzmina L.V.* Calculation of Roots and There Multiplicity for Nonlinear Equation // Mathematical Models and Computer Simulations. 2011. vol. 3. no. 1. pp. 65–80.
22. *Kalitkin N.N.*, *Kuzmina L.V.* The Method of Seconds with Extrapolation for Accurate Calculation of Manifold Roots // Mathematical Models and Computer Simulations. 2011. vol. 23. no. 6. pp. 33–58.
23. *Бычков Ю.А. и др.* Математическое моделирование и анализ нелинейных систем // СПб.: Изд-во СПбГЭТУ «ЛЭТИ». 2015. 302 с.
24. *Бычков Ю.А., Соловьева Е.Б., Щербаков С.В.* Непрерывные и дискретные нелинейные модели динамических систем // СПб.: Лань. 2018. 420 с.
25. *Gofen A.M.* Fast Taylor-Series Expansion and the Solution of the Cauchy Problem // USSR Computational Mathematics and Mathematical Physics. 1982. vol. 22. no. 5. pp. 74–88.
26. *Гофен А.М.* Интегрирование обыкновенных дифференциальных уравнений методом Тейлора и проблема шага // М.: ИПИ АН СССР. 1991. 29 с.
27. *Козин Р.* Программирование алгоритмов численных методов линейной алгебры // New York: LAP Lambert Academic Publishing. 2014. 188 с.
28. *Волков Е.А.* Численные методы: учеб. пособие // СПб.: Лань. 2008. 248 с.
29. *Grewal B.S.* Numerical Methods in Engineering and Science C, C++, and MATLAB // Dulles: Mercury Learning and Information LLC. 2018. 1597 p.
30. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы // М.: БИНОМ. Лаборатория знаний. 2018. 640 с.
31. *Калиткин Н.Н.* Численные методы: учеб. пособие // СПб.: БХВ-Петербург. 2014. 592 с.
32. *Турчак Л.И., Плотников П.В.* Основы численных методов // М.: ФИЗМАТЛИТ. 2003. 304 с.

Бычков Юрий Александрович — д-р техн. наук, профессор, профессор, кафедра теоретических основ электротехники института фундаментального инженерного образования, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В. И. Ульянова (Ленина) (СПбГЭТУ «ЛЭТИ»). Область научных интересов: математическое моделирование динамических систем, численные методы. Число научных публикаций — 260. rimelena@yahoo.com; ул. Профессора Попова, 5, 197376, Санкт-Петербург, Российская Федерация; р.т.: +7(812)-346-17-96.

Соловьева Елена Борисовна — д-р техн. наук, профессор, действительный член Академии электротехнических наук РФ, заведующий кафедрой, кафедра теоретических основ электротехники института фундаментального инженерного образования, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В. И. Ульянова (СПбГЭТУ "ЛЭТИ"). Область научных интересов: математическое моделирование и синтез нелинейных динамических систем, поведенческое моделирование на основе многомерных полиномов и нейронных сетей, цифровая обработка сигналов. Число научных публикаций — 250. selenab@hotbox.ru; ул. Профессора Попова, 5, 197376, Санкт-Петербург, Российская Федерация; р.т.: +7(812)-346-17-96.

Щербаков Сергей Валерьевич — д-р техн. наук, профессор, профессор, кафедра теоретических основ электротехники института фундаментального инженерного образования, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В. И. Ульянова (Ленина) (СПбГЭТУ «ЛЭТИ»). Область научных интересов: математическое моделирование динамических систем, численные методы. Число научных публикаций — 120. gz43@pskovsobranie.ru; ул. Профессора Попова, 5, 197376, Санкт-Петербург, Российская Федерация; р.т.: +7(812)-346-17-96.

DOI 10.15622/sp.2019.18.6.1491-1514

Y.A. BYCHKOV, E.B. SOLOVYEVA, S.V. SCHERBAKOV

ANALYTICAL-NUMERICAL CALCULATION ALGORITHM OF ERRORS OF ALGEBRAIC EQUATIONS ROOTS WITH SPECIFIED LIMITS

Bychkov Y.A., Solovyeva E.B., Scherbakov S.V. **Analytical-Numerical Calculation Algorithm of Algebraic Equations Roots with Specified Limits of Errors.**

Abstract. This paper proposes an algorithm for calculating approximate values of roots of algebraic equations with a specified limit of absolute errors. A mathematical basis of the algorithm is an analytical-numerical method of solving nonlinear integral-differential equations with non-stationary coefficients. The analytical-numerical method belongs to the class of one-step continuous methods of variable order with an adaptive procedure for choosing a calculation step, a formalized estimate of the error of the performed calculations at each step and the error accumulated during the calculation. The proposed algorithm for calculating the approximate values of the roots of an algebraic equation with specified limit absolute errors consists of two stages. The results of the first stage are numerical intervals containing the unknown exact values of the roots of the algebraic equation. At the second stage, the approximate values of these roots with the specified limit absolute errors are calculated. As an example of the use of the proposed algorithm, defining the roots of the fifth-order algebraic equation with three different values of the limiting absolute error is presented.

The obtained results allow drawing the following conclusions. The proposed algorithm enables to select numeric intervals that contain unknown exact values of the roots. Knowledge of these intervals facilitates the calculation of the approximate root values under any specified limiting absolute error. The algorithm efficiency, i.e., the guarantee of achieving the goal, does not depend on the choice of initial conditions. The algorithm is not iterative, so the number of calculation steps required for extracting a numerical interval containing an unknown exact value of any root of an algebraic equation is always restricted. The algorithm of determining a certain root of the algebraic equation is computationally completely autonomous.

Keywords: Algebraic Equation, Nonlinear Differential Equation, Analytical-Numerical Method, Approximate Root Value.

Bychkov Yuri Alexandrovich — Ph.D., Dr.Sci., Professor, Professor, Department of Theoretical Electrical Engineering of Institute of Fundamental Engineering Education, Saint Petersburg Electrotechnical University 'LETI'. Research interests: mathematical modeling of dynamic systems, numerical methods. The number of publications — 260. rimelena@yahoo.com; 5, Professor Popov str., 197376, St. Petersburg, Russian Federation; office phone: +7(812)-346-17-96.

Solovyeva Elena Borisovna — Ph.D., Dr.Sci., Professor, Full Member of Russian Academy of Electrical Engineering Sciences, Head of Department, Department of Theoretical Electrical Engineering of Institute of Fundamental Engineering Education, Saint Petersburg Electrotechnical University 'LETI'. Research interests: mathematical modeling and synthesis of nonlinear dynamic systems, behavioral modeling based on multidimensional polynomials and neural networks, digital signal processing. The number of publications — 250. selenab@hotbox.ru; 5, Professor Popov str., 197376, Sankt-Peterburg, Russian Federation; office phone: +7(812)-346-17-96.

Scherbakov Sergei Valerievich — Ph.D., Dr.Sci., Professor, Professor, Department of Theoretical Electrical Engineering of Institute of Fundamental Engineering Education, Saint Petersburg Electrotechnical University 'LETI'. Research interests: mathematical modeling of dynamic systems, numerical methods. The number of publications — 120. gz43@pskovsobranie.ru; 5, Professor Popov str., 197376, St. Petersburg, Russian Federation; office phone: +7(812)-346-17-96.

References

1. Ugboh J.A., Esuabana I.M. Marching Method: A New Numerical Method for Finding Roots of Algebraic and Transcendental Equations. *American Journal of Computational and Applied Mathematics*. 2019. vol. 9(1). pp. 6–11.
2. Petkovica M.S., Petkovic L.D., Dzunic J. On an Efficient Simultaneous Method for Finding Polynomial zeros. *Applied Mathematics Letters*. 2014. vol. 28. pp. 60–65.
3. Nedzhibov G.H. A Derivative-Free Iterative Method for Simultaneously Computing an Arbitrary Number of Zeros of Nonlinear Equations. *Computers and Mathematics with Applications.* 2012. vol. 63. no. 7. pp. 1185–1191.
4. Khamisov O.V. Finding Roots of Nonlinear Equations Using the Method of Concave Support Functions. *Mathematical Notes.* 2015. vol. 98. no. 3-4. pp. 484–491.
5. Amosov A.A., Dubinskij Ju.A., Kopchenova N.V. *Vychislitel'nye metody: uchebnoe posobie* [Computational Methods: tutorial]. SPb: Lan'. 2014. (In Russ.).
6. Hamming R.W. Numerical Method for Scientists and Engineers. New York: Dover Publications Inc. 1987. 752 p.
7. Voronova M.E., Simakova M.N., Simakov E.E. [Methods for solving nonlinear equations]. *Junyj uchenyj – Young scientist*. 2016. vol. 3. pp. 102–105. (In Russ.).
8. Alharbi A.R. et al. Higher Order Numerical Approaches for Nonlinear Equations by Decomposition Technique. *IEEE Access*. 2019. vol. 7. pp. 44329–44337.
9. Zafar F., Cordero A., Quratulain R., Torregrosa J.R. Optimal Iterative Methods for Finding Multiple Roots of Nonlinear Equations Using Free Parameters. *Journal of Mathematical Chemistry*. 2017. vol. 56. pp. 1884–1891.
10. Akram S., Zafar F., Yasmin N. An Optimal Eighth-Order Family of Iterative Methods for Multiple Roots. *Mathematics.* 2019. vol. 7. no. 8. pp. 672.
11. Behl R., Cordero A., Motsa S.S., Torregrosa J.R. An Eighth-order Family of Optimal Multiple Root Finders and its Dynamics. *Numerical Algorithms*. 2018. vol. 77. pp. 1249–1272.
12. Soleymani F., Babajee D.K.R., Lotfi T. On a Numerical Technique for Finding Multiple Zeros and its Dynamic. *Journal of the Egyptian Mathematical Society*. 2013. no. 21. pp. 346–353.
13. Winkler J.R., Lao X., Hasan M. The Computation of Multiple Roots of a Polynomial. *Journal of Computational and Applied Mathematics*. 2012. vol. 236. no. 14. pp. 3478–3497.
14. Yun B.I. A Derivative Free Iterative Method for Finding Multiple Roots of Nonlinear Equations. *Applied Mathematics Letters*. 2009. vol. 22. no. 12. pp. 1859–1863.
15. Jaiswal J.P. An Optimal Order Method for Multiple Roots in Case of Unknown Multiplicity. *Algorithms.* 2016. vol. 9. no. 1. pp. 10.
16. Tubol'cev M.F., Matorin S.I., Tubol'ceva O.M. [Heuristic computer algorithm for calculating multiple roots of a nonlinear equation]. *Nauchnye vedomosti BelGU. Serija: Jekonomika. Informatika – Belgorod State University Scientific Bulletin.* 2015. vol. 34. no. 7(204). pp. 78–83. (In Russ.).
17. Liao Z. et al. Solving Nonlinear Equations System With Dynamic Repulsion-Based Evolutionary Algorithms. *IEEE Transactions on Systems, Man and Cybernetics: Systems*. 2018. vol. 7. pp. 1–12.
18. Gong W., Wang Y., Cai Z., Wang L. Finding Multiple Roots of Nonlinear Equation Systems via a Repulsion-Based Adaptive Differential Evolution. *IEEE Transactions on Systems, Man and Cybernetics: Systems*. 2018. vol. 5. pp. 1–15.
19. Neamvonk J., Phuenaree B., Neamvonk A. A new method for finding Root of Nonlinear Equations by using Nonlinear Regression. *Asian Journal of Applied Sciences*. 2015. vol. 3. no. 6. pp. 818–822.
20. Yang X.J., Tenreiro J.A., Srivastava H.M. A new numerical technique for solving the local fractional diffusion equation: Two-dimensional extended differential transform approach. *Applied Mathematics and Computation*. 2016. vol. 274. pp. 143–151.

21. Kalitkin N.N., Kuzmina L.V. Calculation of Roots and There Multiplicity for Nonlinear Equation. *Mathematical Models and Computer Simulations*. 2011. vol. 3. no. 1. pp. 65–80.
22. Kalitkin N.N., Kuzmina L.V. The Method of Seconds with Extrapolation for Accurate Calculation of Manifold Roots. *Mathematical Models and Computer Simulations*. 2011. vol. 23. no. 6. pp. 33–58.
23. Bychkov Ju.A. et al. *Matematicheskoe mo-delirovanie i analiz nelinejnyh sistem. Mathematische model-lierung und analyse nichtlinearer systems* [Mathematical modeling and analysis of nonlinear systems]. Saint-Petersburg: Izd-vo SPbGETU "LETI". 2015. 302 p. (In Russ.).
24. Bychkov Ju.A., Solov'eva E.B., Shherbakov S.V. *Nepreryvnye i diskretnye nelinejnye modeli dinamicheskih sistem* [Continuous and discrete nonlinear models of dynamic systems]. SPb: Lan'. 2018. 420 p. (In Russ.).
25. Gofen A.M. Fast Taylor-Series Expansion and the Solution of the Cauchy Problem. *USSR Computational Mathematics and Mathematical Physics*. 1982. vol. 22. no. 5. pp. 74–88.
26. Gofen A.M. *Integrirovanie obyknovennyh differencial'nyh uravnenij metodom Tejlora i problema shaga* [Integration of ordinary differential equations by the Taylor method and the step problem]. Moscow: IPI AN SSSR. 1991. 29 p. (In Russ.).
27. Kozin R. *Programmirovanie algoritmov chislennyh metodov linejnoj algebry* [Programming algorithms for the numerical methods of linear algebra]. New York: LAP Lambert Academic Publishing. 2014. 188 p. (In Russ.).
28. Volkov E.A. *Chislennye metody: uchebnoe posobie* [Numerical methods: tutorial]. SPb: Lan'. 2008. 248 p. (In Russ.).
29. Grewal B.S. Numerical Methods in Engineering and Science C, C++, and MATLAB. Dulles: Mercury Learning and Information LLC. 2018. 1597 p.
30. Bahvalov N.S., Zhidkov N.P., Kobel'kov G.M. *Chislennye metody* [Numerical methods]. Moscow: BINOM. Laboratorija znanij. 2018. 640 p. (In Russ.).
31. Kalitkin N.N. *Chislennye metody: uchebnoe posobie* [Numerical methods: tutorial]. SPb.: BHV-Peterburg. 2014. 592 p. (In Russ.).
32. Turchak L.I., Plotnikov P.V. *Osnovy chislennyh metodov* [Fundamentals of numerical methods]. M.: FIZMATLIT. 2003. 304 p. (In Russ.).

Signed to print 25.11.2019
Printed in Publishing center GUAP, 67, B. Morskaya, St. Petersburg, 190000, Russia

The journal is registered in Russian Federal Agency for Communications
and Mass-Media Supervision, certificate ПИ № ФС77-41695 dated August 19, 2010 г.
Subscription Index П5513, Russian Post Catalog

Подписано к печати 25.мм.2019. Формат 60×90 1/16. Усл. печ. л. 14,47. Заказ № 523.
Тираж 150 экз., цена свободная.
Отпечатано в Редакционно-издательском центре ГУАП, 190000, Санкт-Петербург, Б. Морская, д. 67

Журнал зарегистрирован Федеральной службой по надзору в сфере связи
и массовых коммуникаций,
свидетельство ПИ № ФС77-41695 от 19 августа 2010 г.

Подписной индекс П5513 по каталогу «Почта России»

РУКОВОДСТВО ДЛЯ АВТОРОВ

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Труды СПИИРАН» http://www.proceedings.spiiras.nw.ru. При регистрации авторам рекомендуется заполнить все предложенные поля данных.

Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше. Объем основного текста – от 20 до 30 страниц включительно. Формат страницы документа – A5 (148 мм ширина, 210 мм высота); ориентация – портретная; все поля – 20 мм. Верхний и нижний колонтитулы страницы – пустые. Основной шрифт документа – Times New Roman, основной кегль (размер) шрифта – 10 pt. Переносы разрешены. Абзацный отступ устанавливается размером в 10 мм. Межстрочный интервал – одинарный. Номера страниц не проставляются.

В основную часть допускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

AUTHOR GUIDELINES

Interaction between each potential author and the Editorial board is realized through the personal account on the website of the journal "Proceedings of SPIIRAS" http://www.proceedings.spiiras.nw.ru. At the registration the authors are requested to fill out all data fields in the proposed form.

The submissions should be prepared using MS Word 2007 text editor or higher versions, at that, only manuscripts in *.docx format will be considered. The text of the paper in the main part of it should be from 20 – 30 pages of A5 size that is 210 X 148 mm; orientation – portrait; all margins – 20 mm. The font of the main paper text is Times New Roman of 10 pt font size. The pages' headers and footers should be empty; indention – 10 mm; line spacing – single; pages are not numbered; hyphenations are allowed.

Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered by the paper template in more detail in journal web.

ISSN 2078-9181

Том 18 № 2

ТРУДЫ СПИИРАН
proceedings.spiiras.nw.ru

С-Пб
2019