

УДК 681.3.06

ИНТЕРНЕТ-СЕРВИС ОЦЕНКИ МЕСТОПОЛОЖЕНИЯ СОБЫТИЙ**В.А. Горбатко¹, С.Л. Беляков²**

¹студент, Южный федеральный университет, г. Таганрог, Россия, e-mail: mediumforse@ya.ru

²доктор технических наук, профессор кафедры информационно-аналитических систем безопасности, Южный федеральный университет, г. Таганрог, Россия, e-mail: beliacov@yandex.ru

Аннотация. В работе рассматривается задача определения местоположения объектов и событий, относящихся к поставленной задаче принятия решения. Сформулирована задача максимизации полезности картографического представления результатов поиска. Приведена архитектура интернет-сервиса для оценки местоположения событий. Указаны достоинства и недостатки методов определения местоположения событий с использованием источников информации, методов геокодирования и классификатора событий.

Ключевые слова: геоинформационные системы, интернет-сервис, геокодирование, поиск местоположения, принятие решений.

THE INTERNET SERVICE FOR THE ASSESSMENT OF THE EVENTS LOCATION**Vladislav Gorbatko¹, Stanislav Belyakov²**

¹Students, Southern Federal University, Taganrog, Russia, e-mail: mediumforse@ya.ru

²Ph.D., Professor of Information and Analytical Systems Security Department, Southern Federal University, Taganrog, Russia, e-mail: beliacov@yandex.ru

Abstract. This paper considers the problem of determining the location of objects and events related to the task of decision-making. The problem of maximizing the usefulness of cartographic presentation of search results. The architecture of the Internet service for the assessment of the events location. Shown advantages and disadvantages of the event location determination methods, using information sources, geocoding methods and classifier of the events.

Keywords: geo-information systems, the internet service, geocoding, search for location, make decisions.

Введение. Принятие решений при информационной поддержке со стороны геоинформационных сервисов представляет несомненный практический интерес [1]. Пространственные, временные и семантические координаты событий, явлений и процессов реального мира дают целостное представление о предметах анализа в задачах, которые приходится решать в условиях неполноты, противоречивости и неоднозначности исходных

данных. Ситуационное управление как метод предусматривает построение ситуационной модели – совокупности существенно важных объектов и отношений между объектами, отражающими суть решаемой проблемы. Полнота и непротиворечивость модели определит достоверность принятых впоследствии решений. Под достоверностью понимается степень соответствия решения реальной ситуации внешнего мира. При недостаточной достоверности возникает ущерб, снизить величину которого можно лишь за счет информационного обогащения картины действительности, которая формируется геоинформационной системой или сервисом для принятия решений.

Картографирование внешнего мира по объективным причинам не выполняется в реальном масштабе времени. Это обусловлено особенностями картографического производства, необходимостью выполнения не автоматизируемых интеллектуальных процедур. По этой причине карты геоинформационных сервисов никогда не являются абсолютно достоверными. Однако, информационное наполнение сети Интернет гораздо динамичнее и может использоваться для актуализации карт при решении конкретных прикладных задач. Для этого необходим механизм оперативного поиска и отображения на картах данных, полученных из некартографических источников – тестов, гипертекстов, новостных лент, сообщений, и т.д.

Современные геоинформационные Интернет-сервисы (Яндекс, Google, 2ГИС, Sputnik и многие другие) уже давно строятся как «учитывающие местоположение». Это выражается в том, что предоставляемая клиенту сервиса информация локализуется рабочей областью карты. Рабочая область ограничена либо контекстом поиска, либо координатами местоположения мобильного устройства, использующего сервис. В любом случае база данных сервиса хранит тематические слои и карты с объектами, которые отображаются по прямому (например, «Показать аптеки») либо косвенному запросу (например, «Показать рядом»). Недостатком подобных сервисов является невозможность поиска событий, явлений и объектов заданной смысловой направленности в указанной рабочей области. Решение подобной задачи позволило бы поднять ситуационный анализ на качественно более высокий уровень.

Цель работы. Цель настоящей работы – анализ формальной модели и архитектуры Интернет-сервиса, предназначенного для поиска и оценки местоположения событий.

Рассмотрим задачу, решаемую Интернет-сервисом в общем виде. Обозначим через Ω множество картографических объектов карты геосервиса, через S_{ext} – множество ссылок на внешние источники

информации в Интернет. Клиент сервиса создает в сеансе рабочую область $R_W \subseteq \Omega$, которая необходима для решения некоторой прикладной задачи. Сервис на основе анализа рабочей области отбирает полезные для решения задачи ссылки $S_c \subseteq S_{ext}$ и отображает их в рабочей области, расширяя ее новыми картографическими объектами посредством процедуры картографирования $H(S_c)$. Картографирование должно быть выполнено так, чтобы максимизировать полезность полученной рабочей области:

$$\begin{cases} I(R_W') \rightarrow \max, \\ R_W' \subseteq \Omega, \\ R_W' = R_W \cup H(S_c), \\ S_c \subseteq S_{ext}. \end{cases} \quad (1)$$

Анализ формальной постановки. Задача (1) является формальной основой работы геосервиса оценки местоположения событий. Анализируя ее, необходимо отметить следующие особенности организации сервиса:

1) результатом работы сервиса является карта, на которой отображены полезные для решения задачи объекты. Функция полезности $I(R_W')$ определяет то, как будут отображены объекты. Очевидно, что решающее значение придается не отдельным элементам рабочей области, а ее общей логичности и завершенности;

2) наличие процедуры картографирования $H(S_c)$ указывает на необходимость решать задачу (1) с привлечением элементов машинного интеллекта [2,3]. Для адекватного картографического отображения должны использоваться знания о корректности отображения отдельных элементов, генерализации рабочей области карты в целом;

3) в процессе решения задачи сервис должен осуществлять поиск и анализ информационного содержания ссылок $S_c \subseteq S_{ext}$, оценивая их полезность. Процесс оценивания является двухшаговым: на первом шаге отбираются наиболее полезные ссылки, исходя из их отношения к решаемой задаче, на втором – с учетом возможностей картографирования информации в рабочей области.

Описание архитектуры сервиса. В предлагаемой модели интернет-геосервиса предполагается использовать два способа нахождения событий: «сверху вниз» и «снизу вверх». Последовательность операций для каждого из них приведена на рис. 1.

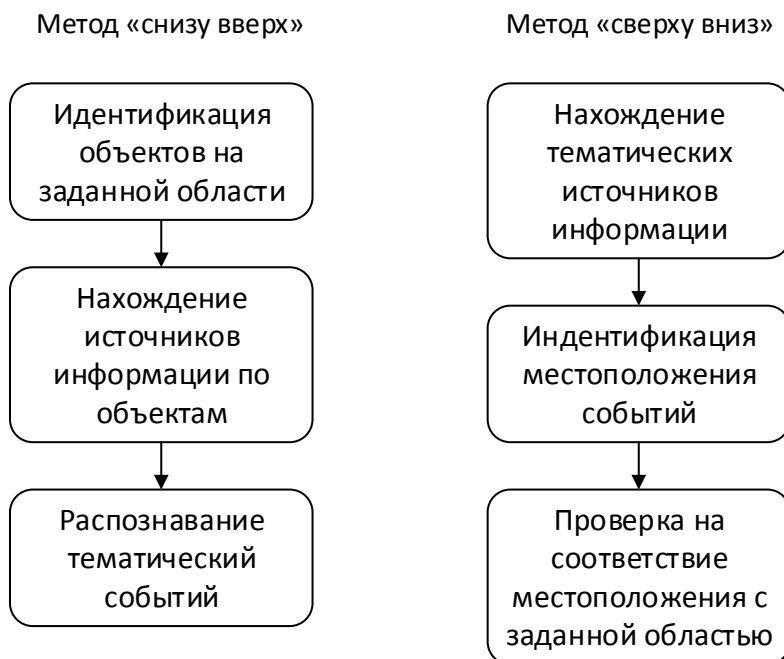


Рисунок 1 – Схема работы методов нахождения событий

Метод «снизу вверх» заключается в нахождении в заданной на карте области, всех объектов, где могут происходить события. После этого для найденных объектов производится поиск источников новостей откуда можно будет извлечь информацию о происходящих на объекте событиях. Среди найденных источников отбираются только те события, которые удовлетворяют заданной тематике. Эти события и можно будет считать конечным результатом работы этого метода.

Метод «сверху вниз» предполагает, что для заданной тематики производится поиск источников событий, из которых может быть извлечена информация о местоположении этих событий. Затем методами прямого геокодирования, соотносятся найденные события с искомой областью и, если их местоположения совпадают – то такие события и будут результатом работы этого метода.

Оба способа имеют свои достоинства и недостатки.

Для метода «снизу вверх» можно отметить следующие достоинства:

- 1) идентификация всех объектов в указанной области. Во время работы метода, перебираются все объекты расположенные на заданной области;
- 2) возможность поиска событий не только по области на карте, но и по отдельным объектам. Это связано с тем что сам поиск событий отталкивается от источников, связанных с отдельными объектами выбранной области на карте;

3) высокая точность определения местоположения событий. Источники событий из которых берутся сами события, изначально уже имеют какое-то местоположение, в процессе их анализа это местоположение уточняется что позволяет увеличить точность.

Среди его недостатков стоит выделить следующие:

1) требуется база данных объектов на карте. Для того чтобы узнать какие объекты присутствуют на заданной области требуется производить обратное геокодирование, перебирая координаты точек. Для снижения времени работы, следует использовать уже готовую карту объектов и дополнять ее в случае необходимости;

2) трудоемкий процесс обхода всех идентифицированных объектов в указанной области на карте. При обработке больших областей на карте время работы метода, резко возрастает. Это связано с поиском взаимосвязей между объектами и источниками информации.

Достоинствами метода «сверху вниз» являются следующие:

1) возможность структурировать особым образом данные, знания и источники информации. Например, данные для поиска могут быть представлены в виртуальных смысловых пространствах со специфическими измерениями и координатными единицами;

2) в отличие от метода «снизу вверх», область поиска не ограничена картографическим представлением объектов в заданной области;

3) ценные сведения могут быть получены из информационных источников, непрерывно накапливающих актуальные знания. Например, таковыми являются профессиональные сообщества социальных сетей.

Недостатками метода можно считать:

1) невысокую точность пространственной привязки результата поиска. Это объективно определяется недостатком информации о пространственных и временных координатах публикуемых в Интернет сведений;

2) игнорирование источников новостей для объектов в анализируемой области. Это связано с тем что поиск событий идет изначально не от объектов, а от источников, не связанных с ними;

3) результат сильно зависит от заданной области на карте. В случае задания области с малым количеством объектов следует использовать метод «снизу вверх».

Поскольку оба метода не являются идеальными, следует использовать их в совокупности, а не давать предпочтение лишь одному из них и реализовывать в системе только его.

В зависимости от входных данных система должна выбрать метод, которому будет отдано предпочтение. Для областей с малым количеством объектов рекомендуется использовать метод «снизу вверх». Для больших областей, следует применять оба метода.

Предлагаемая архитектура геоинформационного интернет-сервиса показана на рис. 2. Используя описанные выше методы нахождения событий, следует включить в архитектуру интернет-сервиса следующие модули:

- 1) математический модуль;
- 2) модуль геокодирования;
- 3) модуль нахождения источников информации;
- 4) модуль распознавания событий.

Математический модуль предназначен для обработки области на карте заданной пользователем. Для этого могут понадобиться алгоритмы, применяемые в компьютерной графике, такие как триангуляция полигона и тесселяция треугольников.

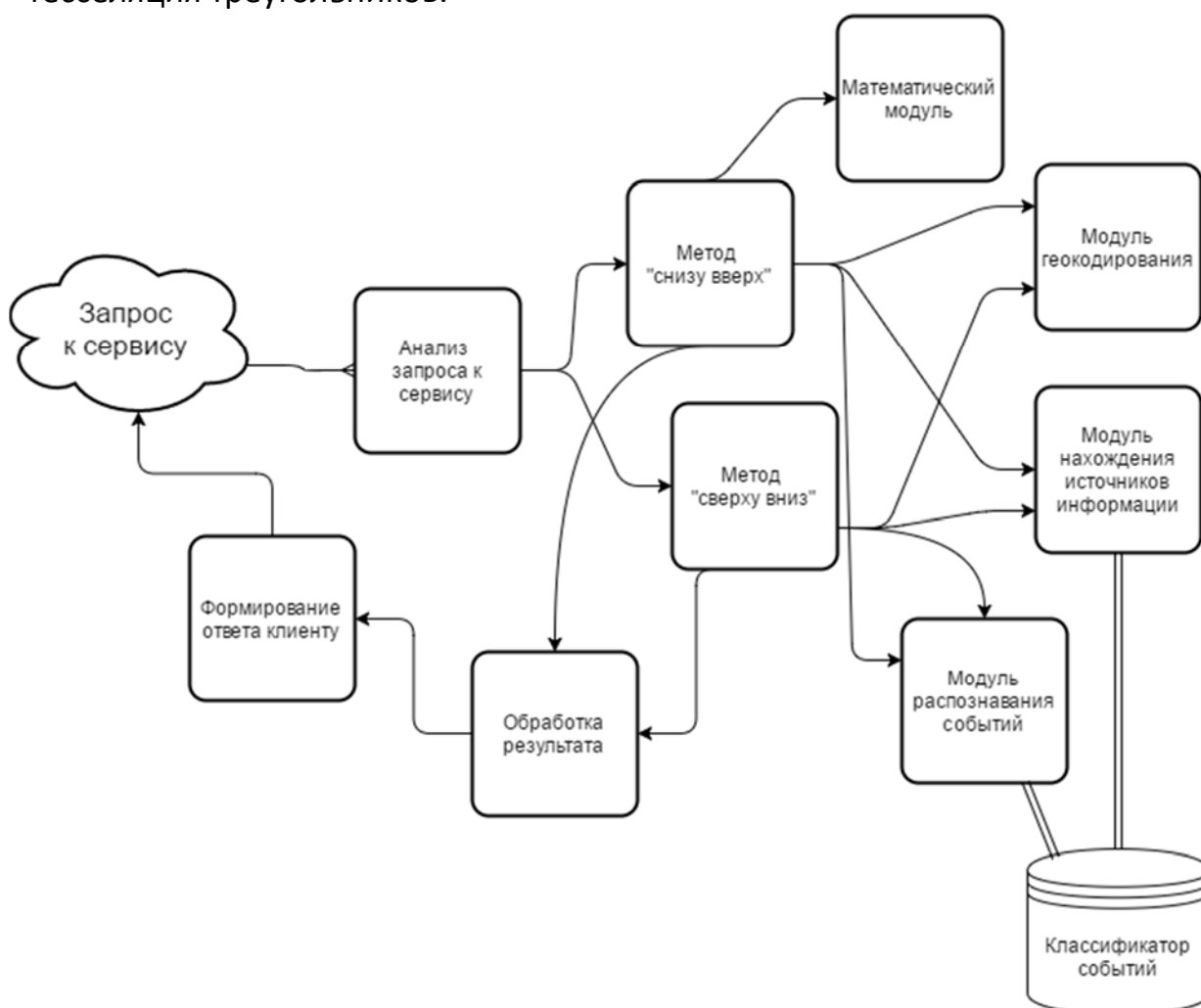


Рисунок 2 – Архитектура интернет-сервиса

Модуль геокодирования предназначен для соотношения географических объектов с их географическими идентификаторами. Это может быть, как обращение к внутренней базе данных в которой уже собрана такая информация, так и обращение к другим геосервисам предоставляющим возможность геокодирования.

Модуль нахождения источников информации позволяет соотносить объекты или тематики событий с источниками информации из сети интернет.

Модуль распознавания событий, позволяет оценивать найденные события из источников информации и фильтровать их по тематикам.

Помимо названных модулей в архитектуру интернет-сервиса должен быть заложен классификатор событий. Классификатор событий — это база данных которая связывает тип события со списком сайтов, где можно найти полезную информацию и списком ключевых слов для поиска информации. Наполнением классификатора событий следует заниматься вручную, поскольку именно от его качества зависит качество и всей системы в целом. В случае автоматизации этого процесса, рекомендуется проводить тщательную премодерацию для исключения ошибок автоматизированного сбора информации.

Вывод. Благодаря описанной в работе архитектуре интернет-сервис имеет следующие достоинства: во-первых, использует разные методы нахождения событий. Система построена так, что сама оценивает и решает какой метод лучше использовать, что дает хорошие результаты на разных задачах принятия решения. Во-вторых, сервис использует картографическое представление результатов, что обеспечит их высокое информационное наполнение благодаря картографической основе.

ЛИТЕРАТУРА

1. Беляков С.Л. Геоинформационный сервис ситуационного центра / Беляков С.Л., Белякова М.Л., Самойлов Д.С. // Информационные технологии. – Москва, 2011. – № 8. – С. 29-32.
2. Беляков С.Л., Розенберг И.Н. Программные интеллектуальные оболочки геоинформационных систем. – М.: Научный мир. – 2010.
3. Belyakov S.L. Model Of Intellectual Visualization Of Geoinformation Service / Bozhenyuk A.V., Belykova M.L., Rozenberg I.N. // Proc. 28th European Conference on Modeling and Simulation ECMS 2014. – Brescia, 2014. – P.326-333.