

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЕСТЕСТВЕННЫХ, ТЕХНИЧЕСКИХ И СОЦИАЛЬНЫХ НАУКАХ

УДК 656.1

М.Е. Елисеев¹, М.Е. Сангалова²

О КЛАССИФИКАЦИИ ОЧАГОВ АВАРИЙНОСТИ

Нижегородский государственный технический университет им. Р. Е. Алексеева¹,
Арзамасский филиал Нижегородского государственного университета им. Н.И. Лобачевского²

Рассматривается классификация очагов аварийности с целью более полного анализа и выявления причин аварий. Стандартные подходы к разбиению очагов аварийности на классы не годятся. В статье для решения этой задачи использована теория графов. Найденная методика позволяет вывить места аварийности схожего типа. При наличии информации об одном очаге из класса автоматически получаем данные о всем классе. Исследование может быть использовано для определения причин аварийности любого крупного города. В частности, результаты работы планируется использовать при разработке подсистемы мониторинга интерактивной карты аварийности.

Ключевые слова: топографический анализ, ДТП, ГИС, геоинформационная система, интерактивная карта.

Объем аналитического материала о дорожно-транспортных происшествиях (ДТП) огромен. Невольно возникает вопрос: нельзя ли на основании строгого математического анализа и современных компьютерных технологий выявить причины аварий? Эта информация могла бы быть использована в двух направлениях:

- 1) предоставление водителям (автошколам, перевозчикам и т.д.) достоверной информации об опасностях на дороге, в удобной для восприятия форме;
- 2) изменение дорожной инфраструктуры в тех районах где это наиболее эффективно.

При анализе аварийности крупного города прежде всего выявляются места концентрации дорожно-транспортных происшествий – очаги ДТП. Эта операция называется топографическим анализом. В работах [1, 2] рассматривались так называемые маркеры очагов – это матрицы строки, элементами которых являются доли ДТП определенного типа. Выделялись основной, временной и причинный маркеры. Данные характеристики очагов позволяют существенно продвинуться в вопросе выявления причин ДТП, но информация получается несколько обрывочной. Например, известно, что в очаге «А» чаще, чем в среднем виновниками аварий являются мужчины возраста 26-40 лет со стажем 3-6 лет, а причиной многих аварий является гололедица. Эти характеристики существенно лучше, чем просто число ДТП, но четкого понимания механизма аварий они не дают [3].

Для более глубокого и полного анализа целесообразно использовать стандартный математический метод – классификацию.

Все очаги прежде всего разделим на типовые и нетиповые. К типовым отнесем очаги, маркеры которых близки к среднему: для каждого очага могут быть найдены доли $x_i = \frac{n_i}{N}$, где n_i – число ДТП соответствующей категории водителей; N – общее число ДТП. Данные доли составляют матрицу-строку, так называемый основной маркер очага.

Маркеры всех очагов сравнивались с маркером среднего, соответствующим всем авариям. Рассмотрим соответствующий график (рис. 1).



Рис. 1

По нему можно судить в целом, насколько часто попадает в аварию та или иная категория водителей. Например, пики графика приходятся на категории 32 (мужчины возраста 26-40 лет со стажем вождения 7 лет и более), 33 (мужчины возраста 41-60 лет, стаж 7 лет и более), 22 (мужчины возраста 41-60 лет, стаж 3-6 лет), 21 (мужчины возраст 26-40 лет, стаж 3-6 лет), 11 (мужчины возраста 21-25 лет, стаж 1-2 года). Для интерпретации этих результатов желательно располагать информацией о реальном нахождении на дорогах всех категорий водителей. Мужчины попадают в аварии чаще, поскольку основную массу водителей все же представляют мужчины. Водители пожилого возраста редко попадают в ДТП, так как их процент от общего числа достаточно мал.

Для каждого очага ввели в рассмотрение сумму квадратов отклонений от среднего $d_{0i} = \sum_{j=1}^{40} (x_j - x_{jcp})^2$. Полученные значения однозначно характеризуют расхождения (большие или меньшие) со средним значением маркера. Для графиков это означает большую или меньшую схожесть с графиком среднего, представленного ранее.

Затем очаги были расположены в порядке возрастания указанного параметра: суммы квадратов разностей d_{0i} .

В верхней части списка оказались типичные очаги, практически не отличающиеся от среднего значения, графики которых имеют незначительные расхождения со средним. Картина ДТП для этих очагов будет полностью ясна из рассмотрения среднего значения.

В нижней части списка находятся очаги, имеющие значительные расхождения со средним, то есть атипичные. Их графики имеют вид, существенно отличающийся от графика среднего.

Далее приведены графики типичного (ул. Московское шоссе д. 302) и атипичного (ул. Народная, д. 38) очага в сравнении со средним (рис. 2).

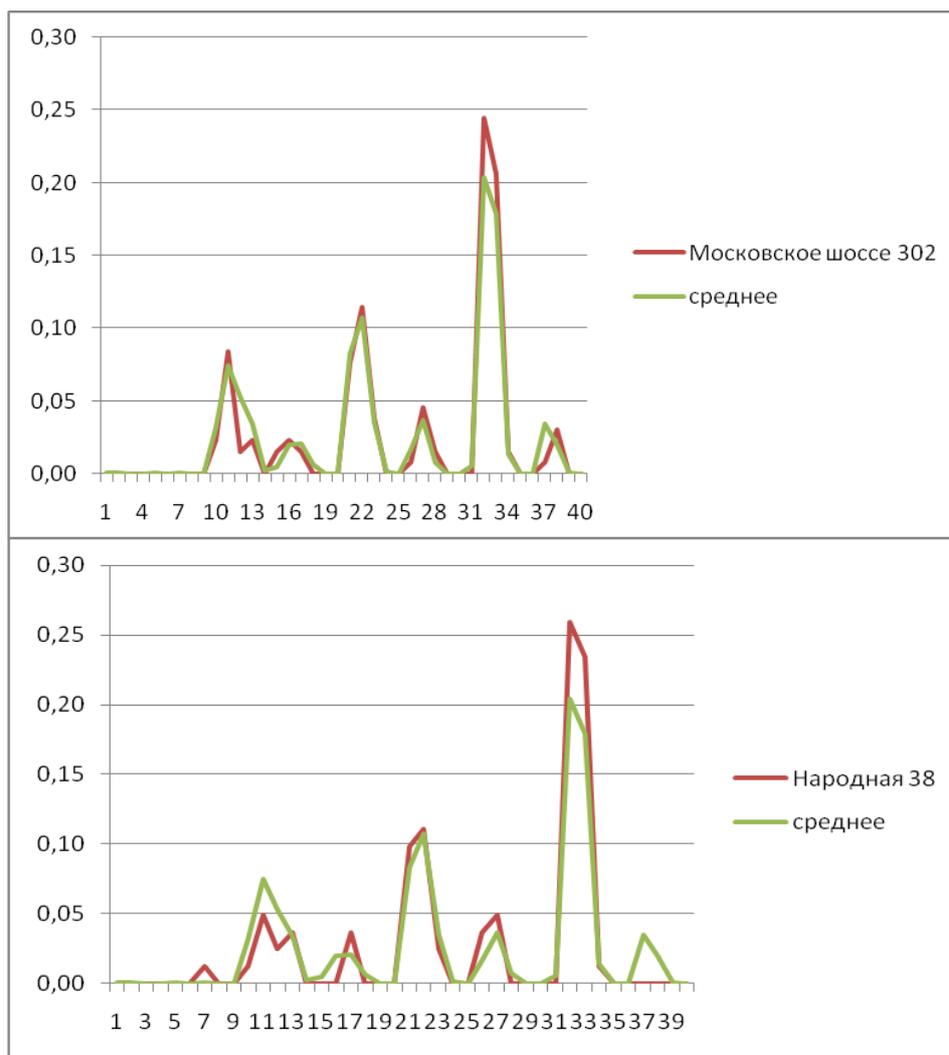


Рис. 2

Далее были посчитаны суммы квадратов отклонений очагов друг от друга: d_{ik} – сумма квадратов отклонений i -го очага от k -го. Анализ показал, что многие очаги имеют малые расхождения друг с другом (рис. 3). Соответствующие графики практически накладываются один на другой. Однако осталось неясным: какое отклонение d_{ik} считать достаточно малым.

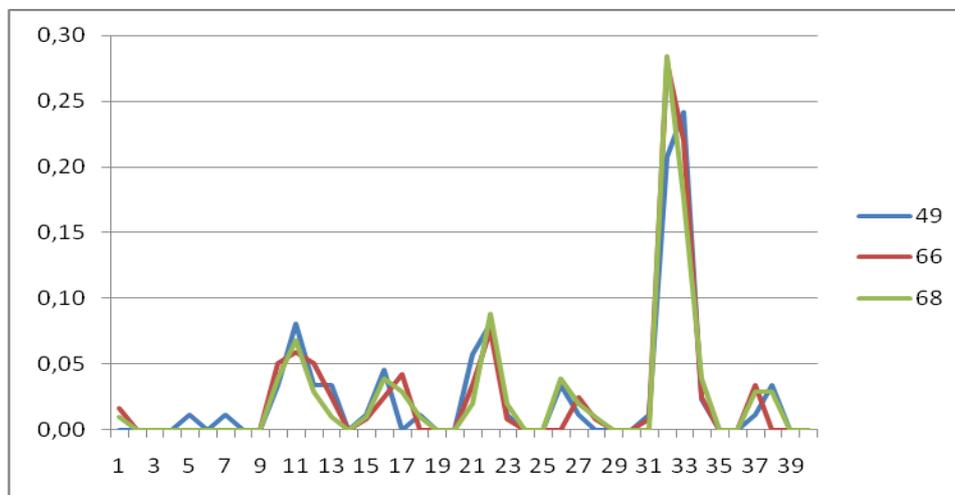


Рис. 3

Для выявления схожих очагов используем аппарат теории графов и стандартную терминологию. Вершинами графа будем считать очаги. Тогда множеству очагов могут быть сопоставлены два графа. Неориентированный: две вершины (i -я и j -я) связаны ребром, если $d_{ji} < d_{0i}$ и $d_{ji} < d_{0j}$. То есть маркеры i -го и j -го очагов отличаются друг от друга менее, чем каждый из них от среднего. И ориентированный: направленное ребро ставится при $d_{ji} < d_{0i}$, при этом необязательно $d_{ji} < d_{0j}$. То есть маркер i -го и ближе к маркеру j -го, чем к среднему.

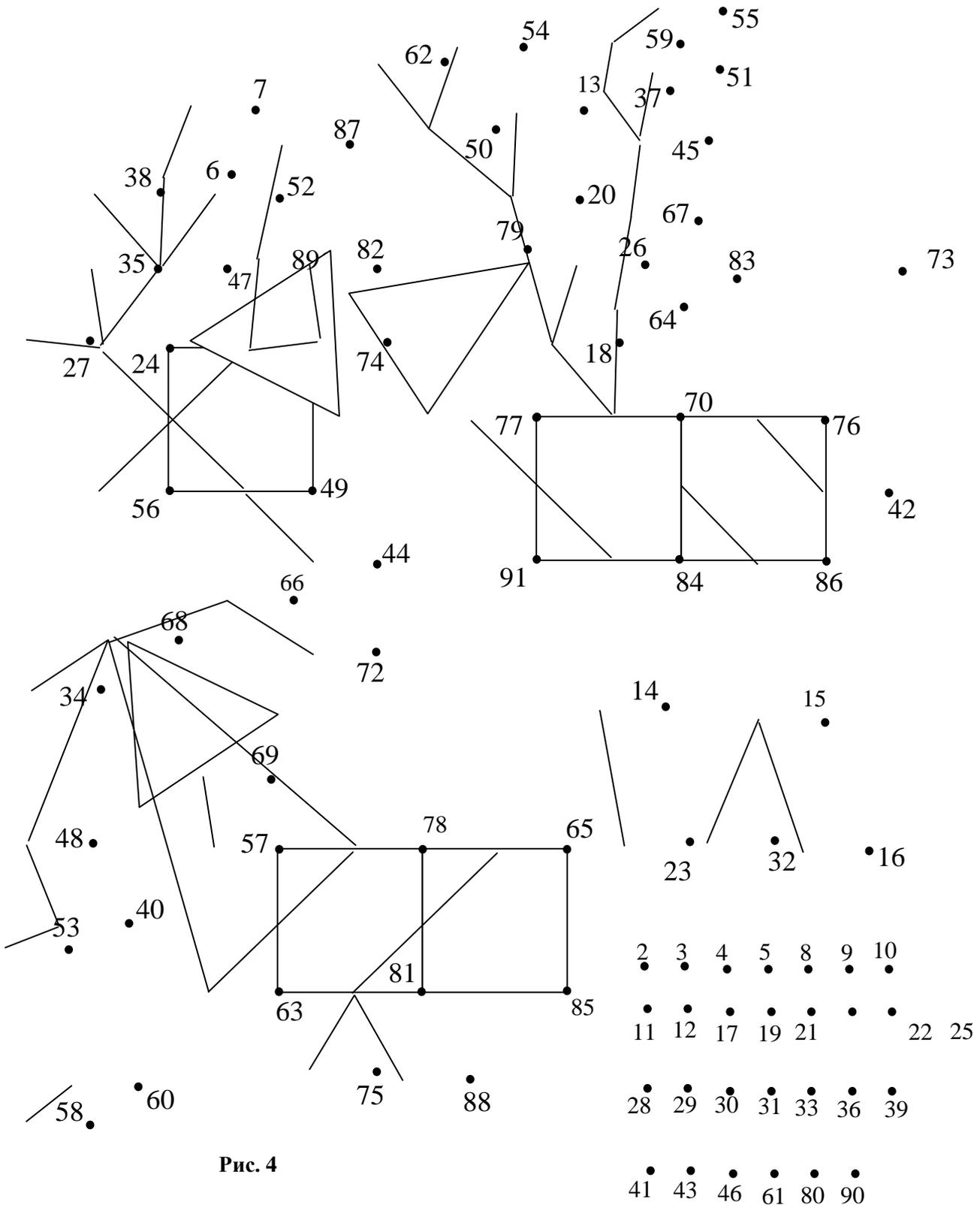


Рис. 4

На рис. 4 приведен неориентированный граф G связности очагов.

Введем в рассмотрение некоторые свойства графа связности, на основе которых можно осуществить деление на группы.

Граф G содержит 31 компоненту, из которых 27 являются изолированными вершинами.

Напомним, что кликой называется полный подграф данного графа. Клика максимальна, если она не является подграфом клики большего порядка, наибольшей, если в графе нет клики с большим числом вершин. В рассматриваемом графе есть клика порядка 4 с вершинами 24, 71, 49, 56 и 12 клик порядка 3, остальные клики имеют порядки 2 и 1.

Подсчитаем кликовое число φ данного графа - число вершин в наибольшей клике. Нетрудно установить, что $\varphi(G)=4$. Выделим доминирующее множество вершин графа - множество, удовлетворяющее условию: любая вершина графа смежна с некоторой вершиной этого множества. Число вершин в наименьшем доминирующем множестве называется числом доминирования и будет обозначаться γ_0 .

На наш взгляд перспективной является идея разделения очагов на группы на основе наименьшего доминирующего множества вершин графа G . Поскольку любая вершина либо входит в доминирующее множество, либо является смежной какой-то его вершине. Значит, будут охвачены все имеющиеся очаги.

В доминирующее множество должны войти вершины графа степени 1 (висячие) или смежные с ними. Чтобы найти наименьшее доминирующее множество, будем записывать вершины смежные с висячими, так как их степень может быть больше 1.

Найдено доминирующее множество $D = \{47, 6, 24, 89, 49, 50, 20, 18, 74, 76, 84, 59, 45, 70, 68, 66, 60, 40, 81, 57, 14, 15, 2, 3, 4, 5, 8, 9, 10, 11, 12, 17, 19, 21, 22, 25, 28, 29, 30, 31, 33, 36, 39, 41, 43, 46, 61, 80, 90\}$. Оно является наименьшим в виду приведенных ранее соображений. Во множество D входят 27 изолированных вершин и 22 неизолрированных. Число доминирования $\gamma_0(G)=49$.

Как видно, граф G имеет достаточно много изолированных вершин, которые:

- ближе к среднему, чем к любой другой вершине (2, 3, 4, 5, 8, 9, 11, 12, 17, 19, 21, 39);
- близки к какой-либо вершине, но она к ним не близка; такие вершины не будут изолированы в орграфе (10, 22, 25, 28, 29, 30, 31, 33, 36, 41, 43, 46, 58, 61, 80, 90).

То есть изолированные вершины можно записать в одну группу, либо присоединить к выделенным группам (на основании ребер орграфа).

Далее будут рассмотрены оба деления на группы.

1. Каждую вершину доминирующего множества (кроме изолированных) можно взять за типичного представителя группы. В группу войдут все смежные с ней вершины. Соответствующие очаги характеризуются меньшим отклонением от типичного представителя, чем от среднего. Группы будут следующего состава (типичный представитель выделен).

47, 6, 24, 38, 52;

6, 7;

24, 27, 35, 47, 49, 56, 71;

89, 71, 87;

49, 24, 44, 56, 71;

50, 20, 54, 62;

20, 13, 18;

18, 26;

74, 71, 77, 79, 82;

76, 42, 83, 73, 70;

84, 86, 91, 77, 70;

59, 55, 37;

45, 51, 67, 37;

70, 64, 84, 86, 77, 76, 18;

68, 34, 66, 78, 48, 63;

66, 72;

60, 58, 40, 63;
 40, 53;
 81, 88, 75, 65, 78, 63, 85;
 57, 69;
 14, 23;
 15, 16, 32.

Изолированные вершины ближе к среднему, поэтому составляют отдельную группу с типичным представителем – «средним» очагом. В данном разбиении 23 группы, причем некоторые очаги одновременно входят в несколько групп.

2. Будет отличаться тем, что изолированные вершины, принадлежащие орграфу, дополняются в группу, с типичным представителем которой они смежны. Добавленные вершины записаны в скобках.

47, 6, 24, 38, 52, (90);
 7, 6, (29), (43), (61), (90);
 24, 27, 35, 47, 49, 56, 71, (61);
 89, 71, 87;
 49, 24, 44, 56, 71;
 50, 20, 54, 62;
 20, 13, 18, (46);
 18, 26, (29), (31);
 74, 71, 77, 79, 82;
 76, 42, 83, 73, 70;
 84, 86, 91, 77, 70, (90);
 59, 55, 37;
 45, 51, 67, 37;
 70, 64, 84, 86, 77, 76, 18, (90);
 68, 34, 66, 78, 48, 63;
 66, 72;
 60, 58, 40, 63;
 40, 53, (58);
 81, 88, 75, 65, 78, 63, 85;
 57, 69;
 14, 23, (46), (90);
 15, 16, 32, (58), (80).

Возможно это разбиение более перспективно для дальнейшего анализа. Получается также 23 группы.

Таким образом, в процессе решения поставленной задачи деления очагов на группы:

- введено в рассмотрение среднее значение маркеров по всем очагам, характеризующих некий «средний» очаг или картину ДТП по всем очагам в целом;
- построена математическая модель с помощью формулы суммы квадратов отклонений от среднего $d_{oi} = \sum_{j=1}^{40} (x_j - x_{jcp})^2$;
- произведена сортировка очагов по возрастанию отклонения и таким образом выделены типичные и атипичные очаги;
- построены графики функций значений маркера, подтверждающие произведенные расчеты;
- построена математическая модель – графы связанности очагов (ориентированный и неориентированный);
- рассмотрены структура и свойства данных графов в контексте приложения их к решению задачи:
 - посчитано число компонент неориентированного графа $k=31$;

- найдены наибольшие клики и кликовое число $\varphi(G)=4$;
- найдено одно из наименьших доминирующих множеств и число доминирования $\gamma_0(G)=49$;

- найдено деление на 23 группы (с выделенным типичным представителем).

Не отвергая возможности иных разбиений очагов на классы или группы, следует отметить преимущества избранного подхода к решению задачи:

- полностью математически обоснован;
- решение наглядно представлено с помощью графа, позволяющего продолжать исследование с использованием богатого инструментария данной теории;
- для построения графа выбрано условие: ребром соединяются вершины более близкие к друг другу, чем к среднему; это условие действительно позволяет по выделенным типичным представителям судить о картине ДТП в других очагах группы;
- количество групп 23 является вполне обозримым, доступным для исследования.

В дальнейшем планируется детально проанализировать типичных представителей выявленных классов очагов и на практике проверить гипотезу о схожести причин ДТП очагов внутри каждого из выделенных классов.

Библиографический список

1. **Елисеев, М.Е.** О интерактивной карте аварийности крупного города / М.Е. Елисеев [и др.] // Сборник трудов Sworld по материалам международной научно-практической конференции. – 2011. Т. 3. №4. С. 40–45.
2. **Елисеев, М.Е.** О статистическом анализе очагов аварийности // Автотранспортное предприятие. 2012. №4. С. 37–38
3. **Елисеев, М. Е.** Подсистема анализа очагов интерактивной карты аварийности / М.Е. Елисеев [и др.] // Труды НГТУ им. Р.Е. Алексеева. 2012. №4. С. 358–362.

*Дата поступления
в редакцию 25.10.201*

М.Е. Eliseev¹, М.Е. Sangalova²

ON CLASSIFICATION OF FOCI OF TRAFFIC ACCIDENTS

Nizhny Novgorod state University Arzamassky (branch)¹,
Nizhny Novgorod state technical university n.a. R.E. Alexeev²

Purpose: In this paper the classification of foci of traffic accidents in order to better analyze and identify the causes of accidents is considered.

Design/methodology/approach: Standard approaches to the partition of foci of accidents on classes unsuitable. In the paper for this task used graph theory. Splitting into classes is in accordance with the data on the accident of different categories of drivers.

Findings: Technique allows find foci accidents of a similar nature. If there is information about an one traffic accident of a class automatically obtain data on the entire class.

Research limitations/implications: This approach can be used in the analyses of traffic accidents of any large city.

Originality/value: The research can be used to determine the causes of accidents and, ultimately, for improve road safety. In particular the results of the study will be used in developing the monitoring subsystem interactive map of the traffic accidents.

Key words: topographical analysis, traffic collision, TRC, traffic accident, TRA, GIS, Geographic Information System, an interactive map.