

СИСТЕМА ГЕОПРОСТРАНСТВЕННОЙ ПРИВЯЗКИ И ВИЗУАЛИЗАЦИИ ДАННЫХ ИЗ ИНФОРМАЦИОННО-НОВОСТНЫХ ИСТОЧНИКОВ

Чубур С. В. (ВКА им. А. Ф. Можайского), **Сабилов Т. Р.** (ВКА им. А. Ф. Можайского)

Научный руководитель – кандидат технических наук Сабилов Т. Р.

(ВКА им. А. Ф. Можайского)

Введение. Объем цифровой информации в современном мире увеличивается в стремительно ускоряющемся темпе. До 95% поток цифровой информации состоит из неструктурированных данных и лишь 5 % составляют различные базы данных, представляющий собой тем или иным образом структурированную информацию [1]. Особенности работы аналитика в настоящих реалиях обязывают его обрабатывать большой объем информации в короткие сроки. Традиционные прикладные аналитические инструменты в настоящее время являются недостаточно эффективными. Разработка программного комплекса, обеспечивающего автоматизированный сбор, хранение, фильтрацию, геопространственную привязку и визуализацию данных из информационно-новостных источников, позволит повысить эффективность работы множества служб и компаний. Создание инструмента для улучшения качества работы информационно-аналитических служб актуально в связи с необходимостью повышения уровня обеспечения безопасности граждан, окружающей среды, улучшения экономических показателей компаний и предприятий, совершенствования механизма изучения обстановки и принятия решений.

Основная часть. С помощью создания программного комплекса решаются следующие типы задач:

- 1) Задача автоматизированного сбора и накопления большого объема информации из различных информационно-новостных источников. Хранение данных за прошедший период позволяет использовать имеющуюся информацию для выявления закономерностей и прогнозирования вариантов дальнейшего развития событий.
- 2) Задача агрегации информационно-новостных сообщений по географическому положению и характерным особенностям. Компоновка по определенным признакам неструктурированных данных большого объема значительно снижает уровень энтропии ситуационной обстановки, тем самым обеспечивает возможность более детального изучения условий изменения настоящего положения [2].
- 3) Задача фильтрации информационно-новостных сообщений с целью уменьшения общего объема данных без потери смысловой нагрузки. Использование определенных алгоритмов для сравнения информационно-новостных сообщений по семантическому признаку позволяет уменьшить количество схожих и повторяющихся данных, тем самым повысить уникальность отображаемой информации [3].
- 4) Задача осуществления геопространственной привязки и визуализации данных, полученных в результате выполнения предшествующих мероприятий. Геопространственная привязка агрегированных по определенному признаку и отфильтрованных данных делает возможным визуализацию сложившейся обстановки. Сопоставление конкретных событий с их положением на местности открывает широкий спектр возможностей в аналитической деятельности.

Осуществление геопространственной привязки информационно-новостных сообщений к точкам на местности осуществляется с применением технологий обработки естественного языка NLP (Natural language processing). Использование морфологического анализатора позволяет осуществить синтаксический анализ данных для сопоставления информационно-новостного сообщения с его географическим положением [4].

Для решения поставленных задач был разработан программный комплекс, имеющий возможность масштабирования. В зависимости от требований пользователя и преследуемых интересов система может быть адаптирована для использования как в разных

информационных сферах, так и на разных по площади участках местности (район, город, регион, страна).

Выводы. Введение в эксплуатацию разработанного программного комплекса обеспечит повышение эффективности деятельности информационно-аналитических служб, существенно повысит уровень ситуационной осведомленности, предоставит возможность выявления закономерностей и сократит время на принятие правильных решений.

Предложенное решение может быть применено в различных областях здравоохранения, обеспечения безопасности граждан, повышения уровня правопорядка, улучшения экономических показателей деятельности компаний и предприятий разной направленности.

Список использованных источников:

1. Информационно технический взрыв и его основные этапы. информационная перегрузка: [Электронный ресурс]. – URL: <https://leally.ru/how-to-open-file/informacionno-tehnicheskii-vzryv-i-ego-osnovnye-etapy/> (дата обращения: 10.02.2023);

2. Методы ситуационного анализа и графической визуализации потоков больших данных: [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/metody-situatsionnogo-analiza-i-graficheskoy-vizualizatsii-potokov-bolshih-dannyh> (дата обращения: 10.02.2023);

3. Подход к устранению избыточности семантически близкой текстовой информации в хранилищах и базах данных: [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/podhod-k-ustraneniyu-izbytochnosti-semanticheskii-blizkoy-tekstovoy-informatsii-v-hranilishchah-i-bazah-dannyh> (дата обращения: 10.02.2023);

4. NLP – обработка естественных языков: [Электронный ресурс]. – URL: <https://cyberleninka.ru/article/n/nlp-obrabotka-estestvennyh-yazykov> (дата обращения: 10.02.2023).

Чубур С. В. (автор)

Сабиров Т. Р. (научный руководитель)