УДК 614.8 + 002.6:004.89
doi:10.31799/1684-8853-2018-6-74-81

Научные статьи
Articles

Поиск кратчайшей траектории социоинженерной атаки между парой пользователей в графе с вероятностями переходов

А. О. Хлобыстова[а, б], *стажер, orcid.org/0000-0002-9811-5476*
М. В. Абрамов[а, б], *канд. техн. наук, научный сотрудник, orcid.org/0000-0002-5476-3025, mva16@list.ru*
А. Л. Тулупьев[а, б], *доктор физ.-мат. наук, доцент, главный научный сотрудник, orcid.org/0000-0003-1814-4646*
А. А. Золотин[в], *канд. физ.-мат. наук, старший инженер-разработчик, orcid.org/0000-0002-1028-4292*
[а]*Санкт-Петербургский институт информатики и автоматизации РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ*
[б]*Санкт-Петербургский государственный университет, Университетская наб., 7–9, Санкт-Петербург, 199034, РФ*
[в]*ЕПАМ Системс ГмбХ, Франклинштрассе, 56, Франкфурт-на-Майне, 60486, Германия*

Введение: *социоинженерные атаки можно разделить на два вида: прямые (одноходовые) и многоходовые, проходящие через цепочку пользователей. Траекторией распространения многоходовых социоинженерных атак между двумя пользователями, как правило, является некоторое непустое множество. Оценки вероятности распространения атаки по разным траекториям будут отличаться.* ***Цель:*** *выявление наиболее критичной (наиболее вероятной) траектории распространения многоходовой социоинженерной атаки между двумя пользователями.* ***Методы:*** *поиск, сопоставление и анализ алгоритмов для выявления наиболее критичной траектории распространения атаки. Методы опираются на сведения, характеризующие интенсивность взаимодействия сотрудников в компании, основанные на данных, извлекаемых из социальных сетей. Указанные алгоритмы сводятся с использованием ряда преобразований исходных данных к алгоритмам поиска наикратчайшего пути в графе. Применяемые оценки вероятности успеха многоходовой социоинженерной атаки рассчитываются с помощью методов построения оценки вероятности сложного события.* ***Результаты:*** *предложен подход к идентификации наиболее критичных траекторий, оценка вероятности успеха прохождения атаки по которым будет наиболее высокой. В простейшем случае задача может быть рассмотрена как задача нахождения в графе пути, в котором произведение весов всех ребер, входящих в данный путь, максимально. Представлен подход к решению задачи сокращения ресурсозатратности алгоритма при поиске наиболее критичной траектории на полном графе с большим количеством вершин. Краткий обзор методов и алгоритмов автоматизированного решения задачи поиска наиболее критичной траектории распространения социоинженерной атаки показал, что она в общем случае при ряде преобразований может быть сведена к задаче поиска наиболее критичной траектории с использованием конфигурации алгоритмов Дейкстры и Беллмана — Форда. Произведена адаптация выбранного алгоритма для указанного контекста, предложен подход к разрежению графа при поиске наиболее критичной траектории. Представленные методы и алгоритмы реализованы в программном коде, для верификации результатов расчетов выполнены численные эксперименты.* ***Практическая значимость:*** *разработанное программное обеспечение, основанное на предложенных в статье методе и алгоритме, дополняет функционал предшествующих версий прототипов программ для анализа защищенности пользователей информационных систем от социоинженерных атак. Оно позволяет учитывать более широкий круг факторов, влияющих на оценку вероятности успеха социоинженерной атаки злоумышленника на пользователя.*

Ключевые слова *— информационная безопасность, социоинженерные атаки, защита пользователей, многоходовые социоинженерные атаки, траектории распространения атак, вероятность успеха поражения пользователя, социальная инженерия, анализ защищенности пользователя, аудит информационной безопасности, мониторинг защищенности, социальный граф компании, интенсивность взаимодействия сотрудников, социальные сети.*

Цитирование: Хлобыстова А. О., Абрамов М. В., Тулупьев А. Л., Золотин А. А. Поиск кратчайшей траектории социоинженерной атаки между парой пользователей в графе с вероятностями переходов. *Информационно-управляющие системы*, 2018, № 6, с. 74–81. doi:10.31799/1684-8853-2018-6-74-81
Citation: Khlobystova A. O., Abramov M. V., Tulupyev A. L., Zolotin A. A. Search for the shortest trajectory of a social engeneering attack between a pair of users in a graph with transition probabilities. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 6, pp. 74–81 (In Russian). doi:10.31799/1684-8853-2018-6-74-81

Введение

В последнее время наблюдается рост числа киберпреступлений, что приводит к существенным убыткам компаний [1, 2]. Характер киберпреступлений также становится многообразнее, в настоящее время для нарушения информационной безопасности компании злоумышленники прибегают не только к поиску программно-технических уязвимостей системы, но и к методам социальной инженерии, т. е. поиску и эксплуатации уязвимостей пользователей. Таким образом, актуальной видится проблема повышения уровня защищенности пользователей информационных систем от социоинженерных атак. Важной составляющей этой проблемы является задача

анализа и мониторинга защищенности пользователей информационных систем от социоинженерных атак.

Социоинженерные атаки по числу задействованных пользователей могут быть одноходовые (прямые, неопосредованные) и многоходовые, совершаемые через цепочку пользователей, где начальный и конечный пользователи не совпадают, пользователи в цепочке не повторяются. Некоторые подходы к оценке защищенности пользователей от прямых социоинженерных атак достаточно подробно изложены в работе [3]. В данной статье исследуется один из аспектов анализа защищенности пользователей информационных систем при многоходовых социоинженерных атаках — выявление наиболее критичных траекторий распространения атаки. Подход к оценке защищенности пользователей при многоходовых атаках, а также к расчету оценок вероятности распространения атаки от пользователя к пользователю был представлен в работе [4]. Отметим, что оценка вероятности успеха многоходовой социоинженерной атаки рассчитывается в предположении, что оценки вероятности распространения атаки от одного пользователя к другому и обратно равны [4]. Однако оценка вероятности распространения атаки от пользователя к пользователю в прямом и обратном направлении может быть разной, соответственно, необходимо рассматривать ориентированный социальный граф. Также в [4] не рассматриваются вопросы поиска наиболее критичных траекторий распространения атаки. Как правило, существует несколько возможных траекторий развития атаки от одного пользователя к другому, и вероятности успеха распространения атаки по каждой из них в общем случае будут иметь разные значения. В связи с этим актуальной видится задача поиска наиболее критичных траекторий, оценка вероятности успеха прохождения атаки по которым будет наиболее высокой. Одна из подзадач в данном контексте связана с выявлением наиболее критичных траекторий (т. е. наиболее вероятных траекторий) распространения атаки от одного пользователя к другому, и целью данной работы является решение этой подзадачи.

Релевантные работы

Заделом для данного исследования послужили работы [3, 4], в которых описаны подходы к оценке защищенности пользователей информационных систем от прямых и многоходовых социоинженерных атак. В [5] представлен комплексный подход к анализу информационной безопасности организации, предложен подход к оценке уязвимости, но при этом большая часть параметров оценивается специалистом по информационной безопасности. В [6] приводятся результаты, связанные с применением обучающих игр для сотрудников компании в целях снижения рисков успешных социоинженерных атак. Эмпирические оценки, полученные в результате исследований [7, 8], могут способствовать построению оценок вероятности успешной социоинженерной атаки на пользователя. Исследование, направленное на разработку методов защиты от нескольких видов социоинженерных атак, представлено в работе [9], однако многоходовые атаки не рассматриваются. В работе [10] представлена многоуровневая модель оценки уязвимости пользователей к социоинженерным атакам, основанная на используемых злоумышленником способах коммуникации с пользователем, состоянии системы в определенный момент времени и на известных сценариях атак, однако не учитываются психологические особенности и влияние самих пользователей информационной системы. В [11] изучается возможность защиты конфиденциальных данных, получаемых из социальных сетей. Основным объектом исследования является социальный граф взаимодействия пользователей и способы обработки нежелательных (с точки зрения конфиденциальности) связей в социальных сетях. В [12] предлагается подход к повышению уровня информационной безопасности организации за счет анализа защищенности посредством автоматизированного сбора данных о сотрудниках компании из открытых источников и их последующего изучения и оценки в целях выявления наиболее уязвимых мест. В [13] исследуются факторы, влияющие на уязвимости пользователей и причины подверженности социоинженерным атакам. Существенный пласт исследований в области защиты пользователей от социоинженерных атак посвящен идентификации и анализу защищенности от атак типа фишинг [14–20]. Результаты таких работ полезны и могут быть использованы при разработке систем упреждающей диагностики и выработки рекомендаций лицам, принимающим решения.

Формализация задачи выявления наиболее критичной траектории распространения социоинженерной атаки

Анализ возможных траекторий распространения социоинженерной атаки предлагается производить на ориентированном социальном графе сотрудников компании. Под социальным графом сотрудников компании будем понимать граф, вершины которого соответствуют сотрудникам компании, а ребра — связям между сотрудниками. Формализуя описанное, зададим граф $G = (U, E)$,

где $U = \{\text{User}_i\}_{i=1}^{n}$ — множество вершин (пользователей); $E = \{u_i, u_j, p_{i,j}\}_{1 \le i,j \le n, i \ne j}$ — множество упорядоченных троек с заданной оценкой вероятности распространения атаки от пользователя к пользователю — $p_{i,j}$. Заметим, что равенство $p_{i,j}$ и $p_{j,i}$ в общем случае не предполагается. То есть вероятность распространения атаки от первого пользователя ко второму может отличаться от вероятности распространения атаки в другую сторону — от второго к первому.

В рассматриваемой модели оценка вероятности успеха распространения атаки от пользователя к пользователю зависит от интенсивности взаимодействия между пользователями. Согласно [2], она может быть рассчитана следующим образом: $p_{i,j} = 1 - \prod_t (1 - p_t^{i,j})^{n_t}$, где $p_t^{i,j}$ — оценка вероятности успеха социоинженерной атаки злоумышленника на пользователя по t-й связи; n_t — число эпизодов взаимодействия. В рассматриваемом частном случае модели, учитывающей сведения, извлекаемые из социальных сетей, $p_{i,j} > 0$; ребра, где оценка вероятности $p_{i,j} = 0$, исключаются из итогового социального графа.

Таким образом, задача поиска наиболее критичной траектории многоходовой социоинженерной атаки от User_i до User_j сводится к задаче нахождения в графе элементарного пути (простого и без циклов) между этими вершинами. Причем путь должен быть таким, что произведение оценок вероятностей переходов от пользователя к пользователю, входящих в него, максимально. Будем называть оценку вероятности успеха многоходовой социоинженерной атаки, которая представляет собой произведение оценок вероятностей распространения атаки от пользователя к пользователю и прямой атаки на первого пользователя, длиной пути.

Подход к выявлению наиболее критичной траектории распространения социоинженерной атаки

Для упрощения, не умаляя общности, рассмотрим граф $G = (U, E')$, где $U = \{\text{User}_i\}_{i=1}^{n}$ — множество вершин (пользователей); $E' = \left\{\left(u_i, u_j, \frac{1}{p_{i,j}}\right)\right\}_{1 \le i,j \le n, i \ne j}$ — множество упорядоченных троек, где каждой паре пользователей сопоставлено число $\frac{1}{p_{i,j}}$. Заметим, что в этом случае, если $p_{i,j} \ge p_{l,k}$, то $\frac{1}{p_{i,j}} \le \frac{1}{p_{l,k}}$, а длина пути будет вычисляться следующим образом: $\frac{1}{p_{ml}} = \frac{1}{p_m} \prod_{i=m}^{l-1} \frac{1}{p_{i,i+1}}$, где p_{ml} — оценка вероятности успеха прохождения атаки от пользователя m до пользователя l; p_m — оценка вероятности успеха прямой социоинженерной атаки злоумышленника на пользователя; $p_{i,i+1}$ — соответствующая оценка вероятности распространения атаки на пользователя через другого пользователя. Таким образом, от задачи поиска пути с максимальной длиной перейдем к задаче поиска пути с минимальной длиной.

Чтобы применить алгоритмы нахождения минимального пути, необходимо произвести ряд преобразований. Согласно основному логарифмическому тождеству: $\frac{1}{p_{ij}} = e^{\log \frac{1}{p_{ij}}}$. Тогда длина пути будет рассчитываться следующим образом:

$$\frac{1}{p_{ml}} = \frac{1}{p_m} \prod_{i=m}^{l-1} \frac{1}{p_{i,i+1}} = e^{\log \frac{1}{p_m}} \prod_{i=m}^{l-1} e^{\log \frac{1}{p_{i,i+1}}} = \exp\left\{\log \frac{1}{p_m} + \sum_{i=m}^{l-1} \log \frac{1}{p_{i,i+1}}\right\}.$$

Поскольку оценка p_m успеха прямого социоинженерного атакующего воздействия на пользователя m будет одинакова для всех траекторий, начинающихся с пользователя m, то задача сводится к поиску пути, в котором $-\sum_{i=m}^{l-1} \log p_{i,i+1}$ минимальна среди всех возможных траекторий, начинающихся с пользователя m и заканчивающихся пользователем l, или, что то же, $\sum_{i=m}^{l-1} \log p_{i,i+1}$ максимальна. Таким образом, задача представляет собой стандартный поиск наикратчайшего пути в ориентированном графе без ребер отрицательного веса.

Пусть n — число вершин в социальном графе (число сотрудников), m — число дуг в социальном графе. Для решения задачи поиска наиболее критичной траектории на социальном графе были рассмотрены алгоритмы Беллмана — Форда, Левита, Флойда — Уоршелла, Дейкстры и его модификации, топологическая сортировка, A* [21–23]. Алгоритмы Левита и Флойда — Уоршелла имеют высокую вычислительную сложность в условиях нашей задачи, поэтому их применение нецелесообразно. Для применения алгоритма топологической сортировки исходный граф должен быть ацикличным. Социальный граф сотрудников компании в большинстве случа-

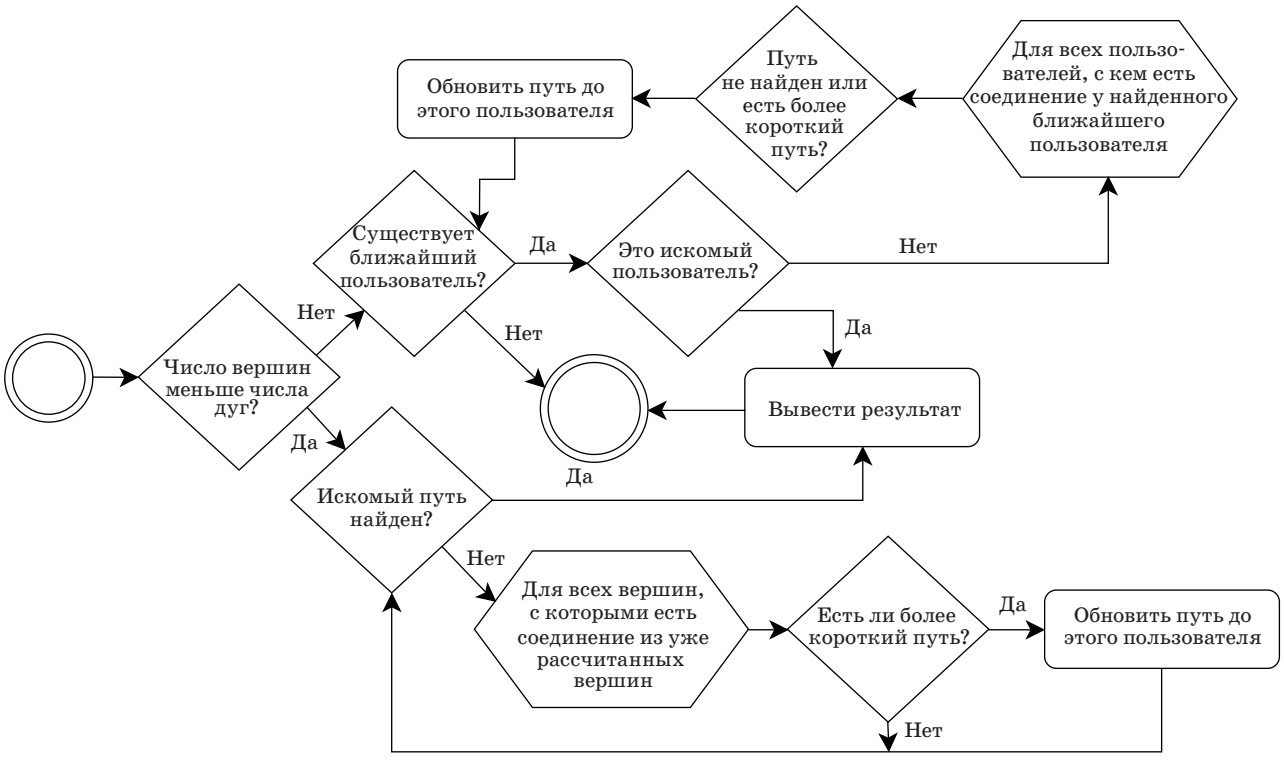
ев не обладает этим свойством, в связи с чем указанный алгоритм не может быть применен. Алгоритм A* удобен тем, что осуществляет поиск кратчайшего расстояния только между двумя вершинами, а не между всеми. Однако трудность использования данного алгоритма заключается в подборе правильной эвристической функции. Кроме того, алгоритм A* требует большого объема памяти при работе, в связи с чем его применение нецелесообразно. Алгоритм Дейкстры подходит для решения нашей задачи, обладает вычислительной сложностью $O(n^2)$. При этом оптимальная сложность для алгоритмов, основанных на алгоритме Дейкстры, составляет $O(n\log n + m)$ и достигается при представлении данных в виде куч Фибоначчи. Однако константы, скрытые в асимптотических оценках трудоемкости упомянутой модификации, зачастую на практике велики. С другой стороны, данные можно хранить в двоичной куче, тогда сложность составит $O(n\log n + m\log n)$. Но заметим, что время работы модификаций сократится по сравнению с классическим алгоритмом Дейкстры только при условии $m \ll n^2$, т. е. в случае разреженного графа. В рамках нашей задачи не всегда предполагается работа с разреженными социальными графами.

Также в случае, если число дуг в графе меньше числа вершин ($m < n$), то для поиска наикратчайшего пути используется алгоритм Беллмана — Форда. Его вычислительная сложность $O(mn)$, и в этом случае она меньше, чем у алгоритма Дейкстры.

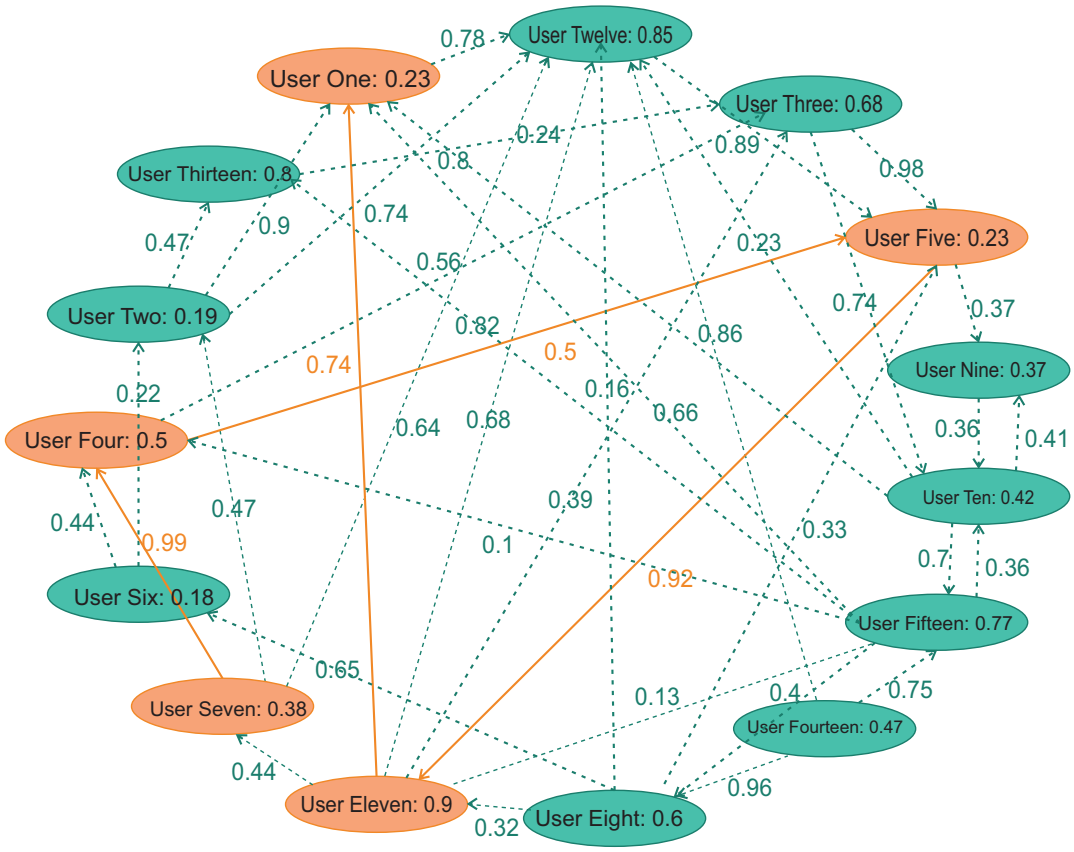
Таким образом, с учетом указанных особенностей наиболее подходящими вариантами являются алгоритмы Дейкстры и Беллмана — Форда. Данные алгоритмы позволяют обеспечить работу при ожидаемых вариациях исходных социальных графов сотрудников организаций. Для того чтобы добиться более быстрой работы алгоритма без ущерба для точности, были введены дополнительные условия. А именно, для уменьшения вычислительной сложности алгоритма установлено пороговое значение: если оценка вероятности успешного распространения социоинженерной атаки от начальной вершины до обрабатываемой в данный момент вершины становится меньше заданного порога, то он исключается из рассмотрения.

Реализация

Описанный подход был реализован в качестве дополнительного модуля комплекса для автоматизированного анализа защищенности пользователей информационных систем от социоинженерных атак, схема и описание которого представлены в работе [24]. Блок-схема алгорит-



■ ***Рис. 1.*** Блок-схема алгоритма поиска наиболее критичной траектории распространения социоинженерной атаки
■ ***Fig. 1.*** The flowchart of the search algorithm for the most critical trajectory of the social engineering attack spread



■ ***Рис. 2.*** Скриншот результата работы программного модуля поиска наиболее критичной траектории реализации социоинженерной атаки (выделенная траектория — наиболее критичная)

■ ***Fig. 2.*** Screenshot of the software works result for search module of the implementations most critical trajectory of a social engineering attack (the selected trajectory is the most critical)

ма, заложенного в реализации данного модуля, представлена на рис. 1. В качестве входных параметров используются идентификационные номера двух пользователей, наиболее критичную траекторию распространения социоинженерной атаки между которыми необходимо найти, а также социальный граф, получаемый с помощью одного из модулей указанного комплекса программ на основе данных, извлекаемых из контента, публикуемого пользователями в социальной сети ВКонтакте [4]. Результатом работы программного модуля является наиболее критичная траектория распространения социоинженерной атаки, оценка вероятности успеха прохождения по которой между двумя пользователями максимальна. Пример работы программного модуля представлен на рис. 2. Для визуализации результата работы модуля используется библиотека JGraphX.

Заключение

В статье предложен подход к выявлению наиболее критичной траектории распространения многоходовой социоинженерной атаки злоумышленника, оценка вероятности прохождения по которой максимальна. Произведен обзор алгоритмов, применение которых возможно при решении данной задачи, осуществлена адаптация выбранных алгоритмов к решению задачи поиска наиболее критичной траектории распространения социоинженерной атаки. Разработан алгоритм выявления наиболее критичной траектории и его реализация.

Практическая значимость полученных результатов заключается в расширении возможностей существующего программного комплекса для анализа защищенности пользователей информационных систем от социоинженерных атак. Разработанные метод и алгоритм лягут в основу решения задачи поиска наиболее критичных траекторий реализаций атак в информационной системе с учетом уровней доступа пользователей и критичности документов. Перспективы дальнейшего исследования также заключаются в разработке подходов к формализации и решению задач бэктрекинга инцидентов. Также перспективы дальнейших исследований могут быть связаны

с разработкой подходов к моделированию и оценке вероятностей успеха распространения многоходовых социоинженерных атак, в частности, может быть использован аппарат байесовских сетей [25].

Работа выполнена в рамках проекта по государственному заданию СПИИРАН № 0073-2018-0001 при финансовой поддержке РФФИ, проект № 18-01-00626 «Методы представления, синтеза оценок истинности и машинного обучения в алгебраических байесовских сетях и родственных моделях знаний с неопределенностью: логико-вероятностный подход и системы графов»; проект № 18-37-00323 «Социоинженерные атаки в корпоративных информационных системах: подходы, методы и алгоритмы выявления наиболее вероятных траекторий».

Литература

1. *По следам CyberCrimeCon 2017: Тенденции и развитие высокотехнологичной преступности*. https://habr.com/company/group-ib/blog/341812/ (дата обращения: 16.04.2018).
2. *ЦБ ожидает роста активности мошенников, использующих социальную инженерию.* https://ria.ru/economy/20171213/1510861611.html (дата обращения: 07.05.2018).
3. **Азаров А. А., Тулупьева Т. В., Суворова А. В., Тулупьев А. Л., Абрамов М. В., Юсупов Р. М.** Социоинженерные атаки: проблемы анализа. СПб., Наука, 2016. 349 с.
4. **Абрамов М. В., Тулупьев А. Л., Сулейманов А. А.** Задачи анализа защищенности пользователей от социоинженерных атак: построение социального графа по сведениям из социальных сетей. *Научно-технический вестник информационных технологий, механики и оптики*, 2018, т. 18, № 2, с. 313–321. doi:10.17586/2226-1494-2018-18-2-313-321
5. **Jaafor O., Birregah B.** Multi-layered graph-based model for social engineering vulnerability assessment. *Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on*, IEEE, 2015, pp. 1480–1488. doi:10.1145/2808797.2808899
6. **Yasin A., Liu L., Li T., Wang J., Zowghi D.** Design and preliminary evaluation of a cyber Security Requirements Education Game (SREG). *Information and Software Technology*, 2018, vol. 95, pp. 179–200. doi:10.1016/j..infsof.2017.12.002
7. **Junger M., Montoya L., Overink F. J.** Priming and warnings are not effective to prevent social engineering attacks. *Computers in Human Behavior*, 2017, vol. 66, pp. 75–87. doi:10.1016/j.chb.2016.09.012
8. **Li H., Luo X. R., Zhang J., Sarathy R.** Self-control, organizational context, and rational choice in Internet abuses at work. *Information & Management*, 2018, vol. 55, no. 3, pp. 358–367. doi:10.1016/j.im.2017.09.002
9. **Olifer D., Goranin N., Kaceniauskas A., Cenys A.** Controls-based approach for evaluation of information security standards implementation costs. *Technological and Economic Development of Economy*, 2017, vol. 23, no. 1, pp. 196–219. doi:10.3846/20294913.2017.1280558
10. **Bhakta R., Harris I. G.** Semantic analysis of dialogs to detect social engineering attacks. *Semantic Computing (ICSC), 2015 IEEE International Conference on*, IEEE, 2015, pp. 424–427. doi:10.1109/ICOSC.2015.7050843
11. **Cai Z., He Z., Guan X., Li Y.** Collective data-sanitization for preventing sensitive information inference attacks in social networks. *IEEE Transactions on Dependable and Secure Computing*, 2018, vol. 15, no. 4, pp. 577–590. doi:10.1109/TDSC.2016.2613521
12. **Edwards M., Larson R., Green B., Rashid A., Baron A.** Panning for gold: automatically analysing online social engineering attack surfaces. *Computers & Security*, 2017, vol. 69, pp. 18–34. doi:10.1016/j.cose.2016.12.013
13. **Albladi S. M., Weir G. R. S.** User characteristics that influence judgment of social engineering attacks in social networks. *Human-centric Computing and Information Sciences*, 2018, vol. 8, no. 1, pp. 5. doi:10.1186/s13673-018-0128-7
14. **Curtis S. R., Rajivan P., Jones D. N., Gonzalez C.** Phishing attempts among the dark triad: Patterns of attack and vulnerability. *Computers in Human Behavior*, 2018, pp. 174–182. doi:10.1016/j.chb.2018.05.037
15. **Dou Z., Khalil I., Khreishah A., Al-Fuqaha A., Guizani M.** Systematization of Knowledge (SoK): A systematic review of software-based web phishing detection. *IEEE Communications Surveys & Tutorials*, 2017, vol. 19, no. 4, pp. 2797–2819. doi:10.1109/COMST.2017.2752087
16. **Chiew K. L., Yong K. S. C., Tan C. L.** A survey of phishing attacks: their types, vectors and technical approaches. *Expert Systems with Applications*, 2018, pp. 1–20. doi:10.1016/j.eswa.2018.03.050
17. **Chin T., Xiong K., Hu C.** Phishlimiter: a phishing detection and mitigation approach using software-defined networking. *IEEE Access*, 2018, vol. 6, pp. 42516–42531. doi:10.1109/ACCESS.2018.2837889
18. **Gupta B. B., Tewari A., Jain A. K., Agrawal D. P.** Fighting against phishing attacks: state of the art and future challenges. *Neural Computing and Applications*, 2017, no. 12, vol. 28, pp. 3629–3654. doi:10.1007/s00521-016-2275-y
19. **Algarni A., Xu Y., Chan T.** An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook. *European Journal of Information Systems*, 2017, vol. 26, no. 6, pp. 661–687. doi:10.1057/s41303-017-0057-y

20. **Junger M., Montoya L., Overink F. J.** Priming and warnings are not effective to prevent social engineering attack. *Computers in Human Behavior*, 2017, vol. 66, pp. 75–87. doi:10.1016/j.chb.2016.09.012
21. **Levitin A.** *Introduction to the design & analysis of algorithms*. USA, Addison-Wesley, 2012. Pp. 304–337.
22. **Rassel S., Norvig P.** *Artificial intelligence: A modern approach*. London, Prentice-Hall International, 2009. Pp. 92–93.
23. **Cormen T. H., Leiserson C. E., Rivest R. L., Stein C.** *Introduction to algorithms*. Second Ed. MIT Press and McGraw-Hill, 2001. Pp. 580–642.
24. **Абрамов М. В.** *Методы и алгоритмы анализа защищенности пользователей информационных систем от социоинженерных атак: оценка параметров моделей*: дис. ... канд. техн. наук. СПб., СПИИРАН, 2018. 232 с.
25. **Харитонов Н. А., Березин А. И.** Синтез математического представления ациклической алгебраической байесовской сети. *Сборник докладов Международной конференции по мягким вычислениям и измерениям (SCM-2018)*, СПб., 2018, т. 1, с. 141–143.

UDC 614.8 + 002.6:004.89
doi:10.31799/1684-8853-2018-6-74-81

Search for the shortest trajectory of a social engeneering attack between a pair of users in a graph with transition probabilities

A. O. Khlobystova[a,b], Trainee, orcid.org/0000-0002-9811-5476
M. V. Abramov[a,b], PhD, Tech., Research Fellow, https://orcid.org/0000-0002-5476-3025
A. L. Tulupyev[a,b], Dr. Sc., Phys.-Math., Associate Professor, Senior Researcher, orcid.org/0000-0003-1814-4646
A. A. Zolotin[c], PhD, Phys.-Math, Senior Software Engineer, orcid.org/0000-0002-1028-4292
[a]Saint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation
[b]Saint-Petersburg State University, 7–9, Universitetskaya Emb., 199034, Saint-Petersburg, Russian Federation
[c]EPAM Systems GmbH, 56, Franklinstraße, 60486, Frankfurt am Main, Germany

Introduction: Social engineering attacks can be divided into two types: direct (one-way) and multi-pass ones, passing through a chain of users. Normally, there are several propagation paths for a multi-pass social engineering attack between two users. Estimates of the probabilities of an attack to spread along different trajectories will differ. **Purpose:** Identification of the most critical (most probable) trajectory for a multi-pass social engineering attack between two users. **Methods:** Methods of searching, matching and algorithm analysis are used to identify the most critical trajectory of attack propagation. They apply the information about the intensity of the interaction between employees in companies based on data extracted from social networks. These algorithms are reduced, using a number of transformations of the original data, to the algorithms of finding the shortest path in a graph. The estimates of a multi-path social engineering attack success probability are calculated with the methods of constructing an estimate of a complex event probability. **Results:** We have proposed an approach to identifying the most critical trajectories, whose estimate of the attack success probability is the highest. In the simplest case, the problem can be reduced to finding a path in the graph with the maximum product of the weights of all the edges involved. The resource intensity of the algorithm when searching for the most critical trajectory on a complete graph with a large number of vertices can be reduced with a specially developed technique. A brief overview of the methods and algorithms providing automated search for the most critical propagation path of a social engineering attack showed that in a general case it can be reduced, with some transformations, to the problem of finding the most critical trajectory using the configuration of Dijkstra and Bellman — Ford algorithms. The chosen algorithm was adapted for the specified context, and an approach was proposed to thin out the graph when searching for the most critical trajectory. The presented methods and algorithms are implemented in software code. Numerical experiments were performed to verify the calculation results. **Practical relevance:** The developed software based on the method and algorithm proposed in this article complements the functionality of the previous versions of software prototypes for analyzing the protection of information system users against social engineering attacks. It allows you to take into account a wider range of factors affecting the assessment of social engineering attack success probability.

Keywords — information security, social engineering attacks, user protection, multi-path social engineering attacks, attack trajectories, user's defeat success probability, social engineering, user security analysis, information security audit, security monitoring, company social graph, employee interaction intensity, social networks.

Citation: Khlobystova A. O., Abramov M. V., Tulupyev A. L., Zolotin A. A. Search for the shortest trajectory of a social engeneering attack between a pair of users in a graph with transition probabilities. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 6, pp. 74–81 (In Russian). doi:10.31799/1684-8853-2018-6-74-81

References

1. *Po sledam CyberCrimeCon 2017: Tendentsii i razvitie vysokotekhnologichnoj prestupnosti* [In the footsteps of CyberCrimeCon 2017: Trends and development of high-tech crime]. Available at: https://habr.com/company/group-ib/blog/341812/ (accessed 16 April 2018).
2. *CB ozhidaet rosta aktivnosti moshennikov, ispol'zuyushchih sotsial'nuyu inzheneriyu* [Central Bank expects growth in activity of fraudsters using social engineering]. Available at: https://ria.ru/economy/20171213/1510861611.html (accessed 07 May 2018).
3. Azarov A. A., Tulupyeva T. V., Suvorova A. V., Tulupyev A. L., Abramov M. V., Usupov R. M. *Sotsioinzhenernye ataki: problemy analiza* [Social engineering attacks: the problem of analysis]. Saint-Petersburg, Nauka Publ., 2016. 349 p. (In Russian).

4. Abramov M. V., Tulupyev A. L., Sulejmanov A. A. Problem of analysis of user protection from social engineering attacks: construction of the social graph on information from social network websites. *Nauchno-tekhnicheskij vestnik informatsionnyh tekhnologij, mekhaniki i optiki*, 2018, vol. 18, no. 2, pp. 313–321 (In Russian). doi:10.17586/2226-1494-2018-18-2-313-321
5. Jaafor O., Birregah B. Multi-layered graph-based model for social engineering vulnerability assessment. *Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on*, IEEE, 2015, pp. 1480–1488. doi:10.1145/2808797.2808899
6. Yasin A., Liu L., Li T., Wang J., Zowghi D. Design and preliminary evaluation of a cyber Security Requirements Education Game (SREG). *Information and Software Technology*, 2018, vol. 95, pp. 179–200. doi:10.1016/j..infsof.2017.12.002
7. Junger M., Montoya L., Overink F. J. Priming and warnings are not effective to prevent social engineering attacks. *Computers in Human Behavior*, 2017, vol. 66, pp. 75–87. doi:10.1016/j.chb.2016.09.012
8. Li H., Luo X. R., Zhang J., Sarathy R. Self-control, organizational context, and rational choice in Internet abuses at work. *Information & Management*, 2018, vol. 55, no. 3, pp. 358–367. doi:10.1016/j.im.2017.09.002
9. Olifer D., Goranin N., Kaceniauskas A., Cenys A. Controls-based approach for evaluation of information security standards implementation costs. *Technological and Economic Development of Economy*, 2017, vol. 23, no. 1, pp. 196–219. doi:10.3846/20294913.2017.1280558
10. Bhakta R., Harris I. G. Semantic analysis of dialogs to detect social engineering attacks. *Semantic Computing (ICSC), 2015 IEEE International Conference on*, IEEE, 2015, pp. 424–427. doi:10.1109/ICOSC.2015.7050843
11. Cai Z., He Z., Guan X., Li Y. Collective data-sanitization for preventing sensitive information inference attacks in social networks. *IEEE Transactions on Dependable and Secure Computing*, 2018, vol. 15, no. 4, pp. 577–590. doi:10.1109/TDSC.2016.2613521
12. Edwards M., Larson R., Green B., Rashid A., Baron A. Panning for gold: automatically analysing online social engineering attack surfaces. *Computers & Security*, 2017, vol. 69, pp. 18–34. doi:10.1016/j.cose.2016.12.013
13. Albladi S. M., Weir G. R. S. User characteristics that influence judgment of social engineering attacks in social networks. *Human-centric Computing and Information Sciences*, 2018, vol. 8, no. 1, p. 5. doi:10.1186/s13673-018-0128-7
14. Curtis S. R., Rajivan P., Jones D. N., Gonzalez C. Phishing attempts among the dark triad: Patterns of attack and vulnerability. *Computers in Human Behavior*, 2018, pp. 174–182. doi:10.1016/j.chb.2018.05.037
15. Dou Z., Khalil I., Khreishah A., Al-Fuqaha A., Guizani M. Systematization of Knowledge (SoK): A systematic review of software-based web phishing detection. *IEEE Communications Surveys & Tutorials*, 2017, vol. 19, no. 4, pp. 2797–2819. doi:10.1109/COMST.2017.2752087
16. Chiew K. L., Yong K. S. C., Tan C. L. A survey of phishing attacks: their types, vectors and technical approaches. *Expert Systems with Applications*, 2018, pp. 1–20. doi:10.1016/j.eswa.2018.03.050
17. Chin T., Xiong K., Hu C. Phishlimiter: A phishing detection and mitigation approach using software-defined networking. *IEEE Access*, 2018, vol. 6, pp. 42516–42531. doi:10.1109/ACCESS.2018.2837889
18. Gupta B. B., Tewari A., Jain A. K., Agrawal D. P. Fighting against phishing attacks: state of the art and future challenges. *Neural Computing and Applications*, 2017, vol. 28, no. 12, pp. 3629–3654. doi:10.1007/s00521-016-2275-y
19. Algarni A., Xu Y., Chan T. An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook. *European Journal of Information Systems*, 2017, vol. 26, no. 6, pp. 661–687. doi:10.1057/s41303-017-0057-y
20. Junger M., Montoya L., Overink F. J. Priming and warnings are not effective to prevent social engineering attack. *Computers in Human Behavior*, 2017, vol. 66, pp. 75–87. doi:10.1016/j.chb.2016.09.012
21. Levitin A. *Introduction to the design & analysis of algorithms*. USA, Addison-Wesley, 2012, pp. 304–337.
22. Rassel S., Norvig P. *Artificial Intelligence: A Modern Approach*. London, Prentice-Hall International, 2009, pp. 92–93.
23. Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. *Introduction to Algorithms*. Second Ed. MIT Press and McGraw-Hill, 2001, pp. 580–642.
24. Abramov M. V. *Metody i algoritmy analiza zashchishchennosti pol'zovatelej informacionnyh sistem ot socioinzhenernyh atak: ocenka parametrov modelej*. Dis. kand. tehn. nauk [Methods and algorithms for analyzing users' protection of information systems from social engineering attacks: estimation of model parameters. PhD tech. sci. diss.]. Saint-Petersburg, SPIIRAN Publ., 2018. 232 p. (In Russian).
25. Haritonov N. A., Berezin A. I. Acyclic Algebraic Bayesian network maths presentation synthesis. *Sbornik dokladov Mezhdunarodnoj konferencii po myagkim vychisleniyam i izmereniyam (SCM-2018)*, [XXI International Conference on Soft Computing and Measurement (CSM'2018)], Saint-Petersburg, 2018, vol. 1, pp. 141–143 (In Russian).