Analysing the influence of the pandemic on crime patterns in North Macedonia*

Aleksandra Dedinec[1][0000−0002−2050−2951], Sonja Filiposka[1][0000−0003−0034−2855], and Anastas Mishev[1][0000−0001−7271−6655]

Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje, North Macedonia
{aleksandra.kanevche, sonja.filiposka, anastas.mishev}@finki.ukim.mk

Abstract. Reducing national crime rate is an extremely important, but also difficult problem. For solving it, it is necessary to discover patterns of its occurrence, the various factors that influence it and the connection between criminal actions, which can help forecast future events, especially violent crimes where the police should act immediately. However, some major events, such as the COVID-19 pandemic may significantly affect these models. Therefore, this paper focuses on analyzing the impact of the pandemic on the crime rates and patterns and the way the crime forecasting models are affected by these changes, using North Macedonia's crime records as case study. The results show significant change in the rate and types of crimes during the pandemic period, when compared to the pre-pandemic period. Due to these changes in the crime patters, the crime forecasting models are also different, in terms of their accuracy and in terms of the importance of the input features that are used for the prediction.

Keywords: crime pattern · crime forecasting · pandemic correlation

1 Introduction

The analysis of crime rates on the national or local level has always been considered as a valuable action that can define the trends in the society and gauge the changes in criminal behaviour as influenced by the social, economic, political events in the region of interest [1]. In the recent years this type of analysis empowered with additional forecasting that can help optimise future use of police resources is increasingly being done using machine learning techniques [2]. This advances in large historical data sets analysis has enabled national police forces to obtain meaningful insight in the recorded events that provide the opportunity for reduced response time by intelligently redistributing the available staff and resources in the areas with highest probability of upcoming crime events.

Taking into consideration that forecasting is based on patterns found in the past events, it is of great interest to analyse how this forecasting is influenced by

* Supported by Cost Action CA17124 - Dig4ASP

major events that produce great changes in the general societal behaviour. The recent COVID-19 pandemic is such an occurrence. Recent studies have shown that the pandemic has had a major influence on the newly emerging patterns in crime rates in many nations. However, the most interesting observation is that this influence does not follow the same rules everywhere. For instance, in the US, recent studies [3] show that the pandemic has influenced on considerable drops in the major categories of crimes, while in India the effect has been reversed and the pandemic has triggered a large increase in certain crimes in this country [4]. However, in both cases, deeper analysis show that there is a shift in the types of crimes committed, with the pandemic having a large effect on lesser demeanor and little to no effect on serious crime categories. Another question that is being risen is how this shift in criminal behaviour influences the forecasting models in the post-pandemic era? Is the shift going to be a transient behaviour or will it have a long lasting effect on crime rate?

The focus of this paper is to analyse the impact of the pandemic on the crime rates in the case study of North Macedonia aiming to compare the pre- and Covid time periods and find meaningful trends in the changes in criminal behaviour. For these purposes, using machine learning we have curated a data set that contains all police reported crime events in the period of 2011-2021 and applied classification and time series analysis and forecasting to extract and compare the past and current trends in crime rates, emphasizing the most important changes.

The rest of the paper is structured as follows: in section II we discuss the crime events data set generated from public police records, in section III we provide a general analysis of the data set in the pre- and pandemic period, while in section IV we discuss the impact of the pandemic on crime rates modeling and forecasting. Section V concludes the paper.

2 North Macedonia crime events data set

The Ministry of Internal Affairs of the Republic of North Macedonia publishes an e-bulletin that describes the daily criminal events in the country in natural native language ever since 2011. Using Natural language processing (NLP) on the information that can be extracted from the e-bulletin [5] an initial crime events data set [1] is created where events are classified into broad crime categories, region and geo-location (latitude and longitude). This classification is done using keyword tagging in combination with analysis of the text preceding and following the identified keyword groups. There are 6 general crime categories defined as armed offences, violent crimes, theft and burglary, drug-related crimes, documents/identification-related crimes and traffic accidents. A seventh category - other crimes - is a combination of all reported events that do not fall into these 6 categories including migrants smuggling, fires, cyber bullying, dog attacks and, lately, pandemic-related offences such as breaking curfew, large gatherings, not wearing protective masks and similar offences.

[1] http://crimemap.finki.ukim.mk/home/en

The initial data set has been carefully curated by analysing ambiguous records that do not contain enough information to be automatically classified correctly in order to clean any erroneous data and significantly increase the quality of the data set. In addition to the data, exact geo-location, city, crime category, and detailed description, Information about the municipality has also been added based on the geographical location in order to obtain a regional view of the information as it is perceived on the level of the local municipality government.

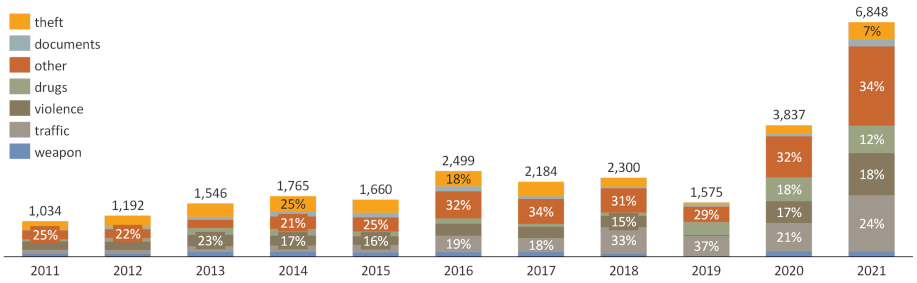

Fig. 1. Number of crimes per type and year

Based on the curated data set, in Fig. 1 the number of crimes per type and year is shown for the span of 11 years that are available in the data set. It is fairly easy to ascertain that in the years of the COVID-19 pandemic, 2020 and 2021, the crime has seen significant increase in the country almost doubling each year since 2019. In this paper we aim to analyse this information further in order to understand if this is the beginning of a serious trend in rising criminal behaviour and does the pandemic have any direct influence on the recorded events.

3 Crime data analysis

The available historical data on crime events in North Macedonia has been used as input to classification and time series analysis. While geographically speaking most of the events are continuously based in the municipalities that make up the general territory of the capital city that daily hosts over half of the country population, from the point of view of types or categories of crime, there seems to be a significant shift in the period before and during the COVID-19 pandemic. In Fig. 2 the relative change in the cumulative percentage for each identified crime category is represented and compared to the overall total. Please note that in this figure the time period before COVID-19 covers 8 years, from 2011 till 2019, while the during COVID-19 period is a 2 year period from 2020 till 2021. It can be seen that the pandemic has had drastic influence on different crime categories, which is very interesting when compared to the total number of reported crime events presented in Fig. 1 where crime is spiking during the pandemic.

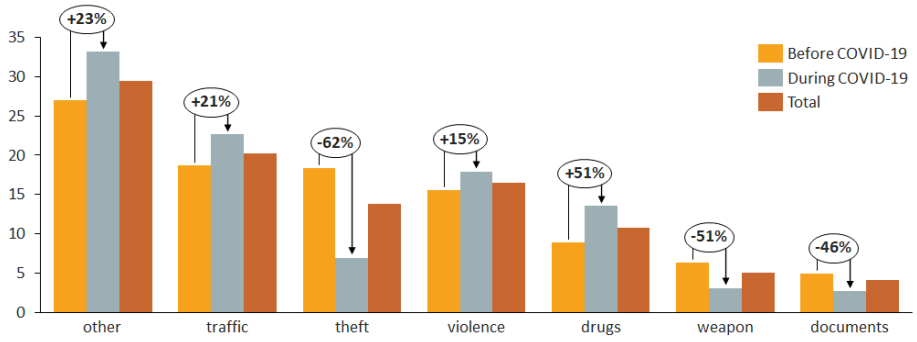

Fig. 2. Change in types of crime before and during the COVID-19 crisis

There is major drop in crimes related to documents, such as forgery of documents and false identification, (-46%) which is even a more significant drop when taking into account that part of these events are also pandemic related (i.e. forged certificates of vaccination). Another significant drop in crime is related to armed offences and weapons (-51%). The most significant drop is recorded for theft and burglary (-62%) and these are mostly related to the fact that travel has been very restricted in the period of the pandemic so opportunities for theft of homes have decreased significantly and most of the thefts reported are of stores and individuals on the street. This observed pandemic induced drop seems to be consistent with reports from other countries such as [7] when it comes to thefts, but the very opposite when considering armed offences where in [7] the authors report significant increase in homicides and gun assaults.

On the other hand, in North Macedonia the pandemic has managed to significantly increase the drugs related crimes (+51%) which is mostly due to the large increase in the number of small drug demeanour (possession, use) and small quantity dealers. Together with the increase in violence related crime (+15%) and traffic accidents and violations (+21%) it can be concluded that the lockdown and social separation has had a tremendous impact on the mental health on people increasing their anxiety and diminishing their tolerance. Again, this observed changes in trends during the pandemic are only partially consistent with other reports, where in drug crime rates are showing significant drops, while violent acts are rising [7].

The crime category that has seen a tremendous rise in the pandemic period is the "other" category that shows a relative increase of 23% compared to the pre-pandemic period. While only 15% of these events are strictly related to breaking the pandemic rules, the same period marks a high rise in number of fire events due to the hot, draughty weather. A particularly interesting case are the increasing events, making 27% of the total other events, that are related with incidents (biting) related to stray dogs. The sharp increase in these types of events is due to the change in regulations related to monetary compensations to the victim payed by the municipality.

Analysing the correlation between the number of active COVID cases and crime events occurrences over time it has been concluded that there is no straightforward relationship between the two. Notable are the short periods of sharp decrease in crime during the strict lockdown in the first half of 2020, which corresponds to reports from other countries such as the UK [8]. However, the resurgence is not gradual in this case, although frequency is increasing over time, especially after the short winding down at the end of the first wave in late 2020.

4 Crime forecasting

This paper also analyzes if the pandemic has an impact on prediction whether a certain type of crime will occur on a certain day and in a certain location, based on historical data. For this purpose, crime forecasting is considered as a classification problem in this research and gradient boosted decision trees based models have been developed. In the feature calculation process, crime categories are grouped into the following types: violent (which includes violence, weapon and drugs), property (which includes theft) and other (which includes other, documents and traffic). The goal of the developed models is to forecast the violent types of crime, which we have previously shown are affected by the pandemic to a large extent and where the police should act immediately. Therefore, besides the geographic location (longitude and latitude), the following information is calculated and used as input features: number of crimes per type (violent, property, other) in the last 120, 30, 7 days and in the last day. Two models were developed, one for the pre-pandemic period and one for the pandemic period.

A summary of the results of three cases is presented in Table 1. In the first two cases, a threshold of 0.38 is set for the fraction of positive labels that we would like to obtain (which is usefully if, for example, the police can patrol only 38% of the analysed regions). When comparing the two cases, although the overall accuracy of the model for the pandemic period is higher, what is more important is that the recall is reduced. That means that, while the police patrols the same percentage of regions, in the corona period 90% of the positive cases will be predicted, while during the pre-pandemic period a higher rate of 94% is correctly predicted. On the other hand, if we set the recall to the same values for the pandemic and pre-pandemic models (case number II and III), we can see that during COVID-19 period the police should patrol an increased number of regions, i.e. 46%, compared to the 38% in the case II. This result is mainly due to the increased rate of violent crimes during the pandemic period. What is also interesting to note, is that 15% of the predicted crimes in the case III will actually happen (the precision) and only 5% of the positive predicted crimes will happen in the case II. The forecasting models are also affected by the pandemic period in terms of feature importance. During the pre-pandemic period, the forecasting of whether crime will happen is mostly based on the corresponding location, while for the pandemic period, the most important features are the sum of other types of crimes in the last 120 days, followed by longitude and the sum of violent types of crimes in the last 120 days.

	case I	pandemic	period	case II	pre-pand.	period	case III	pandemic	period
	precision	recall	f1-score	precision	recall	f1-score	precision	recall	f1-score
0	0.99	0.66	0.79	1	0.63	0.77	0.99	0.58	0.73
1	0.18	0.9	0.3	0.05	0.94	0.09	0.15	0.94	0.26
accuracy			0.68			0.64			0.6
macro avg	0.58	0.78	0.55	0.52	0.78	0.43	0.57	0.76	0.49
weighted avg	0.93	0.68	0.76	0.98	0.64	0.76	0.93	0.6	0.69
fraction of positive labels		0.38			0.38			0.46	
threshold		0.65			0.25			0.52	

Table 1. Results of the forecasting models for three case studies

5 Conclusion

As a conclusion regarding the ability to predict crime, the results show that with the same percentage of regions covered by the police, in the corona period a lower percentage of positive cases can be detected. In addition, due to the change in crime rates and patterns in the pandemic period, it is important to mention that there is a difference in the developed models for the pre-corona and corona period, in terms of the importance of the input characteristics that affect the accuracy of the predictions.

References

1. Phillips, J., Land, K. C.: The link between unemployment and crime rate fluctuations: An analysis at the county, state, and national levels. Social science research, 41(3), 681-694 (2012)
2. Chauhan, C., Sehgal, S.: A review: crime analysis using data mining techniques and algorithms. In 2017 International Conference on Computing, Communication and Automation (ICCCA), pp. 21–25, IEEE (2017)
3. Boman, J. H., Gallupe, O.: Has COVID-19 changed crime? Crime rates in the United States during the pandemic. American journal of criminal justice, **45**(4), 537–545, (2020)
4. Bhagtani, H., Roj, S., Raj, K.: Criminal psychology and impact of recent global pandemic: Analysis of criminal mind with a sudden surge in crime rate. PalArch's Journal of Archaeology of Egypt/Egyptology, **17**(6), 10800–10828, (2020)
5. Temelkovski, D., Jovanovik, M., Mishkovski, I., Trajanov, D.: Towards Open Data in Macedonia: Crime Map Based on Ministry of Internal Affairs' Bulletins, 9th Conference for Informatics and Information Technology (2012)
6. Kodinariya, T. M., Makwana, P. R.: Review on determining number of Cluster in K-Means Clustering. International Journal, **1**(6), 90–95 (2013)
7. Rosenfeld, R., Lopez, E.: Pandemic, social unrest, and crime in US Cities. Council on Criminal Justice (2020)
8. Langton, S., Dixon, A., Farrell, G.: Six months in: pandemic crime trends in England and Wales. Crime science, **10**(1), 1-16 (2021)