Decision Support for Road Safety: Development of Key Performance Indicators for Police Analysts

Katherina Meißner and Julia Rieck

Abstract In 2017, five out of 100,000 people were killed by road accidents in Europe. In order to reduce this number with appropriate measures, the police nowadays manually defines combinations of accident attributes (e. g., accidents on slippery road surfaces at night), which then form the basis for tracking the number of accidents over time. The aim of this paper is to combine the following data analysis approaches in order to detect interesting attribute combinations, also referred to as "itemsets", relevant for current and future observations. The resulting combinations are proposed to the police as new key performance indicators and can also be used directly for planning police measures to increase road safety. A four-stage decision support system is introduced that employs frequent itemset mining in the first stage. The temporal aspect of traffic accident data is illustrated by time series containing, for each itemset, the relative frequencies of accidents with the corresponding attribute combination. In the second step, the time series are grouped according to their shape by time series clustering and classification. In the third step, we determine

Katherina Meißner · Julia Rieck
University of Hildesheim, Institute of Economics and Information Systems
Samelsonplatz 1, 31141 Hildesheim, Germany
✉ meissner@bwl.uni-hildesheim.de
✉ rieck@bwl.uni-hildesheim.de

Archives of Data Science, Series A
(Online First)
KIT Scientific Publishing
Vol. 6, No. 2, 2020

DOI: 10.5445/KSP/1000098012/12
ISSN 2363-9881

the optimal forecasting method for each generated cluster of time series. Based on the prediction of future frequencies, we identify the most interesting attribute combinations in the last step. These are displayed geographically so that a police analyst can easily identify current and developing hot spots.

1 Introduction

Every year, more than 1.35 million people worldwide are killed in road accidents. For children and young adults, injuries on roads are the leading cause of death. Therefore, the World Health Organization encourages countries to participate in a "Decade of Action" to reduce the number of (fatal) road accidents significantly (cf. WHO, 2018). In order to be able to define the field of action, the major accident causes, which may change over time and differ from location to location, have to be identified. Based on the accident circumstances (e. g., number of vehicles involved, wind conditions etc.), the police is able to derive measures to enhance road safety, e. g., adjusting patrol routes, new speed limit reductions, stop signs or investments in walking and cycling infrastructure.

Nowadays, police analysts perform the analysis of accidents manually, where major circumstances are predefined based on experience. Those circumstances (given as *attribute combination*s with their values, e. g., number of casualties = 1, surface condition = ice or snow) are used as *key performance indicators* (KPIs) to track the frequencies of the corresponding accidents. Please note that the predefinition can easily lead to interesting attribute combinations not being observed by the police and changes in frequencies will thus remain undetected. Hence, the aim of the paper is to uncover "interesting" combinations that either should be addressed by police measures or at least be added to the number of tracked KPIs. As soon as a KPI behaves atypically over time (e. g., shows an unexpectedly high value), the accidents with the corresponding combination of attributes are observed. A geographical visualization can then be performed in a decision support system (DSS) to identify current hot spots and hot spots that will become relevant in the future.

Publications in the context of data mining and road accident analysis mainly focus on the identification of patterns within the high dimensional road accident data. However, most of the following authors neglect the temporal aspect of the data. Supervised *classification* approaches are often applied to identify attribute combinations with specific values (i. e., features) that lead, in particular, to serious accidents. In order to explain the severity level of specific features, neural networks (Abdelwahab and Abdel-Aty, 2001; Delen et al., 2006), logistic regression approaches (Al-Ghamdi, 2002), Bayesian networks (De Oña et al., 2010) or decision trees (De Oña et al., 2013a; Sowmya and Ponmuthuramalingam, 2013) can be used. With classification according to severity, we cannot make any conclusions about the extent to which arbitrary attributes occur together. Unsupervised *clustering* approaches overcome this drawback and are used to identify accidents with similar features. Geurts et al. (2003b), Depaire et al. (2008) and De Oña et al. (2013b) present clustering methods for road accident data using several similarity measures. In addition to pure classification or clustering approaches, different data mining techniques can be combined to gain new insights. Starting with clustering of attributes and further applying classification (Sohn and Lee, 2003) or *association rule learning* (Prati et al., 2017; Janani and Devi, 2018; Kumar et al., 2017) on each cluster, the results are improved by reducing the heterogeneity of the data. With association rule learning as an unsupervised method, candidates for causal relations between features can be extracted. By combining geographical clustering with association rule learning, Geurts et al. (2003a) evaluate the causal relations of accidents in high frequency locations in contrast to those in other locations.

Bergel-Hayat and Zukowska (2015) provide an overview of statistical approaches to study time series related to road safety in European countries. The authors mention that in most cases only a few attributes are considered. The temporal aspect of road accidents in combination with data mining approaches is first introduced by Kumar and Toshniwal (2015). The authors cluster accidents by their features and perform a trend analysis on the monthly number of accidents. Moreover, Kumar and Toshniwal (2016) cluster time series of accident frequencies in 39 different regions in India by the similarity of their trends (i. e., industrial areas show similar numbers of accidents over time), without taking into account the different accident features.

The contribution of this paper is to find and forecast ex ante (and not ex post) “interesting” attribute combinations by combining data analysis methods sequentially. The resulting interesting attribute combinations can then be added to the police KPIs. In contrast to the methods in the publications described above, our methods are suitable for very large data volumes resulting from a wide geographical range and a detailed definition of attributes. The procedure can be embedded into a decision support system which is then able to help police analysts to identify hot spots varying over time. After the presentation of the entire DSS in Section 2, conclusions are made in Section 3.

2 Decision Support System

In order to support the police in planning their road safety measures in a targeted manner, we propose a system that analyses the extensive data available and identifies frequently occurring accident features. The framework of the decision support system is shown in Figure 1.

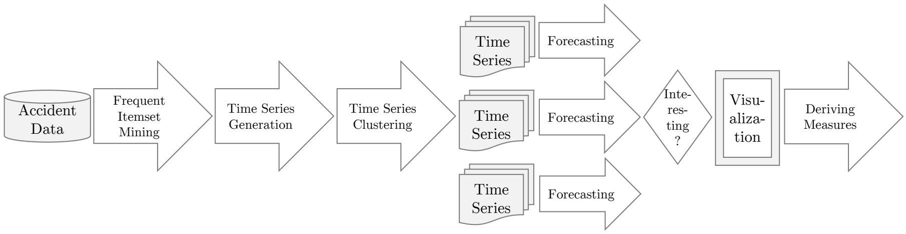

Figure 1: Decision Support System.

Starting from preprocessed data (cf. Subsection 2.1), we apply *frequent itemset mining* to extract frequent attribute combinations and generate time series of the frequency values for each itemset (cf. Subsection 2.2). The heterogeneity of the time series is reduced by time series clustering (cf. Subsection 2.3). In our analysis, only promising configurations are used in the time series clustering step. Promising configurations were determined by an approach proposed by Meißner and Rieck (2019). For each cluster, we describe approaches to identify suitable forecasting methods (cf. Subsection 2.4). A visualization is performed for the most “interesting” time series, (cf. Subsection 2.5), where hot spots are shown in detail (cf. Subsection 2.6).

2.1 Data Preprocessing

For our analysis, we consider the freely available road accident data set provided by the British Department for Transport (Department for Transport, 2019). The accidents include at least one casualty with slight personal injuries and are characterized by several features. Please note that Moradkhani et al. (2014) also use the accident data of Great Britain, where association rule learning on the data is performed without considering temporal aspects. In order to keep the number of combinations to be considered for testing our approach low at first and to make a practicable selection at the same time, we initially refrain from using personal (e. g., female, male) or vehicle-related (e. g., car, motorcycle) features and focus on external circumstances. Table 1 summarizes the selected attributes and attribute values.

Table 1: Attributes and corresponding values.

Attribute	**Attribute value**
Number of vehicles	1; 2; more than 2
Number of casualties	1; 2; more than 2
Accident severity	Crucial (fatal or serious); Slight
Day	Weekday; Weekend
Time	Morning (05:00–09:59); Day (10:00–14:59); Evening (15:00–20:59); Night (21:00–04:59)
Light condition	Darkness; Daylight
Weather condition	Fine; Rain; Other
Wind condition	No wind; Wind; Other
Surface condition	Dry; Frost, Ice or snow; Wet or damp
Special conditions (e. g., roadworks)	Yes; No
Carriageway hazard (e. g., prev. accident)	Yes; No
Road type	Dual carriageway; One way street; Roundabout; Single carriageway; Other
Urban or rural	Urban; Rural
Speed limit	< 30; 30; $30 - 60$; 60; > 60
Junction control	Controlled; Give way or uncontrolled
Junction detail	Crossroads; Roundabout; T or staggered junction; Other junction
Crossing facility (pedestrian)	Pedestrian phase at traffic signal junction; Pedestrian light crossing; Other
1st road class	A; B; C; Motorway; Unclassified
2nd road class	A; B; C; Motorway; Unclassified

In general, measures to increase road safety are applied in particularly defined regions (areas of affiliation). Our analyses, therefore, focus on a single geographical area, covering over approximately 51,000 accidents in Scotland (in the northern part of Great Britain) for the years 2012 to 2017. Scotland's geography is very diverse. In addition to urban areas, where many accidents tend to occur, there are vast rural regions with less traffic and, therefore, fewer accidents. This diversity allows to transfer the results from the Scottish data set to many other areas with similar geographical structure (e. g. Wales or the Midlands). Furthermore, partial results, e. g., for urban areas, can also be transferred to London or Manchester.

2.2 Frequent Itemset Mining and Time Series Generation

In order to detect frequent patterns (i. e., feature combinations occurring frequently), we apply frequent itemset mining to the preprocessed data, since we are only interested in the co-occurrence of features and not in their causal relationship. For the mining, the Eclat algorithm (cf. Goethals, 2010) is applied to the six years data set in order to generate itemsets (features or feature combinations) $I = \{i_1, \ldots, i_k\}$ with $k = 1, \ldots, n$ elements. To restrict the algorithm, we set the minimum support value to $\mathrm{supp}_{\mathrm{min}} = 0.02 = 2\,\%$ as a threshold. In this way, we significantly reduce the number of frequent itemsets from about 72 million to 50,600 and return only those itemsets that are present in at least $2\,\%$ of all 51,000 Scottish accidents. With an additional lower bound on the *all-confidence* value (cf. Equation 1), we obtain the itemsets that occur together with a certain probability, similar to Omiecinski (2003). Please note that we cannot use the regular *confidence* measure, since we do not analyze association rules. However, the all-confidence can estimate a lower confidence bound for all association rules that could be generated from a certain itemset.

$$\mathrm{conf}_{\mathrm{all}} = \frac{\mathrm{supp}(\{i_1, \ldots, i_n\})}{\max\limits_{k=1,\ldots,n} \{\mathrm{supp}(i_k)\}} = 0.1 \tag{1}$$

This further reduces the number of itemsets to 5,705, which will be used in what follows. The maximum number of features within an itemset is now

$n = 9$, while the number of itemsets with four or five features is largest, i. e., $|\{i_1, \ldots, i_4\}| = 1{,}563$ or $|\{i_1, \ldots, i_5\}| = 1{,}496$.

After frequent itemset mining, we generate a *time series* for each itemset in order to take the temporal aspect into account (cf. Liu et al., 2001; Böttcher et al., 2009). For this, we use the *relative frequency* of the itemset within the monthly accident data to obtain the ratio between the number of accidents with and without the itemset. A time series X_I for itemset I always consists of $T = 72$ points in time. A frequency value $x_t, t = 1, \ldots, 72$, lies between 0, if I is not present in the monthly data, and 1, if all monthly accidents contain the itemset ($0 \leq x_t \leq 1$).

2.3 Time Series Clustering

To find itemsets that should be considered by police analysts as KPIs, the current and future trends of the corresponding time series must be estimated. For this purpose, we first perform time series clustering to group time series with similar shapes and then identify a suitable forecasting method for each time series automatically.

A comprehensive overview of *time series clustering* is given by Aghabozorgi et al. (2015). Paparrizos and Gravano (2017) compare different clustering methods as well as similarity measures and introduce advanced approaches for both. The publications point out that several options to adjust the clustering of time series data can be taken into account, in particular, *clustering method* (e. g., partitional, hierarchical with single or complete linkage), *distance measure* (e. g., Euclidean, dynamic time warping), *scaling* (e. g., amplitude scaling, normalizing, centering) or *number of clusters* (e. g., $2, \ldots, 8$ clusters). In order to identify the configuration (a certain combination of options) that produces the clearest clustering results (i. e., similar time series shapes are assigned to the same cluster), we perform several analysis steps. Please note that visualizations and performance indices can be used to compare the quality of different configurations with respect to the resulting clusters. Performance indices such as C-Index, Calinski-Harabasz, and Gamma are usually based on compactness and/or separation which refer to intra- and inter-cluster distances (see for an overview Charrad et al., 2014).

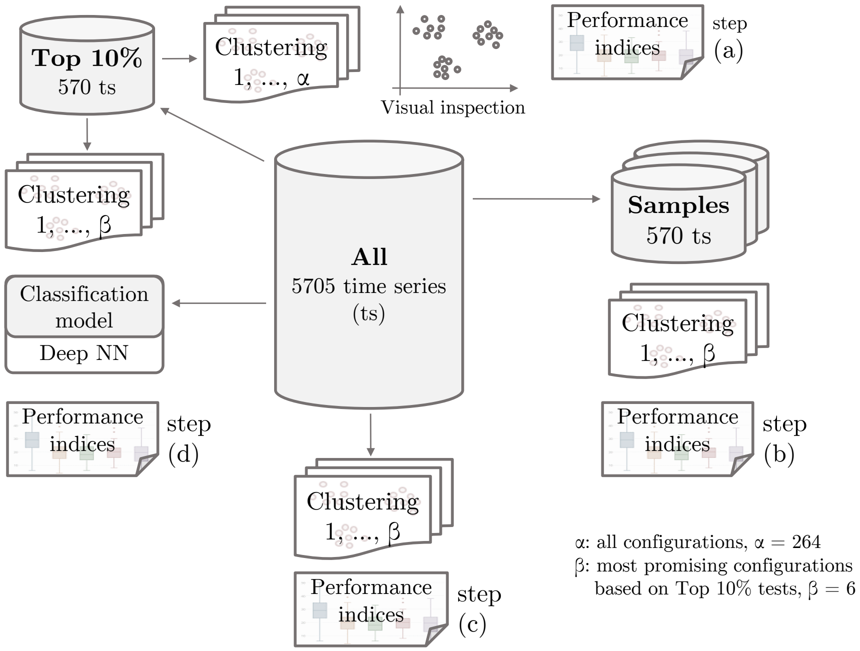

Figure 2: Analysis steps to determine the best clustering approach.

Figure 2 illustrates our analysis steps. In step (a), we select the "Top 10 %" of all underlying 5,705 time series regarding the overall support values. All α configurations (i. e., 264 configurations, cf. Meißner and Rieck, 2019) are processed with the "Top 10 %" data. The clustering results are visually examined in order to get an impression of the quality of each configuration. Moreover, we compare all performance indices with the visualization in order to select the indices that best reflect our data structure (these were C-Index and Gamma). For the remaining analysis steps, the six most promising configurations (i. e., best index values) are considered, which are, in particular, hierarchical clustering with complete/ward linkage, shape-based k-means, Euclidean distance, scaled/normalized data and 4 clusters. In step (b), we validate the results by applying the β configurations to 25 random data samples. In this way, it can be determined that clustering with the selected configurations is still suitable for time series with a lower support value. The evaluations until now are based on a subset of time series. Since we need to cluster all time series, we then apply the configurations to the complete data set in step (c). However, it should be noted that the performance indices did not show very good results for the complete data set. Therefore, we decided to use the very well clustered “Top 10 %” sample as training set for a time series classification model based on a *deep feed-forward neural network*. This model can then be applied to all other time series to classify them accordingly in step (d). As expected, the procedure in step (d) produces better results than the procedure in step (c) with respect to clustering quality, execution time, and memory requirements.

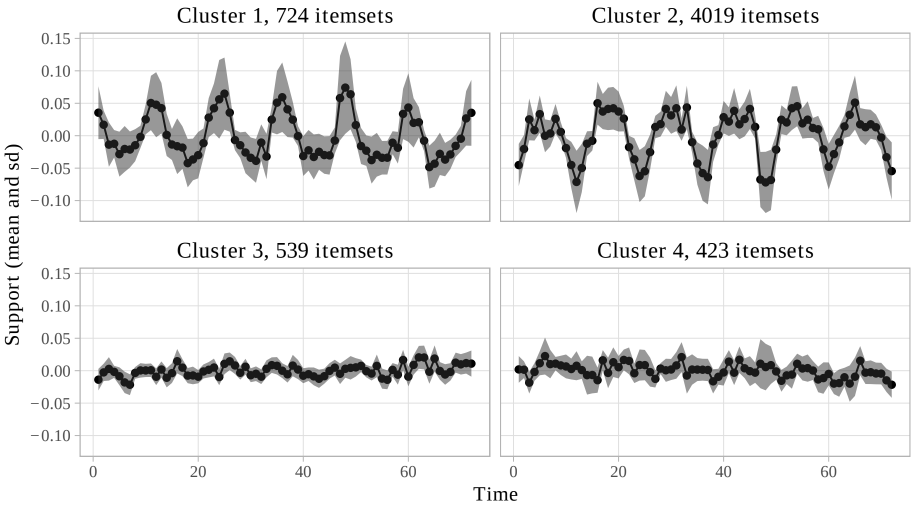

Figure 3: Results of time series clustering and classification with the following configuration: scaled data, Euclidean distance, hierarchical clustering with complete linkage, four clusters.

In order to visualize the final clustering and classification results, we plot all (centered) time series of each of the four determined clusters. Figure 3 shows the results, where the cluster means are given by a black line and the standard deviation is represented by grey areas. Two clusters have a seasonal pattern, with the respective maxima in cluster 1 lying in summer, while the peaks in cluster 2 occur in winter. Cluster 3 contains time series with a slightly increasing trend, while the time series in cluster 4 show no trend and seem to vary more over time.

2.4 Forecasting

Once we have separated the time series by their shape, we determine a suitable *forecasting* method for each cluster. Since our goal in clustering is to group the time series with similar patterns, and these similar patterns can exist at different support levels, centering is necessary before establishing the mean time series (cf. Figure 3). Based on the R-package “forecast” by Hyndman and Khandakar (2008), we train and test several forecasting methods (descriptions of all methods can be found in Hyndman and Athanasopoulos (2019), the abbreviations given in brackets are used later on):

- Simple methods:
  Naïve (naive) and seasonal naïve (snaive) method, average method (mean), random walk with drift (rw)
- Statistical methods:
  Seasonal and Trend decomposition using Loess (stl), exponential smoothing (ets), AutoRegressive Integrated Moving Average (arima), Trigonometric Box-Cox transform, ARMA errors, Trend, and Seasonal components (tbats)
- Advanced techniques:
  Neural network (nn), combination of arima, ets, nn, stl, and ets (combined) where the mean of all five forecasts is calculated as in Hyndman and Athanasopoulos (2019, Sec. 12.4)

Initially, we identify the most promising forecasting method for each cluster by taking the first 66 months of the mean time series of each cluster as training set for those methods that either need training (i. e., neural network) or need to be parameterized like, e. g., exponential smoothing and arima. Please note that we use Hyndman’s implementations in the “forecast” package to estimate all parameters (e. g., order of the moving average model, seasonality factor). In order to predict the relative frequencies $\hat{x}_t, t = 67, \ldots, 72$, of the mean time series, all aforementioned methods are applied and the predicted values are used to check the forecasting accuracy. The forecasting error is determined by the root mean square error

$$\mathrm{RMSE} = \sqrt{\frac{1}{6} \sum_{t=T-6+1}^{T} (\hat{x}_t - x_t)^2} \tag{2}$$

which is based on the deviation between the predicted values $\hat{x}_t$ and the observed values x_t. As Figure 4 shows, the forecasting methods perform differently on the individual cluster's mean time series. In order to finally determine the appropriate forecasting method for each cluster, we select the combined method as well as the two methods with the least test error (taking at most one simple method into account). Consequently, for cluster 1 snaive and arima, for cluster 2 snaive and tbats, for cluster 3 mean and tbats as well as for cluster 4 mean and neural network are considered. Then, the selected methods are applied to all time series of the individual clusters to identify the best performing method. This procedure is time consuming (approx. 1 hour for clusters 1, 3, 4 and 7 hours for cluster 2), even for our relatively small set of 5,705 time series. However, it should be noticed that the procedure does not need to be executed again. Only one pass to set up the system for the police is necessary.

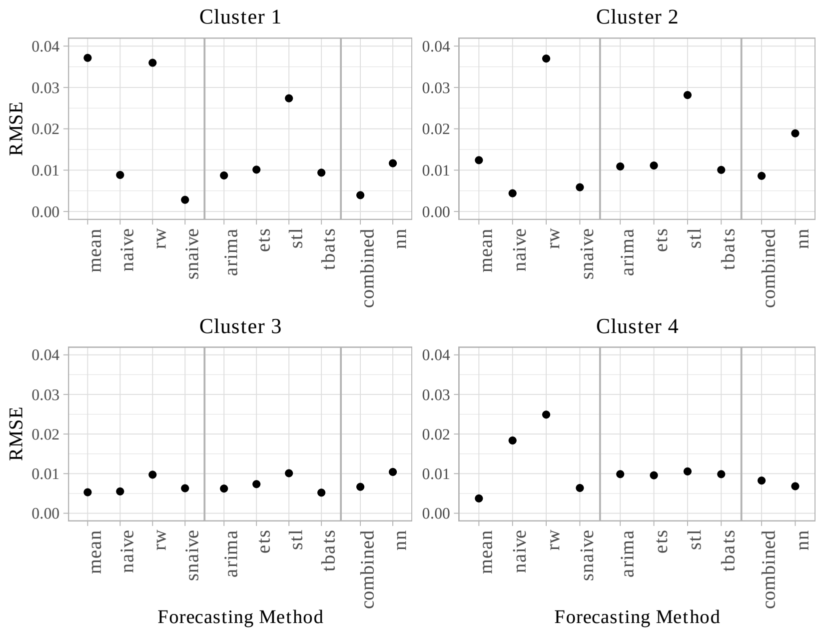

Figure 4: Forecasting error (RMSE) for the test data (6 months) for each cluster's mean time series.

Figure 5 visualizes the RMSE results for all time series in the clusters. We present a boxplot for each method that indicates the median error value of all time series by the line in the middle of the box, while the box height shows the interquartile range (i. e., range between the 1st and 3rd quartile). A low box height signals that many values are found in the area of the median. Hence, there are minor error fluctuations and the results of the respective forecasting method can be regarded as "stable". Particularly for the combined method, this is the case in all clusters. If we only consider the median value then seasonal naïve is preferable for clusters 1 and 2. This observation is plausible, as the time series of clusters 1 and 2 are seasonal (cf. Figure 3) and thus a method that reflects seasonality has been chosen. However, tbats (which also has a seasonal component) performs well for cluster 2 and should be preferred to snaive, taking into account the height of the box. For clusters 3 and 4, the average method is best when considering the median error value. Since the boxes of the average method in the boxplot are relatively large, tbats would also be a good choice for cluster 3 and the neural network for cluster 4. Please note that the neural network would need more training and tuning to achieve good results. Due to the short time series and the resulting small training sets (66 points in time for training), the results of the neural network are not as good as expected.

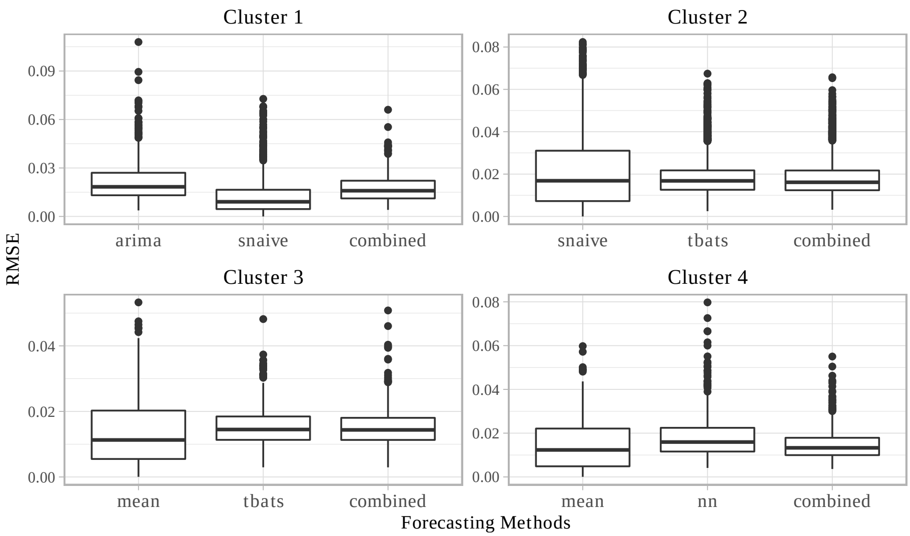

Figure 5: Forecasting error (RMSE) for selected methods based on all time series within each cluster.

2.5 Determining Interesting Itemsets

Using the above mentioned forecasting methods, we predict the future relative frequency values for all time series. Then, the most "interesting" ones have to be selected in order to visualize and present them to police analysts. We choose a forecast horizon of $h = 12$ months to obtain a reasonable time frame for the implementation of road safety measures. To identify the itemsets and corresponding time series that are "interesting", we cumulate three indicators (Equation 3 – 4) describing different aspects of "interestingness" to a (weighted) *single indicator*. The larger the resulting indicator, the more interesting the time series. A time series that is assigned a value close to 1 using Equation (3)

$$\frac{1}{h}\left|\{t \mid \hat{x}_t > \epsilon, t = T+1, \ldots, T+h\}\right| \tag{3}$$

typically obtains many prediction values (i. e. the predicted relative frequency) above a certain threshold ϵ. Such a time series should be considered as an additional KPI, since the corresponding itemset occurs in a large number of accidents. A variation between the previous year's values and the forecasted values can be detected using Equation (4)

$$\frac{1}{h}\sum_{t=T+1}^{T+h}\left|x_{t-12} - \hat{x}_t\right| \ . \tag{4}$$

The formula reveals whether the general course of the time series has changed between the last available period and the same period in the forecast horizon. Time series with strong fluctuations in the forecast values should be taken into account, if they receive a value close to 1 with Equation (5)

$$\frac{1}{h}\sum_{t=T+1}^{T+h}\left|\hat{x}_t - \left(\frac{1}{h}\sum_{t=T+1}^{T+h}\hat{x}_t\right)\right| \cdot \frac{\mathrm{var}(X-S)}{\mathrm{var}(X)} \ . \tag{5}$$

The first term relates the forecasted values to the mean values of the forecast horizon. Please note that a time series with a regular seasonal curve is assigned a lower value than a time series with an irregular curve. This is ensured by the last term, where the information about seasonality is calculated by the ratio of the variance of the time series without and with seasonality. We describe the forecasted time series with X and the seasonal component with S (cf. Wang et al., 2006).

2.6 Visualization and Deriving Measures

The items (i. e., attribute values) of the time series with the highest indicator values are stored and observed by the police as new KPIs. Moreover, accidents featuring such an interesting itemset can be visualized on a map (as shown in Figure 6).

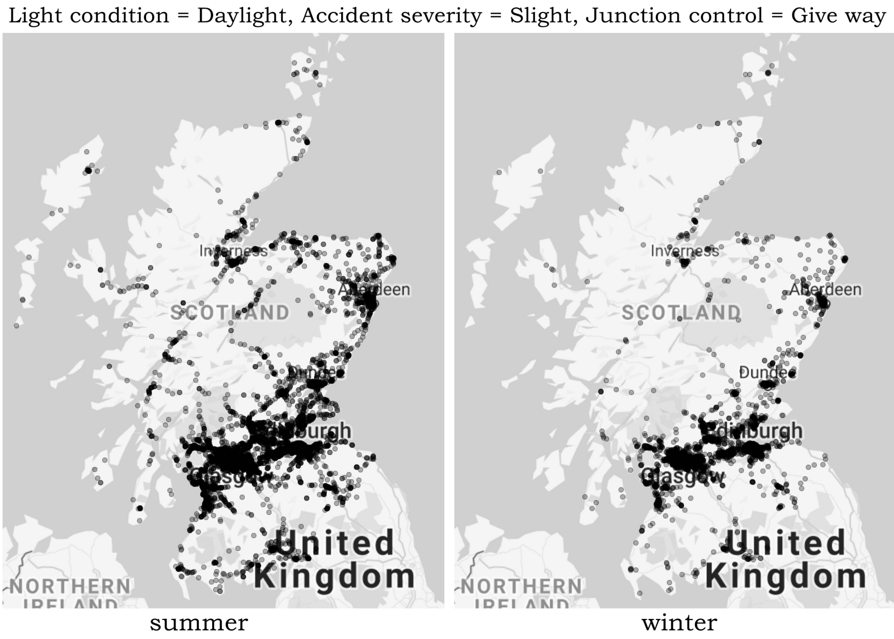

Figure 6: Visualization of a sample itemset for police analysts.

The geographical representation makes it possible to reconstruct the spatial distribution of accidents. In our example, accidents of slight severity are plotted, which additionally occurred during daylight hours at junctions with control type “give way”, as this itemset showed the highest indicator value in the previous step. Police analysts can now easily identify hot spots for the corresponding combination of attributes. Since this itemset belongs to cluster 2 (cf. Figure 3), we decide to visualize the hot spots for summer and winter separately. By comparing the hot spots in summer and winter, more accidents can be detected in summer, especially in rural areas. During this time of the year, many tourists travel to the Scottish Highlands and hence increase the traffic density in this area. In addition to that, tourists are not familiar with the traffic rules for left-hand drive and, therefore, make mistakes at junctions, which then lead to accidents.

3 Discussion

We have presented a decision support system to identify interesting attribute combinations for the police. Those combinations can either be added to the police KPIs or the police can use them to optimally plan their measures to increase road safety. The system consists of four stages, namely: frequent itemset mining, time series clustering, forecasting, and visualization. For each stage, we have identified the essential strategies for their refinement, based on the Scottish data set. To finalize the DSS and transform it into a fully automated system, it is necessary to have the KPI presentation and geographical visualization generated automatically. At present, the geographical representation is still prepared manually.

The methods presented in the individual stages need to be fine-tuned carefully. Therefore, we plan to verify our results using larger test scenarios. First, the system will be transferred to other geographical locations (e. g., Wales, the Midlands or London). This allows us to determine which stages of the system, in particular, need to be adapted during the transfer. When the system is deployed to an area with other geographical conditions, such as London, adjustments can be expected solely on the basis of accident frequency. Additionally, further features (e.g. personal or vehicle-related features) can be included in the decision support system. This will enable us to draw conclusions about the stability of the procedures with regard to changes in content.

References

Abdelwahab HT, Abdel-Aty MA (2001) Development of Artificial Neural Network Models to Predict Driver Injury Severity in Traffic Accidents at Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board 1746(1):6–13. DOI: 10.3141/1746-02.

Aghabozorgi S, Seyed Shirkhorshidi A, Ying Wah T (2015) Time-series Clustering – A Decade Review. Information Systems 53:16–38. DOI: 10.1016/j.is.2015.04.007.

Al-Ghamdi AS (2002) Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. Accident Analysis & Prevention 34(6):729–741. DOI: 10.1016/S0001-4575(01)00073-2.

Bergel-Hayat R, Zukowska J (2015) Road Safety Trends at National Level in Europe: A Review of Time-series Analysis Performed during the Period 2000–12. Transport Reviews 35(5):650–671. DOI: 10.1080/01441647.2015.1030005.

Böttcher M, Spott M, Kruse R (2009) A Condensed Representation of Itemsets for Analyzing Their Evolution over Time. In: Buntine W, Grobelnik M, Mladenić D, Shawe-Taylor J (eds.), Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2009), Vol. 5781. Springer, Berlin, Heidelberg (Germany), pp. 163–178. DOI: 10.1007/978-3-642-04180-8_28.

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014) NbClust : An R Package for Determining the Relevant Number of Clusters in a Data Set. Journal of Statistical Software 61(6). DOI: 10.18637/jss.v061.i06.

De Oña J, López G, Abellán J (2013a) Extracting Decision Rules From Police Accident Reports Through Decision Trees. Accident Analysis & Prevention 50:1151–1160. DOI: 10.1016/j.aap.2012.09.006.

De Oña J, López G, Mujalli R, Calvo FJ (2013b) Analysis of Traffic Accidents on Rural Highways Using Latent Class Clustering and Bayesian Networks. Accident Analysis & Prevention 51:1–10. DOI: 10.1016/j.aap.2012.10.016.

De Oña J, Mujalli RO, Calvo FJ (2010) Analysis of Traffic Accident Injury Severity on Spanish Rural Highways Using Bayesian Networks. Accident Analysis & Prevention 43(1):402–411. DOI: 10.1016/j.aap.2010.09.010.

Delen D, Sharda R, Bessonov M (2006) Identifying Significant Predictors of Injury Severity in Traffic Accidents Using a Series of Artificial Neural Networks. Accident Analysis & Prevention 38(3):434–444. DOI: 10.1016/j.aap.2005.06.024.

Depaire B, Wets G, Vanhoof K (2008) Traffic Accident Segmentation by Means of Latent Class Clustering. Accident Analysis & Prevention 40(4):1257–1266. DOI: 10.1016/j.aap.2008.01.007.

Department for Transport (2019) Road Safety Data. URL: `https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data`. Publication under Open Government Licence, accessed May 10th 2019.

Geurts K, Wets G, Brijs K, Vanhoof K (2003a) Profiling High Frequency Accident Locations using Association Rules. Transportation Research Record: Journal of the Transportation Research Board 1840:123–130. DOI: 10.3141/1840-14.

Geurts K, Wets G, Tom B, Vanhoof K (2003b) Clustering and Profiling Traffic Roads by Means of Accident Data. Proceedings of the European Transportation Conference. URL: http://alpha.uhasselt.be/~brijs/pubs/ETC2003.pdf.

Goethals B (2010) Frequent Set Mining. In: Maimon O, Rokach L (eds.), Data Mining and Knowledge Discovery Handbook. Springer US, Boston (USA), chap. 16, pp. 321–338. DOI: 10.1007/978-0-387-09823-4_16.

Hyndman RJ, Athanasopoulos G (2019) Forecasting: Principles and Practice, 2nd edn. OTexts, Melbourne (Australia). URL: OTexts.com/fpp2.

Hyndman RJ, Khandakar Y (2008) Automatic Time Series Forecasting: The Forecast Package for R. Journal of Statistical Software 27(3). DOI: 10.18637/jss.v027.i03.

Janani G, Devi NR (2018) Road Traffic Accidents Analysis Using Data Mining Techniques. Journal of Information Technology and Applications 14(2):84–91. DOI: 10.7251/JIT1702084J.

Kumar S, Toshniwal D (2015) A Data Mining Framework to Analyze Road Accident Data. Journal of Big Data 2(1):26. DOI: 10.1186/s40537-015-0035-y.

Kumar S, Toshniwal D (2016) A Novel Framework to Analyze Road Accident Time Series Data. Journal of Big Data 3(1). DOI: 10.1186/s40537-016-0044-5.

Kumar S, Toshniwal D, Parida M (2017) A Comparative Analysis of Heterogeneity in Road Accident Data Using Data Mining Techniques. Evolving Systems 8(2):147–155. DOI: 10.1007/s12530-016-9165-5.

Liu B, Ma Y, Lee R (2001) Analyzing the Interestingness of Association Rules From the Temporal Dimension. In: Cercone N, Lin TY, Wu X (eds.), Proceedings of the 2001 IEEE International Conference on Data Mining, Institute of Electrical and Electronics Engineers, Piscataway (USA), pp. 377–384. DOI: 10.1109/ICDM.2001.989542.

Meißner K, Rieck J (2019) Data Mining Framework to Derive Measures for Road Safety. In: Perner P (ed.), 15th International Conference on Machine Learning and Data Mining (MLDM 2019), ibai publishing, Leipzig (Germany), Vol. 2, pp. 625–639. ISBN: 978-3-942952-63-7, URL: https://dblp.org/db/conf/mldm/mldm2019-2.html.

Moradkhani F, Ebrahimkhani S, Bigham B (2014) Road Accident Data Analysis: A Data Mining Approach. Indian Journal of Scientific Research 3(3):437–443. URL: https://www.researchgate.net/publication/262186265_ROAD_ACCIDENT_DATA_ANALYSIS_A_DATA_MINING_APPROACH.

Omiecinski E (2003) Alternative Interest Measures for Mining Associations in Databases. IEEE Transactions on Knowledge and Data Engineering 15(1):57–69. DOI: 10.1109/TKDE.2003.1161582.

Paparrizos J, Gravano L (2017) Fast and Accurate Time-Series Clustering. ACM Transactions on Database Systems 42(2):1–49. DOI: 10.1145/3044711.

Prati G, De Angelis M, Puchades VM, Fraboni F, Pietrantoni L (2017) Characteristics of Cyclist Crashes in Italy Using Latent Class Analysis and Association Rule Mining. PLoS ONE 12(2):e0171484. DOI: 10.1371/journal.pone.0171484.

Sohn SY, Lee SH (2003) Data Fusion, Ensemble and Clustering to Improve the Classification Accuracy for the Severity of Road Traffic Accidents in Korea. Safety Science 41(1). DOI: 10.1016/S0925-7535(01)00032-7.

Sowmya M, Ponmuthuramalingam P (2013) Analyzing the Road Traffic and Accidents with Classification Techniques. International Journal of Computer Trends and Technology (IJCTT) 5(4):183–188. URL: http://ijcttjournal.org/archives/ijctt-v5n4p135.

Wang X, Smith K, Hyndman R (2006) Characteristic-Based Clustering for Time Series Data. Data Mining and Knowledge Discovery 13(3):335–364. DOI: 10.1007/s10618-005-0039-x.

WHO (2018) Road Traffic Injuries. World Health Organization, Geneva (Switzerland). URL: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries. Accessed May 10th 2019.