На правах рукописи

ГИБАДУЛЛИН Руслан Фаршатович

СИСТЕМА БАЗ ДАННЫХ КАРТОГРАФИИ С АССОЦИАТИВНОЙ ЗАЩИТОЙ

Специальность: 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени кандидата технических наук

Уфа – 2011

Работа выполнена на кафедре компьютерных систем
ГОУ ВПО «Казанский государственный технический университет
им. А. Н. Туполева»

Научный руководитель	д-р физ.-мат. наук, проф. **Райхлин Вадим Абрамович**
Официальные оппоненты	д-р тех. наук, проф. **Васильев Владимир Иванович,** каф. вычислительной техники и защиты информации, Уфимский государственный авиационный технический университет
	д-р физ.-мат. наук, доц. **Ишмухаметов Шамиль Талгатович,** каф. системного анализа и информационных технологий, Казанский федеральный университет
Ведущая организация	**Институт информатики Академии наук Республики Татарстан, г. Казань**

Защита состоится «25» марта 2011 г. в 10 часов
на заседании диссертационного совета Д-212.288.07
при Уфимском государственном авиационном техническом университете
по адресу: 450000, Уфа-центр, ул. К. Маркса, 12.

С диссертацией можно ознакомиться в библиотеке университета.
Автореферат разослан «24» февраля 2011 г.

Ученый секретарь
диссертационного совета
д-р техн. наук, проф. **С. С. Валеев**

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

В настоящее время во всем мире широко используются средства пространственного анализа данных различными структурами и учреждениями. В западных странах цифровые модели местности создаются в больших количествах и используются повсеместно. В нашей стране с вступлением в силу федерального закона «О навигационной деятельности» от 14.02.2009, снимающего ограничения на точность определения координат объектов навигационной деятельности, общедоступные картографические данные в удобных для использования цифровых форматах только начинают появляться. С распространением навигационных устройств усилится процесс внедрения и распространения ГИС, рынок станет больше насыщаться цифровыми данными картографии.

Картографическая продукция имеет свои особенности. Она отличается высокой себестоимостью работ по ее получению: топографами и геодезистами затрачиваются немалые усилия на формирование этой продукции. Кроме того, карты могут содержать конфиденциальные сведения, например, новые места локализации нефтегазовых и урановых месторождений, цветных металлов. Поэтому задача их защиты актуальна.

Использование цифровых карт сопряжено с определенными трудностями. Оно требует векторизации растровых изображений, устранения дефектов топологической структуры при создании таких карт, организации хранения цифровой модели местности в реляционных базах данных.

На сегодняшний день существуют СУБД, обладающие встроенными механизмами защиты баз данных. Ведущие места среди них занимают: *Oracle, Microsoft SQL Server, Sybase Adaptive Server, Gupta SQLBase Treasury Edition*. Перечисленные СУБД достаточно универсальны. Построение специализированных СУБД, ориентированных на работу с защищенными БД картографии, может существенно повысить эффективность управления такими БД по критерию быстродействия при требуемом уровне стойкости защиты.

Объект исследования – специализированная система управления базами данных картографии с ассоциативной защитой.

Предмет исследования – архитектура, стегопараметры, алгоритмическое и программное обеспечение системы управления базами данных картографии с ассоциативной защитой.

Цель диссертационной работы – разработка модели, метода и алгоритмов управления базами данных картографии с ассоциативной защитой.

Основные задачи диссертационной работы

Для достижения поставленной цели в диссертационной работе решаются следующие задачи:

1. Разработка фреймовой модели системы баз данных картографии с ассоциативной защитой.
2. Разработка алгоритма формирования стегоконтейнера.
3. Оптимизация значений стегопараметров по критерию быстродействия.
4. Разработка схемы защищенной БД картографии, метода локальной обработки запросов без раскрытия защищенной БД картографии в целом и алгоритмов интер-

претации пользовательских запросов системы баз данных картографии с ассоциативной защитой.

5. Реализация исследовательского прототипа СУБД *Security MapPointCluster* для проверки на его основе рекомендаций к построению системы баз данных картографии с ассоциативной защитой.

Методы исследования

Исследования проводились с привлечением теории баз данных, теории случайных процессов, защиты информации, элементов теории ассоциативной защиты стилизованных бинарных изображений, проектирования кластерных архитектур, компьютерного моделирования.

Основные научные результаты, полученные автором и выносимые на защиту

1. Фреймовая модель системы баз данных картографии с ассоциативной защитой.
2. Алгоритм генерации стегоконтейнера.
3. Результаты оптимизации стегопараметров по критерию быстродействия.
4. Схема защищенной БД картографии, метод локальной обработки запросов без раскрытия защищенной БД картографии в целом и алгоритмы интерпретации пользовательских запросов системы баз данных картографии с ассоциативной защитой.
5. Исследовательский прототип СУБД *Security MapPointCluster*.

Научная новизна работы состоит в следующем:

1. Предложена модель системы баз данных картографии с ассоциативной защитой, которая позволяет повысить эффективность управления защищенными БД картографии в сравнении с известными СУБД по критерию быстродействия.
2. Разработан алгоритм формирования стегоконтейнера, на его основе подтверждена гипотеза о принципиальной возможности достижения безусловной стойкости ассоциативного метода стегозащиты. Оптимизированы значения стегопараметров алгоритма формирования стегоконтейнера по критерию быстродействия и показана предпочтительность использования в системе баз данных картографии с ассоциативной защитой разработанного алгоритма в режиме безальтернативного выбора гаммы в сравнении с ГОСТ 28147-89.
3. Предложен эффективный метод локальной обработки запросов к защищенной БД картографии и на его основе разработаны алгоритмы интерпретации пользовательских запросов, отличающиеся от существующих тем, что позволяют обрабатывать запросы к защищенной БД картографии без ее полного раскрытия.

Обоснованность и достоверность результатов диссертации

Обоснованность результатов, полученных в диссертационной работе, базируется на использовании апробированных научных положений и методов исследования.

Достоверность исследуемой модели и справедливость сформулированных утверждений подтверждены экспериментально на специально разработанном для этой цели инструментальном средстве.

Практическая ценность работы

Предложенная модель системы баз данных картографии с ассоциативной защитой позволяет повысить эффективность управления защищенными БД картографии в

сравнении с известными СУБД по критерию быстродействия.

Предложенный метод локальной обработки запросов к защищенной БД картографии и разработанные на его основе алгоритмы интерпретации пользовательских запросов позволяют обрабатывать запросы к защищенной БД картографии без ее полного раскрытия.

Даны практические рекомендации к построению системы баз данных картографии с ассоциативной защитой. Для проверки этих рекомендаций разработан исследовательский прототип СУБД *Security MapPointCluster*.

Результаты исследования внедрены в учебный процесс КГТУ им. А.Н.Туполева и использованы в ООО «Геодезическая компания «ЗЕНИТ».

Апробация работы

Основные результаты работы докладывались и обсуждались на Международной молодежной научной конференции «Туполевские чтения» (Казань, 2006, 2008–2010), республиканском научном семинаре АН РТ «Методы моделирования» (Казань, 2007–2010), Всероссийской научной конференции «Техническая кибернетика, радиоэлектроника и системы управления» (Таганрог, 2008), Международной научно-методической конференции «Информатика: проблемы, методология, технологии» (Воронеж, 2008), Международной конференции «Высокопроизводительные параллельные вычисления на кластерных системах» HPC–2008,2009 (Казань, 2008; Владимир, 2009), семинаре кафедры радиотехнических и медико-биологических систем Марийского ГТУ (Йошкар-Ола, 2010).

Публикации

Результаты диссертационной работы отражены в 15 публикациях, в том числе в 6 научных статьях, в 1 статье в рецензируемом журнале из списка периодических изданий, рекомендованных ВАК, в 8 материалах конференций.

Структура и объем работы

Диссертационная работа состоит из введения, четырех глав, заключения, приложений, библиографического списка и изложена на 117 страницах машинописного текста. Библиографический список включает 94 наименования литературы.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы диссертации, определяются цель и задачи исследования, приводится перечень основных результатов, выносимых на защиту. Дается структура диссертации.

В первой главе подчеркивается целесообразность организации защиты картографических данных на базе *SQL*-серверов. Проводится анализ современного состояния исследований по защищенным картографическим СУБД. Строится фреймовая модель системы баз данных картографии с ассоциативной защитой. Выделяются объекты исследований.

Известно, что система защиты картографических данных на базе *SQL*-сервера потенциально более производительна по сравнению с традиционной, когда защита строится созданием программной надстройки для ГИС.

Согласно общим сведениям о ГИС Панорама, разработанной с учетом широкого круга пользовательских потребностей, векторная карта может содержать несколько

тысяч листов. Всего один лист карты может содержать до 4 миллиардов объектов, а объем векторной карты может достигать до нескольких терабайт. Задачи, связанные с организацией работ с такими объемами защищенных БД картографии, чрезвычайно ресурсоемки. Их решения «в реальном времени» ассоциируются с малоизученным вопросом построения соответствующих СУБД кластерного типа.

Систему баз данных картографии с ассоциативной защитой целесообразно разделить на два уровня: нижний уровень (серверная сторона), где будут решаться задачи генерации, модификации защищенной БД картографии и передачи выборочных частей БД на верхний уровень (клиентскую сторону), на котором будет происходить обработка пространственных запросов к полученным частям БД картографии. Такая организация работы с защищенной БД картографии позволит избежать избыточной вычислительной нагрузки сервера.

Построение фреймовой модели системы баз данных картографии с ассоциативной защитой в диссертационной работе ограничено случаем защиты точечных объектов картографии, за исключением таких фреймов, как схема и принципы формирования ЗБДК. Исследование этих фреймов расширяется на случай защиты линейных и площадных объектов картографии. Модель системы баз данных картографии с ассоциативной защитой представляется иерархией фреймов (рис. 1). Такая модель допускает упрощение с учетом мирового опыта построения машин баз данных, тенденций использования перспективных инструментальных средств, таких как язык реляционного исчисления *SQL*, СУБД *MySQL*, сеть *Ethernet* с коммутатором, технология параллельного программирования *MPI*, геоинформационная система *MapInfo*, программа очистки следов работы пользователя *Acronis Privacy Expert*. Соответственно часть исходной модели исключается из рассмотрения. На рисунке во фреймах-«листьях» подчеркнуты лишь те фреймы, раскрытие которых составляет предмет исследований в диссертации.

Фрейм «Нижний уровень» составляют следующие дочерние фреймы.

Языковый интерфейс (ЯИ). Язык запросов включает ограниченное подмножество SQL-запросов (селекция, добавление, удаление, изменение) к пользовательской схеме базы данных.

Реляционная модель данных (РМД) включает схему базы данных (СБД) и принципы формирования базы данных (ПФБД).

Информационная безопасность (ИБ) содержит компоненты обеспечения безопасности данных от угрозы несанкционированного доступа: кластеризация (Класт), добавление «пустых» объектов (ДПО), перемешивание (Перемеш), идентификация и аутентификация пользователей в системе (Идентиф/Аутент), разграничение доступа пользователей к тематическим слоям карты (РДП), сокрытие и раскрытие. Сокрытие включает компоненты: двумерно-ассоциативный принцип сокрытия информации (маскирование), алгоритмы *Stegomask* и *Stegomask alternativeless*, оптимизация стегопараметров (ОС). Аутентификация и идентификация пользователей в системе достигается средствами СУБД *MySQL*.

Организация обработки запросов (ООЗ) включает компоненты: синтаксический и семантический анализ запросов (СиСАЗ), стратегия параллельной обработки запросов (ПОЗ), интерпретатор пользовательского языка запросов, локальная обработка запро-

сов без раскрытия защищенной базы данных картографии в целом (ЛОЗ без раскрытия ЗБДК в целом). Для синтаксического и семантического анализа запросов применяется СУБД *MySQL*. Для параллельной обработки запросов используется стратегия *MPP* с технологией обмена сообщениями *MPI*-1 на базе сети *Gigabit Ethernet* с коммутатором в операционной среде *Windows*.

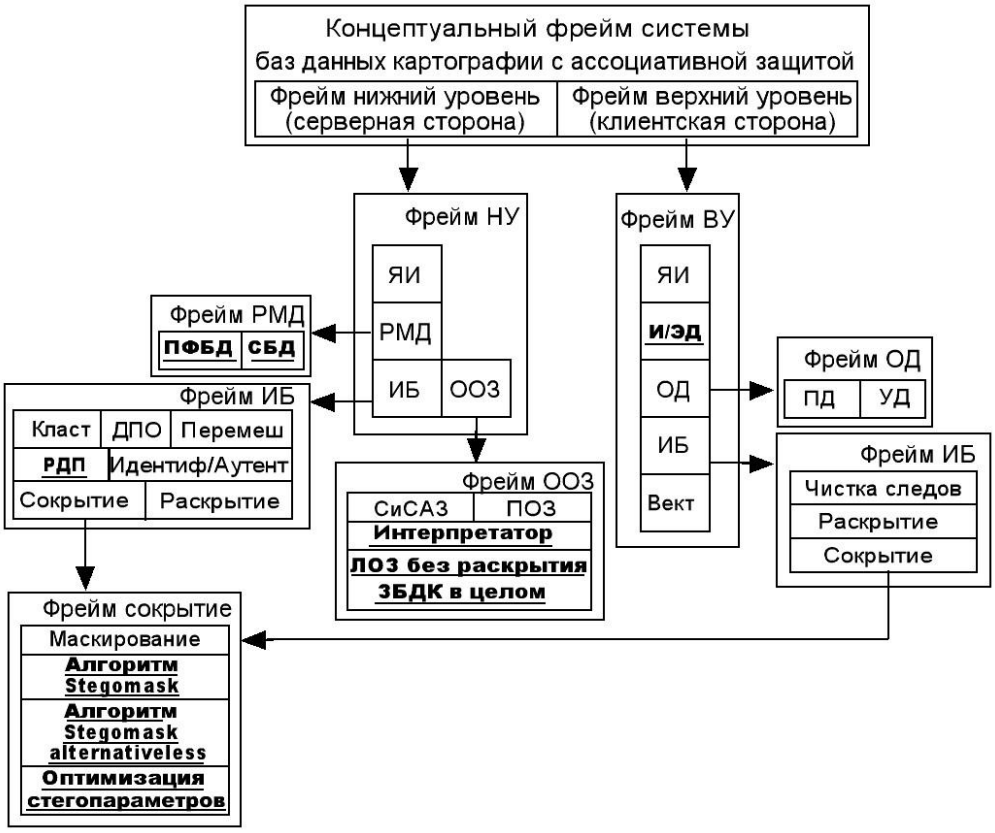


Рисунок 1 – Фреймовая модель системы баз данных картографии с ассоциативной защитой

Фрейм «Верхний уровень» составляют следующие дочерние фреймы.

*Я*зыковый интерфейс (*ЯИ*). *Я*зык запросов *SQL-MM*, расширенных пространственными запросами типа площадь, расстояние, периметр, координаты, вложенность, соседство, пересечение и т.д.

Импорт и экспорт данных (И/ЭД). Импорт – это преобразование векторных карт из файловой модели данных верхнего уровня в реляционную модель данных нижнего уровня. Экспорт – это преобразование векторных карт из реляционной модели данных нижнего уровня в файловую модель данных верхнего уровня.

Организация данных (ОД) включает компоненты представление данных (ПД) и управление данными (УД). Пользовательские данные хранятся в файлах формата *MID*/*MIF* (формат файлов *MapInfo*). Их структура описана в руководстве пользователя ГИС *MapInfo*. Данные выдаются пользователю в виде окон: карта, список, график и отчет. Обработка данных на клиентской машине пользователя осуществляется с помощью обработчика, встроенного в ГИС *MapInfo*.

Информационная безопасность (ИБ) включает такие компоненты, как чистка следов работы пользователя, сокрытие, раскрытие. Очищение следов работы на ком-

пьютере происходит с помощью программы *Acronis Privacy Expert*. В ней пользователю на выбор предоставляется список строгих алгоритмов гарантированного уничтожения конфиденциальной информации, которые соответствуют наиболее известным национальным стандартам, например, американские национальные стандарты *DoD* 5220.22-*M* и *NAVSO P*-5239-26 (*RLL*), немецкий национальный стандарт *VSITR*, российский национальный стандарт ГОСТ Р50739-95 и др.

Векторизация (Вект). Для векторизации растровых слоев карты с выделением конфиденциальных объектов используется инструментарий ГИС *MapInfo*.

Фреймы модели системы баз данных картографии с ассоциативной защитой, составляющих предмет научных исследований в диссертации, являются: алгоритм формирования стегоконтейнера; оптимизация значений стегопараметров по критерию быстродействия; схема и принципы формирования ЗБДК; локальная обработка запросов без раскрытия ЗБДК в целом; интерпретация пользовательских запросов в системе.

Во второй главе предлагается алгоритм генерации стегоконтейнера, приводятся результаты исследований по оптимизации его стегопараметров по критерию быстродействия и оценки эффективности реализаций алгоритма по сравнению с ГОСТ 28147-89.

Известно, что идеальный способ картографической защиты обладает свойством безусловной стойкости (совершенной секретности). Условие совершенной секретности К. Шеннон определяет следующим образом: для всех передаваемых сообщений их апостериорные вероятности должны быть равны априорным вероятностям независимо от величины последних. В таком случае перехват сообщения не дает противнику никакой информации. В работе принята логическая трактовка критерия К. Шеннона: если в итоге применения всевозможных ключей к любому сокрытому кластеру получаем неединичное подмножество равноправдоподобных результатов раскрытия, то использованный метод стегозащиты безусловно стоек.

Достаточным условием его удовлетворения является выбор такого контейнера (гаммы), что при полноте множества кодов объектов (координат) и ограниченном переборе ключей любой сокрытый в нем код объекта (координата) представит все это множество.

Выбор сравнительно малых оснований исчисления γ при кодировании сообщений повышает быстродействие метода, но таит угрозу раскрытия ключа несанкционированным пользователем, знакомым с участком местности. Поэтому за основу рассмотрения принимается случай $\gamma = 10$.

В ранее проделанных работах осталась неподтвержденной гипотеза о принципиальной возможности достижения безусловной стойкости выбранного подхода к защите при $\gamma = 10$: подходящий контейнер для обеспечения безусловной стойкости в случае больших оснований ($\gamma = 10$) всегда существует.

Для подтверждения этой гипотезы разработан алгоритм формирования стегоконтейнера с выбором подходящей гаммы для случая $k = 3$ и $\gamma = 10$, где k – разрядность кодового слова.

Алгоритм формирования стегоконтейнера (алгоритм *Stegomask*):

Инициализация алгоритма:

С помощью алгоритма формирования масок генерируется подмножество различных стегоключей $R = \{R_1, R_2, \ldots, R_K\}$.

Для множества эталонов

$E = \{E^1, E^2, \ldots, E^{10}\}$, $E^t = \left| e^t_{pq} \right|$, $e^t_{pq} \in \{0, 1\}$,

$p = 1 \ldots m$, $q = 1 \ldots (2m-1)$, $t \in \{\overline{1, 10}\}$,

генерируется набор масок (секретный стегоключ)

$S = \{S^1, S^2, \ldots, S^{10}\}$;

$$S^t = \left| s^t_{pq} \right|; \quad s^t_{pq} = \begin{cases} 1, & e^t_{pq} - \text{значимый бит эталона;} \\ 0, & e^t_{pq} - \text{незначимый бит эталона.} \end{cases}$$

Где значимость битов в эталонах определяется алгоритмом формирования масок псевдослучайно.

Формирование стегоконтейнера:

Шаг 1. Вектор состояния ГПСП инициализируется псевдослучайно на основе текущего системного времени.

Шаг 2. Матрицы $G^r = | g^r_{pq} |$, $r \in \{\overline{1, 10}\}$ инициализируются битами ПСП построчно слева направо.

Шаг 3. Для заданного трехразрядного десятичного кода, формируется стегоконтейнер $C = \{C^{t_2}; C^{t_1}; C^{t_0}\}$, где t_0, t_1, t_2 – нулевой, первый и второй разряды десятичного кода соответственно.

$$C^{t_i} = | c^{t_i}_{pq} |, \; c^{t_i}_{pq} = s^{t_i}_{pq} \& e^{t_i}_{pq} \vee \bar{s}^{t_i}_{pq} \& g^{t_i}_{pq}, \; i = 0,1,2.$$

Шаг 4. НАЧАЛО ШАГА
| НАЧАЛО ЦИКЛА для j от 1 до K
| | buf_id[Ident(C,E,R_j)] = 1
| | ЕСЛИ все $\gamma^k = 1000$ элементов буфера идентификаций
| | buf_id = 1
| | ТО перейти в КОНЕЦ ШАГА
| КОНЕЦ ЦИКЛА
| ЕСЛИ какой-либо из элементов buf_id = 0
| ТО перейти к шагу 1
КОНЕЦ ШАГА

Замечание: функция Ident(C,E,R_j) распознает код, сокрытый в стегоконтейнере C, на основании заданных эталонов E и стегоключа R_j.

Шаг 5. Полученный стегоконтейнер C считать результатом. КОНЕЦ.

Процесс формирования стегоконтейнера заключается в следующем. Для стегоконтейнера с псевдослучайно подобранной гаммой организуется перебор подмножества ключей мощностью *K* (т.е. полный перебор ключей не проводится) и осуществляется идентификация кода в стегоконтейнере по каждому ключу.

Если на полученном множестве идентификаций для данного стегоконтейнера не будет выявлен хотя бы один из всевозможных кодов, то текущий стегоконтейнер не

удовлетворяет логической трактовке критерия К. Шеннона. В таком случае осуществляется генерация стегоконтейнера с иной гаммой.

Этот процесс продолжается до тех пор, пока не будет выявлен подходящий стегоконтейнер.

В разделе 2.2 «Оптимизация значений стегопараметров по критерию быстродействия» отмечено, что эффективность работы алгоритма *Stegomask* по критерию быстродействия с удовлетворением энтропийного критерия К. Шеннона в его логической трактовке зависит от следующих параметров:

1) m – задает размер эталона.
2) K – задает максимальное число случайных выборок ключей из полного их множества для выявления приемлемости контейнера (гаммы).

Согласно проведенным исследованиям по определению параметров m и K, при которых скорость формирования стегоконтейнера максимальна, формулируется нестрого индуктивное утверждение.

Утверждение 1. Формирование стегоконтейнера при $k = 3$ и $\gamma = 10$ с удовлетворением энтропийного критерия К.Шеннона в его логической трактовке занимает минимальное время при $40 \le m \le 60$ и $10^5 \le K \le 3\times10^5$.

При выборе ГПСП для генерации контейнеров в принятом методе защиты учитывались требования: высокая скорость работы ГПСП, большие период и длина вектора состояния ГПСП, хорошие статистические свойства генерируемой ПСП. Необходимость учета этих требований при выборе ГПСП применительно к системам защиты от несанкционированного доступа отмечено Б.Я. Рябко, А.Н. Фионовым. На основании данных требований среди линейных генераторов выделены генераторы *DX*-1597-2-7, *MT*19937, *Brent-xor*4096s. Отметим, что из них только *MT*19937 ("Вихрь Мерсенна") успешно прошел тесты *DIEHARD*, включающий набор жестких критериев Дж. Марсальи. Именно он и использован в работе. Кроме выделенных ГПСП, эффективность алгоритма *Stegomask* достигается и при выборе нелинейного генератора ГОСТ 28147-89 в режиме гаммирования.

Основное время формирования стегоконтейнера по алгоритму *Stegomask* определяется временем выполнения четвертого шага, как наиболее трудоемкого по времени из всех шагов алгоритма. При его исключении время формирования стегоконтейнера существенно сократится. Но в таком случае выбор гаммы (контейнера) безальтернативен, ибо для сокрытия кода берется первый попавшийся. Далее алгоритм формирования стегоконтейнера в режиме безальтернативного выбора контейнера будет называться *Stegomask alternativeless*.

Чтобы ответить на вопрос, с какой долей вероятности стегоконтейнер представит все множество кодов при ограниченном переборе ключей, если при формировании этого стегоконтейнера использовать подход безальтернативной стеганографии, проведены следующие исследования.

Было сокрыто 1000 различных трехразрядных кодов при $K = 200000$ и различных значениях m. При каждом сокрытии кода подмножество ключей, необходимых для определения приемлемости контейнера, всякий раз генерировались заново. В ходе эксперимента подсчитывалось число событий A, при которых первый попавшийся контейнер был успешным. Оценка вероятности события A определялась по формуле:

$$P(A) = P^* = f/j;$$

где f – число опытов, в которых произошло событие A;

j – число проведенных опытов.

Результаты исследований отражены на рис. 2. На их основе формулируется утверждение.

Утверждение 2. С увеличением значения m растет оценка вероятности того, что стегоконтейнер, полученный безальтернативным выбором гаммы, покроет полный словарь сообщений при ограниченном переборе ключей. В частности, при $m = 60$, $K = 2\times10^5$ значение $P^* = 0{,}999$.

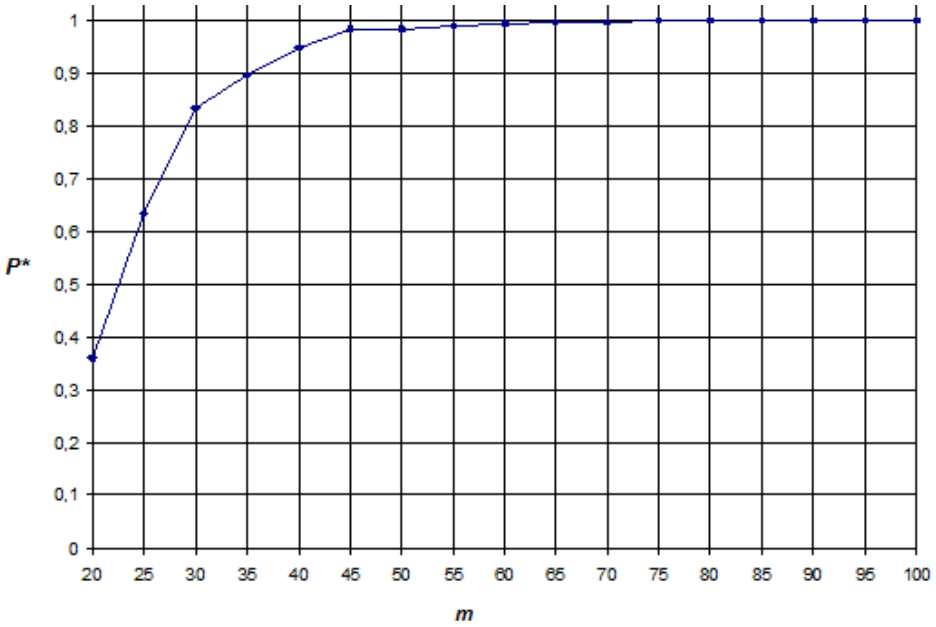


Рисунок 2 – Оценка вероятности события A при разных значениях m

Энтропийная стойкость защиты ассоциируется с существованием неединичного подмножества равноправдоподобных результатов анализа сцены при полном переборе ключей. Полученная на основе набранной статистики оценка вероятности того, что одноразовая рандомизация не обеспечит полноты покрытия, составляет 10^{-3} для $m = 60$, т.е. при таком выборе m в среднем лишь 1 опыт из 1000 окажется неудачным. Такая неполнота практически не должна влиять на энтропийные свойства защиты.

Причина в следующем. Цифровая карта содержит множество стегоконтейнеров, рандомизируемых по-отдельности. «Зернистый шум» на раскрытом изображении, как следствие ошибок ввода может наблюдаться и на истинном ключе. Поэтому наличие такого шума из-за нарушений полноты не нарушает возможной на некотором ложном ключе правдоподобности получаемой картины. Так что значение $m = 60$ для алгоритма Stegomask alternativeless следует считать оптимальным.

В разделе 2.3 «Оценка эффективности ассоциативного метода сокрытия в сравнении с ГОСТ 28147-89» выявлены положительные черты алгоритмов *Stegomask* и *Stegomask alternativeless* сравнительно с ГОСТ 28147-89:

1) Внушительный размер стегоключа. В частности, при $m = 60$ размер стегоключа равняется 5280 бит.

2) Используемая гамма внедряется в передаваемое сообщение и не влияет на процесс санкционированного восстановления данных, вследствие чего скорость санкционированного раскрытия данных в два раза превышает скорость расшифрования данных в ГОСТ 28147-89 при практически неизменной скорости шифровании (случай *Stegomask alternativeless*).
3) Искажения в стегоконтейнере, которые не затрагивают стеговложения, не влияют на процесс санкционированного раскрытия. При $m = 60$ среднее отношение количества бит в стегоконтейнере к числу значащих бит стеговложения равно 160. В ГОСТ 28147-89 искажение любого бита зашифрованного блока влияет на все биты открытого блока.

Выделенные черты алгоритмов потенциально свидетельствуют о высокой вычислительной стойкости и помехоустойчивости используемого метода защиты.

В третьей главе предлагается схема БД для защищенного хранения точечных, линейных и площадных объектов картографии. Изложены принципы формирования защищенной БД картографии. Предлагается метод, позволяющий обрабатывать запросы к защищенной БД картографии без ее полного раскрытия. Приводятся алгоритмы интерпретации пользовательских запросов без нарушения целостности защищенной БД картографии.

Схема БД для защищенного хранения точечных, линейных и площадных объектов картографии представлена на рис. 3.

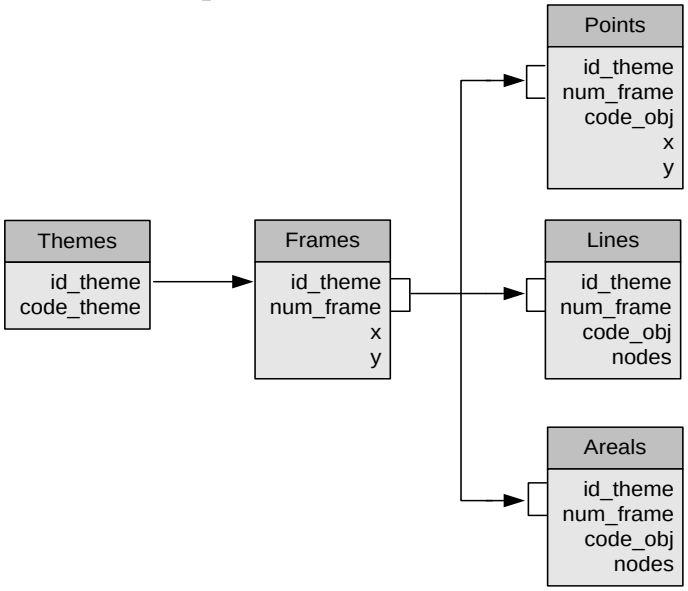


Рисунок 3 – Схема БД для защищенного хранения точечных, линейных и площадных объектов картографии

На рисунке:

Таблица *Themes* – привязывает к кодам тематических слоев уникальные идентификаторы.

Таблица *Frames* – хранит информацию о расположении фрагментов относительно глобальной координатной сетки тематического слоя. Здесь x и y – глобальные координаты левого нижнего угла фрагментов (по глобальной координатной сетке слоя).

Таблицы *Points*, *Lines* и *Areals* – хранят информацию о расположении точечных, линейных и площадных объектов относительно локальной координатной сетки фрагментов соответственно.

Атрибуты *x* и *y* в таблице *Points* предназначены для указания координат точечного объекта. По атрибуту *nodes* таблиц *Lines* и *Areals* хранятся коды и координаты узловых точек объектов в следующей форме:

[код_точки_1 *x y* код_точки_2 *x y* ... код_точки_N *x y*],

где *N* – число узловых точек составляющих объект. Направление обхода узловых точек линейного или площадного объекта для выявления его очертаний определяется порядком возрастания кодов узловых точек.

В разделе 3.2 «Локальная обработка запросов без раскрытия защищенной базы данных картографии в целом» описана реализация методов обработки запросов к защищенной БД картографии посредством сервера СУБД без предварительного раскрытия всей БД на НМД:

1. Использование в *SQL*-запросах стандартных функций, операторов СУБД для раскрытия сокрытых данных.
2. Добавление функции раскрытия в СУБД посредством интерфейса *MySQL UDF* (интерфейса *MySQL*, позволяющего добавлять новую определяемую пользователем функцию).

Для сравнения скорости обработки запросов этими двумя методами, получены временные оценки раскрытия различного числа сокрытых трехразрядных кодов при размерности эталона $m = 18$ и $\gamma = 10$ посредством выполнения однотипных *SQL*-запросов (табл. 1).

Таблица 1 – Время раскрытия БД разными методами

Число стегоконтейнеров	Среднее время раскрытия БД, сек.	
	по первому методу	по второму методу
10000	0,078	0,015
100000	0,641	0,266
1000000	6,235	2,516

По таблице видно, что обработка запроса первым методом происходит медленнее, чем вторым методом. Это связано с тем, что в первом методе для раскрытия кода в стегоконтейнере вызывается огромное число стандартных функций СУБД, когда как во втором методе для данной цели вызывается лишь одна пользовательская функция.

Таким образом, в нашем случае добавление функции раскрытия в СУБД посредством интерфейса определяемых пользователем функций является необходимостью с точки зрения повышения быстродействия.

Алгоритмы интерпретации пользовательских запросов (добавление, удаление, изменение, селекция) без нарушения целостности защищенной БД картографии описаны ниже.

Алгоритм процедуры добавления объекта.

1. Задаются атрибуты добавляемого объекта *q*: код тематического слоя, код объекта, координаты *x* и *y*.

2. Проверить, не совпадают ли атрибуты добавляемого объекта *q* с атрибутами существующих объектов. Если совпадают, то выдать информацию об ошибке и прервать процедуру.
3. Определить фрагмент *Q*, в который добавляется *q*. Если такого фрагмента не существует, то его требуется создать, заполнить необходимым числом пустых объектов и непустым объектом *q* и завершить алгоритм.
4. Добавить *q* во фрагмент *Q*.
5. Проверить, имеется ли в *Q* хотя бы один пустой объект. Если имеется, то удалить в *Q* любой пустой объект. Иначе дополнить каждый фрагмент, кроме *Q*, одним пустым объектом. Данный шаг необходим для соблюдения равного числа объектов в каждом фрагменте.

Алгоритм процедуры удаления объекта.

1. Задание атрибутов удаляемого объекта *q*: код тематического слоя, код объекта, координаты *x* и *y*.
2. Определить фрагмент *Q*, в котором находится удаляемый объект *q*. Если такого фрагмента не существует, то выдать информацию об ошибке и прервать процедуру.
3. Проверить, является ли *q* единственным непустым объектом в *Q*. Если является, то фрагмент *Q* удаляется. Иначе *q* заменяется пустым объектом.

Алгоритм процедуры изменения объекта.

1. Задание атрибутов изменяемого объекта *q* (код тематического слоя, код объекта, координаты *x* и *y)* и атрибутов объекта *q'* (код объекта, координаты *x* и *y*), которым требуется заменить *q*.
2. Определить фрагмент *Q*, в котором находится изменяемый объект *q*. Если такого фрагмента не существует, то выдать информацию об ошибке и прервать процедуру.
3. Проверить наличие изменяемого объекта *q* во фрагменте *Q*. Если такого объекта не существует, то выдать информацию об ошибке и прервать процедуру.
4. Определить фрагмент *Q'*, в который может переместиться изменяемый объект *q*. Если фрагмент *Q'* по новым координатам объекта *q* уже имеется, то перейти к шагу 7.
5. Если новые координаты объекта *q* не выходят за пределы фрагмента *Q*, изменить атрибуты объекта на новые атрибуты, завершить процедуру.
6. Создать фрагмент *Q'* с необходимым числом пустых объектов и непустым объектом *q'*, выполнить 8-ой шаг алгоритма и завершить процедуру.
7. Проверить на совпадение атрибутов объекта *q'* с атрибутами какого-либо объекта во фрагменте *Q'*. Если имеет место совпадение, то выдать информацию об ошибке и прервать процедуру.
8. Проверить, является ли объект *q* единственным непустым объектом во фрагменте *Q*. Если является, то удаляется весь фрагмент *Q*. Иначе объект *q* заменяется пустым объектом.
9. Добавить объект *q'* во фрагмент *Q'*.
10. Проверить, имеется ли во фрагменте *Q'* хотя бы один пустой объект. Если имеется, то удалить в *Q'* любой пустой объект. Иначе дополнить каждый фрагмент, кроме

Q', одним пустым объектом. Данный шаг необходим для соблюдения равного числа объектов в каждом фрагменте заданного тематического слоя.

Алгоритм процедуры селекции.

1. Задаются: код тематического слоя, условие выборки на языке *SQL* (*WHERE predicate*).
2. На каждом узле кластера запускается подчиненный процесс.
3. Управляющий процесс, запущенный на управляющем узле, создает коммуникационную группу, в которую входят процессы тех узлов, на которых хранятся нужные фрагменты слоя карты.
4. Процессы коммуникационной группы параллельно выполняют *SQL*-запросы с *WHERE predicate* со встроенными функциями раскрытия, обращаясь к локальным защищенным БД картографии. Процессы, не вошедшие в группу, завершаются.
5. Главный процесс собирает результаты обработки запросов со всех процессов созданной группы.

В четвертой главе рассматривается разработанный программный проект СУБД *Security MapPointCluster* и приводятся результаты тестирования программного комплекса, реализованного на базе данного проекта.

На рис. 4 показана его упрощенная структура. Система построена на базе СУБД *MySQL* с интегрированным в нее двумерно-ассоциативным методом защиты и принципами параллельных вычислений. Динамические аспекты поведения процессов на серверной части СУБД *Security MapPointCluster* представлены в виде диаграмм деятельности языка *UML*, на клиентской части – в виде блок-схем. Реализованный в системе комплекс программ прошел успешные испытания на множестве репрезентативных запросов.

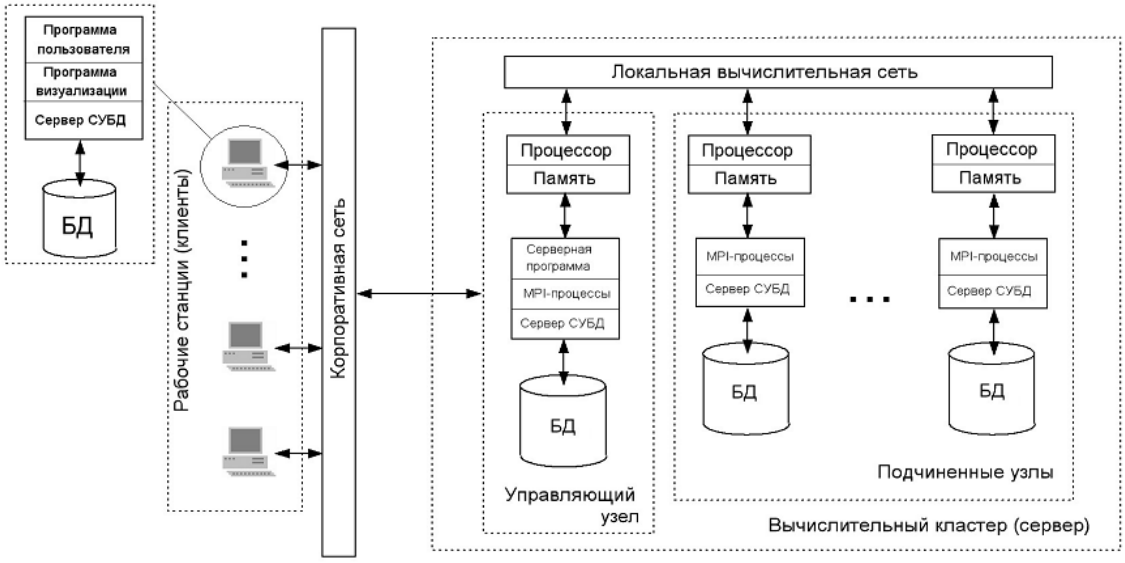


Рисунок 4 – Структура СУБД *Security MapPointCluster*

Для тестирования комплекса программ, составляющих СУБД *Security MapPointCluster*, выбраны:

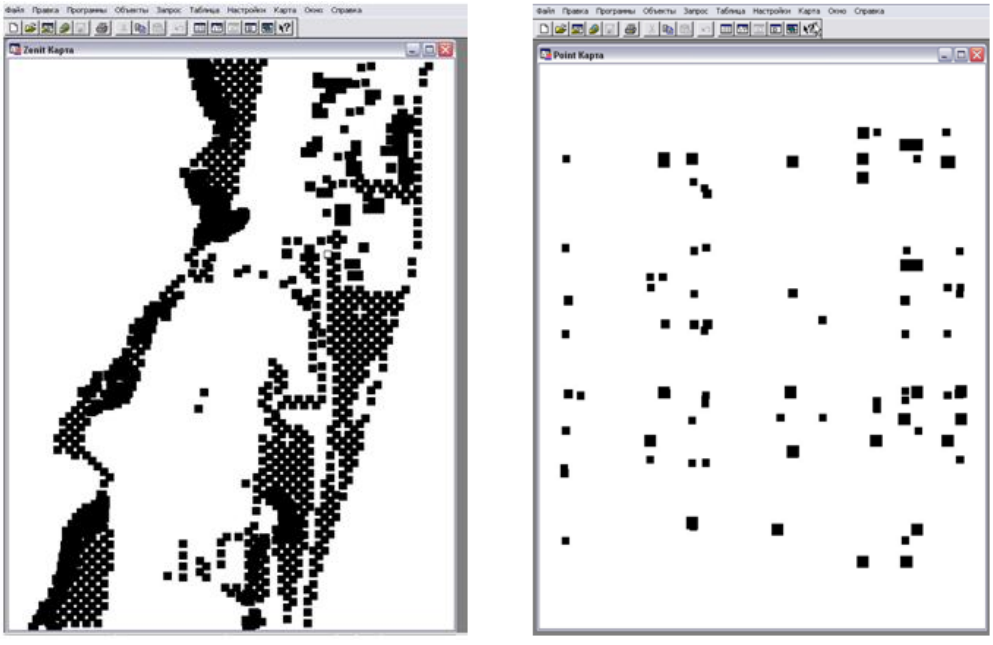
1) Аппаратная платформа. Двенадцать вычислительных узлов, объединенных сетью *Gigabit Ethernet* посредством коммутатора *D-LINK DGS*-1016*D*. Каждый из узлов имеет двухъядерный процессор *Intel(R) Core(TM)*2 *CPU* частотой 1,87 ГГц, опера-

тивную память *DDR*2 3 Гб, дисковый накопитель *Western Digital* 150 *GB* (с интерфейсом *SATA*).

2) Программное обеспечение. ОС семейства *Microsoft Windows XP Professional*, СУБД *MySQL* версии 5.1.45-*win*32, ГИС *MapInfo Professional* 10, интегрированная среда разработки *MS Visual Studio* 2008, библиотеки расширения языка *C++*: *MPICH* 1 (*MPI*), *Boost* 1.43.
3) Тестовая карта размером 300×300 км^2 участка местности республики Чувашии предоставленная ООО «Геодезической компанией «Зенит» г. Казани (рис. 5, a). Выбранная карта содержит один тематический слой и 1035 точечных объектов четырех различных типов.

При формировании ЗБДК установлено, что сокрытие базы данных алгоритмом *Stegomask* на одном вычислительном ядре происходит в n раз дольше, чем на вычислительном кластере, состоящем из n таких ядер. Использование вычислительного кластера для сокрытия БД картографии алгоритмом *Stegomask alternativeless* также дает выигрыш в производительности. При интерпретации запросов добавление и изменение в отдельных случаях кластер демонстрирует хорошую производительность сравнительно с одной машиной. Ожидается, что применение кластера для интерпретации запроса селекция будет актуальным в ходе работы с защищенными БД картографии с объемами в несколько гигабайт и более.

При визуализации тестового картографического слоя на ложном наборе масок получено неединичное подмножество правдоподобных результатов раскрытия, одно из которых представлено на рис. 5, b.



(a) (b)

Рисунок 5 – Результат визуализации картографического слоя
(a) – на истинном ключе; (b) – на ложном ключе

В заключении сформулированы основные результаты диссертационной работы.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В диссертации решена научная задача разработки и исследования компонентов фреймовой модели системы баз данных картографии с использованием ассоциативного метода стегозащиты цифровых карт и получены следующие результаты:

1. Предложена двухуровневая модель системы баз данных картографии с ассоциативной защитой, которая позволяет повысить эффективность управления защищенными БД картографии в сравнении с известными СУБД по критерию быстродействия.
2. Разработан алгоритм формирования стегоконтейнера, основанный на применении двумерно-ассоциативного механизма защиты. На его основе подтверждена гипотеза о принципиальной возможности достижения безусловной стойкости ассоциативного метода стегозащиты.
3. Оптимизированы значения стегопараметров алгоритма формирования стегоконтейнера по критерию быстродействия. Установлено, что время формирования стегоконтейнера предложенным алгоритмом *Stegomask* минимально при $40 \le m \le 60$ и $10^5 \le K \le 3\times10^5$. Показана предпочтительность использования алгоритма *Stegomask* в режиме безальтернативного выбора гаммы при $m = 60$. Выявлены положительные черты двумерно-ассоциативного механизма маскирования сравнительно с ГОСТ 28147-89.
4. Разработана схема БД, которая реализует ассоциативно-защищенное хранение данных картографии. Предложен эффективный метод локальной обработки запросов к защищенной БД картографии и на его основе разработаны алгоритмы интерпретации пользовательских запросов, отличающиеся от существующих тем, что позволяют обрабатывать запросы к защищенной БД картографии без ее полного раскрытия.
5. Разработан исследовательский прототип СУБД *Security MapPointCluster*, на основе которого обоснованы практические рекомендации к построению системы баз данных картографии с ассоциативной защитой.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В рецензируемом журнале из списка ВАК:

1. Развитие единообразного формализма защиты точечных, линейных и площадных объектов картографии / Гибадуллин Р.Ф. // Вестник КГТУ им. А.Н. Туполева. 2010. №2. С. 102–107.

Другие публикации:

2. Реализация и анализ «быстрого» алгоритма идентификации объектов бинарных изображений / Гибадуллин Р.Ф. // Туполевские чтения: Материалы 14-й Междунар. молод. научн. конф. – Казань: КГТУ, 2006. Т. 4. С. 44–45.
3. Проблемы организации параллельной системы управления защищенными базами данных картографии / Вершинин И.С., Гибадуллин Р.Ф. // Распознавание образов и анализ изображений: новые информационные технологии (PRIA-8-2007): Материалы 8-й Междунар. конф. – Йошкар-Ола: МарГТУ, 2007. Т. 2. С. 220–222. (Статья на англ. яз.).

4. Параллельные алгоритмы защиты бинарных объектов картографии / Вершинин И.С., Гибадуллин Р.Ф., Земцов П.Е. // Методы моделирования: Труды Респ. научн. семинара АН РТ. – Казань: КГТУ, 2007. Вып. 3. С. 96–108.
5. Параллельная реализация защищенной векторной модели данных ГИС / Вершинин И.С., Гибадуллин Р.Ф. // Информатика: проблемы, методология, технологии: Материалы 8-й Междунар. научно-методич. конф. – Воронеж: Изд.-полиграф. центр Воронежского гос. университета, 2008. Т. 2. С. 118–122.
6. Параллельная система управления защищенными картографическими базами данных / Гибадуллин Р.Ф., Прохоров А.Е // Туполевские чтения: Материалы 16-й Междунар. молод. научн. конф. – Казань: КГТУ, 2008. Т. 3. С. 48–50.
7. Управление защищенными картографическими базами данных на вычислительном кластере / Гибадуллин Р.Ф., Прохоров А.Е., Пыстогов С.В. // Техническая кибернетика, радиоэлектроника и системы управления: Материалы 9-й Всерос. науч. конф. – Таганрог: ТТИ ЮФУ, 2008. С. 102–103.
8. Распределенное управление защищенными картографическими базами данных / Вершинин И.С., Гибадуллин Р.Ф., Прохоров А.Е. // Высокопроизводительные параллельные вычисления на кластерных системах (HPC-2008): Материалы 8-й Междунар. конф. – Казань: КГТУ, 2008. С. 216–221.
9. Исследование реализаций двумерно-ассоциативного алгоритма шифрования / Гибадуллин Р.Ф., Мишин А.С. // Туполевские чтения: Материалы 17-й Междунар. молод. научн. конф. – Казань: КГТУ, 2009. Т. 4. С. 59–60.
10. Обработка зашифрованных данных посредством СУБД MySQL / Гибадуллин Р.Ф., Пыстогов С.В. // Туполевские чтения: Материалы 17-й Междунар. молод. научн. конф. – Казань: КГТУ, 2009. Т. 4. С. 60–62.
11. Использование кластерных технологий при решении задач защиты картографических данных / Вершинин И.С., Гибадуллин Р.Ф., Пыстогов С.В. // Высокопроизводительные параллельные вычисления на кластерных системах (HPC-2009): Материалы 9-й Междунар. конф. – Владимир: ВлГУ, 2009. С. 68–72.
12. Моделирование процессов управления кластерами защищенных картографических баз данных / Гибадуллин Р.Ф. // Методы моделирования: Труды Респ. научн. семинара АН РТ. – Казань: «ФЭН» (Наука), 2010. Вып. 4. С. 101–115.
13. Конструктивное моделирование систем в приложении к защите данных картографии / Райхлин В.А., Вершинин И.С., Гибадуллин Р.Ф. // Методы моделирования: Труды Респ. научн. семинара АН РТ. – Казань: «ФЭН» (Наука), 2010. Вып. 4. С. 68–95.
14. Реализация файл-сервером функций генерации защищенной картографической базы данных / Гибадуллин Р.Ф., Мишин А.С. // Туполевские чтения: Материалы 18-й Междунар. молод. научн. конф. – Казань: КГТУ, 2010. Т. 4. С. 89–91.
15. Клиентская часть системы управления защищенными картографическими базами данных / Гибадуллин Р.Ф., Пыстогов С.В. // Туполевские чтения: Материалы 18-й Междунар. молод. научн. конф. – Казань: КГТУ, 2010. Т. 4. С. 97–99.

Диссертант  Р. Ф. Гибадуллин

ГИБАДУЛЛИН Руслан Фаршатович

СИСТЕМА БАЗ ДАННЫХ КАРТОГРАФИИ С АССОЦИАТИВНОЙ ЗАЩИТОЙ

Специальность: 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Формат 60 x 84 1/6. Бумага офсетная. Печать офсетная.
Печ. л. 1,0. Усл. печ. л. 0,93. Усл. кр.-отт. 1,16. Уч.-изд. л. 1,0.
Тираж 100. Заказ О17.

Типография Издательства Казанского государственного технического университета
420111, Казань, К. Маркса, 10.