



фото: www.avatars.mds.yandex.net

«ЦИФРА» РЕШАЕТ ВСЕ?

АССОЦИАЦИЯ ТРАНСПОРТНЫХ ИНЖЕНЕРОВ ПРОВЕЛА ВЕБИНАР «ИСПОЛЬЗОВАНИЕ ОТКРЫТЫХ ДАННЫХ ДЛЯ АНАЛИЗА АВАРИЙНОСТИ В РФ»

Вебинар был посвящен обработке информации по совершенным ДТП в России. Выступавшие с докладами на вебинаре Петр Атаев, директор по развитию ООО «Дорнадзор», и Роман Геллер, руководитель отдела геоинформационных технологий ООО «Дорнадзор», рассказали об алгоритме работы с базой данных ГИБДД и привели примеры автоматизации анализа этих данных при транспортном планировании, организации дорожного движения и других исследованиях.

ЖУРНАЛ «АВТОМОБИЛЬНЫЕ ДОРОГИ» ПРЕДСТАВЛЯЕТ КЛЮЧЕВЫЕ ТЕЗИСЫ ИХ ДОКЛАДОВ

– Смертность на дорогах по-прежнему занимает первое место в России среди внешних причин смерти. Поэтому сейчас этой теме уделяется внимание на самом высоком уровне. Государственные программы по данной проблеме, в частности, Стратегия безопасности дорожного движения в Российской Федерации на 2018–2024 годы, имеют целевой показатель снижения социального риска, то есть количества смертей на дорогах на 100 тыс. человек населения, на 2024 год снижение числа погибших до 4 человек. Сейчас оно составляет 13 жертв различных ДТП на 100 тыс. жителей. На 2030 год ставится еще более амбициозная задача – добиться нулевой смертности от ДТП. Конечно, мы все понимаем, что это сегодня выглядит как декларация, но сама поставленная цель – кардинально снизить смер-

ность на дорогах – никем не оспаривается. Сегодня очень большие контрасты наблюдаются в регионах – от 2–3 жертв ДТП в год на 100 тыс. жителей в мегаполисах, например, в Екатеринбурге, Белгороде, Самаре, до 50 и более. В основном это прилегающие к крупным агломерациям районы, например, Серпуховской район, один из «лидеров» Московской области по этому печальному показателю. Еще один тревожный факт – отрицательные показатели по ряду параметров, в частности, рост аварийности на десятки процентов у такси. Так, в 2019 году этот показатель по стране вырос сразу на 60%. И если в 2017 году показатель риска при поездке на такси из расчета на пассажиро-километр был в 2,5 раза ниже, чем у тех, кто пользовался личным автотранспортом, то сейчас уже можно говорить, что ездить на такси становится опасней, чем на личном авто. Устойчиво растет количество ДТП с участием иностранных водителей – до 19% в год, в прошлом году в

таких авариях погибли порядка тысячи человек.

Есть и достаточно серьезные региональные различия. Например, в Санкт-Петербурге в последние годы активизируется велосипедное движение и, как следствие, растет аварийность с участием велосипедистов, в прошлом году в таких ДТП погибли 8 человек. В 2018 году был только один погибший из данной категории участников дорожного движения. В этой связи важно изучать именно пространственные особенности распространения аварийности. Так, по нашим данным, сегодня наибольшее количество ДТП с пострадавшими приходится на дороги местного значения, около 30% на региональные и 35% – на федеральные. Учитывая протяженность федеральных трасс, можно понять, что аварийность на них примерно в 8 раз выше, чем на региональных и местных дорогах, а смертность в ДТП – в 10 раз. Это можно объяснить высокой интенсивностью движения по федеральным дорогам, которая нивелирует разницу в качестве их содержания по сравнению с местными и региональными. Плотный трафик и его высокая скорость оказывают свое фатальное воздействие на статистику происшествий.

Для квалифицированного анализа ситуации с аварийностью на дорогах полезно воспользоваться возможностями, предоставляемыми базой данных ГИБДД. В то же время она имеет ряд недостатков, которые, собственно,



Петр Атаев, директор по развитию ООО «Дорнадзор»

и сподвигли нас создать свою систему учета и изучения аварийности на наших дорогах – в частности, это то, что данные по месту ДТП не всегда точны. Порой инспектор, который их заносит в документы, служащие источником информации и для электронных баз данных ГИБДД, не всегда подходит к этой работе ответственно, видимо, считая эту информацию не представляющей большой ценности. Нами была разработана геоинформационная система, которая позволяет упорядочить и уточнить эти и другие данные, касающиеся автоаварий. Алгоритм этой системы имеет одной из своих ключевых задач предоставление возможности корректного использования данных

о ДТП. Но, прежде чем рассказать об этом поподробней, следует обратить внимание на существующие базы данных по дорожным происшествиям.

Сайт stat.gibdd.ru – это единственный официальный источник данных о ДТП, который консолидирует в себе всю информацию о них со всей территории страны. Он имеет достаточно богатые встроенные инструменты по формированию статистической информации, обладает инструментарием фильтрации, и недавно в тестовом режиме разработчики ресурса запустили технологию выявления очагов концентрации ДТП. Информация выгружается по запросам, также у ресурса красивая и удобная визуализация. Но, к сожалению, пока не визуализируются данные по месторасположению ДТП и отсутствует геопространственный анализ.

Еще один ресурс данных, сайт безопасныедороги.рф, также покрывает всю территорию страны, есть визуализация статистики, визуализация ДТП на карте, имеются инструменты фильтрации. И у него тоже есть свои три но: неполная валидация данных, отсутствие инструментов выявления очагов аварийности и нет работы с геопространственным анализом. Соответственно, все эти недостатки вышеназванных ресурсов вошли в основу функциональных требований к нашей работе по разработке геоинформационной системы и модуля анализа ДТП. Исходные данные те же, что используют наши коллеги. Это открытая



бесплатная база данных stat.gibdd.ru. В качестве геокодера мы используем «Яндекс.геокодер», так как он имеет наиболее полное покрытие на всей территории страны. И еще мы используем геометрические параметры перекрестков.

Первая проблема, которую необходимо решить с данными stat.gibdd.ru, – это валидация данных, то есть проверка того, правильно ли результирующий продукт действует на практике, в нашем случае – фиксирует месторасположение ДТП. В исходной информации эти координаты подчас могут быть перепутаны между собой местами, могут быть неполными или вообще отсутствовать. В таких ситуациях становится крайне сложно или вообще невозможно определить точное место ДТП. Поэтому нами был создан алгоритм по проверке валидации данных. Первым делом мы разбираем весь полученный массив ДТП – тех, что были совершены на территории населенных пунктов, и тех, что произошли вне границ городов. Работая с первым массивом ДТП, мы находим в каждой карточке происшествия адрес его совершения, который состоит из города, улицы и дома. После чего мы передаем этот информационный массив в геокодер «Яндекса», который возвращает нам информацию о центре тяжести, географической координате дома. И если в радиусе 300 метров от центра тяжести дома находятся региональные координаты ДТП, то мы их оставляем, если нет, то переносим ДТП по адресу дома. Что касается ДТП вне населенных пунктов, то мы сравниваем геокодированные адреса, то есть переводим их в широту и долготу и далее поступаем аналогично первому варианту. Бывает, что сквозь весь этот алгоритм все равно проходят ДТП с неверными адресами. Поэтому поверх этого алгоритма мы наносим еще один этап – смотрим, в какой район попало ДТП и соответствует ли этот район указанному в карточке ДТП, попадает ли авария во вводный массив. Действуя таким образом и подытоживая всю статистику базы данных на текущий момент, округлив значения до ближайшего целого, мы можем сказать, что в целом по стране более 20% всех ДТП не имеют истинного определения места совершения – либо отсутствуют и координаты, и адрес дома, а также кило-



Роман Геллер,
руководитель отдела
геоинформационных технологий
ООО «Дорнадзор»

метраж, или адрес дома указан неверно. Остается задаться вопросом о том, как вносится информация в базу данных. И, наверное, еще и таким вопросом: как мы можем поспособствовать тому, чтобы в эту базу вносилась корректная и достоверная информация.

И самое интересное. У нас есть алгоритм валидации данных, и теперь возникает вопрос: а что делать дальше? А дальше необходимо прорабатывать гипотезы и анализ данных, искать необходимую корреляцию. Перед тем как разрабатывать систему, мы используем такой продукт, как FME. С его использованием можно, имея минимальный опыт разработчика, начать работу по анализу данных. То есть FME – это большая платформа для интеграции пространственных данных. Это универсальный инструмент конвертации и манипулирования графическими – векторными и растровыми – и семантическими данными. FME позволяет выполнять конвертацию, преобразование, трансляцию координат и просмотр различных форматов ГИС, САПР и СУБД. Например, с помощью FME можно создавать такие системы, каждое звено которых будет определенным обработчиком данных. На старте работы такой системы мы через трансформер, он же звено, подключаемся к данным stat.gibdd.ru, чтобы скачать информацию. За это у нас отвечает инструмент под названием HTTP-Caller, который формирует запрос на сайт ГИБДД stat.gibdd.ru и получает от него ответ. На-

пример, информацию по конкретным городу и району в нем. Нередко эта информация оказывается неполной. Действуя дальше, мы подключаем геокодер «Яндекса», преобразуя наши данные так, чтобы он понимал, что мы от него хотим, – то есть вытащим подробный точный адрес и в ответ получим координаты домов. А дальше соединим оригинальные координаты и информацию, которую нам вернул геокодер. Обычно итоговая по этой работе информация показывает, что примерно каждые 25 из 100 ДТП оказываются не на тех координатах, которые дает сайт ГИБДД.

Далее можно провести анализ всех запрошенных ДТП, например, по такому параметру, как тип освещения на дороге. Анализируя эти данные, мы можем видеть, где на исследуемом территориальном участке есть освещение, где его нет и как это влияет на аварийность. При этом, используя инструменты FME, мы станем сами разработчиками таких анализирующих систем, которые в любой момент могут начать исследовать статистику или те или иные причинно-следственные взаимосвязанности аварийности.

Следует отметить, что ни в каком официальном сайте пока еще нет инструмента геоинформационного анализа. У нас же как разработчиков аналитических систем есть информация о геометрических параметрах перекрестков и дорог и об интенсивности и характере движения на них, например, о доле в трафике общественного транспорта, и вообще есть возможность проводить корреляции между любыми данными, которые есть в базе ГИБДД. Поэтому мы можем допустить и то, что с использованием такого подхода вполне возможно прогнозирование потенциального развития ситуации на исследуемых дорогах по аварийности и риску ДТП. Этот перспективный участок работы ждет объединения усилий всех занятых в сфере транспортного анализа специалистов. И, конечно же, такие аналитические инструменты позволяют существенно экономить время на подготовку документов и не тратить его на дополнительные запросы информации и ожидание ее получения.

Подготовил Леонид Григорьев
Фото Ассоциации транспортных инженеров