

УДК 355.58; 519.25
doi: 10.34987/vestnik.sibpsa.2021.62.47.010

ПРОБЛЕМА ПРИМЕНЕНИЯ БОЛЬШИХ ДАННЫХ В ИНТЕРЕСАХ ВЫПОЛНЕНИЯ ЗАДАЧ, СТОЯЩИХ ПЕРЕД МЧС РОССИИ

*Рыбаков А.В., д. т. н., профессор; Иванов Е.В., к. т. н.; Дмитриев А.В.; Сидоров В.С.
ФГБВОУ ВО Академия гражданской защиты МЧС России*

Аннотация. В статье рассматриваются перспективные пути применения в практике деятельности МЧС России «больших данных». Подчеркивается актуальность необходимости формирования перечня баз данных и разработки единого инструментария, обеспечивающего корректную обработку статистических данных за счет применения методов Big Data. Приводится краткое описание методов Big Data и перспективные темы научно-исследовательских работ, в рамках выполнения которых, данные методы могут найти применение.

Ключевые слова: Big Data, тематика НИР, направления исследований, базы данных, методы обработки данных.

THE PROBLEM OF USING BIG DATA IN THE INTERESTS OF FULFILLING THE TASKS FACING THE EMERCOM OF RUSSIA

*Rybakov A.V. Holder of an Advanced Doctorate (Doctor of Science) in Engineering Sciences, Full Professor;
Ivanov E.V., Ph.D. of Engineering Sciences; Dmitriev A.V.; Sidorov V.S.
Civil Defense Academy EMERCOM of Russia*

Annotation. The article raises the problem of the need to apply “big data” in the practice of the EMERCOM of Russia. It is emphasized that it is necessary to form both database sources and unified tools that ensure the correct processing of statistical data through the use of Big Data methods. A brief description of the methods of Big Data and promising topics of research works in the framework of which these methods can be used are given.

Key words: Big Data, research topics, research areas, databases, data processing methods.

Введение

Большие данные в настоящее время являются одним из ключевых ресурсов, обеспечивающих устойчивое развитие экономики и страны в целом. Данное обстоятельство обусловило широкое внедрение информационных систем в практику государственного и муниципального управления в нашей стране [1].

Несмотря на то, что практически перед всеми министерствами и ведомствами страны поставлены амбициозные цели по формированию центров хранения, обработки данных, до конца не совсем понятно, какие именно функции и за счет реализации каких механизмов планируется в данных центрах решать [2].

Следует иметь в виду, что ошибки в определении целей применения больших данных, могут привести к напрасной трате ресурсов и что самое главное – времени. В то же время правильно определенные цели и адекватно сформулированные задачи позволят решить ряд наиболее актуальных задач стратегического развития Российской Федерации в условиях цифровой трансформации мировой экономики.

В полной мере данное утверждение касается и задач, решаемых МЧС России. Значительные объемы информации, характеризующей состояние защищенности населения и территорий от чрезвычайных ситуаций, поступающей из различных источников, в настоящее время аккумулируется в Главном управлении «Национальный центр управления в кризисных ситуациях». При этом информация зачастую не структурирована, слабо коррелирует между собой, иными словами может применяться только для текущей оценки складывающейся ситуации. С точки зрения решения задач мониторинга и прогнозирования, большой эффект может быть достигнут посредством глубокого анализа поступившей информации, с получением выводов как о текущем состоянии системы защиты, так и ее прогнозируемом состоянии при принятии или непринятии управляющих мер.

Целью настоящей статьи является описание возможной области применения Big Data в деятельности МЧС России и их возможных ресурсных источников.

Постановка задачи

В общем случае термин Big Data предусматривает по собой не просто некоторые объемы данные, но и то в каком виде эти данные структурированы и в каком формате представлены. Зачастую речь идет даже не о самой информации и ее представлении, а о методах, с помощью которых данная информация обобщается и анализируется. Эти методы можно применить как к большим по объемам массивам данных (таким как например, как база данных по пожарам на всей территории Российской Федерации – ИСДМ Рослесхоз), так и к малым (таким как например, как сведения по пожарам за конкретное лесничество).

Для больших данных выделяют традиционные определяющие характеристики, выработанные Meta Group ещё в 2001 году, которые называются «Три V»:

Volume – величина физического объёма;

Velocity – скорость прироста и необходимости быстрой обработки данных для получения результатов;

Variety – возможность одновременно обрабатывать различные типы данных [3].

В настоящее время количество признаков расширилось до четырех, пяти и даже семи «V» – Veracity, Variability, Visualization, Value [4]. С точки зрения нормативно-правовых документов Большие данные представляют собой большие массивы данных по таким характеристикам как объем, разнообразие, скорость обработки и/или вариативность, которые требуют использования технологии масштабирования для эффективного хранения, обработки, управления, анализа [5].

МЧС России в настоящее время заключило целый ряд договоренностей о сотрудничестве, в части касающейся доступа к базам данных федеральных органов исполнительной власти для решения задач защиты населения и территорий от чрезвычайных ситуаций. В частности, можно отметить доступ к базам данных Росгидромета и Рослесхоза. При этом сам по себе доступ к базам данных не решает проблем с их использованием, поскольку традиционные методы математического анализа не всегда могут дать желаемый результат. Кроме того, необходимо определить четкий круг задач, которые могут быть решены с применением методов Big Data

Решение

С целью определения направлений применения методов Big Data в МЧС России на 2021 год спланировано проведение трех научно-исследовательских работ:

1. Разработка научно обоснованных моделей прогнозирования природных пожаров на основе данных дистанционного зондирования Земли;
2. Разработка научно обоснованных подходов по моделированию уровня подъема паводковых вод, вызванных весенним половодьем (на примере реки Лена) на основе анализа больших данных;
3. Разработка предложений по применению перспективных методов «Аналитики больших данных» в деятельности МЧС России.

Первые две работы носят в основном прикладной характер и направлены на решение конкретных, частных задач по определению порядка работы с источниками больших данных, направленных на мониторинг и прогнозирование конкретных чрезвычайных ситуаций (паводковая и лесопожарная обстановка).

Третья работа представляет из себя по большей части поисковую работу, направленную на определение перспективных тем научно-исследовательских и опытно-конструкторских работ для включения их в план НИОКР МЧС России по проблематике применения методов «Аналитики больших данных» в интересах выполнения задач, возложенных на МЧС России.

При этом в качестве задач, поставленных для достижения, указанной цели рассматриваются следующие:

- проведение анализа ресурсных источников «больших данных» (Big Data), а также комплекса технических средств, предназначенных для их обработки, в том числе обоснование критериев для выбора «открытых данных» в МЧС России;
- проведение анализа методов, применяемых при работе с «большими данными» (Big Data), которые могут быть использованы в интересах МЧС России;
- разработка алгоритма создания информационных технологий, реализующих обработку «больших данных» для их применения в создаваемых и эксплуатируемых информационных системах МЧС России.

Таблица 1. Соответствие методов Big Data и предложения по перспективным темам НИОКР МЧС России

№.	Применяемый метод Big Data	Решаемая задача	Перспективная тема
1	Кластерный анализ	Сбор данных, содержащих информацию о выборке объектов, упорядочивание объектов в однородные группы	1. Определение факторов пожарной опасности, на основе обработки данных о лесопожарной обстановке, с применением методов кластерного анализа; 2. Определение мест рационального размещения сил и средств пожарно-спасательного гарнизона при реагировании на природные пожары в пожароопасном периоде
2	Регрессионный анализ	Оценка степени взаимосвязи между переменными для моделирования некоторой зависимости	1. Разработка модели прогнозирования параметров паводковой обстановки на основе анализа факторов на нее влияющих; 2. Оценка эффективности выполняемых инженерно-технических мероприятий по гражданской обороне
3	Искусственные нейронные сети	Решение задач аппроксимации больших данных	1. Система мониторинга и анализа данных по деятельности МЧС России; 2. Система мониторинга и анализа состояния технически сложных объектов и систем
4	Генетические алгоритмы	Решение задач оптимизации и моделирования путём случайного подбора, комбинирования и вариации искомых параметров	1. Выявление предвестников ЧС техногенного характера на основе обработки данных, поступающих с систем мониторинга; 2. Формирование рационального состава группировки сил РСЧС при ликвидации последствий ЧС.
5	Распознавание образов	Определение принадлежности данного объекта к одному из заранее выделенных классов объектов	1. Применение механизмов распознавания образов при ведении поисково-спасательных работ с применением средств воздушной разведки; 2. Выявление возможных заторных (зажорных) явлений, в период прохождения весеннего паводка путем реализации механизма распознавания образов.
6	Прогнозная аналитика	Прогнозирование будущего поведения объектов и субъектов	1. Перспективные направления развития аварийно-спасательных средств МЧС России; 2. Оценка степени защищенности населения и территории субъектов Российской Федерации от чрезвычайных ситуаций на очередной период программного планирования.
7	Статистический анализ	Выявление характерных зависимостей, тенденций, трендов для решения задач целеполагания	1. Определение трендов научно-технического развития в области техносферной безопасности; 2. Оценка соответствия направлений научно-технического развития МЧС России, направлениям определяемым документами стратегического планирования.
8	Имитационное моделирование	Целенаправленное исследование структуры и функций реальной системы в режиме ее «имитации», оптимизация некоторых параметров системы	1. Создание математической модели оценки устойчивости территорий к воздействию поражающих факторов чрезвычайных ситуаций; 2. Моделирование обстановки при чрезвычайных ситуациях биолого-социального характера, вызванных широко распространяющимся инфекционным заболеванием людей.

В настоящее время идет параллельная работа по определению источников больших данных и методов, которые могут быть реализованы при решении задач, возложенных на МЧС России. В результате проведенного первичного анализа определены следующие основные направления перспективных тем НИОКР (таблица 1).

Данный перечень является одной из первых попыток систематизации задач, стоящих перед МЧС России в эпоху глобальной цифровизации.

В качестве возможных источников больших данных для выполнения научно-исследовательских работ в интересах МЧС России могут выступать уже существующие отечественные, такие как: ЕМИСС (Росстат); ИСДМ-Рослесхоз; АСУНП (Росгидромет); idp-cs (контроль сейсмособытий); firenotification (МЧС России); ЕОС (Росатом); СДМЗСХН (Минсельхоз). Зарубежные: база данных НЦСПЗ (индексы засушливости); база данных ВМО (индекс осадков); база данных Европейской обсерватории (комбинированный показатель засушливости) и т.п.

Кроме того, в настоящее время перед всеми министерствами и федеральными агентствами стоит задача по формированию не менее двух ведомственных наборов данных, т.е. тех самых источников Big Data [6]. Стоит отметить, что поставленные задачи несомненно важны, но в директивном порядке качественно выполнены быть не могут. И связано это прежде всего с отсутствием целеполагания. Иными словами, стоит задача осуществить сбор баз данных, но что делать с этими данными и что самое главное для решения каких задач понятно не до конца.

Вывод

На основе проведенного первичного анализа можно заключить, что перед МЧС России в настоящее время стоит актуальная проблема формирования источников «больших данных» и определения перечня задач, решаемых посредством реализации методов Big Data. Определена необходимость проведения научно-исследовательских работ по определению основных направлений применения методов Big Data в интересах решения задач, стоящих перед МЧС России.

Литература

1. Чаннов С.Е. Большие данные в государственном управлении: возможности и угрозы // Журнал Российского права. 2018. – №10 (262). – С.111-122.
2. Байков С.А. Технология Big Data в российской политике: эволюция или угроза национальной безопасности? // Журнал политических исследований. 2019. – т.3, №1. – С.59-68.
3. Медведев А.А., Сухина Н.Ю., Удовик Е.Э. Big Data и блокчейн - прорыв в области анализа данных // World Science: Problems And Innovations. Сборник статей XIX Международной научно-практической конференции: в 3 ч. 2018. – С.267-269.
4. Назаренко Ю.Л. Обзор технологии «большие данные» (Big Data) и программно-аппаратных средств, применяемых для их анализа и обработки // European Science. 2017. – №9 (31) – С.25-30.
5. ГОСТ Р ИСО/МЭК 20546-2019 Большие данные. Обзор и словарь: [Электронный ресурс]. Доступ из справочно-правовой системы «Кодекс» URL: file:///C:/Users/IVANOV~1.V/AppData/Local/Temp/564847474.pdf (дата обращения 02.03.2021 г.).
6. Дмитрий Чернышенко провёл заключительную стратегическую сессию по искусственному интеллекту для руководителей цифровой трансформации федеральных ведомств: [Электронный ресурс]. Доступ материалов сайта «Госновости.РФ» URL: <https://gov-news.ru/news/1168791> (дата обращения 02.03.2021 г.).