

Шульдешов Ю.Л., Ларионов Д.Ю.

МЕТОДЫ ОБРАБОТКИ ПРОСТРАНСТВЕННЫХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ЭКСПЕРТНОЙ СИСТЕМЫ

Шульдешов Ю.Л., Ларионов Д.Ю. Методы обработки пространственных данных с использованием экспертной системы.

Аннотация. Предложена модель и основные алгоритмы обработки пространственных данных с использованием экспертной системы. Показана принципиальная возможность повышения эффективности функционирования геоинформационных систем на их основе.

Ключевые слова: метод, гармонизация, интеграция, слияние, пространственные данные, интеллектуальные ГИС, модель хранения данных.

Shuldeshov U.L. Larionov D.U. Spatial data handling methods using expert systems

Abstract. A model and basic algorithms for processing spatial data using an expert system is proposed. The principal possibility of increasing the efficiency of geo-information systems based on these models and algorithms is shown.

Keywords: method, harmonization, integration, fusion, spatial data, intelligent GIS, data storage model.

1. Введение. В настоящее время геоинформационные системы (ГИС) все чаще начинают применяться для решения различных задач моделирования процессов и ситуаций, что позволяет говорить о новом классе интеллектуальных геоинформационных систем. Интеллектуальные ГИС осуществляют комплексный анализ и интерпретацию разнотипных данных и, как правило, включают в себя средства поддержки принятия решений.

Современный этап развития ГИС характеризуется следующими негативными тенденциями:

- увеличением количества участников, занимающихся сбором одних и тех же данных и их распространением;
- постоянным расширением круга геоинформационных приложений, типов ГИС-продуктов и форматов;
- дублированием усилий из-за трудностей доступа к существующим данным, разнородным качеством данных, собранных в разных проектах;
- ростом проблем при обмене и использовании данных, созданных различными организациями.

Эти тенденции определяют новые требования к ГИС по качеству и оперативности обработки данных.

Для удовлетворения этих требований предлагается совершенствовать процессы гармонизации, интеграции и слияния пространственных данных, под управлением экспертной системы.

Исследуемая задача ставится в следующем виде: в ГИС поступают данные от различных источников информации, в форматах, соответствующих спецификациям Open Geospatial Consortium, (OGC).[1] Эти форматы, в общем случае, не совпадают друг с другом и внутренним форматом хранения данных в ГИС. Необходимо осуществить преобразование поступающих данных во внутренний формат. Кроме этого, требуется предоставить данные из хранилища в виде, отвечающем требованиям конкретной категории пользователей, с устранением их избыточности и неоднозначности.

Необходимо разработать обобщенную модель обработки пространственных данных (ПД) в геоинформационной экспертной системе и алгоритмы основных преобразований над ними, обеспечивающие повышение качества данных, предоставляемых пользователям.

2. Модель обработки пространственных данных.

Поставленную задачу предлагается решать на основе расширения функций возлагаемых на экспертную систему интеллектуальной ГИС.

Под геоинформационной экспертной системой понимается совокупность взаимодействующих программных средств, использующих экспертные знания для обеспечения эффективного решения задач в области геоинформатики. Отличительной особенностью рассматриваемой ГИС является то, что экспертная система (ЭС) не является надстройкой над ГИС, позволяющей решать прикладные задачи, а встроена в схему обработки данных для повышения качества обработки данных.

Обработку пространственных данных с использованием такой ГИС рекомендуется осуществлять согласно схеме, приведенной на рис. 1.

Особенность этой схемы – модели обработки пространственных данных в том, что процесс обработки подразделяется на этапы гармонизации, интеграции и слияния.[2] При этом, на каждом из этих этапов реализуются ряд дополнительных новых функций позволяющих успешно выдавать пользователю требуемую информацию.

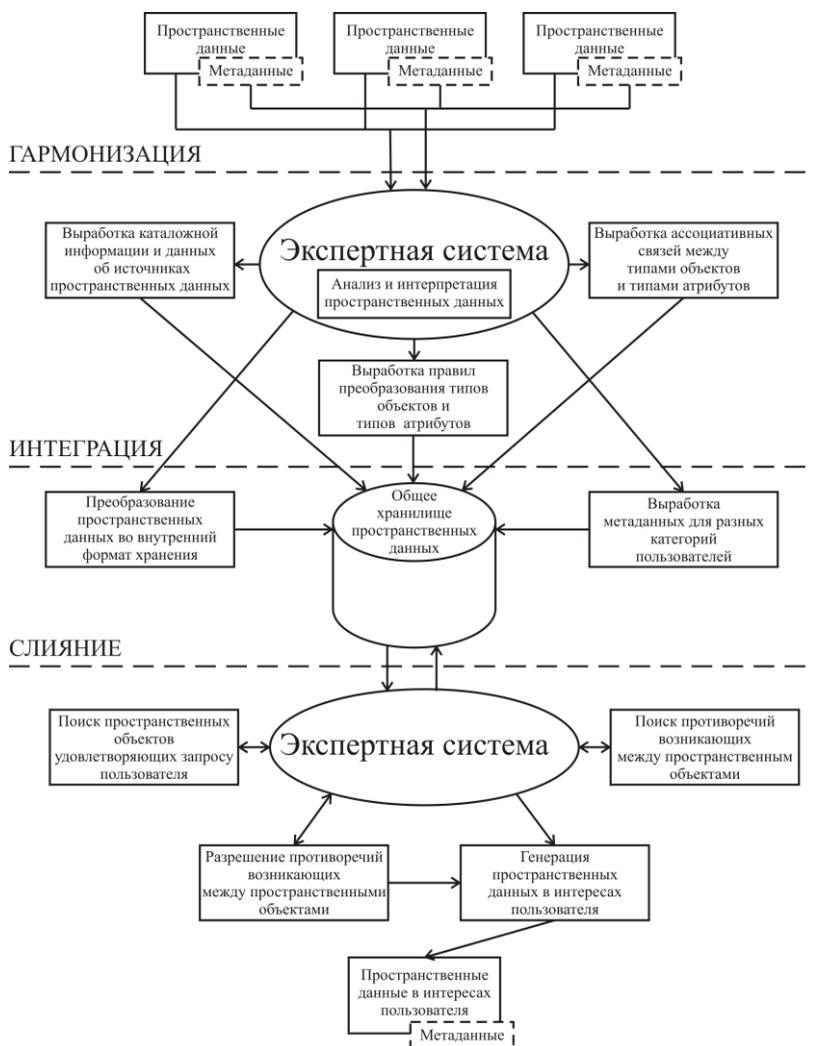


Рис. 1 Схема обработки пространственных данных под управлением экспертной системы.

К таким функциям на этапе гармонизации относятся выработка каталожной информации, выработка ассоциативных связей между объектами и их атрибутами, выработка правил преобразования типов

объектов и типов атрибутов.

В качестве новых функций на этапе интеграции выступают функция преобразования пространственных данных (ПД) во внутренний формат хранения и функция выработки метаданных для различных категорий пользователей.

В интересах совершенствования слияния предлагается дополнительно использовать функции поиска пространственных объектов, поиска противоречий между пространственными объектами, разрешения найденных противоречий и генерализации пространственных данных.

Рассмотрим более детально особенности алгоритмов, реализующих основные из этих функций.

3. Алгоритм гармонизация данных. Под гармонизацией пространственных данных понимается приведение их семантических характеристик, метрической системы в соответствие с другими пространственными данными для обеспечения взаимозаменяемости, взаимного понимания информации, содержащейся в пространственных данных.[3]

Процесс гармонизации (первый этап обработки пространственных данных) предполагает определение основных понятий и их взаимоотношений (онтологии) по соответствующим предметным областям и/или сферам ответственности. Обобщенный алгоритм гармонизации показан на рис. 2.

Гармонизация пространственных данных позволяет решать следующие основные задачи:

- обеспечение доступа к возможно большему числу первичных источников информации;
- возможность преобразования информации в удобный для пользователя вид (декодирование, распознавание, перевод);
- обеспечение доступа к существующим информационным ресурсам.

В ходе процесса гармонизации с использованием ЭС на основе анализа результатов интерпретации пространственных данных под управлением инженера знаний и экспертов создаются процедуры, стратегии, эмпирические правила. Эти знания встраивают в экспертную систему и на их основе вырабатывают каталожную информацию с систематизацией данных в структуре разрабатываемого каталога

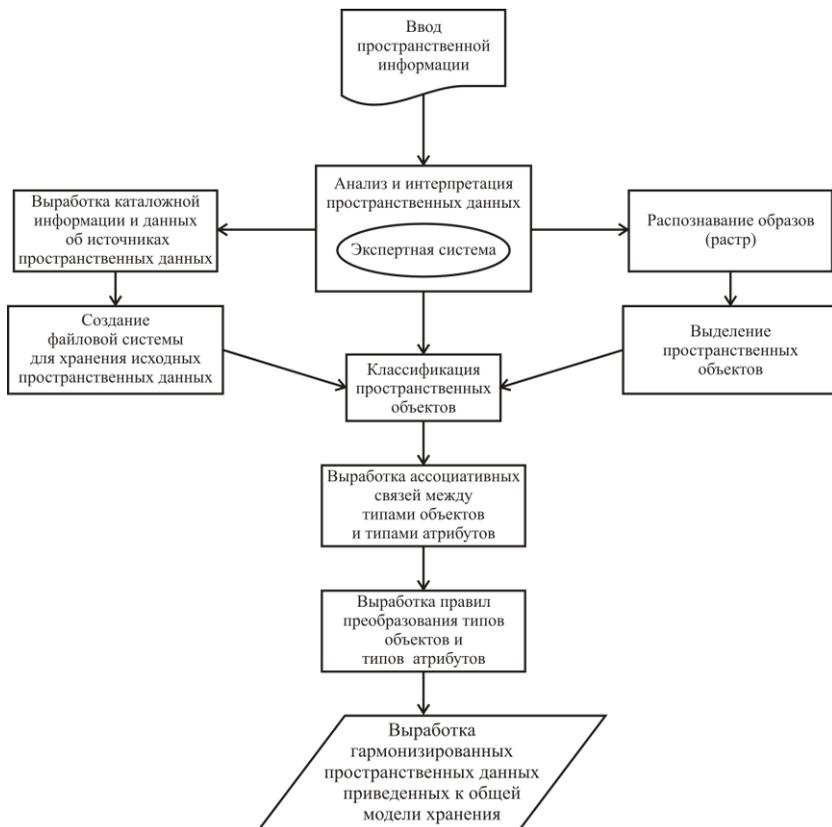


Рис. 2 Алгоритм гармонизации пространственных данных под управлением экспертной системы.

Структура каталога определяется исходя из: метаданных - идентификационной информации, данных о системе координат, о классификаторе, о распространителе и способе получения данных, об ограничениях на распространение метаданных и о периодичности их обновления.

По ассоциативным связям между типами объектов и их типами атрибутов выстраивают зависимость между пространственными объектами различных типов метаданных.

Ассоциативные связи строятся на основании анализа идентификационной информации, данных о системе координат и о классификаторе. При этом эксперты определяют правила и методы

преобразования пространственной информации, правила и методы преобразования атрибутивной информации. Таким образом создаются правила преобразования типов объектов и типов атрибутов

4. Алгоритм интеграция пространственных данных. Под интеграцией пространственных данных понимается разработка комплексных решений по автоматизации сбора и обработки поступающей пространственной информации от различных источников.[4,5] Обобщенный алгоритм интеграции пространственных данных показан на рис. 3.

Интеграция информации предполагает решение ряда задач:[6,7]

1. Определение источников и потребителей данных, выбор систем получения данных, определение их взаимосвязи и правил обновления информации.

2. Реализация и развитие форматов данных как на транспортном, так и на интерфейсном уровнях. Разработка и внедрение системы производства, распределения, защиты и корректуры данных в заданном формате.

3. Разработка лицензионных соглашений, авторских прав и статусов данных, разделение общей информации, организация защиты, копирования и защиты интеллектуальной собственности.

4. Организация финансирования различных работ и определение стоимости информации и предоставляемых услуг. Определение информационного рынка и стоимости, а также получаемой прибыли и ее распределение.

В ходе интеграции ЭС, также под управлением инженера знаний и экспертов, осуществляется:

– преобразование к внутреннему формату хранения данных, которое осуществляется по правилам, выработанным на этапе гармонизации;

– организация хранения данных на основании используемого каталога данных;

– выработка метаданных для различных категорий пользователей, которая производится на основании накопленных экспертных знаний в зависимости от характера деятельности категории пользователей, основной тематики осуществляемой ими работ, наиболее часто используемого масштаба, необходимой информационной насыщенности карты и т.д.



Рис. 3 Алгоритм интеграции пространственных данных под управлением экспертной системы.

5. Алгоритм слияния пространственных данных. На основании разработанных метаданных для каждой категории пользователей и построенных ассоциативных связей между типами объектов, их атрибутами, ЭС производит создание пространственных данных категории пользователей.

Рассмотрим процесс создания пользовательских данных подробнее:

Существуют наборы пространственных данных, полученных из различных источников и от различных производителей. Экспертной системой были подготовлены метаданные для каждой конкретной категории пользователей. По существующим правилам преобразования объектов и атрибутивной информации подготавливается набор пространственных данных для пользователя. При выполнении преобразования данных могут возникнуть следующие трудности:

– при сшивке по границам данных из разных источников или из разных наборов пространственный объект может «дробиться», как показано на рис. 4

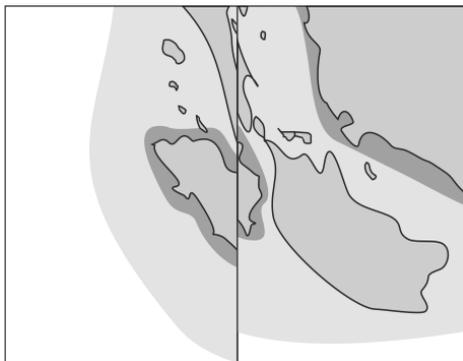


Рис. 4 Пример «дробления» пространственного объекта.

– при слиянии наборов данных, созданных для разных тематических приложений, объекты могут «дублироваться», как показано на рис. 5

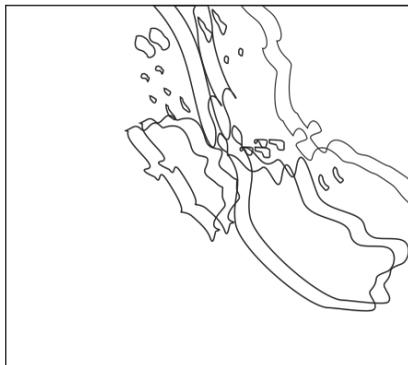


Рис. 5 Пример «дублирования» пространственного объекта.

– при совместном использовании растровой и векторной информации также может проходить как «дублирование», так и «дробление»;

Эти недостатки устраняются на следующем заключительном этапе обработки - слиянии, на котором пространственные данные приобретают новые качественные и количественные свойства.

Суть процесса слияния заключается в поиске нескольких пространственных объектов, которые описывают один реальный географический объект, и обработке найденных объектов таким образом, чтобы получился один пространственный объект, содержащий наиболее полную атрибутивную информацию и имеющий наиболее полно передающий реальный объект пространственную модель.

Обобщенный алгоритм слияния пространственных данных приведен на рис. 6.

Все имеющиеся пространственные данные сортируются по геометрическому типу («точечный», «линейный», «площадной», «коллекция геометрий»). Также выделяется массив данных, не имеющий геометрии и являющийся описательной или служебной информацией.

На этапе сортировки объектов по атрибутам ЭС производится отбор пространственных данных, у которых совпадают значения атрибутивной информации. Формируются массивы данных.

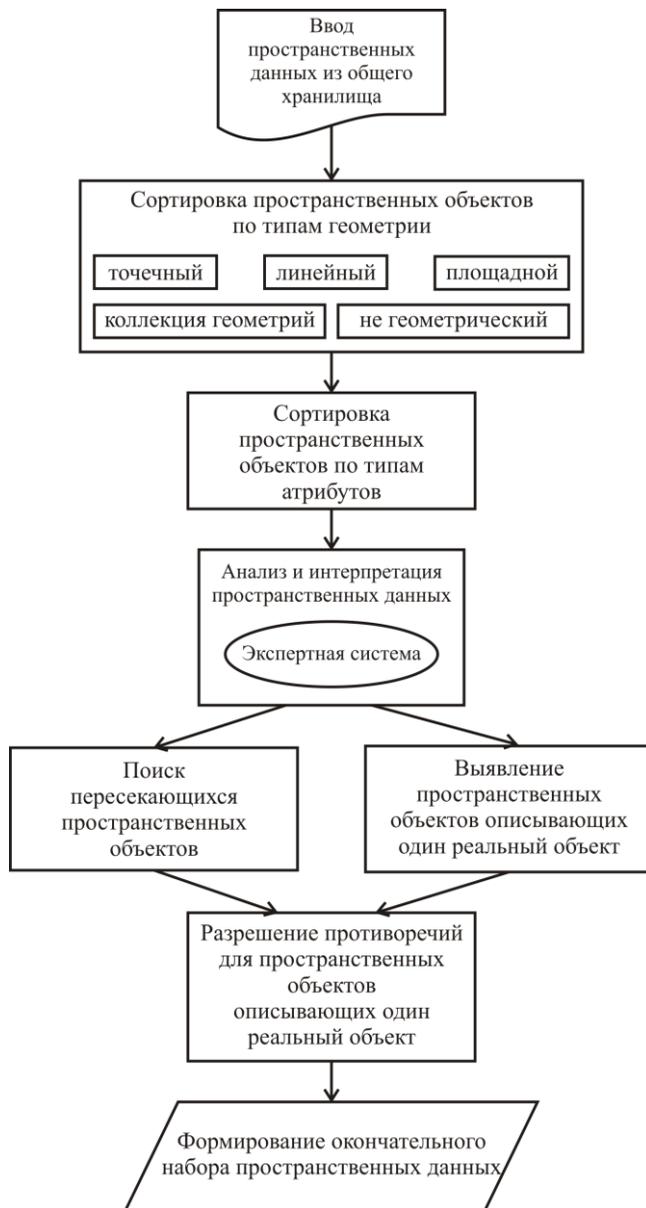


Рис. 6 Алгоритм слияния пространственных данных под управлением экспертной системы

Массивы данных, отсортированные по признаку геометрии, проходят дальнейший анализ с целью выявления таких объектов, которые различным образом взаимодействуют в пространстве. Они могут пересекаться, соприкасаться, накрываться друг другом и т.д. Отсортированные таким образом объекты группируются в отдельные массивы данных.

Для сокращения объема данных, находящихся в процессе обработки, необходимо выделить объекты, которые гарантировано, не должны подвергаться слиянию. Это можно сделать путем сравнения исходного массива пространственных данных и массивов данных, сформированных на первых этапах. Если объект из исходного массива данных не присутствует в сформированных массивах данных, можно утверждать, что данный объект является уникальным и единственным. Выделенные таким образом объекты ЭС может передавать на этап формирования окончательного набора данных.

Массивы данных, сформированные на третьем этапе (по пространственным признакам), проходят дальнейший анализ с целью выявления объектов, описывающих один реальный пространственный объект. Основу алгоритма составляет поиск по идентификационной и атрибутивной информации. Отсортированные объекты сканируются с целью выявления информации, пересекающейся в атрибутивных характеристиках. Информация об объектах анализируется, определяются признаки и характеристики, позволяющие выявить «дублированные» или «дробленные» объекты. Проводится поиск таких объектов, обработка их атрибутов и геометрий с целью объединения информации.

Массивы данных, отсортированные по атрибутивной информации, в свою очередь проходят обработку с целью устранения «дублирования». Такой обработке подвергаются также и данные, сформированные на третьем этапе (по пространственным признакам).

Подготовленные таким образом данные анализируются с целью контроля качества и их целостности.[8] Затем формируется окончательный набор данных для конкретной категории пользователей.

6. Результаты и их анализ. Предложенная схема и алгоритмы обработки пространственных данных использовались при разработке информационной системы обработки и распространения картографической информации для военно-морского флота «Багрень». Введение рассмотренных алгоритмов в процесс создания картографических наборов позволило сократить объем циркулирующей в системе информации до 75% от поступающего потока. Удалось оптимизировать выдачу данных в системы

навигационно-гидрографического и гидрометеорологического обеспечения. Объем выдаваемой информации уменьшился до 65% без потери качества. Сократилось время поиска и выдачи потребителю требуемой информации.

7. Заключение. Таким образом, внедрение в ГИС элементов искусственного интеллекта для обработки пространственной информации на всех этапах (гармонизации, интеграции и слияния), позволяет:

- повысить качество и оперативность обработки данных за счет реализации новых функций;
- совершенствовать алгоритмы обработки пространственных данных;
- в наибольшей мере соответствовать требованиям пользователя по составу и скорости представления запрашиваемой информации.

Литература

1. В. Андрианов, Стандарты в ИПД, ArcReview № 2 (37) 2006
2. Попович В.В., Воронин М.Н., Морской Сборник, №5, Май 2007, Гармонизация, интеграция и слияние данных – три источника и три составных части геоинформационных технологий
3. Kristine Asch, Boyan Brodaric, and etc. An International Initiative for Data Harmonization in Geology. //10th EC-GI&GIS Workshop, Abstracts, Warsaw, 23-25 June, 2004, p.9.
4. L. Valet, G. Mauris, and Ph Bolon, “A statistical overview of Recent Literature in Information Fusion,” Fusion 2000. IEEE AES March 2001.
5. White, F.E., “A Model for Data Fusion”, Proc. 1st National Symposium on Sensor Fusion, 1988
6. E., Blasch, “Fundamentals of Information Fusion and Applications”, Tutorial, TD2, Fusion 2002.
7. James Llinas, et al. “Revising the GDL Data Fusion Model II” 2005.
8. Осипов В.Ю., Кох С.А. «Управление избыточностью данных в геоинформационных системах» XII Санкт-Петербургская международная конференция «Региональная информатика (РИ-2010)». Санкт-Петербург, 20-22 октября 2010 г.: Труды конференции. СПОИСУ.-СПб.

Шульдешов Юрий Леонидович — канд. техн. наук, старший научный сотрудник лаборатории объектно-ориентированных геоинформационных систем Учреждения Российской Академии наук, Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН). Область научных интересов: геоинформационные

системы, интеллектуальные информационные технологии. Число научных публикаций — 55, shuldeshov@oogis.ru, <http://oogis.ru>; 14-линия В.О., д.39, Санкт-Петербург, 199179, РФ; р.т. +7(812)328-0179, факс +7(812)329-0863.

Shuldeshov Yuri Leonidovich — PhD. Tekhn. Sciences, Senior Researcher, Laboratory of object-oriented geo-information systems Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation (SPIIRAS). Research interests: Geographic Information Systems, Intelligent Information Technology. The number of scientific publications - 120, shuldeshov@oogis.ru, <http://oogis.ru>; 39, 14-th line V.O., St.Petersburg 199179, Russia; office phone +7 (812) 328-0179, fax +7 (812) 329-0863.

Ларионов Денис Юрьевич — программист ЗАО «СПИИРАН-НТБВТ». Область научных интересов: геоинформационные системы, интеллектуальные информационные технологии. larionov@oogis.ru, <http://ntb.oogis.ru>; 14-линия В.О., д.39, Санкт-Петербург, 199179, РФ; р.т. +7(812)328-0179, факс +7(812)329-0863. Научный руководитель — Ю.Л. Шульдешов.

Larionov Denis Yur'evich — programmer ZAO SPIIRAS-NTBVT Research interests: geographic information systems, intelligent information technology larionov@oogis.ru, <http://ntb.oogis.ru>; 39, 14-th line V.O., St. Petersburg 199179, Russia; office phone +7 (812) 328-0179, fax +7 (812)329-0863. Supervisor — Y.L. Shuldeshov.

РЕФЕРАТ

Шульдешов Ю.Л., Ларионов Д.Ю. **Методы обработки пространственных данных с использованием экспертной системы.**

В связи с расширением областей применения ГИС появилась насущная потребность в разработке методов обработки пространственных данных под управлением экспертной системы. Их рекомендуется применять поэтапно в следующей последовательности:

1 этап – гармонизация (позволяет привести пространственные данные к общей модели хранения).

2 этап – интеграция (позволяет автоматизировать сбор и обработку пространственной информации от различных источников).

3 этап – слияние (позволяет исключить дублирование информации, не снижая ее качества и целостности).

При этом, на каждом из этих этапов реализуются ряд дополнительных новых функций, позволяющих успешно выдавать пользователю требуемую информацию.

Приведена схема обработки пространственных данных под управлением экспертной системы, а также алгоритмы обработки пространственных данных на всех рассматриваемых этапах.

Выводы: Таким образом, внедрение в ГИС элементов искусственного интеллекта для обработки пространственной информации, позволяет:

- повысить качество и оперативность обработки данных за счет реализации новых функций;
- совершенствовать алгоритмы обработки пространственных данных;
- в наибольшей мере соответствовать требованиям пользователя по составу и скорости представления запрашиваемой информации.

Данный подход является универсальным и может быть применим во всех областях принятия решения с использованием ГИС. Подтверждением этого является использование моделей и алгоритмов обработки пространственных данных при разработке информационной системы картографической информации для Военно-морского флота - шифр «Багрень-2». Введение рассмотренных алгоритмов в процесс создания картографических наборов позволили сократить объем циркулирующей в системе информации до 75% от поступающего потока. Удалось оптимизировать выдачу данных в системы навигационно-гидрографического и гидрометеорологического обеспечения. Объем выдаваемой информации уменьшился до 65% без потери качества. Сократилось время поиска и выдачи потребителю требуемой информации.

SUMMARY

Shuldeshov U.L., Larionov D.U. **Spatial data handling methods using expert systems.**

In connection with the extension of GIS application domain there is a vital need to develop methods for processing spatial data managed by an expert system. They are recommended for use in the following sequence:

Stage 1 - Harmonization (allows you to bring spatial data to a common storage model).

Stage 2 - Integration (allows you to automate the spatial data processing from different sources).

Stage 3 - Fusion (eliminates duplication of information without decrease of quality and integrity).

Each of these stages allows to implement a number of additional new features that allow the user to successfully use the required information.

The scheme for processing spatial data managed by an expert system, as well as algorithms for processing spatial data at all stages is proposed.

Conclusions: Thus, the introduction of the GIS elements of artificial intelligence for spatial information processing, allows to:

- improve the quality and timeliness of data processing through the implementation of new functions;
- develop algorithms for processing spatial data;
- meet user requirements on the composition and rate of submission of the requested information.

This approach is versatile and can be applicable in all domains of decision making using GIS. Proof of this is the use of models and algorithms for spatial data processing in the development of information system mapping information for the Navy - code "Bagren-2." The application of the algorithms in the process of creating of map sets has reduced the amount of information circulating in the system to 75% of the incoming stream. The presented approach allows to optimize the issuance of data in the system of navigational, hydrographic and hydrometeorological services. The volume of the release information has decreased to 65% without loss of quality. Time of searching of information and its delivery to the consumer has been diminished.