Classification Of Road Accidents Using Fuzzy Techniques

Anastasios Katsoukis  
Department of Civil Engineering  
Democritus University of Thrace  
Xanthi, Greece  
taskatsou@gmail.com

Lazaros Iliadis  
Department of Civil Engineering  
Democritus University of Thrace  
Xanthi, Greece  
liliadis@civil.duth.gr

Avrilia Konguetsof  
Department of Civil Engineering  
Democritus University of Thrace  
Xanthi, Greece  
akogkets@civil.duth.gr

Basil Papadopoulos  
Department of Civil Engineering  
Democritus University of Thrace  
Xanthi, Greece  
papadob@civil.duth.gr

Abstract—Nearly 1.3 million people die in road crashes annually in a global scale, on average 3,287 deaths per day [1]. Moreover, 20-50 million are injured or disabled [1]. Road traffic crashes rank as the 9th leading cause of deaths (2.2% of deaths globally) [1]. This research paper presents the contribution of soft computing (fuzzy logic) towards modeling of this huge problem. More specifically, the municipalities of Greece are classified according to their respective road accidents occurrence, by means of a special measure of similarity and consideration of the generating fuzzy transitive closure (FTC).

Keywords—fuzzy transitive closure, tree classification, road accidents

I. PRELIMINARIES

A. Literature Review And Innovation

So far (based in the literature) the problem has not been modeled in depth, although it has a severe social and financial impact in the societies globally. Some valuation reports have been published regarding the casualties in India and UK. European Union has funded the Trace research Program [2], eleven years ago, aiming to find the reasons and the factors affecting the road accidents in a European scale. Though the factors were classified, this research has an overall InterEuropean analysis. However, each specific country and each Region have different characteristics, which should be considered.

An interesting research towards road traffic accidents classification using Artificial Neural Networks, has been conducted in India [3] and another one towards severity prediction employing Decision Trees optimized by Genetic Algorithms, has been performed in Iran [4]. An algorithm for Maximum flow analysis in traffic networks, based on Fuzzy Matrix has been proposed by Zhao and Gu [5] however this research does not deal with accidents.

Fuzzy Logic has been employed in the literature quite frequently towards risk assessment [6].

The innovation of this research is the use of soft computing, Data Mining approaches and fuzzy algebraic tools (FTC) towards a more sophisticated road accidents modeling, in each prefecture of Greece. Obviously, this is a two fold innovation. At first level, it is a research that aims to model the problem beyond the typical descriptive statistics and it is the first one using fuzzy equivalence relations. At second level it is performed locally and at the same time overall. It is obvious that it offers a rational insight which can be used as a tool for the impose of specific actions on site (tailored for each area).

B. Fuzzy Equivalence Relations

Definition 1: Let $X$ be a set. Every function of the form $A : X \rightarrow [0,1]$ is called a fuzzy set or a fuzzy subset of $X$. The symbol $\mu : X \rightarrow [0,1]$ is also used, to denote the fuzzy set $A$. By $\mu(x)$ the membership degree of $x$ in the fuzzy set $A$ is denoted.

Definition 2: The $\alpha$-cut of the fuzzy set $A$ is defined as the crisp set:

$A^\alpha = \{x \in X : \mu(x) \geq \alpha\}.$

Definition 3: Let $X$ and $Y$ be two sets. Then,$R = \{(x,y), \mu_R(x,y) : (x,y) \in X \times Y\}$ is called as fuzzy relation on $X \times Y$.

Definition 4: Let $R \subset X \times Y$, $S \subset Y \times Z$ be two fuzzy relations. The max-min composition $R \circ S$ is defined as follows:

$R \circ S = \{((x,z), \max \{\min \{\mu_R(x,y), \mu_S(y,z)\} : x \in X, y \in Y, z \in Z\} ,

where max is considered for all $y$ and min is considered for all $(x,z)$.

Definition 5: A fuzzy relation $R \subset X \times X$ is called fuzzy equivalence relation on $X$ if it satisfies the following conditions:

Reflexive: $\mu_R(x,x) = 1, \forall x \in X$

Symmetrical: $\mu_R(x,y) = \mu_R(y,x), \forall x, y \in X$
Transitive:
\( \mu_{g}(x, z) \geq \max \{\min \{\mu_{g}(x, y), \mu_{g}(y, z)\}, \forall x, y, z \in X \} \),
where max is considered for all y and min is considered for all \((x, z)\).

**Proposition 1:** If \( R \subseteq X \times X \) is a fuzzy equivalence relation on \( X \), then the \( \alpha \)-cuts \( ^{\alpha}R \) constitute equivalence relations on \( X \). So, with the help of this equivalence relation, a decomposition (a partition in sets, which are per two disjoint) of \( X \) can be achieved. Let this partition be denoted by \( \pi(\alpha R) \). So, the tree classification is accomplished as follows: \( \Pi(R) = \{\pi(\alpha R) : 0 < \alpha \leq 1\} \).

**Proposition 2:** These partitions \( \Pi(R) = \{\pi(\alpha R) : 0 < \alpha \leq 1\} \) are nested in the sense that \( \pi(\alpha R) \) is a refinement of \( \pi(\beta R) \) if \( \alpha \geq \beta \).

**Definition 6:** The transitive closure of \( R_{T} \) of a fuzzy relation \( R \) is the smallest transitive fuzzy relation, which contains \( R \).

**C. Description of the Algorithm**

The algorithm steps calculating the measure of similarity are described below:

i. If \( a_{i}, a_{j} \) is the number of the accidents per 100,000 inhabitants from the prefectures \( P_{i}, P_{j} \) belonging to the set \( P \) of all prefectures in Greece respectively their similarity measure is defined as:

\[ l_{ij} = \min \left\{ \frac{a_{i}}{a_{j}}, \frac{a_{j}}{a_{i}} \right\}, 0 < l_{ij} \leq 1 \]

ii. Using the above similarity measure the following matrix is formed:

\[ R = \begin{bmatrix}
    l_{11} & l_{12} & \cdots & l_{1n} \\
    l_{21} & l_{22} & \cdots & l_{2n} \\
      \vdots & \vdots & \ddots & \vdots \\
    l_{n1} & l_{n2} & \cdots & l_{nn}
\end{bmatrix}, \]

where \( l_{kk} = 1 \), \( n \leq k \leq 1 \) and \( l_{ij} = l_{ji} \).

iii. Since \( R \) is a fuzzy reflexive and fuzzy symmetrical relation on the set \( P \) of all prefectures in Greece, the transitive closure of \( R \) is found following the well-known algorithm:

- \( R' = R \cup (R \circ R) \), where \( \circ \) is the max-min composition
- If \( R' \neq R \), make \( R = R' \) and go to 1
- Stop, \( R' = R_{T} \)

iv. The tree classification procedure is:

\[ \Pi(R_{T}) = \{\pi(\alpha R_{T}) : 0 < \alpha \leq 1\} \],

where \( R_{T} \) is the max-min transitive closure. Thus the \( \alpha \)-cuts of \( R_{T} \), \( ^{\alpha}R_{T} \), are equivalence relations, on the set \( P = \{P_{i} \mid P_{i} \text{ is prefecture in Greece}\} \) so we have a decomposition of this set and hence the tree classification is obtained [4]-[8], [10], [12], [13].

II. THE APPLICATION OF THE ABOVE IN A FUZZY CLASSIFICATION TREE OF ROAD ACCIDENTS (TOTAL AND FATAL) IN GREECE DURING THE YEAR 2016

The official data was extracted from Statistical Yearbook Of Greek Police 2016 [9], [14].

Based on the latest census of residents per prefecture (censuses 2001 and 2011) (Hellenic Statistical Authority) [14] and with respect to the population proportions in prefectures having two Police Departments, such as the prefecture of Evros (the police departments of Alexandroupoli and Orestiada), the prefecture of Aktolakaarnania (the Police Departments of Aetolia and Akarnania). A normalization of the total number of road accidents per 100,000 inhabitants has been carried out.

**TABLE I.**

**TABLE OF TOTAL AND FATAL ROAD ACCIDENTS (F) PER 100,000 INHABITANTS (2016)**

<table>
<thead>
<tr>
<th>DATA PER 100,000</th>
<th>TOTAL</th>
<th>FATAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. D. OF ATTICA</td>
<td>154.287</td>
<td>5.103</td>
</tr>
<tr>
<td>P. D. OF THESSALONIKI</td>
<td>193.414</td>
<td>4.348</td>
</tr>
<tr>
<td>P. D. OF ALEXANDROUPOLI</td>
<td>59.574</td>
<td>5.319</td>
</tr>
<tr>
<td>P. D. OF AITOLIA</td>
<td>91.343</td>
<td>7.797</td>
</tr>
<tr>
<td>P. D. OF AKARNANIA</td>
<td>43.072</td>
<td>5.198</td>
</tr>
<tr>
<td>P. D. OF DODEKANISOS A&amp;B</td>
<td>82.6</td>
<td>14.205</td>
</tr>
<tr>
<td>P. D. OF ARGOLIDA</td>
<td>79.417</td>
<td>8.509</td>
</tr>
<tr>
<td>P. D. OF ARKADIA</td>
<td>73.504</td>
<td>6.86</td>
</tr>
<tr>
<td>P. D. OF ARTA</td>
<td>17.917</td>
<td>7.679</td>
</tr>
<tr>
<td>P. D. OF ACHAIA</td>
<td>81.167</td>
<td>6.505</td>
</tr>
<tr>
<td>P. D. OF VOIOTIA</td>
<td>52.637</td>
<td>10.68</td>
</tr>
<tr>
<td>P. D. OF GREVENA</td>
<td>21.082</td>
<td>7.905</td>
</tr>
<tr>
<td>P. D. OF DRAMA</td>
<td>81.75</td>
<td>3.847</td>
</tr>
<tr>
<td>P. D. OF EVRIA</td>
<td>86.456</td>
<td>12.085</td>
</tr>
<tr>
<td>P. D. OF EVRITANIA</td>
<td>24.958</td>
<td>0</td>
</tr>
<tr>
<td>P. D. OF ZAKYNTHOS</td>
<td>102.524</td>
<td>12.815</td>
</tr>
<tr>
<td>P. D. OF ILIA</td>
<td>42.941</td>
<td>8.795</td>
</tr>
<tr>
<td>P. D. OF IMATHIA</td>
<td>48.044</td>
<td>9.051</td>
</tr>
</tbody>
</table>
In the total of road accidents, it may be seen that the first places respectively are:

- 1st: Traffic police of Thessaloniki with 193.414 per 100,000 inhabitants
- 2nd: Traffic police of Lefkada with 159.95 per 100,000 inhabitants
- 3rd: Traffic police of Attica with 154.287 per 100,000 inhabitants
- 4th: Traffic police of Cyclades with 150.068 per 100,000 inhabitants
- 5th: Traffic police of Corinthia with 139.046 per 100,000 inhabitants
- 6th: Traffic police of Chalkidiki with 116.307 per 100,000 inhabitants

Thus, it seems that Lefkada, the Cyclades islands and Chalkidiki, apart from Athens and Thessaloniki, due to the increased number of tourists, hold some of highest positions among prefectures considering the road accidents. And so is Corinthia, due to the Athens-Corinth network.

The respective catalogue considering fatal road accidents is formed as follows:

- 1st: Traffic police of Lefkada with 22.216 per 100,000 inhabitants
- 2nd: Traffic police of Chalkidiki with 17.16 per 100,000 inhabitants
- 3rd: Traffic police of Corinthia with 15.521 per 100,000 inhabitants
- 4th: Traffic police of Dodecanese’s with 14.205 per 100,000 inhabitants
- 5th: Traffic police of Zakynthos with 12.815 per 100,000 inhabitants
- 6th: Traffic police of Chania with 12.634 per 100,000 inhabitants

Also, it should be noted that the Attica and Thessaloniki Traffic Police Departments are in lower positions with 5.103 and 4.348 fatal accidents per 100,000 inhabitants. Due to the huge magnitude of the actual table (53X53) it is impossible to present it herein. So, to save space we have included a small indicative sample which is only a portion of the actual one.

The following Tables II and III present the Similarity Coefficients of the Total Road Accidents and the Similarity Coefficients of The Fatal Road Accidents, respectively.

<table>
<thead>
<tr>
<th>P. D. OF IRAKLO</th>
<th>26.667</th>
<th>8.889</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. D. OF THESPROTI</td>
<td>69.427</td>
<td>8.678</td>
</tr>
</tbody>
</table>

### TABLE II. PORTION OF THE 53X53 TABLE OF THE SIMILARITY COEFFICIENT OF THE TOTAL ROAD ACCIDENTS

<table>
<thead>
<tr>
<th>TOTAL</th>
<th>P.D OF ATTICA</th>
<th>P.D OF THESSALONIKI</th>
<th>P.D OF ALEXANDROUPOLI</th>
<th>P.D OF AITOLIA</th>
<th>P.D OF AKARNAIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>P.D OF ATTICA</td>
<td>1</td>
<td>0.798</td>
<td>0.386</td>
<td>0.592</td>
<td>0.279</td>
</tr>
<tr>
<td>P.D OF THESSALONIKI</td>
<td>0.798</td>
<td>1</td>
<td>0.308</td>
<td>0.472</td>
<td>0.223</td>
</tr>
<tr>
<td>P.D OF ALEXANDROUPOLI</td>
<td>0.386</td>
<td>0.308</td>
<td>1</td>
<td>0.652</td>
<td>0.723</td>
</tr>
<tr>
<td>P.D OF AITOLIA</td>
<td>0.592</td>
<td>0.472</td>
<td>0.652</td>
<td>1</td>
<td>0.472</td>
</tr>
<tr>
<td>P.D OF AKARNAIA</td>
<td>0.279</td>
<td>0.227</td>
<td>0.723</td>
<td>0.472</td>
<td>1</td>
</tr>
</tbody>
</table>

### TABLE III. PORTION OF THE 53X53 TABLE OF THE SIMILARITY COEFFICIENT OF THE FATAL ROAD ACCIDENTS

<table>
<thead>
<tr>
<th>FATAL</th>
<th>P.D OF ATTICA</th>
<th>P.D OF THESSALONIKI</th>
<th>P.D OF ALEXANDROUPOLI</th>
<th>P.D OF AITOLIA</th>
<th>P.D OF AKARNAIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>P.D OF ATTICA</td>
<td>1</td>
<td>0.852</td>
<td>0.959</td>
<td>0.654</td>
<td>0.982</td>
</tr>
<tr>
<td>P.D OF THESSALONIKI</td>
<td>0.852</td>
<td>1</td>
<td>0.817</td>
<td>0.558</td>
<td>0.836</td>
</tr>
<tr>
<td>P.D OF ALEXANDROUPOLI</td>
<td>0.959</td>
<td>0.817</td>
<td>1</td>
<td>0.682</td>
<td>0.977</td>
</tr>
</tbody>
</table>
So, the 53x53 similarity matrix $R$ is calculated, concerning the total accidents and consequently the transitive closure of $R$, $R_T$. The same holds for the fatal accidents and let be called $R’$ and $R_T’$ the corresponding relations. Finally, the above algorithm, implemented by means of MATLAB is used for the transitive closure of $R$, $R_T$ to be found [7]. A total of 17 different values are calculated and placed in the matrix $R_T$ namely: 0.76 0.83 0.84 0.87 0.88 0.89 0.90 0.91 0.92 0.93 0.94 0.95 0.96 0.97 0.98 0.99 1.

For the fatal accidents 14 different values are calculated and placed in the matrix $R_T$ namely: 0.72 0.77 0.88, 0.90, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 1.

So, for example, the decomposition of $P$ (the set of prefectures of Greece) in level α=0.96 is the following:

### A. Subgroups Of The Partitions Of P (For $A = 0.96$) Related To The Total Accidents

The following values refer to the number of all road accidents in each traffic police department in the territory, normalized per 100,000 inhabitants without any decimal places.

1. {THESSALONIKI=193}
2. {ATTICA=154, CYCLADES=150, LEFKADA=159}
3. {CORINTHIA=139}
4. {CHALKIDIKI=116}
5. {ZAKYNTHOS=102}
6. {AETOLIA=91, DODEKANISA=82, ARGOLIDA=79, ARKADIA=73, ACHAIA=81, DRAMA=81, EOVOIA=86, THESPROTIA=69, KAVALA=92, KERKYRA=68, KEFALONIA=81, KILKIS=70, LAKONIA=78, RODOS=87, SAMOS=75, FOKIDA=93, CHIOS=69}
7. {ALEXANDROUPOLI=59, MESSINIA=59, FTHIOTIDA=60}
8. {VOIOI=52, XANTHI=52, ORESTIADA=54}
9. {IMATHIA=48}
10. {AKARNANIA=43, ILEIA=42, KARDITSA=42, KASTORIA=43, LASITHI=39, LESVOS=44, PELLA=43, PREVEZA=45, SERRES=42, CHANIA=40}
11. {KOZANI=35}
12. {LARISSA=31}
13. {IOANNINA=28}
14. {IKRKLIO=26, PIERRIA=26, TRIKALA=26}
15. {EVRIANIA=24}
16. {GREVENA=21}
17. {ARTA=17, RETHIMNO=18}
18. {MAGNISIA=16}
19. {FLORINA=12}

### B. Subgroups Of The Partitions Of P (For $A = 0.96$) Related To The Fatal Accidents Are Created.

1. {LEFKADA=22.2}
2. {CHALKIDIKI=17}
3. {CORINTHIA=15.5}
4. {DODEKANISA=14.2}
5. {ZAKYNTHOS=12.8, CHANIA=12.6, EOVOIA=12, CYCLADES=11.5, PELLA=12.3, FOKIDA=12.4}
6. {VOIOI=10.6, SERRES=10.9}
7. {LESVOS=10, MESSINIA=10, PREVEZA=10, RETHIMNO=9.7}
8. {AETOLIA=7, ARGOLIDA=8, ARTA=7, GREVENA=7, ILEIA=8, IMATHIA=9, IKAHLIO=9, THESPROTIA=8, KAVALA=7, LASITHI=9, XANTHI=9, ORESTIADA=9, RODOS=8, SAMOS=9, FTHIOTIDA=8, FLORINA=7, CHIOS=7}
9. {ARKADIA=6.86}
10. {ACHAIA=6.5, IOANNINA=6.4}
11. {MAGNISIA=5.8}
12. {ATTICA=5.1, ALEXANDROUPOLI=5.3, AKARNANIA=5.2, KARDITSA=5.4, KEFALLONIA=5, LAKONIA=5, PIERRIA=5.4}
13. {THESSALONIKI=4, KILKIS=4, LARIssA=4}
14. {DRAKA=3.8, KASTORIA=3.7, KERKIRA=3.5, TRIKALA=3.6}
15. {KOZANI=2.5}
16. {EVRIANIA=0}

### III. Conclusions

From the normalized accidents’ values as they are presented in the above analysis, it is clearly shown that Attica and Thessaloniki, are the most risky areas, followed by Cyclades (islands), Corfu, Chalkidiki and Zakynthos. This is a very interesting result and it shows that though the last three areas are not major urban centers, they have a serious problem. Their common characteristic is that they have a huge touristic development to handle. Though their population is not big, they accept dozens of thousands of tourists in a weekly basis during the summer season. This changes their winter profile and it converts it to a similar one of the major urban cities. A future extension of this research should focus on the seasonal study of these areas. A fuzzy tree-classification of all prefectures $P$ in Greece has been achieved, according to the number of accidents occurred in them and depending on the degree of truth aving been chosen. This classification is based on the fact that for each degree of truth a classical equivalence relation can be achieved and therefore certain decomposition is produced. A
remarkable open theoretical question which occurs from the above is the following: Let in a set $P$ we define a similarity ratio between its two elements, based on a relation $T$ (in our case the number of total accidents) and via this relation is generated a fuzzy classification tree. Using this similarity ratio in a subset $F$ (in our case of fatal accidents) is generated another fuzzy classification tree. So which is the relation between those two trees?

REFERENCES

http://www.who.int/gho/road_safety/mortality/en/

[2] Trace EU Research Project No. 027763
http://www.trace-project.org/publication-archives/trace-wp3-d3-1-v3.pdf


