Проектирование архитектурного шаблона кибериммунной ГИС

Выполнил: К.Р.Ивашнев, аспирант кафедры ИМО
Руководитель: Д.Ж.Корзун, к.ф.-м.н., доцент

ПетрГУ 2025

Содержание

Введение

Геоинформационные системы - базы знаний, выполняющие обработку геопространственной информации для поддержки принятия решений.

Функции:

- Пространственный анализ
- Моделирование
- Прогнозирование

Сферы применения:

- Ситуационный мониторинг
- Городское планирование
- Мониторинг окружающей среды
- Транспортная логистика
- Управление чрезвычайными ситуациями

Пример:

- Система мониторинга и интеллектуального анализа тематических источников по арктической теме
- Функции:
  - Сбор, классификация и структурирование Интернет-данных
  - Суммаризация и перевод информации
  - Формирование ежедневных аналитических сводок
- Сфера применения:
  - Поддержка ситуационного мониторинга

Проблематика

Угрозы безопасности ГИС

- Широкая интеграция с различными источниками данных: веб-ресурсы, датчики, сервисы и мобильные устройства
- Проблема доверенности данных и несанкционированного доступа
- Результатом компрометации данных является несвоевременное/некорректное принятие решений

Угрозы безопасности прикладной системы:

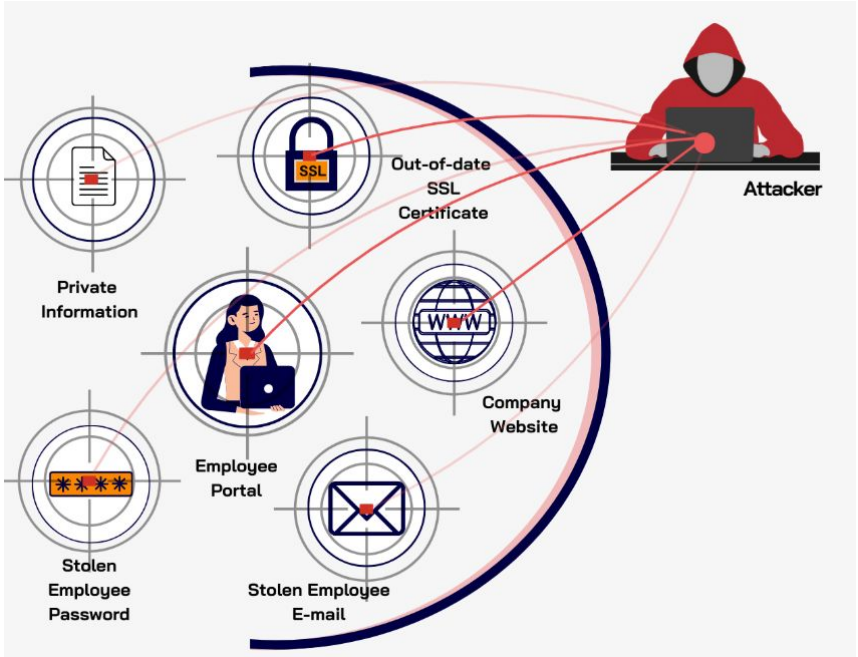
- Перехват данных при их передаче через открытые каналы
- Компрометация и искажение данных, поступающих с веб-ресурсов
- Несанкционированный доступ в веб-приложение и API

Методологические барьеры кибербезопасности

- Высокая сложность и стоимость доказательства свойств безопасности
- Отсутствие единого понимания целей деятельности между заказчиками и разработчиками

Подходы к обеспечению безопасности

- **Применение наложенных средств безопасности в зависимости от субъекта угроз**
- Пример: Firewall для защиты от DDOS-атак
- Невозможность обеспечения абсолютной защиты
- Реактивный подход, неэффективен при обширной поверхности атаки



https://www.zenarmor.com/docs/network-security-tutorials/what-is-attack-surface

Подходы к обеспечению безопасности

- **Внешние проверки и аудиты, мониторинг изменений**
- Привлечение сторонних специалистов, высокие финансовые и ресурсные затраты
- Базовый аудит безопасности:
  Сроки - от 2-х недель
  Стоимость - от 200 000 руб.

https://subcoder.ru/photo/audit-sistemiy-upravleniya-bezopasnostyu/

Подходы к обеспечению безопасности

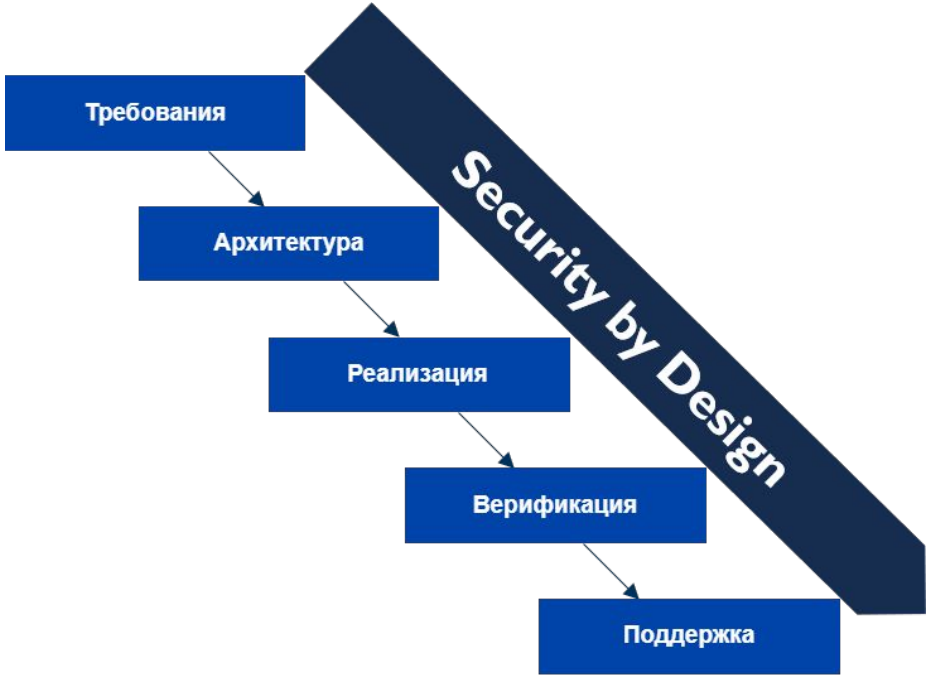
- **Предсказания атак (системный анализ)**
- Требование точного описания угроз,
- Высокие затраты на сбор данных (объемы, открытость)
- Бюджет аналитики данных может превышать бюджет проекта (с т.зр. финансов и трудозатрат)

Подходы к обеспечению безопасности

- **Синтез СЗИ на основе математических моделей**
- Оценивается вероятность возникновения угроз
- Высокие затраты на сбор данных (объемы, открытость)
- Бюджет аналитики данных может превышать бюджет проекта (с т.зр. финансов и трудозатрат)

Основы конструктивной безопасности

- Требования безопасности приравниваются к функциональным требованиям и влияют на выбор архитектуры решения и аппаратной базы
- Меры безопасности интегрированы в архитектуру и программный код
- Применяется в прикладных сферах, традиционно относящихся к критическим
- Подход недостаточно проработан методически, практическое применение сводится к экспертизе конкретных разработчиков



https://cetome.com/training/security-by-design

Kaspersky Blog - Security by Design - https://os.kaspersky.ru/blog/security-by-design/

Принципы кибериммунного подхода. Практический пример

Прикладная система:Умный светофор, получающий команды из центральной системы управления городским наземным транспортом.

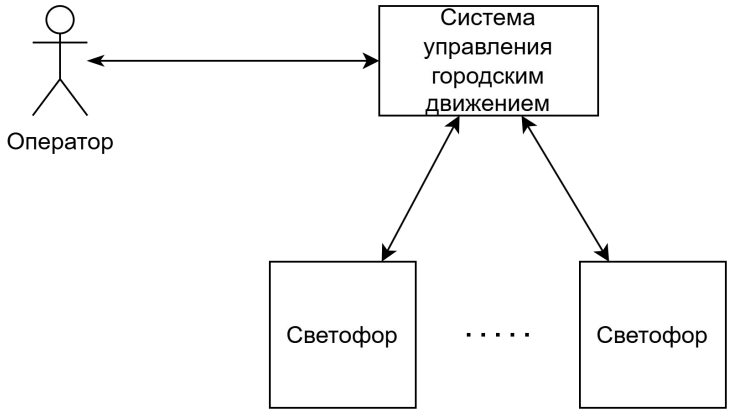
- Светофор управляется центральной системой по беспроводному интерфейсу
- Светофор передаёт состояние в центральную систему

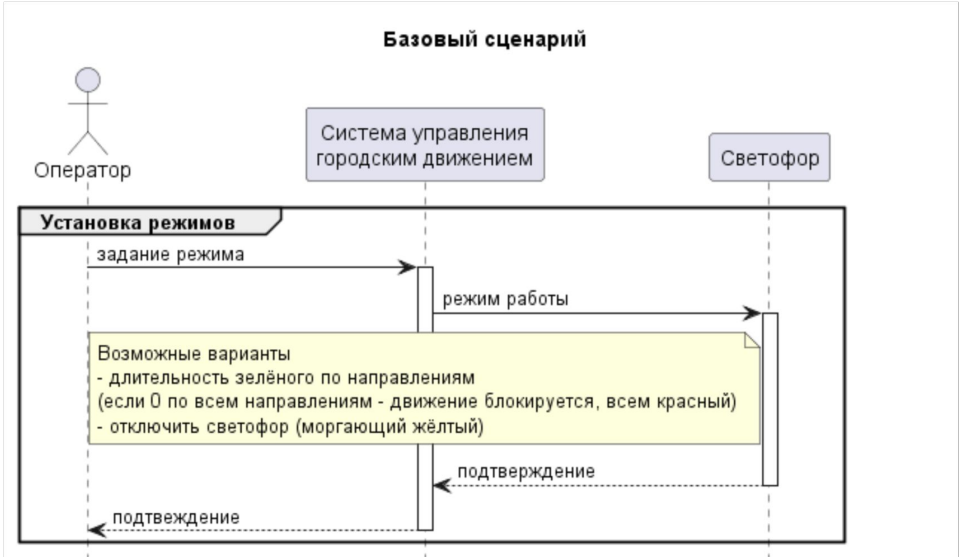
Негативные события в отношении активов системы:

- два разрешающих сигнала в пересекающихся направлениях привели к дтп с пострадавшими
- из-за неправильной работы светофора образовалась пробка

Концепция безопасности продукта "Защищённый светофор" - https://github.com/cyberimmunity-edu/cyberimmune-systems-training-materials-ru/tree/main/traffic-light

Принципы кибериммунного подхода. Практический пример





Принципы кибериммунного подхода. Практический пример

Принципы кибериммунного подхода. Основные свойства

- Применение концепции конструктивной безопасности, обеспечивающие методологии - MILS (multiple independent levels of security) и FLASK (flux advanced separation kernel)
- Защищенные механизмы разделения и контроля взаимодействия между подсистемами
- Актуален в условиях ограниченных ресурсов, основной приоритет - минимизация доверенной кодовой базы (TCB - Trusted Computing Base)
- Активы системы защищены от нежелательных событий при любых условиях, даже под атакой, при условии заданных ограничений.
- Предотвращение эксплуатации уязвимостей при разработке архитектуры

Принципы кибериммунного подхода. Этапы разработки

Ключевые этапы разработки:

1. Определение концепции безопасности системы.

2. Формулирование целей и предположений безопасности (ЦПБ).

3. Разработка политики архитектуры системы.

    3.1 Разработка архитектурной диаграммы, которая иллюстрирует допустимые взаимодействия между различными доменами безопасности.

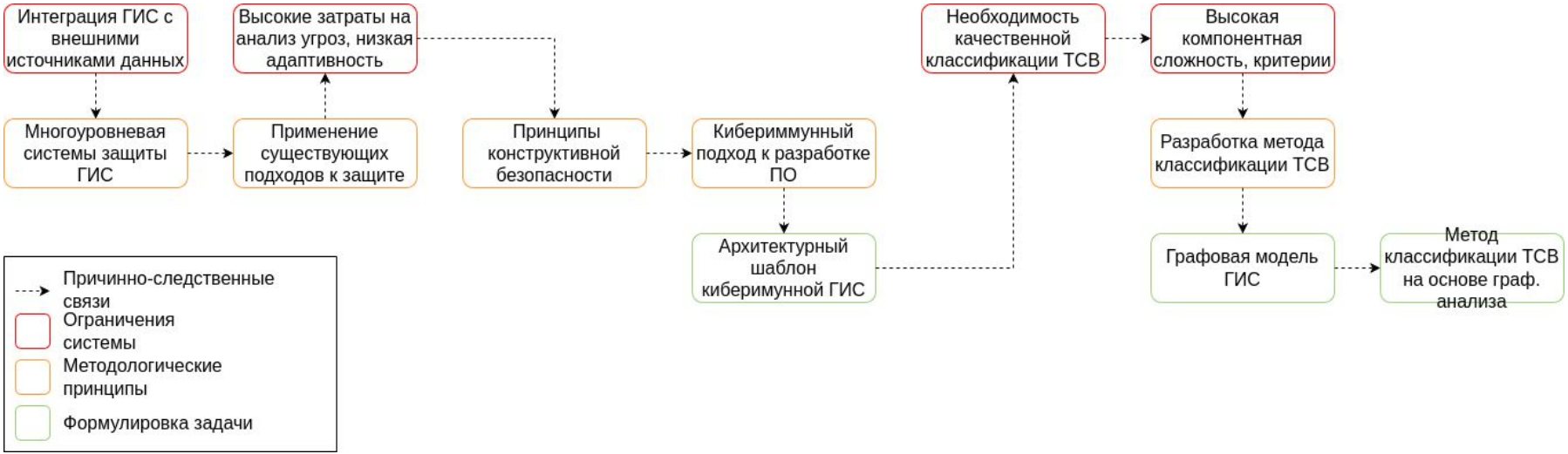
    3.2 Классификация TCB (Trusted Computing Base).

        Предлагается ввести графовую модель системы. Анализ критичности домена на основе графового анализа вершин.

    3.3 Уточнение архитектуры на основе классификации TCB.

Методология исследования

Практическая проблема: Разработка ГИС с функциями мониторинга и интеллектуального анализа. Ограничения существующих подходов не позволяют гарантировать функционирование системы в условиях угроз.
Цель исследования: Разработка архитектурного шаблона кибериммунной ГИС, включающего графовую модель и метод классификации ТСВ.



Предлагаемый архитектурный шаблон кибериммунной ГИС

1. Концепция безопасности ГИС
2. Исходная архитектура ГИС
3. Графовая модель архитектуры ГИС
4. Метод классификации ТСВ на основе алгоритмов графового анализа
5. Правила уточнения архитектуры системы на основе полученных метрик

Концепция безопасности ГИС. Активы системы

Кибериммунный подход предполагает определять концепцию безопасности в терминах **активов и нежелательных событий**.
Определим описание для рассматриваемого класса систем:

Актив	Критичные события
Геопространственные данные, хранимые в БД системы	1. Подмена или нарушение целостности исходных данных. 2. Утечка данных.
Результаты анализа геопространственных данных	1. Компрометация результатов анализа. 2. Утечка результатов анализа. 3. Недоступность данных в момент запроса пользователя

Концепция безопасности ГИС. ЦПБ

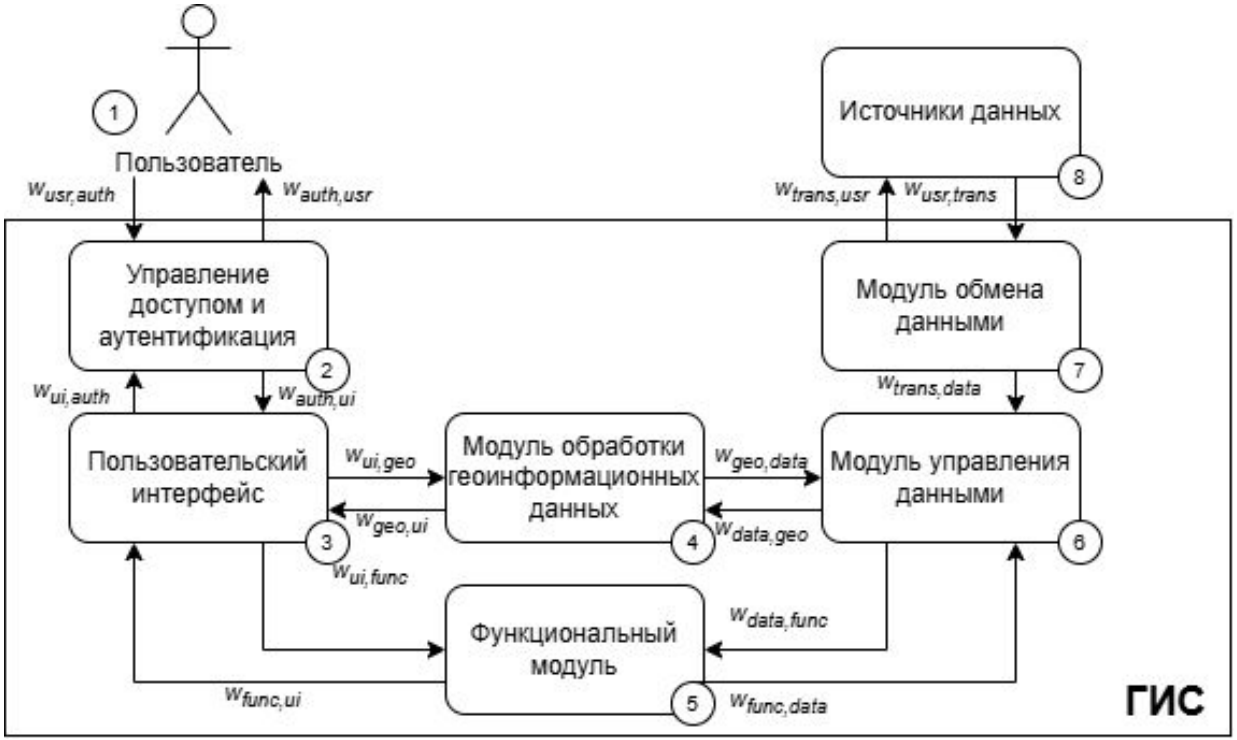
Определим **цели безопасности** (требуемые состояния системы) и **предположения безопасности** (допущения) для рассматриваемого класса систем:

Цели безопасности
- Только аутентифицированный пользователь имеет доступ к критичным данным. - Только авторизованные источники данных могут передавать информацию в систему. - Утеря геопространственных данных или результатов анализа недопустима. - Система предоставляет данные в любое время по запросу пользователей.
Предположения безопасности
- Внешние источники данных нельзя считать доверенными. - Аутентичные пользователи системы являются доверенными лицами и не создают угрозы безопасности в отношении ценностей системы. - Физическая безопасность вычислительного оборудования гарантирована.

Типовая архитектура ГИС (на основе обзора)

Предлагается рассмотреть следующую архитектуру системы:

- Сущности 2-7: основные домены безопасности системы
- W_{xy} , где X и Y – обозначения взаимодействующих доменов - весовые коэффициенты потоков данных [0,1]
- W_{xy} зависит от характеристик:
  – Оперативность
  – Объем данных
  – Уровни обслуживания

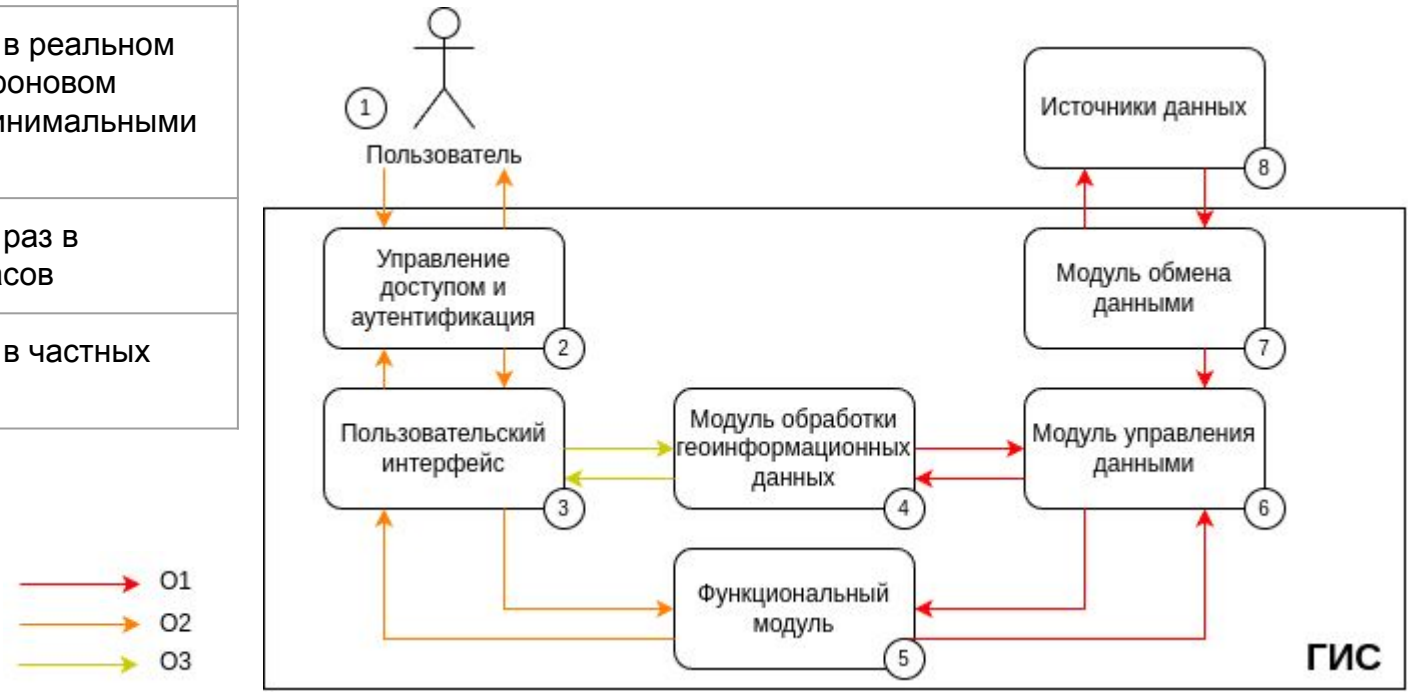


Графовая модель архитектуры ГИС

Компонент архитектуры	Параметры модели	Обозначения
Домены безопасности	Множество вершин *V*	v_{usr} - Пользователь v_{ui} – Пользовательский интерфейс v_{func} – Функциональный модуль v_{geo} – Модуль обработки геоинформационных данных v_{auth} – Управление доступом и аунтентификация v_{trans} - Модуль обмена данными v_{data} - Модуль управления данными v_{src} - Источники данных
Потоки данных	Множество дуг *E*	$E = \left\{ \begin{array}{l} (v_{\text{usr}}, v_{\text{auth}}), (v_{\text{auth}}, v_{\text{usr}}), \\ (v_{\text{auth}}, v_{\text{ui}}), (v_{\text{ui}}, v_{\text{auth}}), \\ (v_{\text{ui}}, v_{\text{geo}}), (v_{\text{geo}}, v_{\text{ui}}), \\ (v_{\text{ui}}, v_{\text{func}}), (v_{\text{func}}, v_{\text{ui}}), \\ (v_{\text{func}}, v_{\text{data}}), (v_{\text{data}}, v_{\text{func}}), \\ (v_{\text{geo}}, v_{\text{data}}), (v_{\text{data}}, v_{\text{geo}}), \\ (v_{\text{trans}}, v_{\text{data}}), (v_{\text{data}}, v_{\text{trans}}), \\ (v_{\text{src}}, v_{\text{trans}}) \end{array} \right\}$
Характеристики потоков данных	Функция весов дуг *W(e)*	$w : E \to [0, 1],$ $w(u, v) = \mu_O(O_{uv}) \cdot \mu_D(D_{uv}) \cdot \mu_L(L_{uv}),$ $\forall (u, v) \in E$

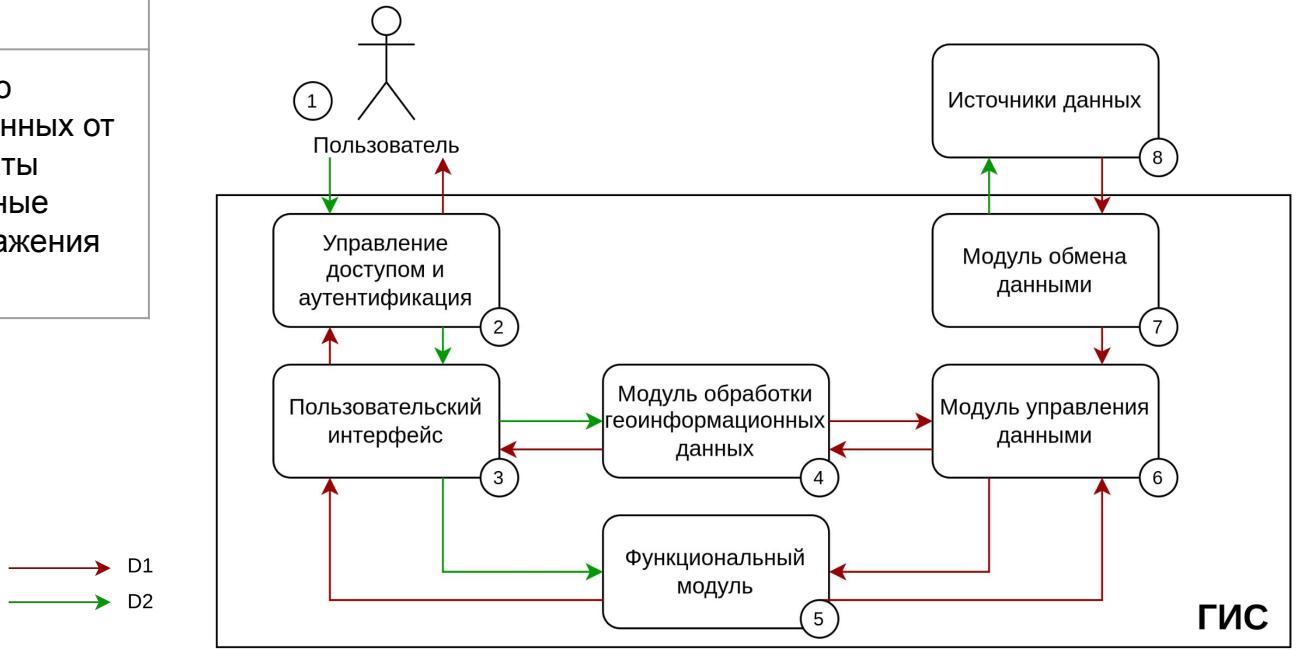
Характеристики потоков данных. Оперативность

Категория	Описание
Очень часто (О1)	Обновление в реальном времени, в фоновом режиме, с минимальными задержками
Часто (О2)	Обновление раз в несколько часов
Редко (О3)	Обновление в частных случаях



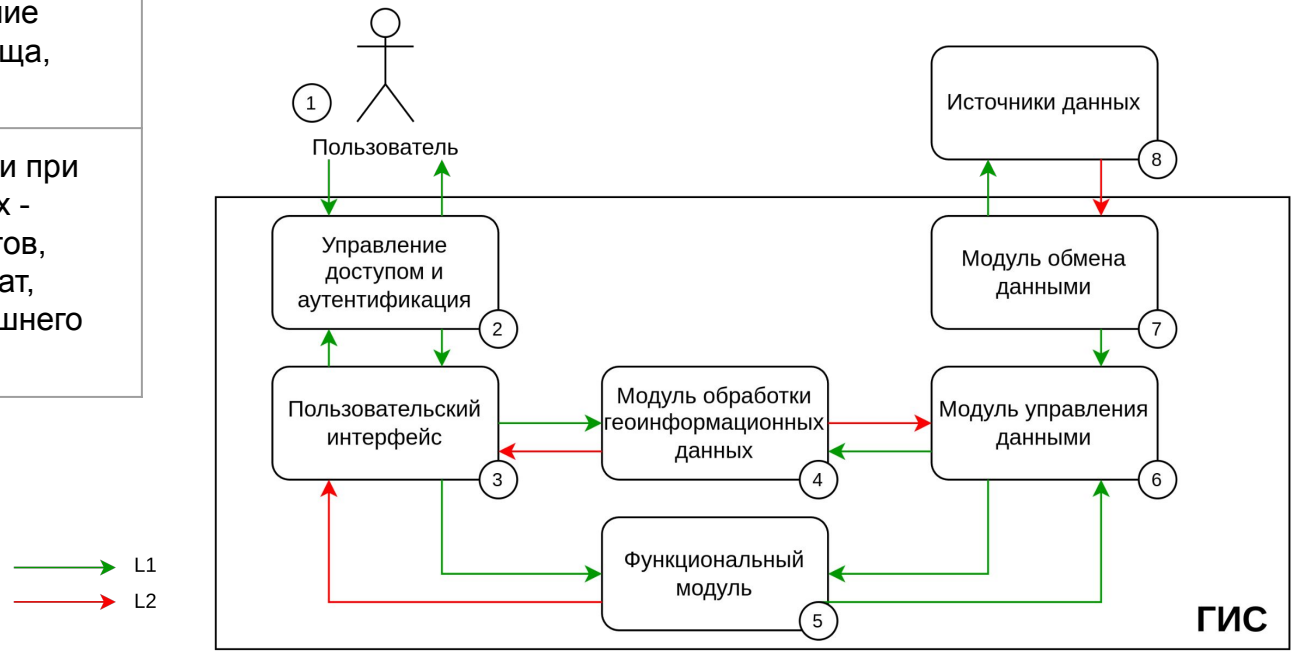
Характеристики потоков данных. Объем данных

Категория	Описание
Запрос (D1)	Запрос данных, например API или modbus
Данные по запросу (D2)	Данные различного объема, пакеты данных от датчиков, результаты обработки, векторные данные для отображения на карте



Характеристики потоков данных. Уровни обслуживания

Категория	Описание
Уровень приоритета L1	Минимальные задержки передачи - получение данных из хранилища, аутентификация
Уровень приоритета L2	Возможны задержки при обновлении данных - выполнение расчетов, обработка координат, недоступность внешнего источника данных



Характеристики потоков данных. Проблематика

- На этапе проектирования отсутствуют точные количественные характеристики потоков данных
- Имеются категориальные характеристики потоков данных:
  - Оперативность → O1,O2,O3 (три градации)
  - Объем данных → D1,D2 (два уровня)
  - Задержка (приоритет) → L1,L2 (два уровня)
- В качестве функции весов не можем использовать линейные и экспоненциальные модели
- Можем задать функцию весов для каждого ребра на основе функций принадлежности категориальных характеристик

Функция весов потоков данных на основе нечеткой логики

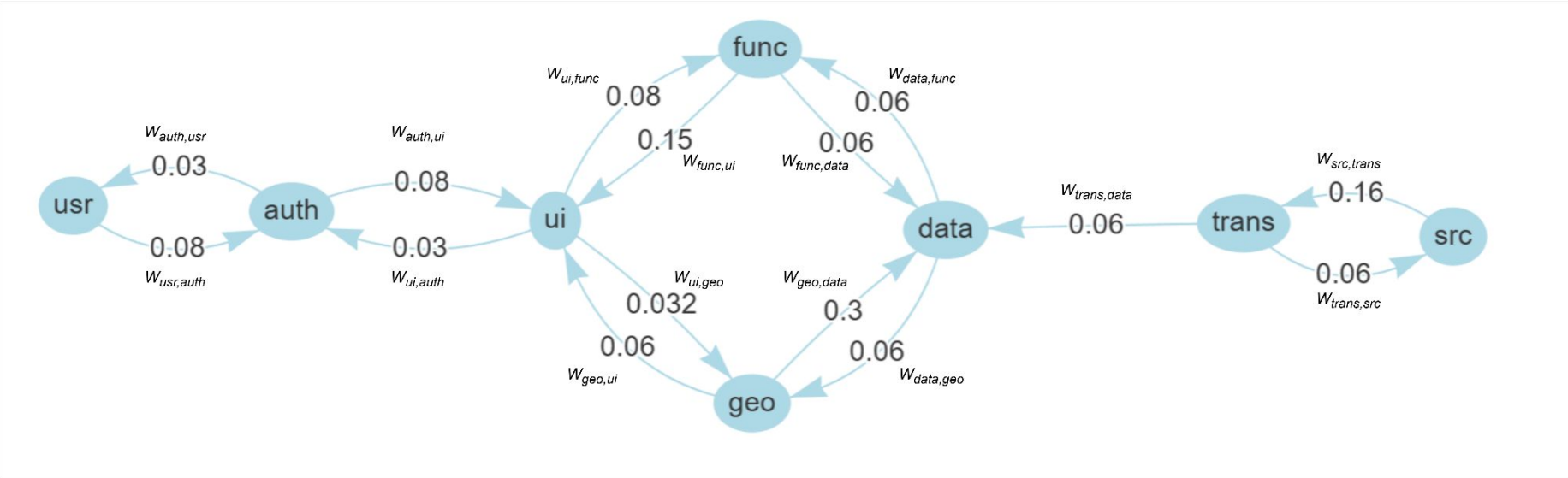
1. Зададим функции принадлежности

$$\mu_O(O) = \begin{cases} 1.0, & O_1 \text{ (Очень часто)} \\ 0.5, & O_2 \text{ (Часто)} \\ 0.2, & O_3 \text{ (Редко)} \end{cases} \quad \mu_L(L) = \begin{cases} 0.2, & L_1 \text{ (Минимальная)} \\ 1.0, & L_2 \text{ (Возможны задержки)} \end{cases} \quad \mu_D(D) = \begin{cases} 0.3, & D_1 \text{ (Запрос)} \\ 0.8, & D_2 \text{ (Ответные данные)} \end{cases}$$

2. Вычисляем итоговый вес связи как произведение степеней принадлежности:

$$w : E \to [0, 1],$$
$$w(u, v) = \mu_O(O_{uv}) \cdot \mu_D(D_{uv}) \cdot \mu_L(L_{uv}),$$
$$\forall (u, v) \in E$$

Построение модели общей архитектуры ГИС



Метод классификации ТСВ. Возможные метрики

Метрика	Формула	Интерпретация
Центральность по степени	$C_D(v) = \deg(v)$	**Активность участия узла в обмене данными.** Учитывает количество подключений узла (число входящих и исходящих дуг).
Центральность по близости	$C_C(v_i) = \frac{1}{\sum_{v_j \neq v_i} d_W(v_i, v_j)}$	**«Близость» к остальным узлам системы.** Узел с высоким показателем метрики быстро получает информацию и может оперативно передавать её другим.
Центральность по посредничеству	$C_B(v_i) = \sum_{s \neq v_i \neq t} \frac{\sigma_{st}^{(W)}(v_i)}{\sigma_{st}^{(W)}}$	**Отражает, насколько часто узел встречается на маршрутах данных между другими узлами.** Узлы с высоким значением центральности часто оказываются критическими точками маршрутизации.
PageRank	$PR(v_i) = \frac{1-d}{N} + \sum_{v_j \in L(v_i)} \frac{W(v_j, v_i) PR(v_j)}{\sum_{v_k \in L(v_j)} W(v_j, v_k)} (4)^d$	**Измеряет, насколько важен узел в сети с учетом важности его соседей.** PageRank — это узел, который получает данные от других важных узлов и является значимой точкой сети.

Метод классификации ТСВ. Итоговая метрика

Введем итоговую метрику сложности домена как взвешенную сумму полученных метрик центральности вершины графа:

$$C(v) = \lambda_1 \sum W(e) + \lambda_2 \cdot C_B(v) + \lambda_3 \cdot C_C(v)$$

- Центральность по степени не включаем в метрику, так как не учитываются веса связей.
- PageRank учитывает только входящие связи, может дублировать центральность по посредничеству.
- Коэффициенты $\lambda_1, \lambda_2, \lambda_3$ подбираются с учетом экспертного ранжирования. Коэффициенты будем выбирать с учетом сценария использования системы.
- **Пример. Если высокая нагрузка в системе связана с передачей и обработкой данных, то локальная нагрузка важнее, чем топология. Зададим $\lambda 1$=0.6,$\lambda 2$=0.3,$\lambda 3$=0.1**

Метод классификации ТСВ. Интерпретация метрики

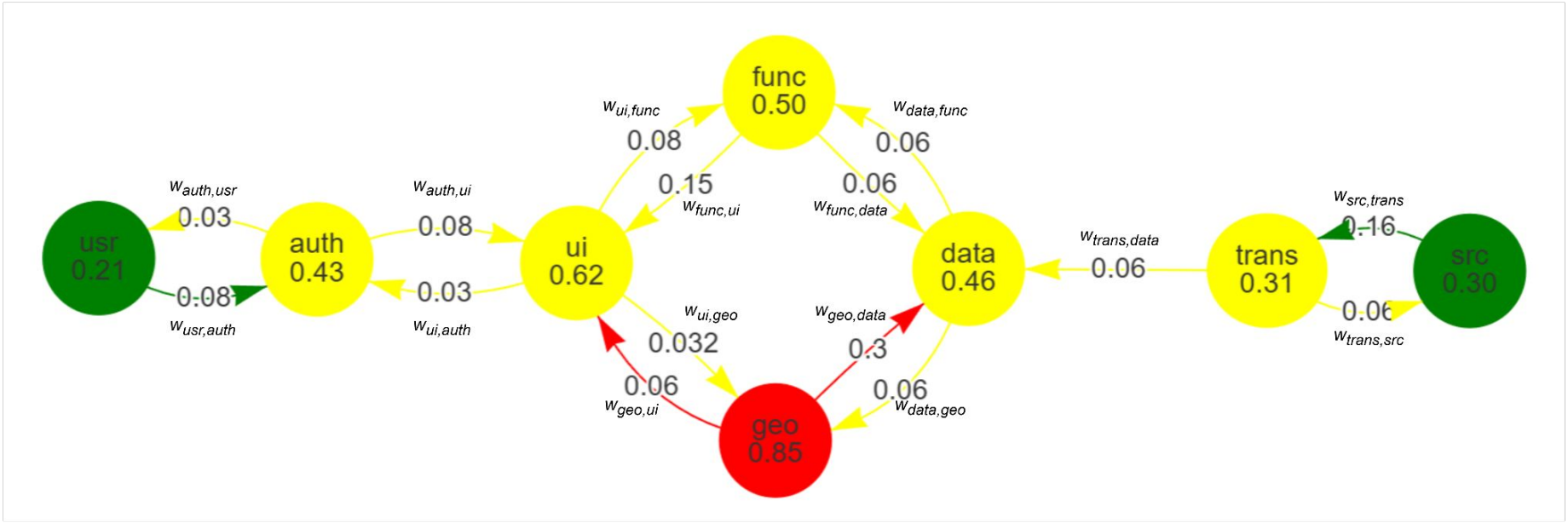
После того как мы вычислили итоговую метрику $C(v)$ для каждой вершины, можем классифицировать домены системы по сложности:

- **Простой (SS):** $C(v) < 0.3$
- **Средний (MV):** $0.3 \leq C(v) < 0.7$
- **Сложный (CL):** $C(v) \geq 0.7$

Классификация доменов системы

Домен	Σw_i	C_B	C_C	*C(v)*	*Оценка домена*
Пользовательский интерфейс	0.14	0.43	0.84	0.62	MM
Управление доступом и аутентификация	0.11	0.24	0.83	0.43	MM
Модуль обработки геоинформационных данных	0.36	0.21	1.0	0.85	CL
Функциональный модуль	0.21	0.1	0.85	0.5	MM
Модуль передачи данных	0.12	0.14	0.12	0.31	MM
Модуль управления данными	0.12	0.26	0.76	0.46	MM

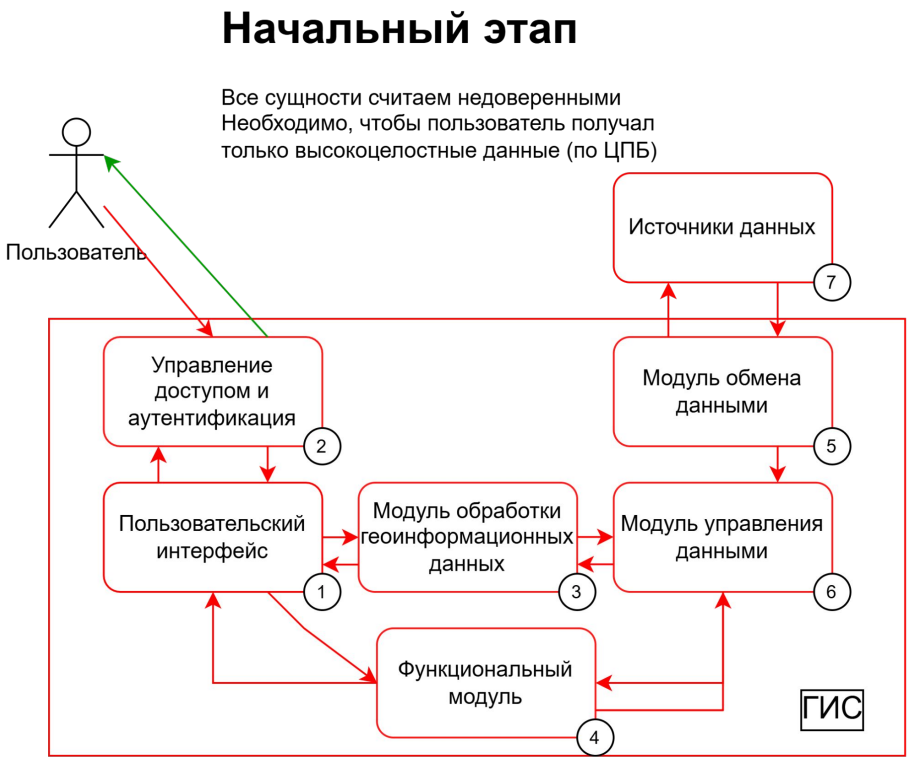
Классификация доменов системы



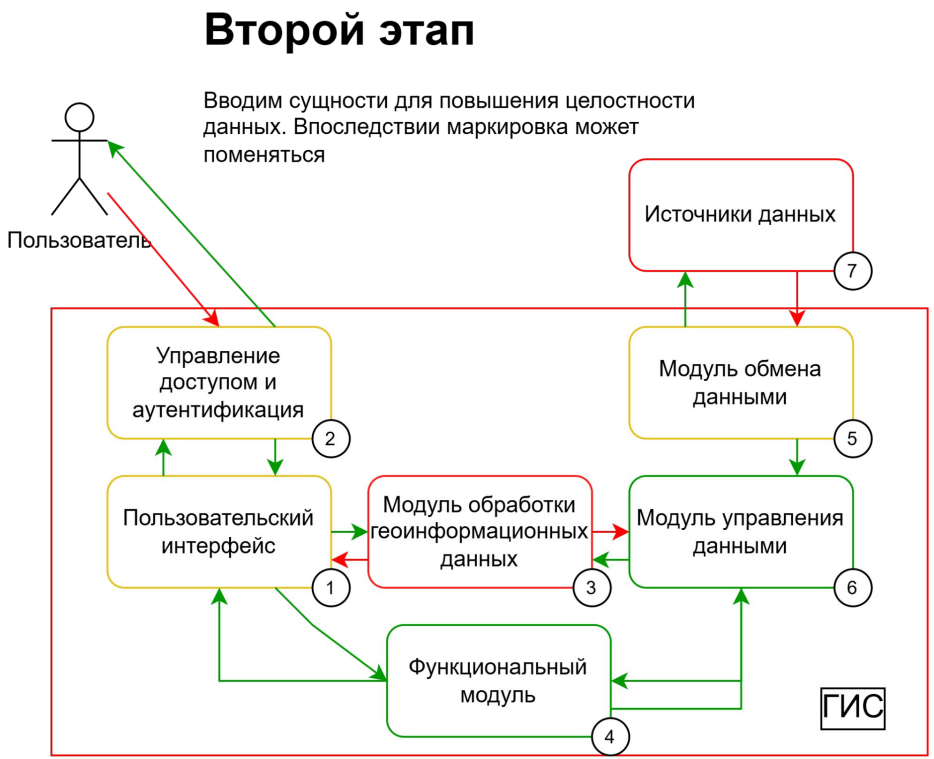
Уточнение архитектуры на основе полученных метрик

- Одна из целей кибериммунного подхода - снижение объемов TCB.
- Для крупных и сложных сущностей допустимы только не критичные данные, поскольку возникают ограничения в их полноценном тестировании и экономическая затратность подобных операций.
- Для обработки потоков данных C,XL доменов в архитектуре вводятся прокси-домены классов S,S. Один из вариантов введения сущностей - декомпозиция на мелкие и простые домены.

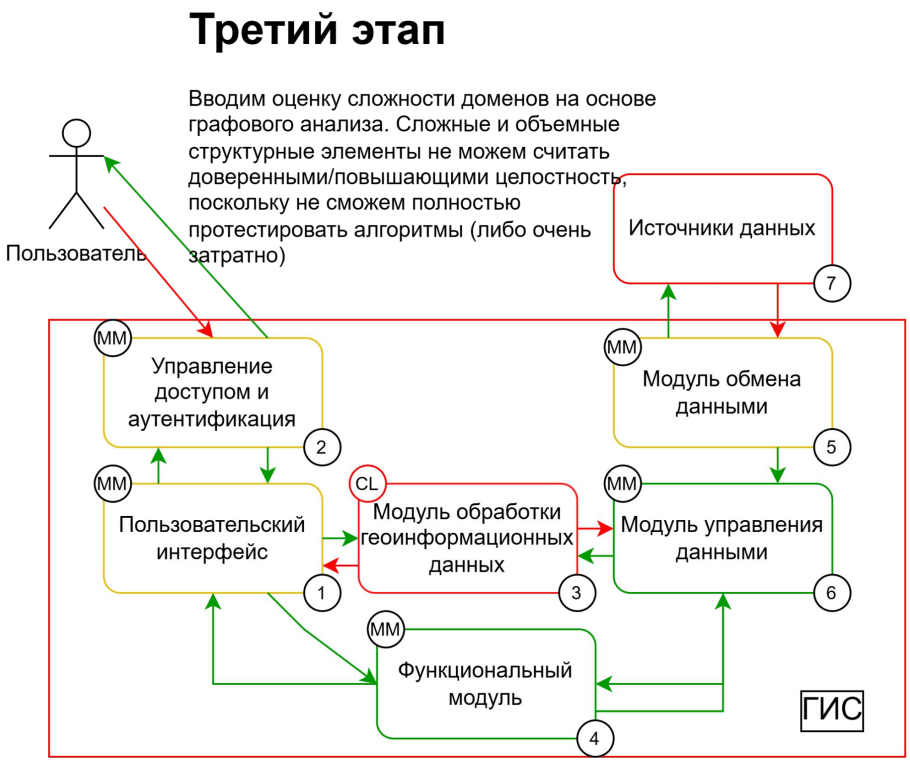
Уточнение архитектуры на основе полученных метрик



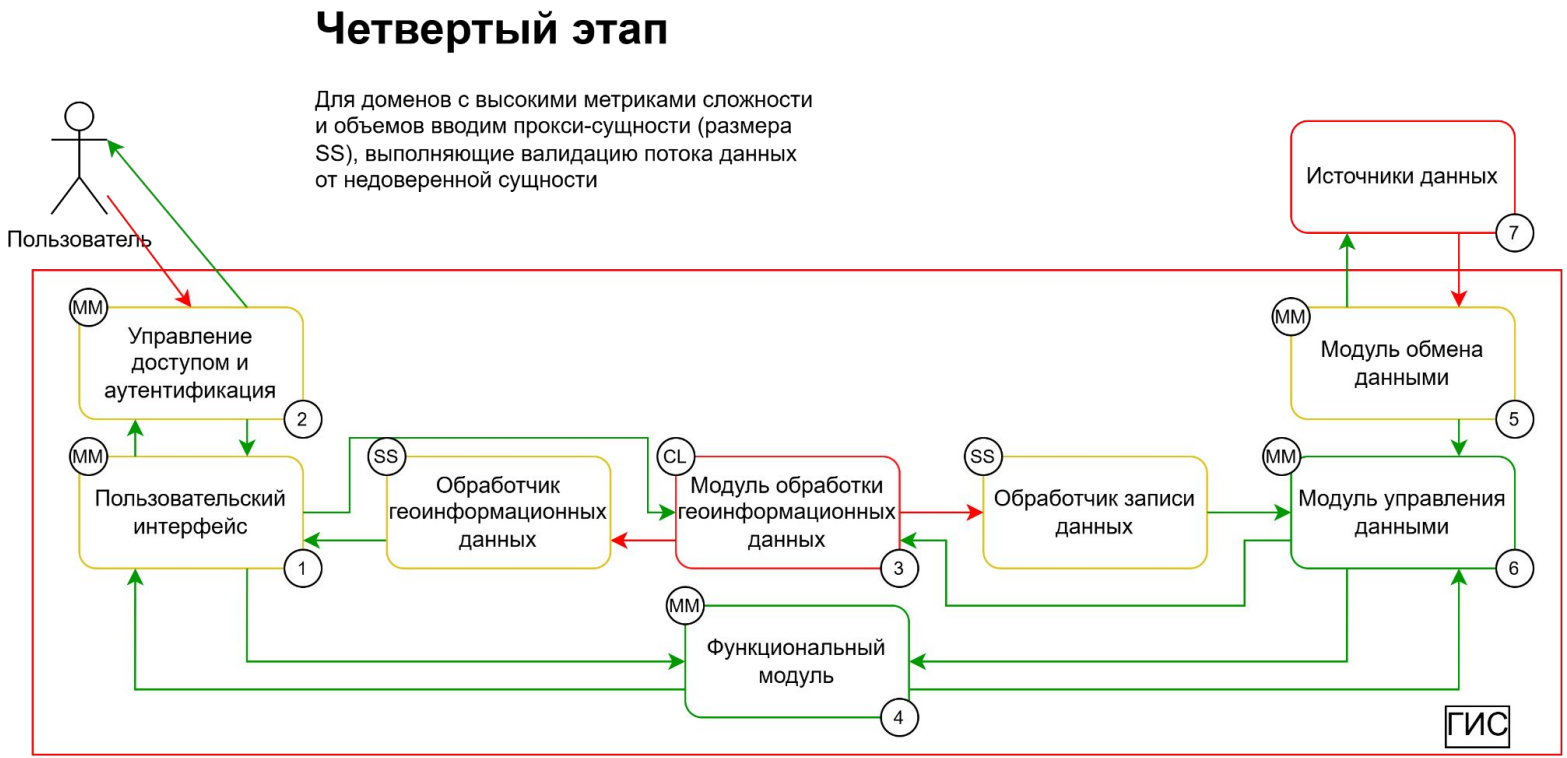
Уточнение архитектуры на основе полученных метрик



Уточнение архитектуры на основе полученных метрик



Уточнение архитектуры на основе полученных метрик



Пример моделирования для реальной системы