ARTIFICIAL INTELLIGENCE

Article

DOI: 10.17803/2713-0525.2022.2.20.329-342

Predictive Policing: High-tech Modeling as a Method to Identify Serial Killers

Lev V. Bertovskiy, Margarita S. Novogonskaya
Lomonosov Moscow State University, Moscow, Russia

Alexey R. Fedorov
National Research University of Electronic Technology (MIET), Moscow, Russia

Abstract: The article considers the development of predictive policing in Russia through the creation of software, based on the use of artificial intelligence (AI) to identify serial killers. Forensic modelling in crime investigation, in particular modern digital twin technology is analyzed. The system of the digital twin is trained on the basis of a set of mathematical models of various level of complexity and specified by results of full-scale experiments. Existing approaches to solving the serial killer portrait problem are investigated. Digital twins in conjunction with machine learning can predict the behavior of the object under study in the future, based on statistical data and accelerate the work of the investigator in the investigation of serial crimes.

Keywords: predictive policing; digital twin; artificial intelligence; neural network; machine learning; ensemble methods; forensics; serial murders

***Cite as*:** Bertovskiy, L.V., Novogonskaya, M.S., and Fedorov, A.R., (2022). Predictive Policing: High-tech Modeling as a Method to Identify Serial Killers. *Kutafin Law Review*, 9(2), pp. 329–342, doi: 10.17803/2713-0525.2022.2.20.329-342.

Contents

I. Introduction

One of the latest trends in law enforcement is the use of computer models that predict crime, based on algorithms capable of self-learning (machine learning) in order to investigate and prevent crimes. Such predictive policing models are used in Europe and the United States to identify a person who is likely to commit a crime or a place in which a crime may be committed.

Predictive policing uses methods of mathematical, predictive analytics as well as machine learning algorithms to identify potential criminal activity. When predicting, they distinguish methods for predicting crimes, methods for predicting the identity of criminals and methods for predicting victims of crimes. Prediction methods cannot predict the future — they can only identify people and places at increased risk of crime, and they are to be used as elements of larger proactive strategies to address crime problems.

At the moment, Russia does not have software solutions based on machine learning algorithms that could be used to solve serial killings. It seems promising to develop predictive policing in Russia by creating models and software based on artificial intelligence in order to investigate and solve serial killings.

II. Predictive Policing

Back in the 1970s, the French philosopher Herbert Marcuse predicted the emergence of some new technologies that could change the world. On the one hand, they will open up new prospects for freedom, but on the other, they will create new forms of alienation and give the

state and corporations new mechanisms for controlling people. To date, his prediction has come true.

Computer models that predict crime fall into two following groups (Shapiro, 2017).

— Type one programs determine who is most likely to commit a crime or, conversely, become a victim. They assess people's profiles in terms of age, criminal history, employment data, dating (conducted through social media pages) and other data. What kind of data is used for this depends on the developers, and they do not always disclose details.

— In models of the second type, the main thing is time and place, that is where and when a crime can be committed. Algorithms divide the territory of the city into small zones with an area of several tens of meters — this can be a specific quarter or intersection — and calculate the probabilities based on the incoming data. For example, one of the market leaders, PredPol, makes forecasts based on statistics on reports of murders, thefts, robberies, and vehicle thefts. But even the weather, opening hours of the surrounding bars and schools can also be taken into account. If the program believes that the risk of committing a crime is high, a police squad is sent to the scene.

There is no legal definition of predictive policing in Russia. However, in doctrine and the media, this term is sometimes used, it refers to a preventive strategy based on computer calculations, with the help of which the police can assess the degree of risk of committing certain crimes in certain places. Nevertheless, it is possible to note the presence of a significant amount of use of systems based on artificial intelligence in predictive policing and related fields.

The modern application of artificial intelligence in all spheres of activity is impossible without working with Big Data. To this end, the Ministry of Internal Affairs of Russia, together with leading research centers and start-ups, is holding large joint conferences on the most relevant breakthrough approaches in the use of artificial intelligence and Big Data in order to combat crime. The Academy of Management of the Ministry of Internal Affairs of Russia housed the second major event of this kind in December 2021.

In Russian practice, the successful use of artificial intelligence technologies in predictive policing can be noted. For example, the use of artificial intelligence algorithms for video surveillance is the Safe City program. Part of this system is the FindFace Security face recognition system, created by the Russian company NtechLab in 2015. As the NtechLab website notes, the main goals of their created program are advanced analytics, search for offenders, search for missing people, ensuring the safety of public events, as well as transport security. That is, as a rule, criminals are searched for using this system. In addition, active cooperation on the use of artificial intelligence to combat cybercrime is carried out in cooperation with large banks, their cyber defense structures and information technologies (IT) structures working in the prevention and investigation of cybercrimes.

So, let us give an example of an anti-fraud system. Fraud Hunting Platform is a Group-IB product that is used by Sberbank, Post Bank, Raiffeisenbank. This program allows to detect payment fraud, combat money laundering, identify fraudsters. According to Sberbank, in 2018, using the introduced anti-fraud system, it was possible to save more than 32 billion rubles belonging to depositors. The technology quality is confirmed by the international Cybersecurity Excellence Awards.

In relation to the programs taught abroad, it should be noted the program based on artificial intelligence, which is actively used in New York (USA). This program recognizes the "handwriting" of robbers. The software is called Patternizer. For AI training, the data collected over a decade was used. To identify common features of robberies, the algorithm draws attention to the method of breaking doors, the list of stolen things and the scene of the incident. The algorithm began to apply back in 2016. Its effectiveness has been proven many times — for example, it was able to link two robberies that were committed in two different areas of New York City within two weeks of each other. The perpetrator was identified by his unusual weapon, a syringe, which he used to intimidate Home Deport shop staff. Thanks to Patternizer, instead of hours of work, the investigators only needed to make one click of the computer mouse.

III. High-Tech Modeling as a Method to Identify Serial Killers

Modeling originates in antiquity. The word "model" comes from the Latin word "module," meaning "measure," "sample." Its intended meaning is connected with the art of building, and in almost all countries it is used to signify an image, a type, or things in some connection with an external thing (Nechaevsky, 2013). To describe logical operations, modeling was first used by Raymundus Lullius (1235–1315) (Luzgin, 1981, p. 4).

In the ancient period, modeling can be characterized as follows:

— there is no concept of "models" and "modeling" in science, however, the presence of relevant objects and phenomena in the practice of human cognition and life;

— appearance of language characters (symbols) as models denoting real objects, phenomena, processes;

— the existence of mental modeling as an integral feature of human thinking;

— reflection of the results of scientific research in the form of models. The emergence of elements of scientific theories as a reflection of the surrounding world;

— the formation of modeling languages: verbal; symbolic, including mathematical; geometric; drawing; schematic; mechanical;

— development of descriptive verbal pedagogical models as a fixation of the scientist's thoughts and as an indicative basis for the upbringing process. The use of models of studied objects, processes, phenomena as means of facilitating the understanding and assimilation of knowledge in the content of education and education (Kotlyarova, 2019, p. 7).

Modelling is also found in cults. A prime example of modelling is the volt (voodoo doll), a doll used in voodoo witchcraft. It is believed that as a result of a special ritual, the doll receives a special kind of connection with a particular person, enabling the possessor of the doll to influence through it the person it symbolises. Building and studying models of real objects, processes or phenomena in order to obtain explanations for these phenomena, as well as to predict the behavior of these objects,

has become one of the main tools of human cognitive activity. This effective method of cognition is used by various sciences and branches of knowledge, and if earlier there was an opinion that the inheritance of modeling is only technical and natural sciences, now its use in public sciences is considered expedient and, moreover, promising.

Recently, modeling has developed most dynamically in the activities of solving and investigating crimes. For the first time, G. Gross (1847–1915) expressed ideas for using the modeling method in forensic science to recreate a specific picture of what happened and search for a criminal. The origins of recommendations for the use of modeling in investigative practice appeared during this period, that is, at the dawn of forensic science. The founders of Russian criminalistics V.I. Gromov and I.N. Yakimov recommended that investigators during the investigation recreate a specific picture of what happened and use it to search for the culprit. The term "modeling" itself appeared in the 1960s thanks to the work of A.R. Ratinov (Ratinov, 1967). The active development of domestic forensic modeling has been taking place since the 1970s and 1980s.

Forensic modeling is a process of perception and processing of source data, which is based on conditional probabilistic syllogism "if, then, probably" and sets itself the task of establishing and using natural connections and relationships in versive and predictive models of the mechanism of criminal activity and investigation models in order to form a system of evidence in a criminal case (Luzgin, 1980). The forensic model allows to gain new knowledge about the original and use it to solve search, cognitive, recognition, identification and other tasks in the process of investigating and solving crimes (Volchetskaya, 1997). Recent years have been distinguished by a truly explosive development of technology. According to L.V. Bertovskiy, the fourth industrial revolution — the logical continuation of computerization, namely, the optimization of automatic and machine processes — is coming. In this regard, the era of high-tech law is coming (Bertovskiy, 2021).

One of the latest trends in law enforcement is the use of computer models based on algorithms capable of self-learning (artificial intelligence and machine learning) in order to investigate, predict and prevent crimes. The most popular modern technology based on

computer modeling is Digital Twin — a synchronized virtual model of any objects, systems, people, processes and environments that simulates internal processes, technical characteristics and behavior of a real object under the influence of interference and the environment. The digital twin tracks the past and predicts the future and is a learning system consisting of a complex of mathematical models of different levels of complexity, refined from the results of field experiments, and is a changing digital profile containing historical and most relevant data about a physical object or process. Digital twins together with machine learning allow to create a reliable model and predict the behavior of the studied object in the future, based on the analysis of large and semi-structured data arrays (Lubin, 2021).

The high-tech modeling is widely discussed in foreign scientific literature. For example, modeling the behavior of offenders who commit serious sexual assaults (Adderley and Musgrove, 2001), predicting criminal recidivism, classification system for serial criminal patterns (Caulkins *et al.*, 1996).

In Russia, the problem of identifying serial killers is especially acute. Serial killings occupy a significant place in the structure of murders, are latent and often remain unsolved. In Russia, statistics of serial crimes are not kept. However, from the data provided by the territorial investigating authorities of the Investigative Committee of Russia, as well as analytical materials of the GUK (CC), it was established that 138 episodes of 1,083 crimes were registered in the first half of 2021. In the study period, 136 episodes of 1,080 crimes were revealed (with crimes of previous years). Also in this period, 96 episodes were registered about 861 crimes committed against minors. 94 episodes in relation to minors have been disclosed. Given the crimes committed in past years, 142 episodes of 552 crimes remain unsolved. 69 episodes of 196 crimes committed against minors remain unsolved. A digital twin can be used in investigative activities to identify a criminal, a potential crime scene, predict the development of a series and solve other problems.

Often a serial killer develops a certain stable "criminal handwriting" — a specific manner through which they can be established by the same type of victims, the stereotype of criminal actions, the instrument, method and situation of its commission, as well as the

nature of the bodily injuries, wounds and other signs displayed in the trace data, and by which it can be established. For example, the serial killer A. Pichushkin in the last nine episodes inserted a bottle, stick or branch into the head of the victims.

Due to the large amount of data on the case, insufficient analysis of the material and the relationship with other cases, the investigation is not always able to quickly identify the coincidence and relationship of the obtained data to establish seriality. In addition, not all investigative units comply with the requirements of paragraph 2.4 of the Order of the Investigative Committee of Russia No. 130 dated August 11, 2011 "On the organization of work in the Investigative Committee of the Russian Federation to investigate crimes with signs of seriality," according to which it is necessary to immediately consider the issue of combining criminal cases in one proceeding on crimes with signs of seriality. Such facts, as a rule, take place in cases of crimes committed on the territory of different subjects of Russia. As a result, the preliminary investigation in such criminal cases is carried out independently without taking into account the data on the commission of crimes by one person on the territory of other subjects of Russia. Based on the results of the investigation, decisions are made to suspend it due to the unidentified person to be brought in as an accused (suspect).

Thus, the problem is that the crime is recorded as isolated, which prevents the rapid disclosure of the crime, which is why the serial killer continues his criminal activity. One of the options for solving this problem is the development of a digital double of a serial killer. This computer model that predicts crime will combine 2 groups of models. With its help, it will be possible to identify a potential offender, as well as time and place, that is, where and when a crime can be committed.

The advantages of an artificial intelligence-based program are that it can analyze a large amount of data (not only data about crimes, but also data about the world around it, such as natural phenomena in a given period, news reports, political situations in the world, *etc.*), and thus identify non-obvious connections between the crime and other events in the world.

For example, the presence of non-obvious connections was proven in 1980 by Phillips. Phillips demonstrated that immediately after a wave

of publications reporting on suicides, the number of people who died during plane crashes is increasing by 1,000 %! Moreover, the increase in the number of deaths from accidents concerns not only deaths in airplanes. The number of road accidents is also increasing sharply. Thus, it seems promising to put a large array of data into the program in order to identify certain patterns that have not been previously identified in science (Cialdini, 2001, pp. 135–136).

When developing a digital twin, several areas of research can be identified, which are carried out in parallel, based on the basic foundations of machine learning, namely data, features and algorithms. The first area — without which, in principle, it is impossible to create software tools — is the creation of a sufficient data set consisting of actually completed cases. Moreover, for the development of predictive policing in Russia, this data must be promptly replenished and available for use.

The next problem to be solved is the separation of the basic signs of the committed crime from many criminal cases, the analysis of the informativity of the signs, the identification of the most important ones, the implementation of the task of reducing the dimension of the space of the signs. Signs are formed from protocols of inspection of the scene, protocols of interrogations of victims, statistical cards, conclusions of a forensic medical examination, indictments. It is important to understand what features the digital twin will represent in the future model, since the entire further process is based on them. Firstly, it needs not to miss important characteristics that describe the object, and secondly, create tough criteria for deciding on the characteristic. The following set of characteristics is proposed:

1) the situation of the crime;
2) the method of committing the crime;
3) typical traces;
4) circumstances in which the crime was committed;
5) data on the identity of the victim;
6) data on the identity of the offender;
7) place of release of objects (guns, victim's belongings, *etc.*);
8) age of the perpetrator;
9) previous convictions of the criminal;

10) the presence of psychological diseases in the criminal;

11) the family and children of the criminal;

12) data on the presence of a connection between the criminal and the victim.

Three following characteristic categories can be distinguished.

1. Boolean (bicategoric), the answer to which is: Yes or No (1 or 0). For example, the answer to the question: is there a connection between the criminal and the victim?

2. Categorical, the answer to which is a specific class. Usually there are more than two classes (multicategory). For example, the age of the criminal.

3. Quantitative, the responses to which are numbers characterizing a specific measure. For example, the distance to the ejection of objects by a criminal.

Significant heterogeneity, weak structuring, small amount of data from the point of view of machine learning methods, high dimension of the feature space, the absence of a hypothesis for the distribution of source data complicates the model of the digital twin. A separate task to be solved is the analysis of natural language texts in order to automatically distinguish signs from the text of protocols for inspecting the scene of the incident, protocols for interrogating victims, conclusions of a forensic medical examination, and indictments.

When building a model, it needs to understand what types of questions the algorithm can deal with. It can be argued that a computer program is trained as experience is gained regarding a certain class of tasks and objective function, if the quality of solving these problems relative to the objective function improves with new experience. Thus, it is not the data that takes center stage in the learning process, but the objective function and the way the results are evaluated. The choice of objective function completely determines all further work, and even in similar tasks, different objective functions can lead to completely different models.

For example, the decision tree is trained on all types of questions, and the neural network receives only numerical inputs and is trained only on quantitative features. Does this mean that we should abandon some questions for the sake of a more advanced model? Not at all,

you just need to properly prepare the data based on the problem being solved. So for each data set, two fundamentally different machine learning problems can be solved: the classification problem and the clustering problem.

The task of classification for detecting serial crimes implies the presence of pre-marked data relating specific cases to the already identified series of crimes. One series is one class of crime. Such markings are carried out by competent investigators based on their personal experience or on the results of indictments. The greatest efficiency of the model can be achieved by using ensemble machine learning methods, where several models are trained to solve the same problem and combined to obtain better results. The main hypothesis is that with the right combination of weak models, we can get more accurate and reliable crime classification results, identify crime series and predict the actions of the offender in the future.

The task of clustering in relation to the problem of predictive policing can be set as follows. When clustering, the number of future clusters is initially unknown, that is we do not know how many series of crimes are contained in the training sample. A dataset is a set of feature vectors of objects, events, processes, or phenomena. For each of the sample examples, additional data or description relating only to the individual examples themselves but not directly involved in the training may be known. It is essential to obtain a model that will perform one or more functions from the list throughout its life cycle:

— relate the vectors supplied to its inputs to one of the already existing clusters closest to the input vector;

— create a new cluster if it does not find a close enough match to existing clusters

— delete clusters that were unused as a result of the last learning era.

This problem can be effectively solved using Cohonen neural networks, graph neural networks or adaptive resonance networks. The input variables of such networks can take both binary and analog values. Their output values can be conditionally considered the distance to the center of each of the existing clusters. In fact, they have no outputs in the understanding familiar to other neural network architectures, since

there is no clear pre-known training goal — an output image or a class of images. At the final stage, the developed architecture is trained, its training parameters are adjusted, as well as the adequacy of the model's conclusions on the trends and similarities of unsolved murders and the identification of the connection between crimes is checked.

IV. Conclusion

Technological progress does not stand still. Predictive policing is actively used all over the world and has already demonstrated its effectiveness. The detection rate of serial murders in Russia at the moment is low. In this connection, we consider it possible to create a program of "digital twin serial killers." With its help it is possible to get data on a potential offender, potential scene of crime, victim and weapon.

The main tasks to be solved in the near future, including on the legislative level, are creation of a dataset of sufficient volume, consisting of completed cases and cases in progress. This data must be promptly updated and made available for use. The next task to be solved is to justify the choice of machine learning algorithms to obtain the most effective digital twin model for the creation of a software tool for the analysis of crimes for seriality. The prospect of application of this program is increase of solvability of serial murders, decrease of load on law enforcement bodies.

References

1. Adderley, R. and Musgrove, P.B., (2001). Data mining case study: modeling the behavior of offenders who commit serious sexual assaults. In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 26 August 2001, New York: ACM. Pp. 215–220.

2. Bertovskiy, L.V., (2021). High-tech law: concept, genesis and prospects. *RUDN Journal of Law*, 25(4), p. 739. (In Russ.).

3. Caulkins, J., Cohen, J., Gorr, W., and Wei, J., (1996). Predicting criminal recidivism: a comparison of neural network models with

statistical methods. *Journal of Criminal Justice*, 24(3), pp. 227–240. http://doi.org/10.1016/0047-2352(96)00012-8.

4. Cialdini, R., (2001). Influence: The Psychology of Persuasion. Translated from English. "Masters of Psychology" series. St. Petersburg: Piter. Pp. 135–136. (In Russ).

5. Digital Twins, (2022). Itenterprise. Available at: https://www.it.ua/ru/knowledge-base/technology-innovation/cifrovoj-dvojnik-digital-twin [Accessed 08.02.2022].

6. Gromov, V.I., (2003). Inquiry and preliminary investigation. Methods of investigation of crimes. *Examination of the crime scene (Collection of scientific works): To the 200th anniversary of the founding of the Lomonosov Moscow State University*. Moscow: LexEst. (In Russ.).

7. Gross, H., (2002). A guide for forensic investigators as a system of criminology. Moscow: LekEst. P. 1088. (In Russ.).

8. Kotlyarova, I.O., (2019). Method of modeling in pedagogical research: history of development and current state. *Bulletin of SUSU, series* "*Education. Pedagogical Sciences,*" 11(1), p. 7. (In Russ.).

9. Lubin, A.F., (2021). Forensics: features of didactics of versive thinking. *Legal science and practice*, 3, p. 10. (In Russ.).

10. Luzgin, I.M., (1980). On the issue of forensic modeling. *Topical problems of Soviet criminalistics*. (In Russ.).

11. Luzgin, I.M., (1981). *Modeling in the investigation of crimes*. Moscow: Yurid. Lit. P. 4. (In Russ.).

12. Marcuse, H., (2003). Eros and Civilization. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Translated from English. Afterword, commentaries by A.A. Yudin. Comp. by V.Yu. Kuznetsov. Moscow: AST. (In Russ.).

13. Nechaevsky, A.V., (2013). History of the development of computer imitation modeling. System analysis in science and education. *System analysis in science and education*, 2, pp. 4. (In Russ.). Available at: http://simulation.su/uploads/files/default/2013-nechaevsky-1.pdf [Accessed 08.02.2022].

14. Phillips, D.P., (1980). Airplane accidents, murder, and the mass media: Towards a theory of imitation and suggestion. *Social Forces*, 58, pp. 1001–1024.

15. Ratinov, A.R., (1967). *Forensic psychology for investigators*. Moscow. (In Russ.).

16. Review of the practice of identifying and investigating by the investigating authorities of the Investigative Committee of the Russian Federation in the first half of 2021 serial crimes (2021). *Bulletin of the Main Directorate of Criminalistics (Forensic Center)*, 9(69), p. 9. (In Russ.).

17. Shapiro, A., (2017). Reform predictive policing. *Nature*. Available at: https://www.nature.com/articles/541458a [Accessed 12.04. 2022].

18. Volchetskaya, T.S., (1997). *Modern modeling problems in forensic science and investigative practice*: *a textbook*. Kaliningrad. (In Russ.).

19. Yakimov, I.N., (1924). A practical guide to the investigation of crimes. Moscow: "Zagotkhoz" of the militia of the Republic. (In Russ.).

Information about the Authors

Lev V. Bertovskiy, Dr. Sci. (Law)**,** Professor, Lomonosov Moscow State University; National Research University of Electronic Technology (MIET), Moscow, Russia
1 stroenie 13 Leninskiye Gory, Moscow 119991, Russia
bgl1980@yandex.ru
ORCID: 0000-0002-2835-288X

Margarita S. Novogonskaya, Postgraduate student, Lomonosov Moscow State University, Moscow, Russia
1 stroenie 13 Leninskiye Gory, Moscow 119991, Russia
novogonskaya.m@mail.ru

Alexey R. Fedorov, Cand. Sci. (Technical Studies), Associate Professor, National Research University of Electronic Technology (MIET), Moscow, Russia
1 ploshchad Shokina, Zelenograd, Moscow 124498, Russia
af123@yandex.ru