Association Rules to Identify Factors Affecting Risk and Severity of Road Accidents

Irina Makarova[1][a], Gulnara Yakupova[1][b], Polina Buyvol[1][c], Eduard Mukhametdinov [1][d] and Anton Pashkevich[2][e]

[1]Kazan Federal University, Syuyumbike prosp., 10a, 423822 Naberezhnye Chelny, Russian Federation
[2]Politechnika Krakowska, Warszawska st., 24, Krakow, Poland

Keywords: Accidents, Association Rules, Factors, Safety, Severity of Consequences, Vehicles.

Abstract: The current increase in automobilization leads to a decrease in road safety. Therefore, the purpose of this research is to analyze and identify the causes that significantly affect the risk and severity of accidents. For this purpose, both histograms plotting and association rules were used. The statistics of traffic accidents in Elabuga town for 2017-2018 were taken as initial information. To identify the most traumatic types of traffic accidents, graphs of the number of accidents by types of accidents and the number of victims for 2017-2018 were constructed. It was found that the greatest number of injured (wounded or dead) is observed in collisions and hitting a pedestrian. Then, road sections of the concentration of accidents were analyzed, the most common types of accidents and the main violations of traffic rules that contribute to the occurrence of accidents were determined. To identify hidden relationships between factors and accidents with severe consequences, association rules were applied. As a result, the influence of weather conditions, quality of road infrastructure and marking was established.

1 INTRODUCTION

The annual growth of the global fleet of vehicles with new units using the existing limited road infrastructure leads to an increase in traffic intensity and the number of congestion on the roads, which, in turn, causes many problems, including those related to a decrease in traffic safety and an increase in the negative impact on environment.

In this regard, in 2010, the European Commission launched the European Smart Cities Initiative, which includes four urban areas: buildings, heating and cooling systems, electricity and transportation (European Initiative on Smart Cities, 2010-2020). The goal associated with transport is to create intelligent public transport systems based on real-time information, traffic management systems (TMS) to prevent traffic collapses, as well as reduce the environmental negative effects of the transport system on the environment. However, as noted in (Djahel, S., Doolan, R., Muntean, G.-M., Murph, J., 2015), existing TMS do not provide sufficient and accurate traffic information to provide detailed and timely traffic monitoring and management.

On the other hand, over the past decade, commercial companies and government organizations have been actively developing research areas related to autonomous vehicles, which, it seems, will soon become widespread on the roads after legal issues regarding the functioning of these new road users are resolved. It seems that one of the main problems that inevitably arise when changing the composition of road users is their safe and efficient operation in difficult road conditions, such as road junctions. At road network locations where traffic conflicts may occur, such as at intersections, it is necessary to ensure that autonomous vehicles operate safely and efficiently, and, more importantly, that

[a] https://orcid.org/0000-0002-6184-9900
[b] https://orcid.org/0000-0001-6822-3700
[c] https://orcid.org/0000-0002-5241-215X
[d] https://orcid.org/0000-0003-0824-0001
[e] https://orcid.org/0000-0002-4066-5440

Makarova, I., Yakupova, G., Buyvol, P., Mukhametdinov, E. and Pashkevich, A.
Association Rules to Identify Factors Affecting Risk and Severity of Road Accidents.
DOI: 10.5220/0009836506140621
In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2020)*, pages 614-621
ISBN: 978-989-758-419-0

conventional people-driven vehicles at least maintain their current level of safety.

In our opinion, a base of the most representative "reference" scenarios of the interaction of autonomous vehicles with traditional ones, especially with the most vulnerable, should be created for this. It is necessary to determine a list of driving situations that should be evaluated in terms of possible conflict avoidance, and then traffic control options for these situations will be modeled and defined. Since there is no information on the interaction of autonomous vehicles with other participants in the movement, but taken into account that their movement will be carried out according to the given algorithms, at the first stage it is necessary to analyze the existing statistics of road traffic accidents, select their concentration, and then use the simulation models to determine the most dangerous scenarios.

2 OVERVIEW OF EXISTING METHODS

Recently, the number of researches devoted to the study of the significance of factors affecting the severity of accidents has increased significantly (Zou, X., Yue, W. L., Vu, H.L., 2018). An active field of research by scientists from different countries is the study of the complex relationship between influencing factors and the severity of accidents using statistical methods and machine learning algorithms: classification and regression trees (Moral-García, S Castellano, J. G., Mantas, C. J., Montella, A., Abellán, J., 2019), neural networks (Theofilatos, A., Chen, C., Constantinos, A., 2019; Zheng, M., Li, T., Zhu, R., Chen, J., Ma, Z., Tang M., et al., 2019), support vector methods (Chen, C., Zhang, G., Qian, Z., Tarefder, R.A., Tian, Z., 2016), naive Bayes classifier (Chen, C., Zhang, G., Yang, J., Milton, J.C., Alcántara, A.D., 2016; Li, Z., Wu, Q., Ci, Y., Chen, C., Chen, X., Zhang, G., 2019), binary (Zhai, X., Huang, H., Sze, N.N., Song, Z., Hon, K.K., 2019; Salon, D., McIntyre, A., 2018; Jalayer, M., Shabanpour, R., Pour-Rouholamin, M., Golshani, N., Zhou, H., 2018; Rezapour, M., Moomen, M. and Ksaibati, K., 2019; Sam, E.F., Daniels, S., Brijs, K., Brijs, T., G. Wets, 2018; Sam, E.F., Daniels, S., Brijs, K., Brijs, T., G. Wets, 2018; Ahmed, M.M., Franke, R., Ksaibati, K., Shinstine, D.S., 2018) and polynomial (Penmetsa, P., Pulugurtha, S.S., 2018) logistic regression, association rules (AR) (Montella, A., 2011; Wu, P., Meng, X., Song, L., Zuo, W., 2019; Weng, J., Zhu, J.-Z., Yan, X., Liu, Z., 2016; Nitsche, P., Thomas, P., Stuetz, R., Welsh, R., 2017; Xu, C., Bao, J., Wang, C., Li, P., 2018). The accelerating growth in computing power of computers and the emergence of more sophisticated methods have contributed to the rapid development of road safety prediction models. Multivariate modeling and mining methods are gradually replacing traditional one-dimensional modeling methods based on the linear model and the Poisson model.

When many researchers identify the relationship of a large number of factors influencing to the severity of the accident consequences, the method of AR is widely used. So, as a result of research, the authors of (Montella, A., 2011) identified the factors leading to accidents at intersections and established the interdependencies between these factors. In general, they identified numerous factors related to road and environmental problems, but not related to pedestrian or vehicle. The most important factors characterizing the geometry of the road were the radius and angle of deviation. The significant role of road markings and signs was also identified.

The authors of (Wu, P., Meng, X., Song, L., Zuo, W., 2019) selected the city crossroads for analysis as places that pose a serious security risk, since most accidents within the city territory occur in places or near junctions. They analyzed safety indicators for six types of intersections and factors affecting the severity of accidents. Fault tree analysis was used to assess the risk of intersections, and AR were used to analyze the nature of the severity of accidents. As a result, four types of urban junctions with a high level of accident risk and more than 4,000 rules describing accidents with severe consequences were identified.

In (Weng, J., Zhu, J.-Z., Yan, X., Liu, Z., 2016), a method based on AR is designed to analyze the characteristics and factors contributing to emergency situations during road repair work. Most AR include conditions such as a speed of more than 40 km / h and the use of traffic control devices.

The authors in the article (Nitsche, P., Thomas, P., Stuetz, R., Welsh, R., 2017) presents a data analysis technique, including the preparation, analysis and visualization of accident data, which allows identifying critical pre-emergency scenarios at T - and X- junctions as a basis for testing the safety of autonomous vehicles. In this methodology, the k-medoid method is used to form homogeneous groups (clusters) among the array of accident records. Subsequently to this clusters AR are applied to generate typical motion scenarios and accident patterns.

In (Xu, C., Bao, J., Wang, C., Li, P., 2018), the method of AR was also used to study the factors contributing to the occurrence of serious traffic accidents in China. The analysis showed that serious accidents with victims are the result of complex interactions between road users, the technical characteristics of vehicles, the geometric parameters of roads and environmental factors. As a result, the authors received a number of reasons for the occurrence of serious accidents with victims.

In Russian practice, as a rule, to analyze the severity of injuries in road traffic accidents, models based on the calculation of the total number of injured (dead and wounded) are applied (Petrov, A., Petrova D., 2016; Evtukov, S., Golov, E., Sazonova, T., 2018), as well as visualization tools for these indicators are used (Open data and traffic feedback, 2020, Accident map, 2020, Dornadzor, 2020). It should also be noted that in the Russian Federation there is no official universal procedure for the comprehensive analysis of road accidents, approved and adopted by government bodies. Despite this, the duty of each subject of the road safety system is a detailed analysis of the factors on the manifestation of which this subject or the agencies involved with him are able to influence.

In connection with the foregoing, today the search for complex relationships between factors that affect the severity of the consequences is an urgent task. Therefore, the aim of this study is to analyze and identify the causes that significantly affect the risk and severity of accidents.

3 THE SOURCE DATA

Real data on accidents plays an important role in improving the safety of road transport, since information on the factors that led to malfunctions and accidents is necessary to understand the causes of malfunctions and how such events can be prevented in the future.

In order to start any assessments of accidents, first of all, it is necessary to have an appropriate information array, the completeness and confidence of which are of paramount importance in ensuring the effectiveness of the analysis. Moreover, as noted in (Imprialou, M., Quddus, M., 2019), insufficient reporting of accidents is a recognized and studied problem of road safety researches worldwide. This fact is connected, firstly, with the failure to notify the relevant state authorities about the accident, since its participants agree to sign private payments for insurance purposes, either there was no other-party involvement (for example, an accident with individual vehicles) or there were no obvious injuries immediately after accident. And secondly, there is an underestimation of the results of the road safety analysis due to the non-presentation of certain categories of injuries in the collected reports.

Consider the procedure for forming an data array about accidents in the Russian Federation.

According to the Rules for the registration of road traffic accidents (Decree of the Government of the Russian Federation of June 29, 1995 N 647), all accidents are divided from the point of view of accounting into three groups.

The first group includes accidents in which people died or were injured. Information about these incidents is recorded in a special accident accounting card and is included in the state statistical reporting. According to the Decree of the Government of the Russian Federation of November 19, 2008 N 859 “On amendments to the rules for the registration of accidents”, killed in the accident is person who injured in an accident and who died within 30 days from its consequences.

The second group includes accidents with material damage without injuries, as well as an accident in which people received bodily injuries that did not cause any tangible harm to their health (while the participants do not fall into the category of “wounded”). Information about such accidents is not included in the state statistical reporting, but is taken into account and analyzed at the level of individual cities and regions.

The third group includes individual incidents, which, according to formal signs, can be qualified as road transport, but information about them is not included in the state statistical reporting, they are not subject to accounting as road accidents.

However, to improve the degree of accounting, it seems necessary to keep a complete record of accidents with their classification, depending on the severity of the consequences, as involving:

- material damage,
- light bodily harm,
- moderate to severe bodily injuries,
- the death of the victim,
- especially grave consequences (4 or more were killed or 15 or more people were injured).

As the initial information, data collected by the State Inspectorate for Road Safety in the Elabuga town for 2017-2018 were used. According to the classification of cities by population, Elabuga refers to medium-sized towns (population 73,913 people).

The initial selection consisted of the following factors, which were divided into four categories:

- Characteristic of an accident - Type of accident, traffic violations, injured and died, children injured and died;
- Driver - Social characteristics of the driver, Experience (years), Sex, Type and Degree of intoxication (mcg / l), Direct traffic violations, Related traffic violations;
- Vehicle - Number of vehicles involved in the accident, Vehicle Type, Vehicle Brand and Model, Faults, Damage;
- Road - Street, Deficiencies in the road conditions, Type of road infrastructure in place, Factors affecting traffic conditions (presence of artificial bumps), Number of lanes, State of the carriageway, Road profile, The lane in which the accident occurred, Carriageway width, Curb Width, Sidewalk Width, Dividing Strip Width, View of the dividing strip;
- Environment - Lighting, Weather, Year, Month, Day of the week, Hour.

4 METHODOLOGY

A graph of the distribution of the accidents for 2017-2018 was built. (Fig. 1). The most common types were a collision, hitting a standing vehicle, hitting an obstacle, hitting a pedestrian and exit from the road. It is worth noting that the number of some accidents in 2018 increased significantly (collisions), and the number of less common ones was small.

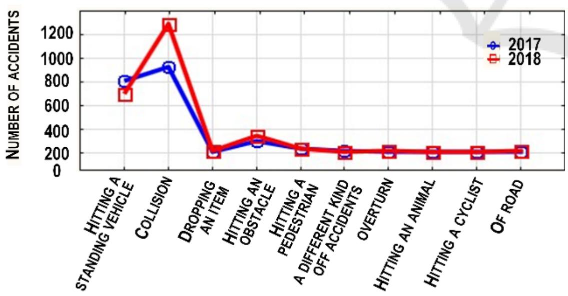

Figure 1: A graph of the accidents grouped by the type of accident for 2017-2018.

As a result of analyzing the histogram of the distribution of the accidents by the number of victims and by the type of accident, it was found that the largest number of injured (wounded or dead) was observed during collisions and hitting a pedestrian (Fig. 2). Consequently, the question of finding the main causes leading to an increase in the number of collisions remains relevant. This is necessary for adjustment the policy of regulating road safety.

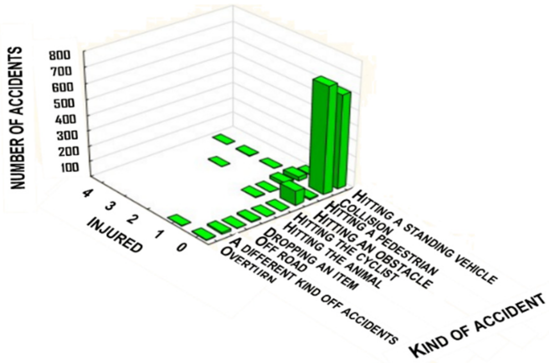

Figure 2: Histogram of the distribution of the accidents by the number of victims and by the type of accident for 2017-2018.

Next, specific road sections of accident concentration in Elabuga town were analyzed.

According to the results of the analysis of accidents for a 2-year period, two of the most dangerous sections of the road network were identified, on which the largest number of accidents occurred during the considered period:

- Mira street, b. 30 - b. 33;
- Neftyanikov street, b. 55 - b. 57.

It is worth noting that, in general, on these streets the number of accidents has increased rapidly from 2017 to 2018 (Fig. 3).

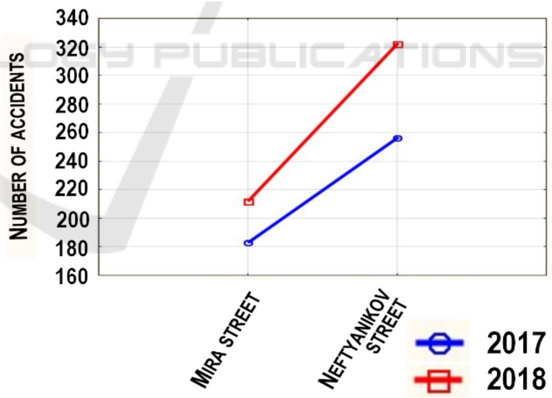

Figure 3: A graph of the accidents in Mira street and Neftyanikov street for 2017-2018.

When analyzing the distribution of accidents occurred in these sections of accident concentration by type, it is clear that on Mira street the most common ones are collisions and hitting a standing vehicle (Fig. 4). The main causes of road accidents are violation of the rules for the vehicle location on the roadway, wrong choice of distance, non-compliance of lateral interval, non-compliance with the order of travel (Fig. 5).

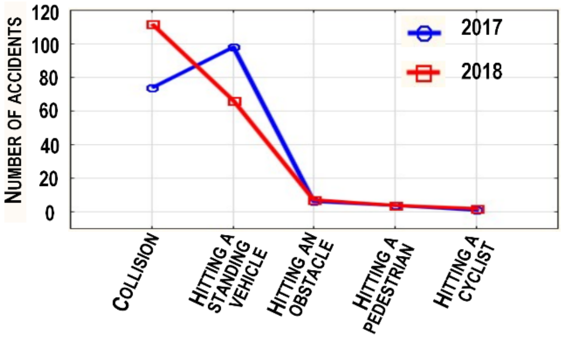

Figure 4: A graph of the number of accidents on Mira street grouped by type of accident for 2017-2018.

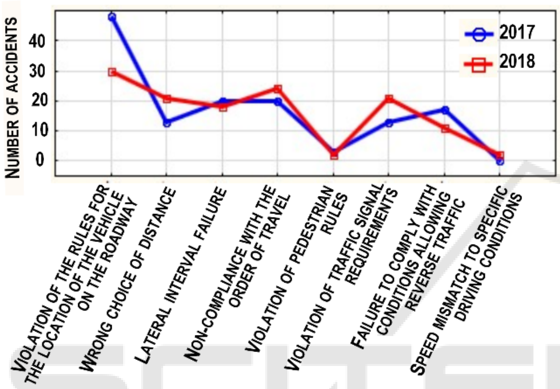

Figure 5: A graph of the number of accidents on Mira street grouped by type of driving offence for 2017-2018.

It is worth noting that hitting a pedestrians with serious consequences for human health were occurred on this section of road. One of the contributing factors is the multi-apartment residential buildings located here, unregulated pedestrian crossing, as well as places of attraction in the form of city malls (Fig. 6).

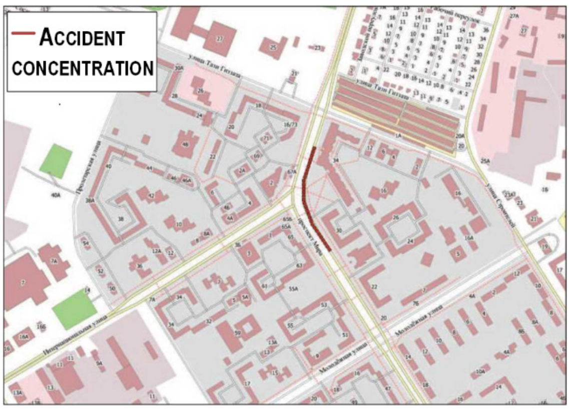

Figure 6: Road section of accident concentration on Mira street, b. 30 - b. 34 for 2017-2018.

On Neftyanikov street for 2017-2018 the largest number of accidents was recorded in Elabuga town for the following types of accidents: collision and hitting a standing vehicle (Fig. 7). The main driving offences, as well as on Mira street, are violation of the rules for the vehicle location on the roadway, wrong choice of distance, non-compliance of lateral interval, non-compliance with the order of travel (Fig. 8).

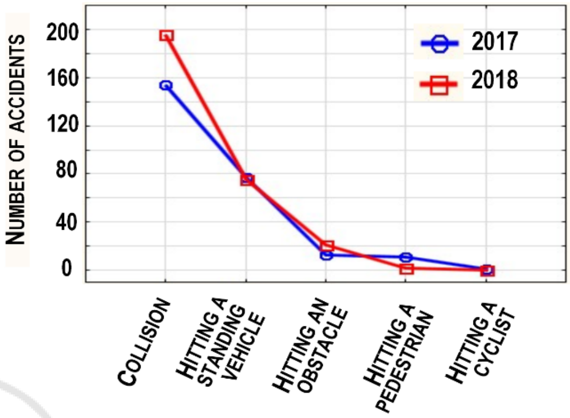

Figure 7: A graph of the number of accidents on Neftyanikov street grouped by type of accident for 2017-2018.

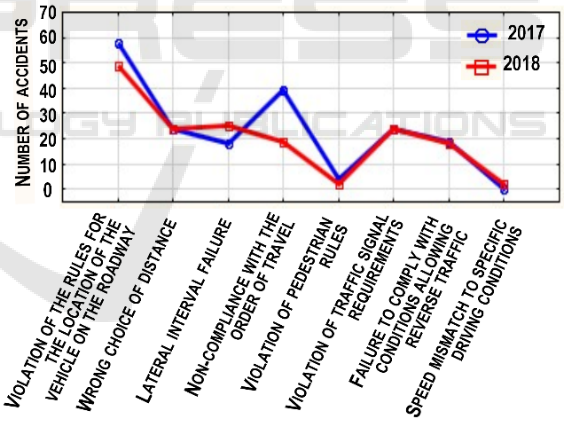

Figure 8: A graph of the number of accidents on Neftyanikov street grouped by type of driving offence for 2017-2018.

On this section there are an unregulated pedestrian crossing, a place of public transport stop, a place of attraction in the form of an emergency department (Fig. 9). The largest number of pedestrian accidents with moderate to severe bodily injuries was recorded here.

Such a type of accident as hitting a standing vehicle at these dangerous sections can be explained by a shortage of parking spaces for permanent storage of the vehicle.

Figure 9: Road section of accident concentration on Neftyanikov street, b. 55 - b. 57 for 2017-2018.

In these sections of Mira street and Neftyanikov street in the morning rush hour, the intensity of vehicle traffic was measured, presented in Table 1.

Analysis of the capacity of these shows that despite their low utilization coefficient (which is equal to the ratio of traffic intensity to crossing capacity), not exceeding 35% of the design capacity, these sections occupy the first positions in terms of accident rate. Therefore, we can assume the presence of other factors that are not related to traffic intensity, leading to a large number of accidents.

Table 1: Traffic intensity in the morning rush hour on Mira street and Neftyanikov street.

Street	Traffic intensity, units / h	Crossing capacity, units / h	Utilization coefficient
Prospect Neftyanikov street (from the side of Proletarskaya street)	1265	3600	0,35
Mira street	705	2400	0,29
Neftyanikov street (from the side of Stroiteley street)	1126	3600	0,31

In this regard, a deeper analysis of the interdependencies between the factors affecting the occurrence of various types of accidents in these sections was carried out. Most of the available accident data used in this study is categorical, that is, it is described by qualitative factors (also called nominal factors). Although categories can be encoded as numbers, for example, 1-woman, 2- man, these numbers will not have mathematical meaning. Therefore, special methods are needed to analyze categorical data.

Recently, interest in the methods of "discovering knowledge in databases" has been growing. One of the common analytical methods of data processing is affinity analysis. The method determines the mutual relations between events occurring jointly.

One of the applications of affinitive analysis is market basket analysis in order to detect associations between different data, for example finding rules to quantify the relationship between two or more data. Such rules are called AR.

The basic concept in the theory of AR is a transaction - a set of events that occur together.

An association rule consists of two sets of objects called a antecedent and a consequent, written in the form $X \rightarrow Y$, which reads "from X follows Y". Thus, an association rule is formulated in the form "If an antecedent, then a consequent." The antecedent is often limited to the content of only one subject.

AR describe the relationship between sets of items matching the antecedent and the consequent. This relationship is characterized by two indicators - support and confidence. Let D be a transaction database, and N - the number of transactions in this database. Each transaction Di represents a certain set of items. S is a support, C is a confidence. Association rule support is the number of transactions containing both an antecedent and a consequent.

For example, it can written for the association $A \rightarrow B$:

$$S(A \rightarrow B) = P(A \cap B) \qquad (1)$$

A confidence of an association rule is a measure of the rule accuracy, which is determined as the ratio of the number of transactions containing both the antecedent and the consequent to the number of transactions containing only the antecedent. For example, it can written for the association $A \rightarrow B$:

$$C(A \rightarrow B) = P(A|B) = P(A \cap B)/P(A) \qquad (2)$$

If the support and confidence are high enough, then this makes it possible to assert that any future transaction that includes an antecedent will also contain a consequent. However, it is also necessary to estimate the degree of independence of the antecedent and the consequent in order to avoid the situation of obtaining "fictitious" rules, when the support and confidence are high.

A lift is the ratio of the frequency of occurrence of a antecedent in transactions, which also contain a consequent, to the frequency of occurrence of the consequent as a whole. A lift is defined as follows:

$$L(A \rightarrow B) = C(A \rightarrow B)/S(B) \qquad (3)$$

In our case, association detection is the identification of a combination of factors leading to an accident. To identify a combination of the most common factors that led to severe injuries or death of road accident participants, 75 records of the accidents were selected. The AR method was applied to these data. By experience, the minimum values of support were established - 4%, confidence - 60%, lift - 1, the power of the rule set - 4, at which 152 rules are formed - the amount acceptable for consideration and interpretation. From a practical point of view the most interesting and useful among the rules obtained are given in table 2.

Table 2: The most informative AR.

Antecedent	Consequent	Support	Confi-dence	Lift
"Zhiguli" VAZ-2108, 09 and modifications	Friday	4,11	60	3,982
November	Okruzhnoe highway	4,11	60	4,867
Front Left Side And Female	Wounded and Female	4,11	60	4,867
Septembr	Rain	4,11	60	5,475
Gas station	Naberezhnye-Chelninskoe highway	4,11	60	6,257
Curb Width – 25 m	Friday	4,11	75	4,977
Ground stripe	**Mira street**	**5,48**	**66,7**	**6,083**
February	Lacks of winter street cleaning and Snowfall and Curb Width – 30 m	5,48	66,7	9,733
Public transport stop And Lack, poor distinguishability of horizontal marking of the carriageway	Improper usage, poor visibility of road signs	5,48	80	8,343
Lack of pedestrian fences in the required places	Adjustable intersection	8,22	60	3,65
Adjustable pedestrian crossing	Lack of pedestrian fences in the required places	11	61,5	4,492

The rules were filtered out by support, confidence and lift.

An analysis of some of the obtained rules shows the direct relationship of accidents with severe consequences and the lacks of winter street cleaning in February, leading to the appearance of snowfall on the road. In three accidents in September it was raining, also worsening road conditions. In eight cases, in the immediate vicinity of the regulated pedestrian crossing, there were no fences in the required places, which could lead to violation of traffic rules by pedestrians. The combination of “an adjustable intersection” and “lack of pedestrian fences in the necessary places” is also quite common - in 6 cases. In 4 cases, grave consequences were caused by poor visibility of the horizontal marking of the roadway and poor visibility of road signs near a public transport stop. On Friday, in three cases, an accident occurred with the participation of VAZ-2108, 09 “Zhiguli” and modifications, as well as with a curb width of 25 meters. In three cases, injuries occurred as a result of a hitting to the left side, that is from the driver. Also, according to the received rules, accidents with severe consequences on Mira street more often occur in places where this street has ground stripe (4 cases), on Naberezhno-Chelninsky highway at a gas station (3 cases) and on Okrugny highway in November (3 cases).

5 CONCLUSION

When developing a road safety policy, the government of any developed or developing country needs to analyze factors that affect not only the likelihood of an accident, but also determine the severity of the consequences.

To identify the most traumatic types of traffic accidents, graphs of the number of accidents by types of accidents and the number of victims for 2017-2018 were constructed. It was found that the greatest number of injured (wounded or dead) is observed in collisions and hitting a pedestrian. Then, road sections of the concentration of accidents in Elabuga town were analyzed, the most common types of accidents and the main violations of traffic rules that contribute to the occurrence of accidents were determined. At the same time, the analysis of the traffic intensity of these streets showed their low load. Therefore, to identify hidden relationships between factors and accidents with severe consequences, AR were applied. As a result, the influence of weather conditions, quality of road infrastructure and marking was established.

The resulting rules will allow to build simulation models, with the help of which, and taking into account the different situations that occur in a real transport system, it can be possible to determine the most effective measures to prevent accidents and mitigate the severity of the consequences.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research: grant No. 19-29-06008\19.

REFERENCES

Accident map. https://dtp-stat.ru/(accessed 05.01.2020).

Ahmed, M.M., Franke, R., Ksaibati, K. and Shinstine, D.S., 2018. Effects of truck traffic on crash injury severity on rural highways in Wyoming using Bayesian binary logit models. In *Accident Analysis & Prevention*, vol.117, pp. 106-113.

Chen, C., Zhang, G., Qian, Z., Tarefder, R.A., Tian, Z., 2016. Investigating driver injury severity patterns in rollover crashes using support vector machine models. In *Accident Analysis & Prevention,* vol.90, pp. 128–139.

Chen, C., Zhang, G., Yang, J., Milton, J.C., Alcántara, A.D., 2016. An explanatory analysis of driver injury severity in rear-end crashes using a decision table/Naïve Bayes (DTNB) hybrid classifier. In *Accident Analysis & Prevention,* vol. 90, pp. 95-107.

Decree of the Government of the Russian Federation of June 29, 1995 N 647. Meeting of the legislation of the Russian Federation, 1995, N 28, art. 2681.

Djahel, S., Doolan, R., Muntean, G.-M., Murph, J., 2015. A Communications-oriented Perspective on Traffic Management Systems for Smart Cities: Challenges and Innovative Approaches. In *IEEE Communications Surveys & Tutorials*, vol. 17, iss. 1, pp.125-151.

Dornadzor. https://dornadzor-sz.ru/(accessed 05.01.2020).

European Initiative on Smart Cities, 2010-2020, http://setis.ec.europa.eu/set-plan-implementation/technologyroadmaps/european-initiative-smart-cities.

Evtukov, S., Golov, E., Sazonova, T., 2018. Prospects of scientific research in the field of active and passive safety of vehicles. In *MATEC Web of Conferences*, vol. 239, paper № 040182018.

Imprialou, M., Quddus, M., 2019. Crash data quality for road safety research: Current state and future directions. In *Accident Analysis & Prevention*, vol.130, pp. 84-90.

Jalayer, M., Shabanpour, R., Pour-Rouholamin, M., Golshani, N., Zhou, H., 2018. Wrong-way driving crashes: A random-parameters ordered probit analysis of injury severity. In *Accident Analysis & Prevention*, vol.117, pp. 128-135.

Li, Z., Wu, Q., Ci, Y., Chen, C., Chen, X., Zhang, G., 2019. Using latent class analysis and mixed logit model to explore risk factors on driver injury severity in single-vehicle crashes. In *Accident Analysis & Prevention,* vol.129, pp. 230–240.

Montella, A., 2011. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. In *Accident Analysis & Prevention,* vol. 43, iss. 4, pp. 1451-1463.

Moral-García, S, Castellano, J. G., Mantas, C. J., Montella, A., Abellán, J. 2019. Decision Tree Ensemble Method for Analyzing Traffic Accidents of Novice Drivers in Urban Areas. In *Entropy*, vol. 21(4), pp.360.

Nitsche, P., Thomas, P., Stuetz, R., Welsh, R., 2017. Pre-crash scenarios at road junctions: A clustering method for car crash data. In *Accident Analysis & Prevention*, vol.107, pp. 137-151.

Open data and traffic feedback. Безопасные дороги.рф. (accessed 05.01.2020).

Penmetsa, P., Pulugurtha, S.S., 2018. Modeling crash injury severity by road feature to improve safety. In *Traf Inj Prev*, vol. 19, iss. 1, pp. 102-109.

Petrov, A., Petrova D., 2016. Assessment of Spatial Unevenness of Road Accidents Severity as Instrument of Preventive Protection from Emergency Situations in Road Complex. In *IOP Conf Ser: Materials Sc Eng*, vol. 142, iss. 1, paper № 012116.

Rezapour, M., Moomen, M. and Ksaibati, K., 2019. Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: A case study in Wyoming. In *Jour Saf Res*, vol. 68, pp. 107-118.

Salon, D., McIntyre, A., 2018. Determinants of pedestrian and bicyclist crash severity by party at fault in San Francisco, CA. In *Accident Analysis & Prevention*, vol. 110, pp. 149-160.

Sam, E.F., Daniels, S., Brijs, K., Brijs, T. and G. Wets, 2018. Modelling public bus/minibus transport accident severity in Ghana. In *Accident Analysis & Prevention*, vol. 119, pp. 114-121.

Theofilatos, A., Chen, C., Constantinos, A., 2019. Comparing machine learning and deep learning methods for real-time crash prediction. In *Journal of the Transportation Research Board,* Vol.2673 issue: 8, pp. 169-178.

Weng, J., Zhu, J.-Z., Yan, X., Liu, Z., 2016. Investigation of work zone crash casualty patterns using association rules. In *Accident Analysis & Prevention*, vol.92, pp. 43-52.

Wu, P., Meng, X., Song, L., Zuo, W., 2019. Crash Risk Evaluation and Crash Severity Pattern Analysis for Different Types of Urban Junctions: Fault Tree Analysis and Association Rules Approaches. In *Trans Res Rec*, vol. 2673, iss. 1, pp. 403-416.

Xu, C., Bao, J., Wang, C., Li, P., 2018. Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. In *Journal of Safety Research*, vol. 67, pp. 65-75.

Zhai, X., Huang, H., Sze, N.N., Song, Z. and Hon, K.K., 2019. Diagnostic analysis of the effects of weather condition on pedestrian crash severity. In *Accident Analysis & Prevention*, vol. 122, pp. 318-324.

Zheng, M., Li, T., Zhu, R., Chen, J., Ma, Z., Tang M., et al. 2019. Traffic accident's severity prediction: a deep learning approach based CNN network. In *IEEE Access PP (99)*, vol.7, pp. 39897-39910.

Zou, X., Yue, W. L., Vu, H.L., 2018. Visualization and analysis of mapping knowledge domain of road safety studies. In *Accident Analysis & Prevention,* vol.118, pp. 131-145.