

# Адаптивное управление доступностью в геоинформационной системе, использующей туманные вычисления

В.Г. Бурлов, В.В. Грызунов, Д.Е. Сипович

**Аннотация**—Геоинформационные системы интегрируются с информационными системами предприятий и государства, являются системами критического применения и работают с большим количеством разнородных и неструктурированных данных. Объём данных, количество пользователей и кибератак увеличивается с каждым годом. Обеспечение информационной безопасности, в свою очередь, требует доступности ресурса системы. Применяемые способы централизованного хранения и обработки данных не справляются с поставленными задачами, не могут точно и в срок обеспечить доступность запрашиваемого ресурса, что в дальнейшем ведёт к нарушению других аспектов информационной безопасности: целостности и конфиденциальности. Один из возможных вариантов решения проблемы обеспечения доступности – применение туманных вычислений.

В статье предлагаются: иерархическая модель геоинформационной системы, использующей туманные вычисления (FIST – Full Infrastructure of Sources Toolkit). Модель включает уровни программного обеспечения, логической структуры, физической структуры. Формализовано взаимодействие между уровнями модели. Приведён пример базовой закономерности, позволяющей объединить отдельные элементы в пул (fog node). Формулируется принцип постепенного распространения задач по геоинформационной системе, ставится задача адаптивного управления производительностью геоинформационной системы как модифицированная задача об упаковке контейнера, раскрывается метод объединения в пулы необходимых ресурсов: вычислителей, каналов связи, устройств ввода/вывода, накопителей. Метод (D-FIST – Dynamic Full Infrastructure of Sources Toolkit) включает 3 шага: выбор элементов-кандидатов, комбинирование пулов, формирование управления на изменение. Доказывается сходимости и завершаемость предложенного метода. Приводится контрольный пример применения D-FIST. Анализируются особенности данных, с которыми работают

геоинформационные системы. Сравняются современные технологии, на базе которых возможна реализация предложенного метода.

**Ключевые слова**—Геоинформационная система, доступность, информационная безопасность, туманные вычисления.

## I. ВВЕДЕНИЕ

Геоинформационные системы стали неотъемлемой частью почти всех сфер человеческой деятельности: от предоставления безопасных условий жизнедеятельности, прогнозирования и предотвращения травм населения [1], функционирования умных городов, раскрытия преступлений [2], внедрения технологий Индустрии 4.0 и автоматизации средств доставки товаров до клиентов [3], [4], до применения оружия и развития закрытых технологий. На основе данных из геоинформационных систем принимаются решения, от которых зависит процветание и обороноспособность страны, распределение экономических благ. А значит, геоинформационные системы являются системами критического применения и нуждаются в защите.

По мнению отдельных исследователей в ближайшее время геоинформационные системы (ГИС) войдут в состав всех информационных систем [5].

ГИС – информационная система, оперирующая пространственными данными [6].

ГИС наполняется данными с помощью мобильных наземных, водных, воздушных, космических комплексов, с помощью метеодатчиков и устройств сбора данных, организованных в виде сеточных массивов. Поток поставляемых данных по самым скромным оценкам в ближайшие несколько лет должен вырасти больше, чем в 200 раз [7]. Пользователи ГИС, мировые и региональные центры метеорологических, радиационных, климатических и других данных находятся в разных частях земного шара на разном расстоянии от пользователей.

Сказанное определяет следующие особенности ГИС:

- гетерогенная архитектура ГИС;
- большое количество разнородных пользователей;
- огромный поток неструктурированных данных, поступающих в реальном масштабе времени;
- растущий интерес к ГИС злоумышленников, а значит, необходимость выделения ресурсов ГИС на информационную безопасность. Согласно ГОСТ Р

Статья получена 14 июня 2021.

Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ) Проект №08/2020.

В. Г. Бурлов, д-р техн. наук, профессор кафедры информационных технологий и систем безопасности, Российский государственный гидрометеорологический университет, Санкт-Петербург, Россия (e-mail: burlovvg@mail.ru).

В. В. Грызунов, канд. техн. наук, доцент кафедры информационных технологий и систем безопасности, Российский государственный гидрометеорологический университет, Санкт-Петербург, Россия (e-mail: viv1313r@mail.ru).

Д. Е. Сипович, старший преподаватель кафедры информационных технологий и систем безопасности, Российский государственный гидрометеорологический университет, Санкт-Петербург, Россия (e-mail: Sipovich@electronic.spb.ru).

ИСО/МЭК 17799-2005 информационная безопасность – это механизм, обеспечивающий целостность, конфиденциальность и доступность. Но целостность и конфиденциальность реализуются отдельными функциональными компонентами ГИС, которые требуют для своей работы ресурса ГИС. Следовательно, предоставив необходимый ресурс (реализовав доступность), ГИС создаёт необходимые условия реализации других аспектов информационной безопасности: целостности и конфиденциальности.

Доступность – это механизм защиты, обеспечивающий доступ к информации и связанным с ней активам авторизованных пользователей по мере необходимости.

Доступность ресурсов ГИС нарушается под действием факторов разной природы:

- детерминированная природа: выполняемые по расписанию задачи, регламентные работы и так далее. Не представляет опасности для ГИС, потому что рассчитывается и учитывается заранее;

- стохастическая природа: множество альтернатив дестабилизирующих факторов и вероятностное описание механизма выбора из этого множества альтернатив полностью известны. К этим факторам относятся сбои и отказы, обусловленные естественными причинами и действиями низкоквалифицированных специалистов, поток пользовательских задач, уже известных и не раз выполняемых геоинформационной системой;

- нестохастическая природа, то есть не стохастическая и недетерминированная: случаи, когда неизвестны множество альтернатив или вероятностное описание, либо действующий фактор имеет целенаправленный характер и не может описываться средствами теории вероятностей. К факторам нестохастической природы принадлежат не известные ранее новые задачи, созданные пользователем, агрессивная целенаправленная активность нарушителя, осуществление кибератак и т.п.

Модели и методы работы современных ГИС ориентированы на централизацию и в теории могут справиться с обозначенными особенностями, например, с помощью декомпозиции процесса доступа к данным [8], особых подходов к разработке программного обеспечения [9] или использовании облачных технологий крупных компаний [10], [11]. Однако применение централизованных методов обработки данных не соответствует распределённой природе ГИС, и на практике сталкивается со следующими проблемами и технологическими ограничениями:

- высокая стоимость серверов и каналаобразующего оборудования, так, например, стоимость высокоскоростных каналов связи в 1 Гбод между серверами ГИС, разнесёнными географически на большие расстояния растёт пропорционально количеству серверов (см. рис. 1). График получен путём опроса трёх самых крупных операторов связи РФ: Ростелеком, Мегафон, МТС;

- доступность некоторого количества серверов может быть нарушена при недостаточной связности в случае аварийной ситуации и/или повреждении физического кабеля. В настоящее время при организации дублирующего канала в автономной системе часто получается, что основной и резервный каналы размещаются в одном физическом кабеле. Другой причиной нарушения доступности может быть выход из строя или потеря доступа к серверам имён, обеспечивающим разрешение доменных имён. В этом случае мастер, отвечающий за балансировку нагрузки, не может ответить на запрос;

- центральные серверы, устройства ввода данных и ресурсы конечных пользователей ГИС разнесены в пространстве. Зачастую даже по территориям с разной юрисдикцией. Это вызывает задержки в передаче данных, потери данных при разрывах соединений [12], что особенно негативно влияет на мобильных пользователей ГИС.

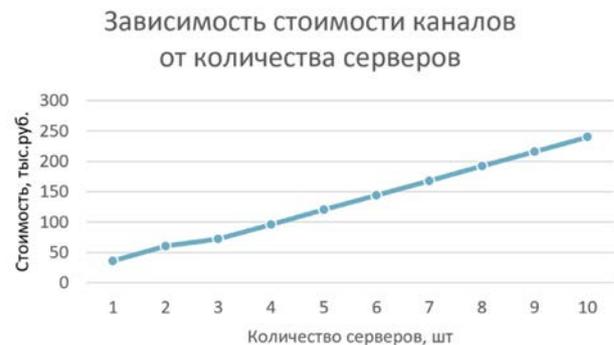


Рис. 1. Стоимость аренды каналов связи в зависимости от количества объединяемых серверов, тыс. руб.

Одним из способов обеспечить доступность сервисов ГИС и разрешить обозначенные проблемы является построение ГИС на базе архитектуры туманных вычислений. Туманные вычисления, в отличие от централизованных методов обработки данных, ориентированы на распределённые информационные системы, каковой и является ГИС. Такой подход применяется флагманами мировой IT-промышленности: Intel, CISCO, Microsoft и др. и позволяет снизить нагрузку на узлы информационной системы и одновременно повысить её живучесть [13].

Идея применения туманных вычислений в ГИС не нова. Этот вопрос изучается в [14]. Однако здесь рассматриваются только web-приложения, использующие туманную архитектуру, при этом не уделяется внимание тому, что выступает в роли узла тумана (fog node), и как именно должно работать системное программно-аппаратное обеспечение ГИС. Узел тумана (fog node) – физический или логический элемент сети, реализующий туманные вычисления [13].

В работе [15] авторы описали экспериментальный подход, который позволяет снизить на 30-60% энергопотребление и задержку обработки запросов к геопространственным данным относительно существующих централизованных ГИС, то есть решили задачу в части энергопотребления. Предоставление

требуемых ресурсов задачам, решаемым ГИС, осталось за рамками исследования.

Интересный подход предлагается в [16]. Здесь задачи узлов тумана классифицируются согласно их вычислительной природе и назначаются узлам тумана. Используется централизованное управление нагрузкой на основе заданных шаблонов, что не совсем подходит для динамично изменяющейся структуры ГИС, построенной на базе туманных вычислений.

Авторы [17] исследуют алгоритм случайной пересылки задач между узлами тумана, но аналогично предыдущей работе основано на централизованном управлении.

В исследованиях, описанных в [18], делается акцент на разнородность ресурсов туманных вычислений, предлагается распределение задач по узлам тумана случайным образом по равномерному закону распределения. Используется децентрализованная модель управления. Есть элементы адаптации алгоритма в зависимости от загрузки узлов тумана. Однако, во-первых, в работе речь идёт только о вычислительном ресурсе, во-вторых, в случае, когда ресурса одного узла тумана или граничного устройства недостаточно, задача получает отказ в обслуживании. Не зависимо от того, что, принципиально, есть возможность решить её, распараллелив на несколько узлов. В-третьих, не предусмотрена возможность преобразования одного ресурса в другой.

Предложим модель ГИС, свободную от недостатков вышеназванных моделей. Рассмотрим представление задач в ГИС.

## II. ПРЕДСТАВЛЕНИЕ ЗАДАЧ В ГЕОИНФОРМАЦИОННОЙ СИСТЕМЕ

Смысловое значение задач, выполняемых ГИС, для самой ГИС безразлично: нанесение слоя, синхронизация данных удалённых пользователей, шифрование, авторизация и т.д. ГИС, как информационная система, всего лишь преобразует участки памяти по заданным алгоритмам и возвращает результат в указанное место памяти: преобразует нолики и единички в нолики и единички. Что, действительно, имеет значение для ГИС, так это тип запрашиваемого пользователем ресурса, объём ресурса и срок, на который ресурс требуется выделить. Сказанное относится и к системе защиты информации, которая сама потребляет ресурс для обеспечения целостности, конфиденциальности и доступности в ГИС.

Ресурс – актив, который используется или потребляется в ходе выполнения процесса (ГОСТ Р ИСО/МЭК 12207-2010).

Следовательно, управление защитой информации в ГИС может формулироваться как управление её ресурсами, то есть как управление расходом или потреблением актива в ходе решения задач. Актив расходуется во времени и задачи решаются во времени, а значит, управление ресурсами можно свести к управлению производительностью.

Производительность – количество задач, решаемых в

единицу времени.

**Требуемой производительностью** ( $\Omega^{треб}$ ) обозначим производительность, которая нужна задаче, чтобы быть решённой точно и в срок. **Текущей производительностью** ( $\Omega$ ) назовём производительность, которая имеется в распоряжении ГИС для решения задачи.

Очевидно, возможны ситуации:

– текущая производительность больше или равна требуемой  $\Omega \geq \Omega^{треб}$ . Здесь нет нужды в изыскании дополнительной производительности;

– текущая производительность меньше требуемой  $\Omega < \Omega^{треб}$ . Необходимо найти недостающую производительность, чтобы задача могла быть решена.

Для решения задач в ГИС существует четыре типа производительностей:

- вычислители ( $C$ );
- каналы связи ( $L$ );
- память ( $Sp$ );
- устройства ввода/вывода ( $Tr$ ).

Следовательно, любая задача существует для ГИС как необходимый для её решения набор четырёх типов производительностей  $\{C, L, Sp, Tr\}$ . В случае нехватки ресурса формируется соответствующий пул: вычислительный, каналов связи, устройств ввода/вывода или пул из накопителей.

**Пул** – временное объединение ресурсов, предназначенное для решения какой-то задачи (группы задач). Пул существует в период решения задачи. После решения задачи пулы могут быть расформированы. В концепции туманных вычислений роль пула выполняет узел тумана (fog node). Наиболее известные технологии работы с пулами представлены компаниями Microsoft, CISCO, Sun Micro systems в продуктах: кластер на базе Windows Server, Dynamic Trunking Protocol (DTP) для формирования транков (пулов из каналов связи); пул ресурсов Sun Solaris 11.x.

Названные технологии статичны, нацелены на централизованную архитектуру и не могут в полной мере в реальном масштабе времени управлять ГИС, использующей технологии туманных вычислений, потому что ГИС работает в условиях нестохастической неопределённости, порождённой как действиями злоумышленников, так и особенностями решаемых задач [19]. С нестохастической неопределённостью можно справиться, синтезируя управление производительностью в реальном масштабе времени.

Построим модель ГИС, позволяющую синтезировать такое управление, и сформулируем задачу управления.

## III. МОДЕЛЬ ГИС (МОДЕЛЬ *FIST*)

В работе [20] исследуется многоуровневая система узлов (N-Tier Fog Deployment) на предмет применимости для туманных вычислений. Система узлов сводится к трёхуровневой модели их взаимодействия: облако->туман->конечные устройства. Создание кластеров возможно лишь на уровне конечных устройств. Межуровневые связи предусмотрены на уровнях облака и тумана. Модель учитывает только

вычислительный ресурс информационной системы, что, как будет показано далее, составляет малую часть возможностей ГИС.

Предложим модель полного набора инструментов инфраструктуры источников (*FIST – Full Infrastructure of Sources Toolkit*). FIST рассматривает ГИС как совокупность узлов, каждый из которых представляет собой отдельную независимую информационную систему, расположенную в местах разной юрисдикции, принадлежащую разным организациям, администрируемую разными людьми, имеющую свои собственные политики безопасности и т.д. Каждый узел имеет в своём составе следующую иерархию:

- уровень физической структуры (уровень ФС, УФС);
- уровень логической структуры (уровень ЛС, УЛС);
- уровень программного обеспечения (уровень ПО, УПО).

На уровне физической структуры присутствуют физические устройства: процессоры, каналы связи, накопители, устройства ввода/вывода. Все устройства связываются между собой физическими каналами связи и предоставляют в распоряжение ГИС свою производительность  $\Omega^{УФС}$ : MIPS, Гбод, Бит/с, FPSи др.

Уровень логической структуры определяет логическую организацию ГИС, которая может отличаться от физической. Например, обычный wi-fi роутер подключая к себе устройства создаёт структуру «общая шина» на физическом уровне и «звезда» на логическом. Производительность этого уровня  $\Omega^{УЛС}$ .

Программное обеспечение образует отдельный уровень, по эффективности работы которого пользователи судят об эффективности работы всей ГИС. Производительность этого уровня ( $\Omega^{УПО}$ ) обычно измеряется в задачах, решённых в единицу времени: количество отправленных и принятых писем в секунду, количество нанесённых объектов на карту в час, скорость создания нового слоя или предоставление аналитического отчёта и т.д.

Устанавливая взаимодействие между собой, отдельные уровни формируют всю геоинформационную систему (ГИС) с производительностью  $\Omega^{ГИС}$ ,  $\omega \in \Omega$  (рис. 2). Можно сказать, что уровни вложены друг в друга, то есть верхний уровень модели является мета системой для вложенных уровней и определяет их требуемые пространственно-временные состояния. Подробнее о зависимости производительности уровня от его пространственно-временного состояния и связях уровней между собой описано в работах [21], [22].

Пусть  $Q^{УФС}$ ,  $Q^{УЛС}$ ,  $Q^{УПО}$ ,  $Q^{ГИС}$  – соответственно, множества требуемых пространственно-временных состояний уровней физической и логической структур, уровня программного обеспечения, всей ГИС,  $q \in Q$ . Уровни соотносятся между собой так:

$$Q^{УФС} \cup P(Q^{УЛС} \cup P(Q^{УПО})) = Q^{ГИС},$$

где  $P(Q)$  – булеан на множестве  $Q$ .

На каждом уровне решается своё количество задач  $K$  за заданное время  $T$ . Будем считать  $K$  показателем эффективности уровня ГИС.

Под эффективностью понимается понятие, введённое в работе [23, с.31] «эффективность – это свойство (качество) процесса функционирования системы, определяемое как его приспособленность к решению поставленных перед системой задач».

В [21] показано, что показатель эффективности верхнего уровня зависит от показателей эффективности вложенных уровней.

$$K^{УФС} (K^{УЛС} (K^{УПО})) = K^{ГИС},$$

где  $K^{УФС}$ ,  $K^{УЛС}$ ,  $K^{УПО}$ ,  $K^{ГИС}$  – соответственно, показатели эффективности применения уровней физической и логической структур, уровня программного обеспечения, всей ГИС.

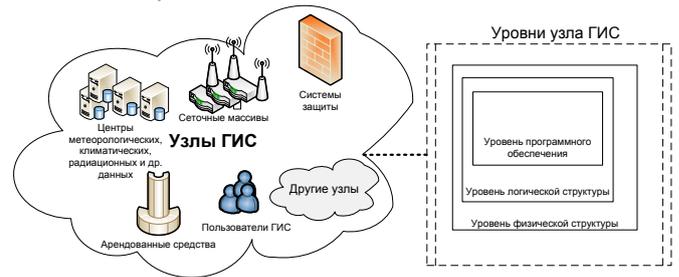


Рис. 2. Иерархическая модель ГИС

Поскольку эффективность работы всей ГИС оценивается по эффективности работы уровня программного обеспечения [21], то именно этот уровень формирует запросы на выделение ресурсов и направляет их на верхний уровень ГИС, то есть уровень, являющийся для него метасистемой (рис. 3).



Рис. 3. Формирование запросов на выделение ресурсов в ГИС

Если метасистема (верхний уровень) может сразу удовлетворить запрос, то происходит выделение ресурса. В противном случае для удовлетворения запроса формируются пулы согласно правилам, характерным для запрошенного ресурса.

Приводимые рассуждения оправданы, если элемент, добавляемый в пул, не уменьшает производительность пула. В противном случае добавлять элемент в пул не имеет смысла.

*Ограничение 1.* Добавление элемента в пул не уменьшает общую производительность пула.

Для каждого типа ресурсов ГИС действуют свои законы формирования пулов (fog nodes). Например, для вычислителей и каналов связи выполняется правило:

$$\left\{ \begin{array}{l} \Omega_p \leq \sum_{j=1}^M \omega_j \\ \text{— параллельное} \\ \text{соединение элементов в} \\ \text{пуле;} \end{array} \right. \quad (1)$$

– последовательное  
соединение элементов в  
пуле

где

$\Omega_p$  – производительность пула (количество задач, решаемых пулом в единицу времени);

$\omega_j$  – производительность  $j$ -ого элемента (вычислителя, канала связи, памяти или устройства ввода-вывода) пула;

$M$  – количество элементов в пуле.

Таким образом, количество задач, выполненных уровнем, зависит от следующих факторов:

- 1) текущего пространственно-временного состояния уровня, то есть от того, сколько элементов присутствует в пуле и каким образом они соединены между собой;
- 2) производительностей вложенных уровней.

$$\omega^{GIS}(q^{GIS}) = \omega^{yFC}(q^{yFC}, \omega^{yLC}(q^{yLC}, \omega^{yPO}(q^{yPO})))$$

где

$\omega^{GIS} \in \Omega^{GIS}$ ,  $\omega^{yFC} \in \Omega^{yFC}$ ,  $\omega^{yLC} \in \Omega^{yLC}$ ,  $\omega^{yPO} \in \Omega^{yPO}$  – соответственно, производительности узла GIS, элементов уровня физической структуры и логической структуры, уровня программного обеспечения.

$q^{GIS} \in Q^{GIS}$ ,  $q^{yFC} \in Q^{yFC}$ ,  $q^{yLC} \in Q^{yLC}$ ,  $q^{yPO} \in Q^{yPO}$  – соответственно, элемент множества пространственно-временных состояний всей GIS, уровня физической и логической структуры, уровня программного обеспечения.

Предлагаемая модель в отличие от модели, лежащей в основе устоявшегося понимания туманных вычислений [13], позволяет расширить варианты предоставления и эксплуатации GIS в режиме Fog-as-a-Service, а именно предоставлять потребителю или заимствовать у него не только вычислительные ресурсы и память, но и каналы связи, устройства ввода/вывода. Например, если при вводе данных в GIS доступен широкий канал связи или на конечном устройстве/датчике слабые вычислительные мощности, то GIS может принимать на вход «сырые» данные. А если ситуация обратная, то есть канал узкий, и окончное устройство имеет достаточные вычислительные мощности и энергозапас, то GIS может перенести предварительную обработку данных на окончное устройство и принимать уже подготовленные данные.

Другая особенность приведённой модели состоит в том, что она работает по принципу фрактала. Если добавляется новый элемент, то он может включаться в любой из уровней GIS.

Таким образом, на вход модели подаются физические элементы, которые не могут в отдельности решать задачи, поставленные перед GIS. Чтобы собрать из этих элементов GIS, обладающую требуемой производительностью, возможны два диаметрально противоположных подхода: централизованное управление и самоорганизация.

Практическое применение модели *FIST* нуждается в разработке соответствующего метода, динамически

формирующего пулы ресурсов в GIS. Основные отличия предлагаемого метода от его ближайших аналогов, описанных в работах [24], [25], состоят в том, что метод позволяет работать со всеми типами ресурсов GIS, предполагает произвольный алгоритм сортировки элементов-кандидатов и предусматривает управление устройствами, входящими в пул.

#### IV. ПОСТАНОВКА ЗАДАЧИ АДАПТИВНОГО УПРАВЛЕНИЯ ПРОИЗВОДИТЕЛЬНОСТЬЮ GIS

Вербальная постановка задачи звучит следующим образом.

Дано: GIS является распределённой системой, и элементы GIS распределены в пространстве, а значит, производительность GIS тоже распределена в пространстве. GIS функционирует в условиях действия дестабилизирующих факторов детерминированной, стохастической и нестохастической природы [19]. Дестабилизирующие факторы влияют на расположение и производительность элементов GIS, а также на поток задач, решаемых GIS.

Требуется: таким образом расположить элементы GIS в пространстве (сформировать из элементов пул) и настроить взаимодействие между ними, чтобы выполняемые задачи получали требуемую производительность, несмотря на действие дестабилизирующих факторов. Говоря в терминах адаптивного управления, необходимо адаптировать GIS к действию дестабилизирующих факторов.

Формирование пула – комбинаторная задача, сводимая к модифицированной задаче об упаковке контейнеров, где

$N^{mpe6}$  – требуемое количество пулов (количество контейнеров),

$N$  – текущее (доступное) количество пулов,

$\Omega_i^{mpe6}$  – требуемая производительность  $i$ -ого пула (размер контейнера),

$\Omega_i$  – текущая производительность  $i$ -ого пула,

$\omega_j \in U_{i=1}^N \Omega_i$ ,  $j = \overline{1, |V|}$  – текущая производительность  $j$ -ого элемента (размер упаковываемого объекта),

$V$  – множество элементов GIS (множество упаковываемых объектов).

Требуется найти функцию со свойствами:

$$f^{yLC} : (\Delta \Omega_i \rightarrow 0),$$

$$\Delta \Omega_i = \Omega_i^{mpe6} - \Omega_i, i = \overline{1, N}$$

при ограничениях:  $\Omega_i \geq \Omega_i^{mpe6}$ ,  $N = N^{mpe6}$

Для решения задачи адаптивного управления производительностью GIS разработан метод динамического формирования пулов (*D-FIST*).

#### V. МЕТОД ДИНАМИЧЕСКОГО ФОРМИРОВАНИЯ ПУЛОВ (*D-FIST – DYNAMIC FULL INFRASTRUCTURE OF SOURCES TOOLKIT*)

Каждая задача, решаемая GIS имеет точку входа, т.е. конкретный узел GIS, который принял задачу к исполнению. Если ресурсов у точки входа достаточно,

задачу целесообразно выполнить в ней. Иначе собирается пул. При этом, чем дальше от точки входа находятся другие элементы-кандидаты на формирование пула, тем больше вариантов требуется перебрать, чтобы решить задачу об упаковке контейнера. Обычно в реальном масштабе времени комбинаторные задачи решаются с помощью динамического программирования или генетических алгоритмов.

Количество вариантов включения элементов в пул растёт нелинейно с удалением от точки входа задачи, что замедляет формирование пула, а значит, решение всей полученной задачи.

Если ГИС предстоит решать задачи и формировать пулы в реальном или близком к реальному масштабу времени, то необходимо сократить количество возможных вариантов, которые анализирует система управления ГИС ( $\mathcal{P}^{JLC}$ ). Это достигается применением **принципа постепенного распространения задач по ГИС**. Следуя принципу, в пул добавляются только те элементы ГИС, которые:

1. подходят для этого (обладают минимально допустимой производительностью);
2. ближе всего расположены к точке входа задачи. Метрикой расстояния выступают такие величины как: количество путей, идущих через элемент; *hop*; количество связей, созданных элементом и пр.

Принцип постепенного распространения задач по ГИС лежит в основе метода динамического формирования пулов.

Введём следующие обозначения:

$V$  – множество элементов ГИС;

$V_r$  – упорядоченное множество элементов-кандидатов, которые могут участвовать в формировании пулов,  $V_r \subset V$ ;

$V_{\Omega_j}$  – множество элементов, запланированных в  $i$ -й пул;

$\omega_i^{\min} > 0$  – минимальная допустимая производительность элемента в  $i$ -м пуле;

$r_{\omega_j}$  – радиус  $j$ -ого элемента в  $i$ -м пуле;

$$r_{\omega_j} = \frac{\omega_i^{\min}}{\omega_j}$$

$x_j$  – количество физических связей  $j$ -ого элемента;

$U_x$  – множество команд на образование физических связей;

*Физическая связь* – связь на физическом уровне или MAC-подуровне канального уровня модели OSI.

$U_l$  – множество команд на образование логических связей;

*Логическая связь* – связь, образованная на LLC-подуровне канального уровня, сетевом или транспортном уровнях модели OSI.

$U_k$  – множество команд на перемещение элементов.

Входными данными для предлагаемого метода динамического формирования пула являются:

количество пулов  $N$ ;

размеры пулов  $\Omega_i^{mpeб}, i = \overline{1, N}$ ;

минимально возможная производительность элемента в  $i$ -м пуле  $\omega_i^{\min} > 0$ ;

текущая производительность  $j$ -ого элемента  $\omega_j > 0$ .

Схема метода приведена на рис. 4. На первом шаге сужается множество возможных комбинаций за счёт выбора и сортировки элементов-кандидатов, подходящих для формирования пула.

Шаг 1. Выбор элементов-кандидатов (рис. 5).

Выбор элементов-кандидатов реализуется на основе радиуса элемента  $r_{\omega_j} \leq 1$ , то элемент имеет нужную минимальную производительностью, и помечается как кандидат на включение в пул. Элементы, имеющие  $r_{\omega_j} \leq 1$ , образуют множество  $V_r$ .

После формирования множества  $V_r$  элементы упорядочиваются, например, по уменьшению текущей производительности элемента  $\omega_j$ , по увеличению количества образованных связей или какому-то другому правилу, лучше всего подходящему для решения загруженной задачи.



Рис. 4. Схема метода динамического формирования пулов

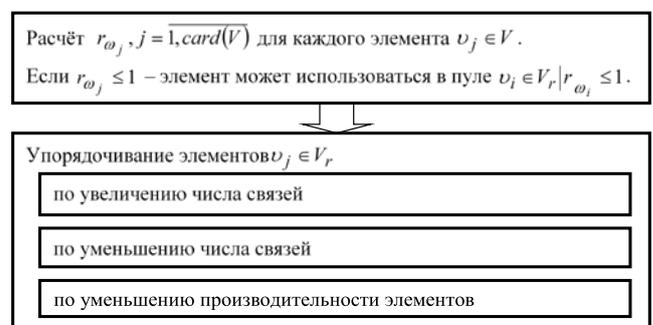


Рис. 5. Шаг 1. Выбор элементов-кандидатов в пул

По итогам первого шага формируется упорядоченное множество элементов-кандидатов ( $V_r$ ), подходящих для

создания пула. Это множество используется в качестве отправной точки для второго шага метода.

Шаг 2. Комбинирование пула (рис. 6).

Второй шаг предполагает дополнительное сужение множества альтернатив при формировании пула. Для этого используется принцип постепенного распространения задач по ГИС, а значит, в первую очередь планируются в пул элементы, смежные с точкой входа решаемой задачи (расстояние до точки входа 1).

Точка входа  $v^*$  – элемент, где задача начинает выполняться/загружается в ГИС. В общем случае является произвольным элементом ГИС.

Для элементов с расстоянием до точки входа, равным единице, решается сформулированная выше модифицированная задача по упаковке контейнеров. На рисунке представлена закономерность для формирования пулов из вычислителей или каналов связи. Если задача решается успешно и пул с требуемой производительностью может быть сформирован, то происходит переход к шагу 3 метода, чтобы физически сформировать пул и передать его для решения задачи.

В противном случае, когда производительности элементов, смежных с точкой входа недостаточно, радиус увеличивается на единицу (расстояние до точки входа равно двум), и производится попытка создать пул на базе элементов, смежных с элементами, имеющими радиус 1 (смежными с точкой входа).

После чего выполняется проверка скомбинированного пула на соответствие предъявляемым требованиям. Скомбинированный пул обладает требуемой производительностью? Переход к шагу 3, иначе – увеличить расстояние до точки входа на единицу (равно 3) и повторить действия по комбинированию пула с использованием новых элементов.

Процесс комбинирования пула выполняется до тех пор, пока не будет скомбинирован пул с требуемой производительностью или не исчерпаются все элементы, имеющие физическую связь.

Если элементы, имеющие физическую связь исчерпаны, а пул не был скомбинирован, то производится попытка создать варианты с элементами, между которыми пока нет физической связи, но она может быть установлена, например, с мобильными устройствами ГИС или устройствами её пользователей (подвижные центры обработки данных, автоматические комплексы разного назначения, беспилотные летательные аппараты и т.п.). После этого реализуются действия по комбинированию пула по правилам шага 2.

Результаты комбинирования пула выдаются системе управления задачами, запросившей ресурс, то есть на уровень ПО. Система решает целесообразно использовать скомбинированный пул или нет.

Если смысл использовать пул есть, то система посылает команды на физическое перемещение/коммутиацию элементов ГИС (шаг 3), образование физических и/или логических связей и создаёт пул, скомбинированный на шаге 2.

Шаг 3. Формирование управления на изменение (рис. 7).

Здесь проверяется физическая связность комбинаций, найденных на шаге 2.

Если физическая связь присутствует между всеми элементами скомбинированного пула, то выполняется физическое формирование пула и элементы связываются между собой логически, т.е. создаются логические связи. Иначе формируются новые связные сегменты по специальному алгоритму.

Если возможно несколько вариантов формирования пула, то выбор конкретного варианта происходит случайно, либо с предъявлением дополнительных требований. Например, к структуре по максимизации живучести [26].

После формирования  $i$ -ого пула шаги 1-3 метода повторяются для  $i+1$  пула.

Если в процессе выполнения метода изменился поток задач, скорректировавший требования к пулам  $(\Omega_i^{mpeб}, \omega_i^{\min}, i = \overline{1, N})$ , либо состав ГИС ( $V$ ) разрушается по каким-то причинам, то на уровне программного обеспечения принимается решение выполнить метод сначала.

Очевидно, метод наиболее эффективен, когда решаемые задачи могут быть распараллелены. Если задача не может быть распараллелена, то она представляется в виде единственного пула  $\Omega$  с минимальной допустимой производительностью.

Метод целесообразно применять, когда задача требует производительности  $(\Omega_i^{треб})$  больше, чем способен предоставить существующий пул  $\Omega_i$ :

$$\Omega_i < \Omega_i^{mpeб}, \Omega_i^{mpeб} = Const.$$

Докажем, что предложенный метод  $D-FIST$  обладает свойствами сходимости и завершаемости.

## VI. СХОДИМОСТЬ

Метод  $D-FIST$  устраняет разницу между требуемой производительностью и производительностью пула  $\Delta\Omega_i \rightarrow 0, \Delta\Omega_i^{mpeб} - \Omega_i$  (см. постановку задачи). Это возможно, когда производительность пула стремится к требуемой. Требуемая производительность определяется решаемыми задачами и является постоянной на протяжении всего периода работы метода. Следовательно, чтобы доказать сходимость метода, достаточно доказать, что производительность пула стремится к требуемой.

Элементы ГИС, используемые для формирования пулов, обладают положительной производительностью  $\omega_z > 0$  и при добавлении не уменьшают производительность пула (см. ограничение 1). А значит, на каждом шаге метода (при каждом включении элемента с производительностью  $\omega_z$ ) производительность пула  $\Omega_i$  не убывает. Метод прекращает работу, когда исчерпаны все элементы или производительность пула становится не меньше требуемой. Следовательно, производительность пула стремится к требуемой  $\Omega_i \rightarrow \Omega_i^{mpeб}$ , и метод сходится.

VII. ЗАВЕРШАЕМОСТЬ

Метод *D-FIST* работает, пока  $\Omega_i \leq \Omega_i^{mpeб}$  либо пока  $V_r \neq \emptyset$ . На каждом шаге  $\Omega_i$  не убывает (см. ограничение 1) и сокращается  $V_r = V_r \setminus \{v\}$ . Т.к.  $V_r$  и  $\{v\}$  ограничены и замкнуты, следовательно,  $|V_r| < \infty$ , а значит, метод завершаем.

Предложенный метод реализован и проверен в имитационной модели, написанной в среде MatLab.

Если за  $\omega$  обозначить пропускную способность канала, а за  $\Omega$  – пропускную способность транка, то метод подойдет для формирования пулов из каналов

связи.

Чтобы применить метод для устройств ввода/вывода или накопителей, необходимо закономерность  $\Omega_i \leq \sum_{z=1}^{|V_{\Omega_i}|} \omega_z$ , заменить на закономерности, соответствующие устройствам ввода/вывода и накопителям.

В случаях критического дефицита ресурсов метод расширяется методикой, позволяющей сохранить количество задач, решаемых ГИС, но пожертвовать их точностью или временем выполнения [27].

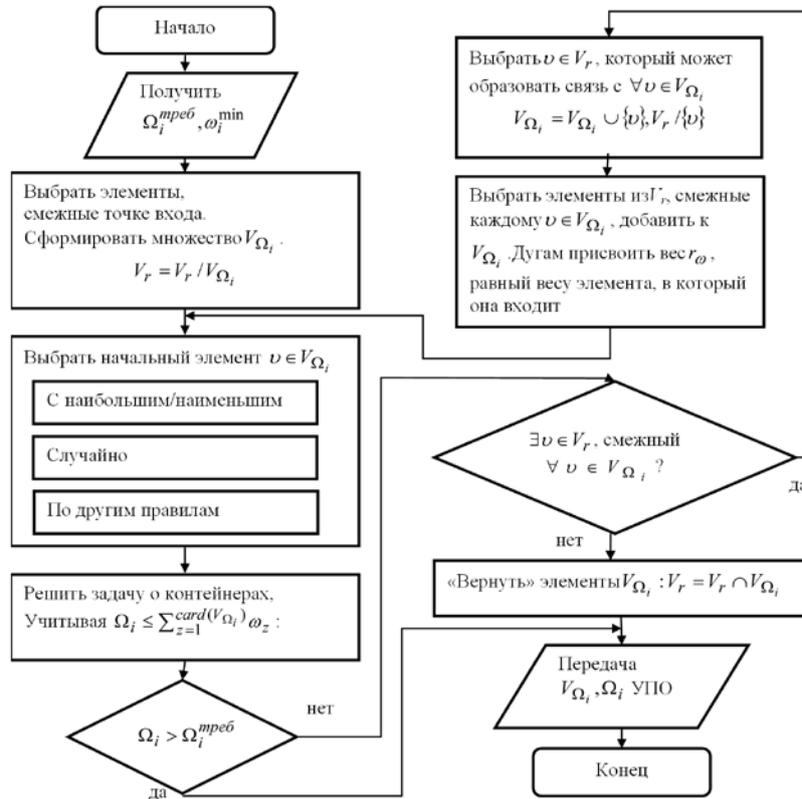


Рис. 6. Шаг 2. Комбинирование пула

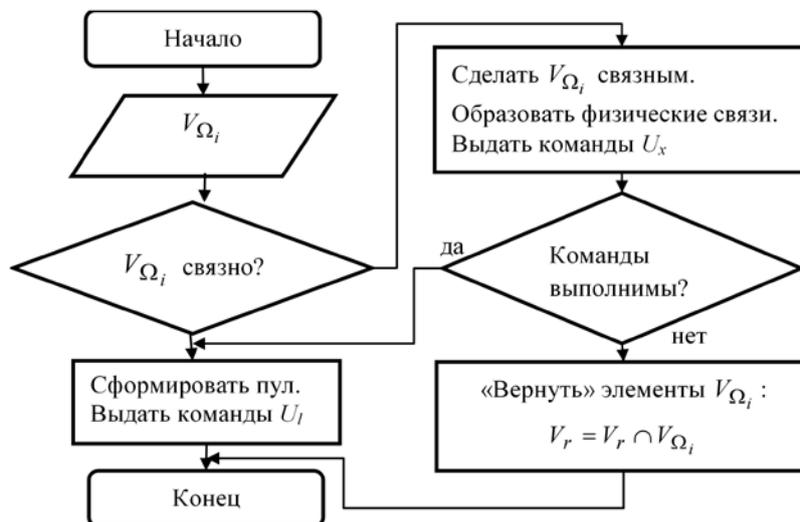


Рис. 7. Шаг 3. Формирование управления на изменение

VIII. ПРИМЕР ПРИМЕНЕНИЯ *FIST* и *D-FIST*

Предположим, что пользователи ГИС захотели вывести: 1) все объекты, находящиеся на определённом расстоянии от объекта  $X$ ; 2) имеющие представителей в объекте  $X$ ; 3) перечень данных, к которым у пользователя есть доступ на чтение; 4) список пользователей, у которых есть доступ к объекту  $X$ .

Задача может быть решена путём декомпозиции на 4 части. Предположим, что минимальная требуемая производительность  $\omega_i^{\min}=3$  условных единиц. Требования к производительности пулов для каждой части  $\Omega = \{ \langle \Omega_i, \omega_i^{\min} \rangle \} = \{ \langle 17,3 \rangle, \langle 19,3 \rangle, \langle 3,3 \rangle, \langle 4,3 \rangle \}$ .

ГИС имеет структуру (рис. 8), где цифры показывают максимальные производительности каждого элемента ГИС. Возможность установления физической связи между элементами обозначена штрих-пунктирной линией.

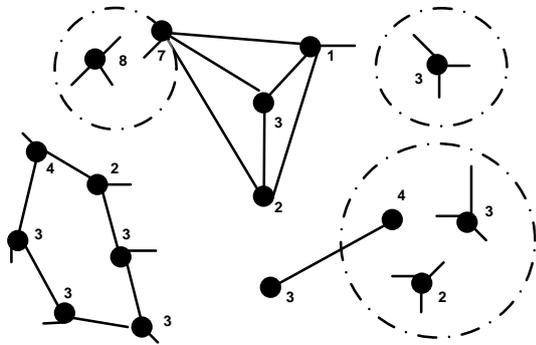


Рис. 8. Производительности элементов ГИС и связи между ними

Создаём первый пул, который должен иметь характеристики  $\Omega = \{ \langle \Omega_i, \omega_i^{\min} \rangle \} = \{ \langle 17,3 \rangle \}$ .

–  $r_\omega$  элементов представлены в виде простой дроби, в числителе которой минимальная требуемая производительность  $\omega_i^{\min}$ , а в знаменателе максимальная возможная производительность элемента. Таким образом получаем множество  $V_r$ . Если круг не закрашен, элемент не может включаться в пул (рис. 9);

– сортируем элементы во множестве  $V_r$  по убыванию числа имеющихся физических связей. В общем виде основание сортировки может быть другим. Указатели содержат номер, под которым элемент входит во множество  $V_r = \{1,2,3,4,5,6,7,8,9,10,11,12\}$  (рис. 9).

Для упрощения примера потребуем производительности других типов (каналов связи, накопителей и устройств ввода/вывода) достаточно для работы ГИС в любой конфигурации. Каждый элемент способен пропускать через себя данные транзитом.

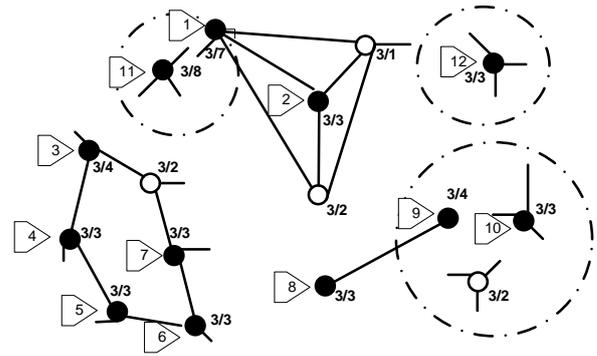


Рис. 9. Выбор элементов ГИС, входящих во множество  $V_r$

Шаг 2. Дополнительно сужаем множество альтернатив. Пусть элемент №1 ( $v_1$ ) – точка входа для задачи. Начинаем формирование  $V_{\Omega_1}$  с включение этого элемента

$$1. V_r = \{2,3,4,5,6,7,8,9,10,11,12\}, V_{\Omega_1} = \{1\}, \Omega_1 \leq 7 < 17;$$

2. Увеличиваем множество  $V_{\Omega_1}$ , включив в него элемент со следующим номером, имеющий физическую связь с первым элементом. И проверяем, выполнено ли требование по достижению требуемой производительности пула  $\Omega_1$ :

$$V_r = \{3,4,5,6,7,8,9,10,11,12\}, V_{\Omega_1} = \{1,2\}, \Omega_1 \leq 7 + 3 = 10 < 17;$$

3. Требования к производительности первого пула  $\Omega_1$  не выполнено. Элементы из  $V_r$ , имеющие физическую связь с элементами, уже вошедшими в пул, отсутствуют. Согласно *D-FIST*, выбираем элемент с наименьшим номером, расположенным ближе всего к точке входа:

$$V_r = \{3,4,5,6,7,8,9,10,12\}, V_{\Omega_1} = \{1,2,11\}, \Omega_1 \leq 10 + 8 = 18 \geq 17.$$

В итоге получили пул с производительностью не ниже требуемой  $\langle \Omega_1, \omega_1^{\min} \rangle = \langle 18,3 \rangle$ .

Далее применяем *D-FIST* для формирования других пулов.

На очереди пул  $i=2$ ,  $\langle \Omega_2, \omega_2^{\min} \rangle = \langle 19,3 \rangle$  Элемент №3 является точкой входа.

$$1. V_r = \{4,5,6,7,8,9,10,11,12\}, V_{\Omega_2} = \{3\}, \Omega_2 \leq 4 < 19;$$

$$2. V_r = \{5,6,7,8,9,10,11,12\}, V_{\Omega_2} = \{3,4\}, \Omega_2 \leq 4 + 3 < 19;$$

3. 7 и 5 элементы равноудалены от точки входа на два хопа. Согласно выбранной логике, начинаем с элемента, имеющего меньший номер – 5,

$$V_r = \{6,7,8,9,10,12\}, V_{\Omega_2} = \{3,4,5\}, \Omega_2 \leq 7 + 3 = 10 < 19;$$

$$4. V_r = \{6,8,9,10,12\}, V_{\Omega_2} = \{3,4,5,7\}, \Omega_2 \leq 10 + 3 = 13 < 19;$$

$$5. V_r = \{8,9,10,12\}, V_{\Omega_2} = \{3,4,5,7,6\}, \Omega_2 \leq 13 + 3 = 16 < 19.$$

Как видно, закончились элементы, которые можно добавить в пул, и сформированный пул не обладает требуемой производительностью. Пусть на уровне ПО решено, что использовать такой пул не имеет смысла. Следовательно, пул должен быть расформирован, и элементы из возвращены в  $V_r$ .



загрузки.

#### СУБД NoSQL в ГИС

Неструктурированные данные могут загружаться и обрабатываться с помощью дополнительных сервисов, мобильных пользователей, подвижных элементов ГИС и т.д. Это обстоятельство требует от СУБД организовать оперативный доступ. С точки зрения скорости доступа к большим массивам данных лучшим вариантом служит организация данных в виде графа. Это реализовано в СУБД NoSQL, что делает СУБД NoSQL предпочтительнее, чем СУБД SQL.

Дополнительным преимуществом применения СУБД NoSQL в ГИС является масштабирование. Добавление нового узла в кластер или пул позволяет увеличить быстродействие обработки запросов на предоставляемые ресурсы и распределить нагрузку. Одним из недостатков СУБД NoSQL является нерешенная задача топологической сортировки данных, т.к. отсутствует критерий выбора и доступности тех или иных данных.

Стоит принимать во внимание, что СУБД NoSQL не использует полностью язык SQL, т.к. данные организованы иначе и команды объединения выполняются крайне медленно или не выполняются вообще. Это означает, что построить полноценные аналитические отчёты средствами СУБД NoSQL не представляется возможным, а значит, сложно работать со слоями ГИС.

Однако существуют программные комплексы, в которых пытаются реализовать полноценную аналитику, например, Vantage [32] или ГИС, использующая Hadoop [33], SpatialHadoop и GeoSpark [34].

Таким образом, современная ГИС, основанная на технологии туманных вычислений, должна сочетать в себе и СУБД SQL, и СУБД NoSQL. Наиболее распространённая технология, которая пытается это осуществить – MapReduce. На сегодняшний день на рынке присутствуют программные продукты трёх компаний. В таблице 1 представлены для сравнения часть модулей, остальные модули изменяются и дополняются от версии к версии, т.к. разработка происходит динамично.

Таблица 1. Сравнение продуктов технологии MapReduce в коммерческих дистрибутивах Hadoop

Компания	Программный продукт	Модуль NoSQL	Модуль SQL	Файловая система	Управление ресурсами
Cloudera	Apache Hadoop ecosystem	HBase	Impala	HDFS	Yarn
HPE	HPE Ezmeral Data Fabric	HBase	Hive	HDFS	Yarn
ArenaData	ArenaData Hadoop	HBase	Hive	HDFS	Yarn

#### Х. ЗАКЛЮЧЕНИЕ

Подытожив сказанное, отметим, что ГИС являются системами критического применения. Предоставляя необходимый ресурс, т.е. реализуя такой аспект защиты информации как доступность, ГИС создаёт необходимые условия для защиты целостности и конфиденциальности.

Смысловое значение решаемых задач безразлично для ГИС. ГИС важно знать, какой ресурс и на какой срок требуется для решения задачи. В ГИС задача представляется в виде совокупности четырёх типов производительностей: вычислителей, каналов связи, накопителей, устройств ввода/вывода. Преобразование производительности физических элементов ГИС (ГГц, боды и т.д.) в производительность, понятную пользователю (задачи в единицу времени) производится с использованием модели FIST.

ГИС функционируют в условиях действия дестабилизирующих факторов разной природы, которые не позволяют решить все поставленные пользователем задачи точно и в срок. Адаптация ГИС к действию дестабилизирующих факторов реализуется с помощью метода D-FIST. Метод базируется на принципе постепенного распространения задач и является сходимым и завершаемым. Пулы, созданные с помощью метода D-FIST, могут использоваться как услуга (Fog-as-a-Service)

В дальнейшем необходимо выявить конкретный вид функций, описывающих производительность пулов, формируемых из устройств ввода-вывода и накопителей. Отдельный интерес представляет возможность замены одного типа производительности другим, то есть возможность преобразовать один тип производительности в другой: производительность вычислителей в производительность каналов связи, производительность каналов связи в производительность устройства ввода-вывода и т.д.

На сегодняшний день существуют технологии, позволяющие реализовать предложенные модель FIST и метод D-FIST.

#### БИБЛИОГРАФИЯ

- [1] R. J. Winchell, J. Broecker, A. J. Kerwin, B. Eastridge, and M. Crandall, "Comparing geographic information system-based estimates with trauma center registry data to assess the effects of additional trauma centers on system access", *Journal of Trauma and Acute Care Surgery*, vol. 89, issue 6, pp. 1131–1135, 2020.
- [2] В. А. Куделькин и Ф. В. Денисов, "Опыт интеграции распределенных информационных систем", *ИТ Стандарт, № 1, С. 24–30, 2017.*
- [3] О. В. Столетов, И. А. Чихарев, О. А. Москаленко, и Д. В. Маковская, "Геоинформационное сопровождение средиземноморской ветви Шелкового пути", *Интеркарто. Интергис*, Т. 25, № 1, С. 102–113, 2019.
- [4] K. Schwab, "Globalization 4.0. New architecture for the fourth industrial revolution", *Eurasian integration: economics, law, politics*, № 1 (27), pp. 79–84, 2019.
- [5] М. О. Колбанёв, И. И. Палкин, и Т. М. Татарникова, "Вызовы цифровой экономики", *Гидрометеорология и экология*, № 58, С. 156–67, 2020.
- [6] *Геоинформационные системы. Термины и определения*, ГОСТ Р 52438-2005, 2018.

- [7] В. В. Грызунов и А. О. Нестерова, “Структура живучей сети метеокомплексов транспортно-логистических систем «Индустрии 4.0»”, *Гидрометеорология и экология*, № 59, С. 111–123, 2020.
- [8] J. Zhang, L. Xu, Y. Zhang, G. Liu, L. Zhao, and Y. Wang, “An On-Demand Scalable Model for Geographic Information System (GIS) Data Processing in a Cloud GIS”, *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, pp. 392, 2019.
- [9] C. Quinde, D. Guillermo, L. Siguenza-Guzman, D. Orellana, and P. A. Pesántez-Cabrera, “Software Architecture Proposal for a Data Platform on Active Mobility and Urban Environment”, in 2020 Conference on Information and Communication Technologies of Ecuador (TICEC 2020), pp. 501–515.
- [10] S. Sumith Sathendran and S. Smitha Chandran, “Development of a Cloud Based Spatial Information System for Karunagappally Municipality, Kerala, The God’s Own Country”, *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 7, issue 7, pp. 2304–2310, 2020.
- [11] H. Broesamle et al., “Assessment of solar electricity potentials in North Africa based on satellite data and a geographic information system”, *Solar Energy*, vol. 70, no 1, pp. 1–12, 2001.
- [12] Official web site of Data Center Companies. Available: <https://www.datacenters.com/directory/companies>.
- [13] *Ieee 1934-2018-ieee standard for adoption of open fog reference architecture for fog computing*, IEEE Standard Association et al., 2018.
- [14] E. Panidi, “Fog computing perspectives in connection with the current geospatial standards”, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, pp. 171–174, 2017.
- [15] J. Das, A. Mukherjee, S. K. Ghosh, and R. Buyya, “Spatio-Fog: A green and timeliness-oriented fog computing model for geospatial query resolution”, *Simulation Modelling Practice and Theory*, vol. 100, pp. 102043, 2020.
- [16] A. Kapsalis, P. Kasnesis, I. S. Venieris, D. I. Kaklamani, and C. Z. Patrikakis, “A cooperative fog approach for effective workload balancing”, *IEEE Cloud Computing*, 4 (2), pp. 36–45, 2017. doi:10.1109/MCC.2017.25.
- [17] A. Yousefpour, G. Ishigaki, and J. P. Jue, “Fog Computing: Towards Minimizing Delay in the Internet of Things”, in *Proceedings - 2017 IEEE 1st International Conference on Edge Computing*, EDGE 2017, pp. 17–24.
- [18] R. Beraldi et al., “Distributed load balancing for heterogeneous fog computing infrastructures in smart cities”, *Pervasive and Mobile Computing*, pp. 101221, 2020.
- [19] В. В. Грызунов, “Концептуальная модель адаптивного управления геоинформационной системой в условиях дестабилизации”, *Проблемы информационной безопасности. Компьютерные системы*, № 1(45), С. 102–108, 2021.
- [20] M. Mukherjee, R. Matam, L. Shu, L. Maglaras, M. A. Ferrag, N. Choudhury, and V. Kumar, “Security and privacy in fog computing: Challenges”, *IEEE Access*, vol. 5, pp. 19293–19304, 2017.
- [21] В. В. Грызунов, “Аналитическая модель целостной информационной системы”, *Доклады Томского государственного университета систем управления и радиоэлектроники*, № 1-1 (19), С. 226–230, 2009.
- [22] В. Г. Бурлов и А. М. Гробицкая, “Управление строительным производством с учетом показателя успешного выполнения производственного задания”, *Инженерно-строительный журнал*, № 3, С. 77–91, 2016.
- [23] Р. М. Юсупов и другие, *Элементы теории испытаний и контроля технических систем*. Ленинград: Энергия. Ленинград. отд-ние, 1978, 191 с.
- [24] Y. Sun, F. Lin, and H. Xu, “Multi-objective optimization of resource scheduling in fog computing using an improved NSGA-II”, *Wireless Personal Communications*, vol. 102, no. 2, pp. 1369–1385, 2018.
- [25] M. Aazam, S. Zeadally, and K. A. Harras, “Offloading in fog computing for IoT: Review, enabling technologies, and research opportunities”, *Future Generation Computer Systems*, vol. 87, pp. 278–289, 2018.
- [26] В. В. Грызунов, “Оценивание живучести неоднородных структур”, *Вестник СибГУТИ*, № 1, С. 28–35, 2011.
- [27] В. В. Грызунов, “Методика решения измерительных и вычислительных задач в условиях деградации информационно-вычислительной системы”, *Вестник СибГУТИ*, № 1, С. 35–44, 2015.
- [28] Федеральный закон от 2 декабря 2019 г. N 405-ФЗ “О внесении изменений в отдельные законодательные акты Российской Федерации”.
- [29] С. Н. Леоненков, “Целевая оптимизация структуры потока задач суперкомпьютеров”, *Вычислительные методы и программирование*, Т. 20, № 3, С. 199–210, 2019.
- [30] В. В. Грызунов и Д. А. “Украинцева, Облик системы защиты аутентичности данных геоинформационной системы”, в *Информационная безопасность регионов России (ИБРР-2019): Материалы XI Санкт-Петербургской межрегиональной конференции (Санкт-Петербург, 23-25 октября 2019 г.)*, СПб., 2019, С. 172–173.
- [31] “Варианты архитектуры для обработки больших данных”. Available: [www.dataved.ru/2014/06/big-data-architecture.html](http://www.dataved.ru/2014/06/big-data-architecture.html).
- [32] *The cloud data analytics platform (Teradata)*. Available: [www.teradata.ru/Products/Software/Vantage/Analyst](http://www.teradata.ru/Products/Software/Vantage/Analyst).
- [33] Z. Zhang H. Sun, Z. Liu, C. Xu, and L. Wang, “Dart: A geographic information system on hadoop”, in *2015 IEEE 8th International Conference on Cloud Computing*, pp. 90–97.
- [34] R. K. Lenka, R. K. Barik, N. Gupta, S. M. Ali, A. Rath, and H. Dubey, “Comparative analysis of SpatialHadoop and GeoSpark for geospatial big data analytics”, in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 484–488.

# Adaptive accessibility management in geographic information systems using fog computing

Vyacheslav Burlov, Vitaliy Gryzunov, Dmitriy Sipovich

**Abstract**—Geographic information systems are integrated with information systems of enterprises and the state, are systems of critical application and work with a large amount of heterogeneous and unstructured data. The volume of data, the number of users and cyberattacks is increasing every year. Ensuring information security, in turn, requires the availability of the system resource. The applied methods of centralized storage and processing of data do not cope with the assigned tasks, cannot ensure the availability of the requested resource accurately and on time, which further leads to a violation of other aspects of information security: integrity and confidentiality. One possible solution to the accessibility problem is the use of foggy computing.

The article proposes: a hierarchical model of a geographic information system using fog computing (FIST – Full Infrastructure of Sources Toolkit). The model includes the levels of software, logical structure, physical structure. Interaction between the levels of the model is formalized. An example of a basic law allowing you to combine individual elements into a pool (fog node) is given. The principle of gradual spreading of tasks in the geographic information system is formulated, the task of adaptive control of the geographic information system performance is posed as a modified task of container packing, the method of pooling the necessary resources into pools is disclosed: computers, communication channels, input / output devices, storage devices. The method (D-FIST – Dynamic Full Infrastructure of Sources Toolkit) includes 3 steps: selection of candidate elements, pools combination, formation of management for change. The convergence and termination of the proposed method is proved. The features of the data with which geographic information systems work, and modern technologies, on the basis of which the proposed method can be implemented, are analyzed.

**Keyword**— Fog computing, geographic information system, information security.

## REFERENCES

- [1] R. J. Winchell, J. Broecker, A. J. Kerwin, B. Eastridge, and M. Crandall, “Comparing geographic information system-based estimates with trauma center registry data to assess the effects of additional trauma centers on system access”, *Journal of Trauma and Acute Care Surgery*, vol. 89, issue 6, pp. 1131–1135, 2020.
- [2] V. A. Kudelkin and V. F. Denisov, “Experience of integration of distributed information systems”, *IT standard*, no. 1, pp. 24–30, 2017.
- [3] O. V. Stoletov, I. A. Chikharev, O. A. Moskalenko, and D. V. Makovskaya, “Geoinformation support of the mediterranean branch of the silk road”, *Intercarto. InterGIS*, part 25, no. 1, pp. 102–113, 2019.
- [4] K. Schwab, “Globalization 4.0. New architecture for the fourth industrial revolution”, *Eurasian integration: economics, law, politics*, № 1 (27), pp. 79–84, 2019.
- [5] M. O. Kolbanev, I. I. Palkin, and T. M. Tatarnikova, “The challenges of the digital economy”, *Hydrometeorology and ecology*, no. 5, pp. 156–167, 2020.
- [6] *Geographical information systems. Terms and definitions*, GOST R 52438-2005, 2018.
- [7] V. V. Gryzunov and A. O. Nesterova, “Survivable structure of network of meteorological complexes of transport and logistics systems “industry 4.0””, *Gidrometeorologija i jekologija*, no. 59, pp. 111–123, 2020.
- [8] J. Zhang, L. Xu, Y. Zhang, G. Liu, L. Zhao, and Y. Wang, “An On-Demand Scalable Model for Geographic Information System (GIS) Data Processing in a Cloud GIS”, *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, pp. 392, 2019.
- [9] C. Quinde, D. Guillermo, L. Siguenza-Guzman, D. Orellana, and P. A. Pesántez-Cabrera, “Software Architecture Proposal for a Data Platform on Active Mobility and Urban Environment”, in 2020 Conference on Information and Communication Technologies of Ecuador (TICEC 2020), pp. 501–515.
- [10] S. Sumith Satheendran and S. Smitha Chandran, “Development of a Cloud Based Spatial Information System for Karunagappally Municipality, Kerala, The God’s Own Country”, *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 7, issue 7, pp. 2304–2310, 2020.
- [11] H. Broesamle et al., “Assessment of solar electricity potentials in North Africa based on satellite data and a geographic information system”, *Solar Energy*, vol. 70, no 1, pp. 1–12, 2001.
- [12] *Official web site of Data Center Companies*. Available: <https://www.datacenters.com/directory/companies>.
- [13] *Ieee 1934-2018-ieee standard for adoption of open fog reference architecture for fog computing*, IEEE Standard Association et al., 2018.
- [14] E. Panidi, “Fog computing perspectives in connection with the current geospatial standards”, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, pp. 171–174, 2017.
- [15] J. Das, A. Mukherjee, S. K. Ghosh, and R. Buyya, “Spatio-Fog: A green and timeliness-oriented fog computing model for geospatial query resolution”, *Simulation Modelling Practice and Theory*, vol. 100, pp. 102043, 2020.
- [16] A. Kapsalis, P. Kasnesis, I. S. Venieris, D. I. Kaklamani, and C. Z. Patrikakis, “A cooperative fog approach for effective workload balancing”, *IEEE Cloud Computing*, 4 (2), pp. 36–45, 2017. doi:10.1109/MCC.2017.25.
- [17] A. Yousefpour, G. Ishigaki, and J. P. Jue, “Fog Computing: Towards Minimizing Delay in the Internet of Things”, in *Proceedings - 2017 IEEE 1st International Conference on Edge Computing*, EDGE 2017, pp. 17–24.
- [18] R. Beraldi et al., “Distributed load balancing for heterogeneous fog computing infrastructures in smart cities”, *Pervasive and Mobile Computing*, pp. 101221, 2020.
- [19] V. V. Gryzunov, “Conceptual model of geoinformation system adaptive control under conditions of destabilization”, *Information security problems. Computer systems*, no. 1 (45), pp. 102–108, 2021.
- [20] M. Mukherjee, R. Matam, L. Shu, L. Maglaras, M. A. Ferrag, N. Choudhury, and V. Kumar, “Security and privacy in fog computing: Challenges”, *IEEE Access*, vol. 5, pp. 19293–19304, 2017.
- [21] V. V. Gryzunov, “The analytical model of the whole information system”, *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki*, no. 1-1 (19), pp. 226–230, 2009.
- [22] V. G. Burlov and A. M. Grobickaja, “Construction management in terms of indicator of the successfully fulfilled production task”, *Inzhenerno-stroitel'nyj zhurnal*, no. 3, pp. 77–91, 2016.

- [23] R. M. Yusupov and others, "Elements of the theory of testing and control of technical systems". Leningrad: Energy. Leningrad. unit, 1978, 191 p.
- [24] Y. Sun, F. Lin, and H. Xu, "Multi-objective optimization of resource scheduling in fog computing using an improved NSGA-II", *Wireless Personal Communications*, vol. 102, no. 2, pp. 1369–1385, 2018.
- [25] M. Aazam, S. Zeadally, and K. A. Harras, "Offloading in fog computing for IoT: Review, enabling technologies, and research opportunities", *Future Generation Computer Systems*, vol. 87, pp. 278–289, 2018.
- [26] V. V. Gryzunov, "The estimation of the survivability of heterogeneous structure", *Vestnik SibGUTI*, no. 1, pp. 28–35, 2011.
- [27] V. V. Gryzunov, "Problem solving method of measuring and calculating tasks under conditions of data computing system degradation", *Vestnik SibGUTI*, no. 1, pp. 35–44, 2015.
- [28] Federal law No. 405-FZ of December 2, 2019 "On amendments to certain legislative acts of the Russian Federation".
- [29] S. N. Leonenkov, "Target optimization of a supercomputer task flow", *Vychislitel'nye metody i programirovanie*, vol. 20, no. 3, pp. 199–210, 2019.
- [30] V. V. Gryzunov and D. A. Ukraineva, "Technical features of the investigation of computer crimes", in *Information security of russian regions (ISRR-2019): XI St. Petersburg interregional conference*, SPb., 2019, pp. 172–173.
- [31] "Architecture choices for big data processing". Available: [www.dataved.ru/2014/06/big-data-architecture.html](http://www.dataved.ru/2014/06/big-data-architecture.html).
- [32] *The cloud data analytics platform (Teradata)*. Available: [www.teradata.ru/Products/Software/Vantage/Analyst](http://www.teradata.ru/Products/Software/Vantage/Analyst).
- [33] Z. Zhang H. Sun, Z. Liu, C. Xu, and L. Wang, "Dart: A geographic information system on hadoop", in *2015 IEEE 8th International Conference on Cloud Computing*, pp. 90–97.
- [34] R. K. Lenka, R. K. Barik, N. Gupta, S. M. Ali, A. Rath, and H. Dubey, "Comparative analysis of SpatialHadoop and GeoSpark for geospatial big data analytics", in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 484–488.