

Исследование дорожно-транспортных происшествий с применением технологий больших данных

Пекова Е.А., магистрант,

ФГБОУ ВО «Московский политехнический университет», Москва, Россия

Суворов В.С., к.э.н., профессор,

ФГБОУ ВО «Московский политехнический университет», Москва, Россия

Царькова Н.И., к.п.н., доцент,

ФГБОУ ВО «Московский политехнический университет», Москва, Россия

Аннотация. В статье рассматривается применение технологий анализа больших данных на основе дорожно-транспортных происшествий, произошедших в США. Целью исследования является повышения безопасности дорожного движения путем анализа причин возникновения дорожно-транспортных происшествий, выявления зависимостей и закономерностей между факторами, влияющими на появление аварийных ситуаций на дорогах. В результате исследования было выявлено влияние на возникновении дорожно-транспортах происшествий погодных условия и внешних факторов, таких как: близость светофоров, пешеходных переходов, перекрестков. Помимо этого, была найдена взаимосвязь между полосой движения автомобиля и фактом возникновения аварийной ситуации.

Ключевые слова: Дорожное движение, большие данные, ДТП, анализ больших данных, OLAP, визуализация, корреляция, имитационное моделирование

A study of road accidents with the use of big data technologies

Peкова E.A., undergraduate student,

Moscow Polytechnic University, Moscow, Russia

Suvorov S.V., Candidate of Economic Sciences, Professor,

Moscow Polytechnic University, Moscow, Russia

Tsarkova N.I., Ph.D., Associate Professor,

Moscow Polytechnic University, Moscow, Russia

Annotation. The article discusses the application of big data analysis technologies based on road accidents that occurred in the United States. The purpose of the study is to improve road safety by analyzing the causes of road accidents, identifying dependencies and patterns between factors that affect the occurrence of accidents on the roads. The study revealed the influence of weather conditions and external factors on the occurrence of road accidents, such as the proximity of traffic lights, pedestrian crossings, and intersections. In addition, the relationship between the car lane and the fact of an emergency was found.

Key words: Traffic, big data, road accidents, big data analysis, OLAP, visualization, correlation, simulation modeling

Введение

Проблемы в образовании, пренебрежительное отношение к правилам и недочеты в организации транспортного регулирования – всё это неизменно оказывает негативное влияние на безопасность дорожного движения. Дорожно-транспортные происшествия являются одной из основных причин травм и смертей по всему миру, а также наносят весомый ущерб экономике. По данным Всемирной организации здравоохранения (ВОЗ):

- Ежегодно в результате дорожно-транспортных происшествий погибает более 1,25 миллиона человек.
- Дорожные травмы – основная причина смерти людей в возрасте от 15 до 29 лет.
- Столкновения на дорогах обходятся большинству стран в среднем в 3 процента их валового внутреннего продукта.

- Если не будут приняты долгосрочные меры, то уже к 2030 году дорожно-транспортные происшествия станут седьмой по значимости причиной смерти во всем мире. [1]

В настоящий момент основным подходом к повышению безопасности дорожного движения являлся метод системы безопасности, который учитывает уязвимость людей и человеческие ошибки. Хотя подход системы безопасности оказался довольно эффективным, он не сделал достаточно для снижения смертности и получения травм.

Очевидно, что смерть является наиболее серьезным последствием дорожно-транспортных происшествий. Однако травмы ежегодно получают миллионы людей. По данным юридической фирмы Marks & Harrison, которая ежегодно рассматривает дела сотен жертв автомобильных аварий, наиболее распространенные травмы включают переломы костей, ожоги, травмы головы и мозга, травмы спины и позвоночника, а также хлыстовые травмы. [2]

Большие данные – это данные, объем, разнообразие и сложность которых требуют новой архитектуры, методов, алгоритмов и аналитики для управления и извлечения ценности и скрытых знаний [3].

В сфере обеспечения безопасности дорожного движения технологии обработки больших данных могут дать всесторонний анализ причин возникновения дорожно-транспортных происшествий, выявить закономерности и зависимости между факторами, оказывающими отдельно, так и в совокупности, влияние на возникновение аварийной ситуации, а также спрогнозировать, на каком участке дороги и при стечении каких обстоятельств, может произойти дорожно-транспортное происшествие.

Всё это позволит повысить безопасность дорожного движения, снизить смертность, количество травм и негативные экономические последствия. Можно выделить три основных варианта применения технологий анализа больших данных к сфере обеспечения безопасности дорожного движения:

1. Сбор данных

Каждый раз, когда происходит авиакатастрофа, все говорят о восстановлении «черного ящика». Черные ящики входят в стандартную комплектацию самолетов и представляют собой небольшие компактные устройства, которые записывают данные о катастрофах.

Поскольку эта технология относительно недорогая, ее можно было бы широко развернуть и в автомобилестроении. Преимущества технологии в том, что появится возможность отслеживать точное время, скорость, положение и другие факторы, связанные с автомобильными столкновениями и авариями. По мере изучения этих данных можно будет лучше понять тенденции и причины автомобильных аварий и использовать эти данные для сокращения и предотвращения инцидентов в будущем.

2. Предиктивные карты сбоев

В течение большей части последних пяти лет многие штаты и муниципалитеты в США использовали передовое программное обеспечение для анализа исторических данных и составления прогнозов в реальном времени о дорогах, автомагистралях и перекрестках с высоким уровнем риска. По мере совершенствования технологии местные правоохранительные органы и сотрудники службы безопасности смогут использовать прогнозные данные о ДТП для более тщательного контроля участков с повышенным риском и принятия дополнительных мер безопасности по мере необходимости, таких как, усиленное присутствие полиции и снижение ограничений скорости.

3. Автоматизированные транспортные средства

По мере того, как общество приближается к массовому производству беспилотных транспортных средств, будет интересно увидеть, как большие данные могут быть использованы для повышения безопасности этих легковых и грузовых автомобилей, а также самим фактом наличия аварий, который в значительной степени является результатом человеческой ошибки. Считается, что прогнозирующие карты аварий и другие данные могут быть подключены к транспортным средствам для повышения безопасности по всем направлениям.

Объектом настоящего исследования является повышение безопасности дорожного движения при помощи технологий больших данных для снижения негативных социальных и экономических последствий.

Цель работы заключается в проведении анализа причин возникновения дорожно-транспортных происшествий, выявления зависимостей и закономерностей между факторами, влияющими на появление аварийных ситуаций на дорогах. Для достижения поставленной цели необходимо выполнить следующие задачи: подготовить и очистить данные; провести анализ; проверить полученные результаты с помощью имитационного моделирования.

Результат исследования

В качестве инструмента для анализа данных в работе используется кубическая база данных, размерностью более пяти миллиона строк, содержащая в себе информацию о ДТП за год наблюдений, охватывающие 49 штатов США. Данные включают в себя такие параметры, как:

- Точное время;
- Штат;
- Полоса движения;
- Время ограничения движения по полосе;
- Погодные условия;
- Окружающее пространство (знаки дорожного движения, перекрестки, остановки общественного транспорта и т.д.).

Благодаря тому, что данные собраны в OLAP-куб, есть возможность проанализировать их в разных разрезах, чтобы лучше понять степень и тесноту взаимосвязи. Первыми факторами для анализа были выбраны данные о погоде (рис. 1).

Больше всего ДТП происходит во время осадков (дождь, снег) и в облачную погоду. Вопреки сложившемуся в обществе мнению, в туманную погоду происходит всего 18,4% аварий от общего количество. Это может быть

обусловлено тем, что в туманную погоду люди более сосредоточены на дороге, а некоторые вообще отказываются от поездок.

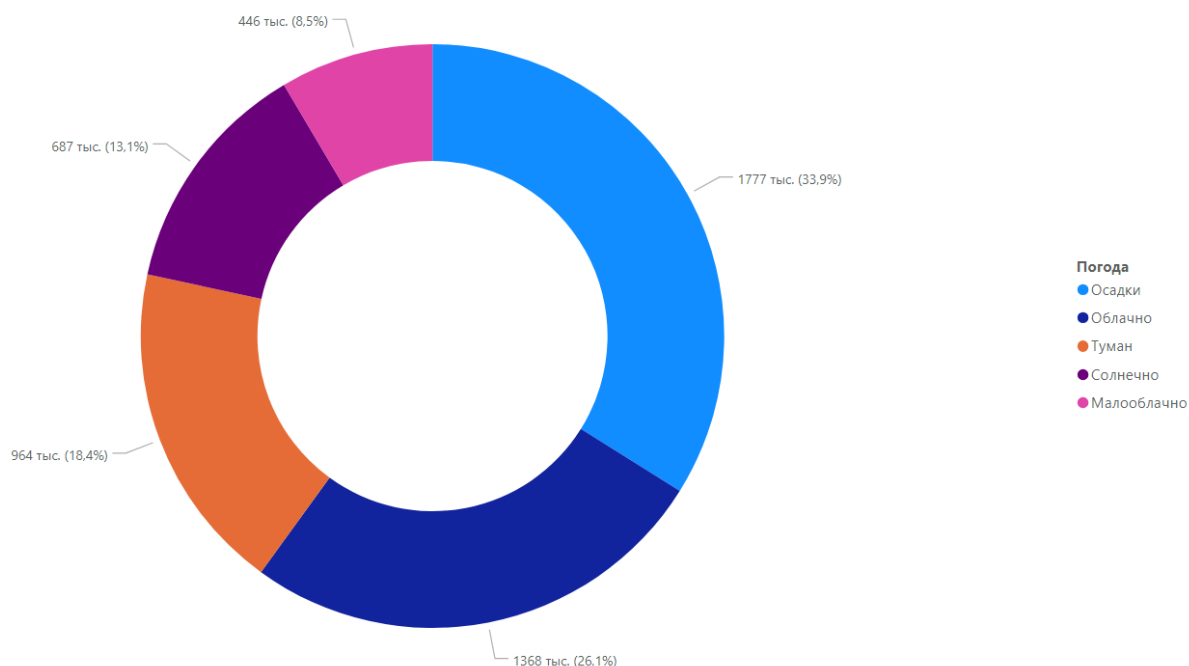


Рис. 1 – Погодные условия и факт наличия ДТП

Затем данные были рассмотрены на наличие корреляции между полосой движения автомобиля (левой или правой) и фактом наличия ДТП (рис. 2). Было выявлено, что аварии чаще случаются (81,2% случаев), когда автомобиль движется по правой полосе.

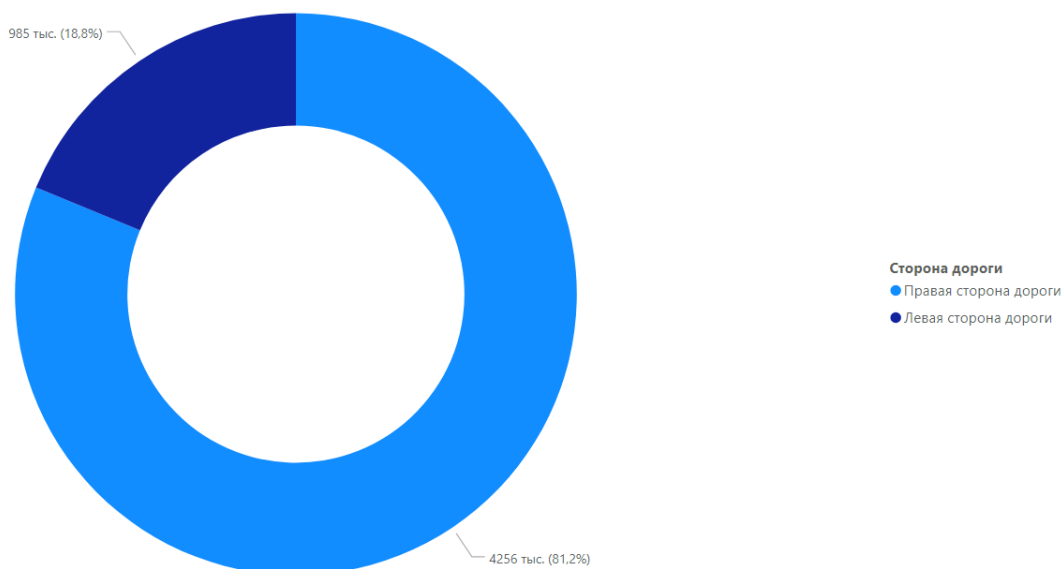


Рис. 2 – Полоса движения автомобиля и факт наличия ДТП

Также был произведен анализ, который помог выявить, в каких штатах США чаще всего происходят аварии (рис. 3). Наиболее часто встречающимся

штатом в рапортах об авариях является Калифорния, занимающая 21,6% от общего числа аварий. Затем Техас и Флорида. Возможность OLAP-куба помогли посмотреть также совокупные данные о факте наличия ДТП, штате и погодных условиях.

Было выявлено, что в данной многофакторной системе наибольшую опасность представляли погодные условия, привычные жителям этих штатов. То есть, например, в штате Алабама, занимающем первое место по количеству осадков за год, наибольшее количество аварий в день происходило в дождливые и туманные дни, а в солнечную погоду ДТП случались на 32,7% реже. Штат Аризона, один из городов которого внесен в книгу рекордов Гиннеса, как самый солнечный город в мире, наиболее подвержен ДТП в солнечные дни, нежели в туманные, облачные или дождливые. [4] Исходя из этого можно сделать вывод, что в более привычных погодных условиях участники дорожного движения менее внимательны, нежели в нестандартных для их местности погодных условиях.

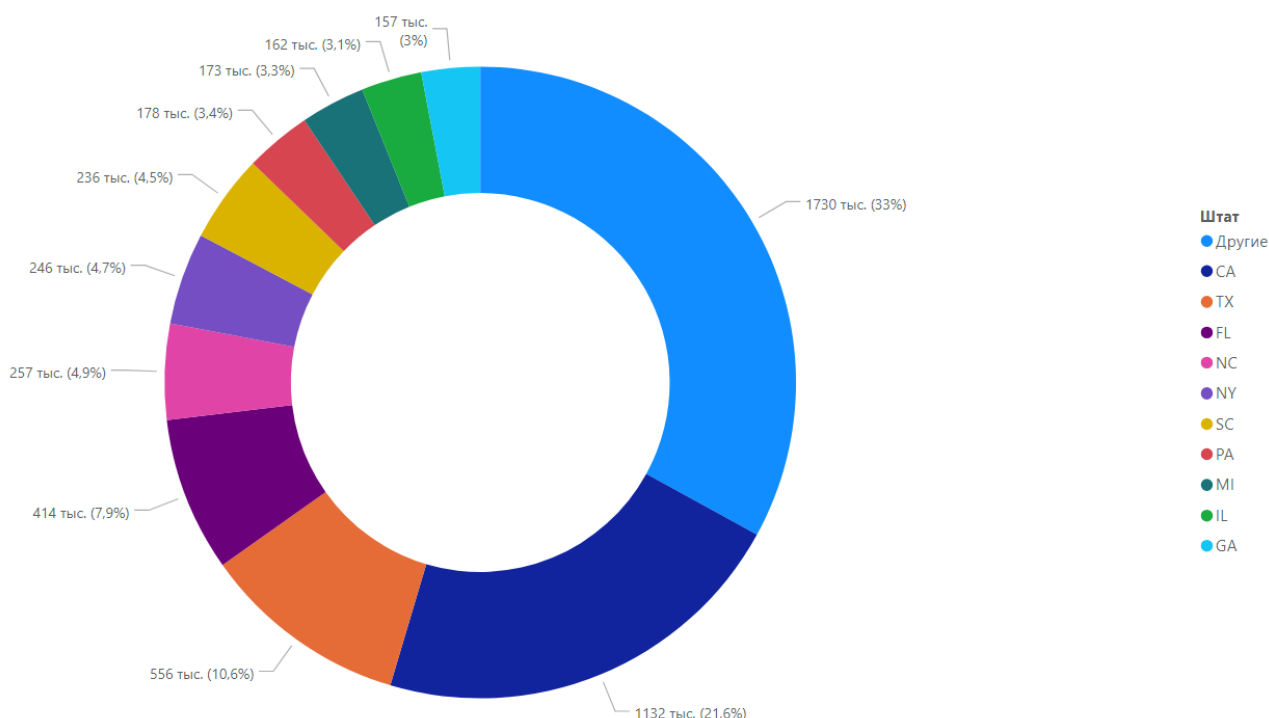


Рис. 3 – Распределение количества аварий по штатам

Затем были выбраны данные, содержащие в себе информацию об окружающей среде (рис. 4), а именно, о факте наличия таких объектов, как:

светофор, пешеходный переход, дорожные знаки, перекрестки, эстакады, искусственные неровности, остановки общественного транспорта и т.д. Наиболее частыми объектами, находящимися около ДТП, а, значит, вероятно оказывающими влияние на факт его происшествия, являются: светофоры (46,1%), пешеходные переходы (22,1%) и перекрестки (14,4%).

Рассмотрение данных об окружающей среде в совокупности с данными о погоде и штатах помогло выявить, что наиболее частыми причинами ДТП (от 34,5% до 48,7% случаев) является совокупное влияния факторов привычной для штата погоды и факта наличия светофоров, пешеходных переходов и перекрестков.

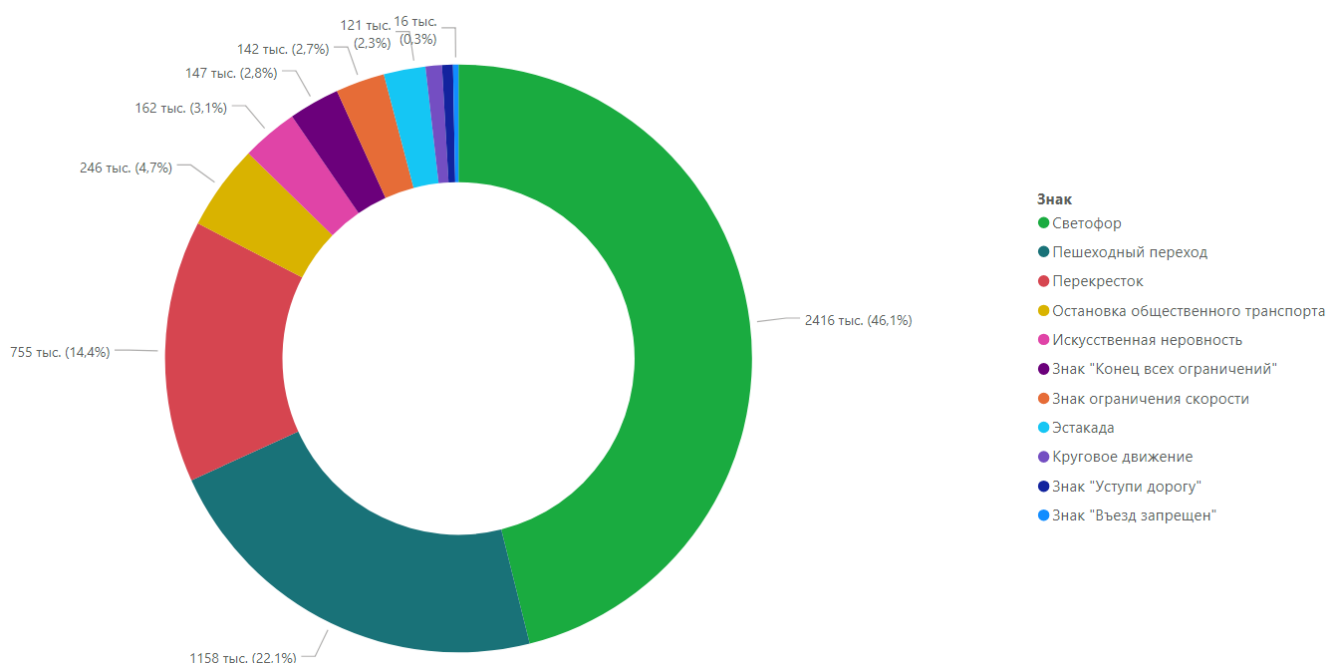


Рис. 4 – Окружающие объекты

Таким образом, можно сделать вывод, что основные сочетания параметров, влияющих на возникновение аварии на дороге, в преимущественном большинстве случаев, состоят из привычных человеку факторов. Проблемы могут возникать и из-за необходимости остановиться или сбавить скорость на светофоре, пешеходном переходе или на перекрестке. Также можно сделать вывод, что при более неординарных сочетаниях параметров участники дорожного движения более сосредоточены на дороге.

Чтобы проверить полученные выводы, были выбраны данные по отдельному штату и на их основе составлен прогноз с помощью метода имитационного моделирования. Данные были выбраны по штату Калифорния, где средняя дневная температура самого холодного месяца – января – составляет +17-18 °С, а самого теплого – августа – составляет +23-24°С. Осадки и туман для данного штата редкость, в основном погода солнечная и малооблачная. [5] Имитационное моделирование данных – это метод, в котором модель представляет корреляционные отношения между одним набором данных и другим набором данных [6]. В данном случае, выходные данные для Калифорнии, с учетом движения машины по правой полосе (более подверженной аварийным ситуациям), получило представление в графическом виде (рис. 5).

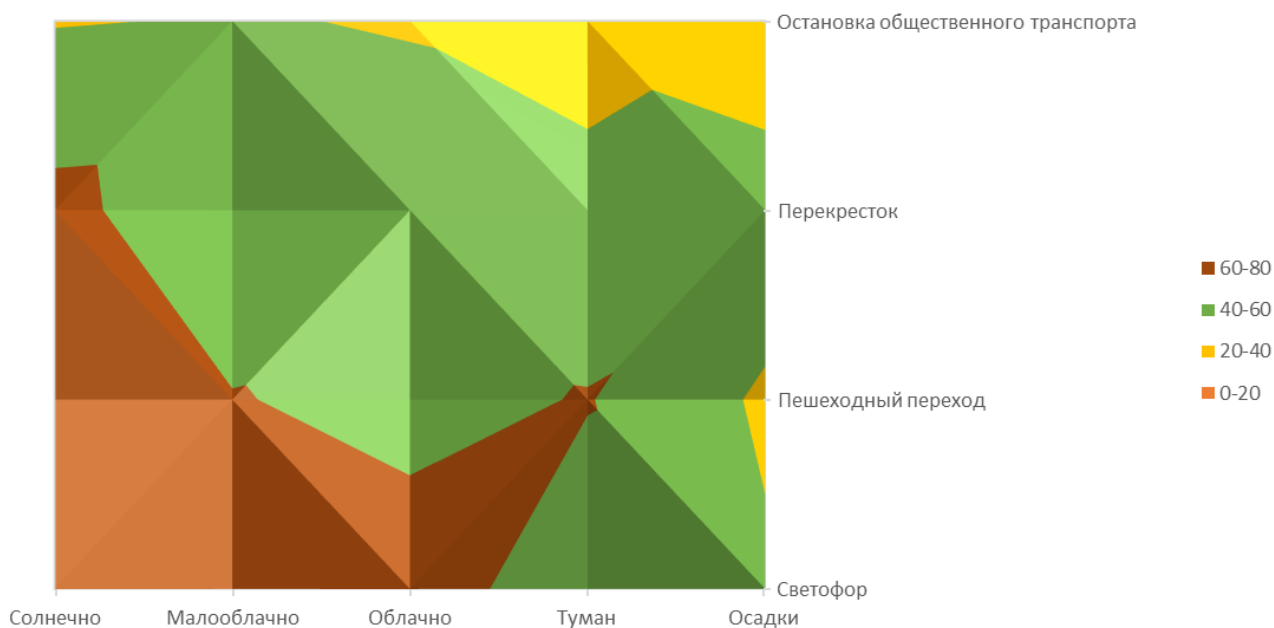


Рис. 5 – Выходные данные имитационного моделирования

В результате чего видно, что при более привычных для участников дорожного движения обстоятельствах, вероятность возникновения ДТП возрастает, а при нестандартных условиях, наоборот, понижается.

Заключение

В результате проделанного исследования были выполнены все поставленные задачи, а именно:

- Подготовлены и проанализированы данные дорожно-транспортных происшествий;
- Выявлены основные причины и факторы, влияющие на появлении аварийных ситуаций.
- Полученные выводы подтверждены с помощью метода имитационного моделирования.

Применение технологий анализа больших данных без сомнения окажет положительное влияние на безопасность дорожного движения. Они помогут проанализировать опасные участки дороги, сделать автомобильные дороги более безопасными и исключить в отдельных случаях риски появления аварийных ситуаций, так как без человеческих усилий и внимания технологий анализа больших данных не сделают дороги абсолютно безопасными. До тех пор, пока отдельные водители не будут в полном объеме руководствоваться всеми правилами дорожного движения, то большие данные помогут, в лучшем случае, избежать 30% аварий от текущего их количества.

Библиографический список

1. ВОЗ — официальный сайт Всемирной организации здравоохранения [Электронный ресурс]. Режим доступа: https://www.who.int/violence_injury_prevention/road_safety_status/report/ru/, свободный – (дата обращения: 10.11.2020).

2. Marks and Harrison – официальный сайт юридическо-аналитической компании Marks and Harrison [Электронный ресурс]. Режим доступа: <https://www.marksandharrison.com/accident-attorney/car-accidents/types-of-car-accidents/>, свободный – (дата обращения: 10.11.2020).

3. Царькова Н.И., Ерисов В.Д., Пекова Е.А. Технология BigData как инструмент управления в межкультурной коммуникации // Управление экономическими системами: электронный научный журнал. – 2019. – № 7.

4. Википедия – свободная энциклопедия [Электронный ресурс]. Режим доступа: <https://ru.wikipedia.org/wiki/Аризона>, свободный – (дата обращения: 10.11.2020).

5. Википедия – свободная энциклопедия [Электронный ресурс]. Режим доступа: <https://ru.wikipedia.org/wiki/Калифорния>, свободный – (дата обращения: 10.11.2020).

6. Fundamental-research – научный интернет-журнал фундаментальных исследований [Электронный ресурс]. Режим доступа: <https://www.fundamental-research.ru/ru/article/view?id=29198>, свободный – (дата обращения: 10.11.2020).

References

1. WHO – the official website of the World Health Organization [Electronic resource]. Access mode: https://www.who.int/violence_injury_prevention/road_safety_status/report/ru/, free – (date of access: 10.11.2020).

2. Marks and Harrison – official site of the legal and analytical company Marks and Harrison [Electronic resource]. Access mode: <https://www.marksandharrison.com/accident-attorney/car-accidents/types-of-car-accidents/>, free — (date of access: 10.11.2020).

3. Tsarkova N.I., Erisov V.D., Pekova E.A. BigData technology as a management tool in intercultural communication // Management of economic systems: electronic scientific journal. – 2019. – № 7.

4. Wikipedia — free encyclopedia [Electronic resource]. Access mode: <https://ru.wikipedia.org/wiki/Arizona>, free – (date of access: 11/10/2020).

5. Wikipedia — free encyclopedia [Electronic resource]. Access mode: <https://ru.wikipedia.org/wiki/California>, free – (date of access: 11/10/2020).

6. Fundamental-research — scientific Internet-journal of fundamental research [Electronic resource]. Access mode: <https://www.fundamental-research.ru/ru/article/view?id=29198>, free – (date of access: 10.11.2020).