



THE ACCURACY OF GEOGRAPHIC PROFILING METHODS BASED ON THE EXAMPLE OF BURGLARIES IN WARSAW

Michał GÓRSKI

*College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Warsaw,
Poland*

Abstract

In the current study different geographic profiling methods were tested on a sample of burglaries in Warsaw. Overall 38 crime series were included in study, accuracy was measured as hit score and error distance. Results show limited obtained accuracy, comparably lower than what can be seen in similar tests conducted in western cities, but each method also shows improvement with series length. This shows that it is possible to use geographic profiling in polish setting, but also that there are significant limitations that one has to consider.

Keywords

Geographic profiling; Accuracy; Series length.

Received 16 July 2021; accepted 10 September 2021

Geographic profiling has enjoyed increasing interest in recent years, and new Polish-language publications touching on the topic are also appearing. They usually cover problems related to geographic profiling, such as the distribution of the distances covered by perpetrators (e.g. Goldschneider, 2012; Górski 2018a), or they are only theoretical publications (e.g. Mordwa, 2019). However, there are no studies testing the accuracy of geographic profiling on a Polish sample. As Mburu and Helbich (2014) point out, such studies do not exist outside highly developed Western countries. The authors mention the USA, Canada, the UK and one study in Germany, while they themselves performed their analysis on the Kenyan population. Thus, such analyses are lacking in continental Europe, and the present study aims to fill this gap.

Geographic profiling is a set of methods aimed at predicting the location of a perpetrator of a series of crimes based on the location of incident sites

(Rossmo, 2000; Canter, Hammond, Youngs, Juszcak, 2012). These are optimization methods indicating the most probable area, which in consequence is supposed to speed up the process of finding the perpetrator. The use of profiling firstly reduces search costs by minimizing the area (or number of potential suspects) that need to be checked until the actual perpetrator is found (Chainey, 2008). Secondly, it makes it possible to find the offender before they commit another crime, which is particularly important in the case of serious crimes (Chainey, Ratcliffe, 2008).

The theories underlying geographic profiling have been described many times in the Polish-language literature (e.g. Mordwa, 2019), including in the journal *Problems of Forensic Sciences* (Górski, 2018a). However, it is necessary to describe how programmes for creating geographic profiles (quantitative methods of geographic profiling) work and methods for assessing the accuracy of predictions made by such tools.

Geographic profiling software

In order to give an idea of how geographic profiling software works, it is necessary to describe the concept of the spatial resistance function. This is a concept derived from analyses of the movement of people for various (non-crime related) purposes (Levine, 2013). The distance decay function is the distribution of distances that individuals cover from their place of residence to a given destination (Levine, 2013). However, in the context of crime, two modifications have been introduced. First – residence has been generalized as the offender's anchor point (Chainey, Ratcliffe, 2008), i.e., the location from which the offender carries out their crimes (not necessarily the place of residence, Rossmo, 2000; Chainey, 2008). Secondly, the concept of a buffer zone was introduced, i.e., an area in the immediate vicinity of the offender's place of residence, where the probability of committing a crime is lower (Canter, Larkin, 1993).

Geographic profiling programmes work as follows: the entire search area is divided into small squares arranged in a checkerboard pattern; for each point in such a grid, the probability of finding the perpetrator is calculated based on the distances to the locations of the event sites and the distance decay function (Chainey, Ratcliffe, 2008). Therefore, based on the distance of the given point to each of the event locations, the probability of the perpetrator traveling the distance from the given place can be calculated. This consequently translates into the probability that the offender's anchor point is located there. Since such calculations concern the entire grid, this produces a probability plane, with the points with the highest values indicating the areas with the highest probability of finding the offender (Chainey, Ratcliffe, 2008). Methods of this type will hereafter be referred to as grid-based or probability plane-based (Mburu, Helbich, 2014).

Some much simpler methods, referred to as centographic methods, rely on identifying a single point from which to start the search (Mburu, Helbich, 2014). These points are, for example, the spatial mean, the spatial median, or the first or last known event in the series. In order to assess which methods are more accurate, it is first necessary to discuss the different methods for assessing the accuracy of profiles.

Assessment of the accuracy of the geographical profile

Two basic methods have been adopted in the literature to assess the accuracy of geographic profiling

methods. One is based on determining the distance of the point with the highest probability to the offender's actual place of residence (error distance; Rich, Shively, 2004). This allows comparing methods based on creating a probability plane with methods based on indicating only one point as the place to start the search. However, since the probability does not decrease the same in every direction when using methods based on a distance decay function, the error distance measure does not allow this to be taken into account (Rossmo, 2005b). It is even possible for two (or more) areas to be indicated as predicted perpetrator anchor locations. In such a situation, if the perpetrator is located in a slightly less probable area, then intuitively the profile could be judged as accurate (Rossmo, 2005b). Contrary to this intuitive inference, the value of the error distance could be large and consequently the profile judged as inaccurate. An example of this situation was presented by Rossmo (2005b).

Among other things, to address this problem, a second method of assessing the accuracy of profiles, the so-called hit score percentage, is used. It consists in dividing the area that needs to be checked (from the most probable point until the actual perpetrator is found) by the entire search area (Rich, Shively, 2004; Rossmo, 2000, 2005b). This approach solves the problem described earlier, but also better defines the degree to which the profile helped to limit the initial search area. E.g. a hit score percentage of 5% means that if the perpetrator was searched for according to probability, 5% of the total search area would have to be checked until the right person was found. At this point there is no consensus on this issue: there are advocates of all the mentioned approaches, but also other methods that are a combination of the above ones (Rich, Shively, 2004; Paulsen, 2006).

Minimum length of the event series

Rossmo (2000) posits a minimum of five related events for the predictions made to be accurate and for the geographical profile to be able to be used. He bases the value of these five events on the results of computer simulations. This threshold is disagreed with by Snook, Taylor and Bennell (2005a), who point out the weaknesses of Rossmo's method and argue that there is insufficient justification for setting a threshold of a minimum of five events. Furthermore, the authors argue that with the methodology adopted by Rossmo (2000) the correct minimum series length should be 17 events. They also point out that setting such a high threshold may result in geographic profiling only being useful in very few situations.

Research on the accuracy of geographic profiling

There is no research in Poland on the accuracy of algorithmic methods of geographic profiling. This is most likely due to the limited availability of both the software and the lack of appropriate data to conduct such analyses. Among foreign studies it is worth noting the work of Paulsen (2006) and Snook, Zito, Bennell and Taylor (2005b). In the former work, the author compared the accuracy of the applied methods with evaluation by previously trained individuals and their predictions without the support of computer programmes. The data in these studies consisted of 25 series of different types of events; when the predictions of the trained individuals were taken into account, the actual perpetrators were found in the most probable 5% of the area less frequently than in the corresponding area calculated from the algorithms (8% accuracy for the trained individuals' predictions, 48% for the exponential distribution; Paulsen, 2006). This advantage of the software disappeared when such analysis was performed for the top 15% of the area – under these conditions, trained individuals' indications were comparable to some of the results of using computer programs.

Snook et al. (2005b) in their study showed that centrographic methods for shorter series can give the same (or better) results as methods based on probability grids. The authors tested the accuracy of different methods depending on series length. Under most of the conditions tested, the most accurate method was to determine the centre of a circle whose diameter is the line connecting the two events that are most distant from each other. Only when series length reached nine events were predictions that were based on the point whose sum of distances to all events is the smallest more accurate, while for series of ten events, predictions based on the linear distance decay function were the most accurate (Snook et al., 2005b). Thus, it was only from a series of ten events onwards that the use of methods calculating the entire probability plane was optimal. For less events, other simpler methods were optimal. To apply most centrographic methods, no software is needed – just a map, a pencil and a ruler. So if these methods are sufficient, then the need for software is questionable. In a test conducted on a sample of a series of burglaries, Canter and Hammond (2007) showed that depending on the method used, between 72% and 46% of perpetrators are in the top 5% of the probability plane. This is significantly higher than obtained by Paulsen (2006), but, in contrast to the results

obtained by Snook et al. (2005b), using the centre of the circle proved to be the least accurate method.

It is worth noting that in Paulsen's (2006) study, the use of a probability grid calculated from an exponential distribution was one of the best methods. Using this method, most perpetrators were in the top 5% of the geographic profile, while in the top 15% of the profile, only predictions using Rigel and Dagnet software performed better. The exponential function also gave some of the best results in the study by Canter and Hammond (2007).

Current research

The main objective of the current study was to determine how well different algorithmic methods of geographic profiling work in the Polish population of offenders, depending on the length of the analysed series of events. As mentioned earlier, this type of research is rare, and has not been conducted in Poland before. The following research hypotheses were put forward:

Hypothesis 1. Most of the geographic profiling methods tested will produce better accuracy results with increasing series length.

Hypothesis 2. The exponential distribution will be the best method of predicting the place of residence of an offender.

Hypothesis 3. For the best geographic profiling methods, the average hit score percentage will be less than 15%.

The exponential function was predicted to give the best results because of the research of Canter and Hammond (2007), who analysed a similar group of offenders, and also based on the results of Paulsen (2006), in which it turned out to be one of the best methods. The 15% threshold was also chosen because of Paulsen's (2006) study, a comparison with which will allow an assessment of whether the tested methods give better accuracy results than those methods in which results are obtained without the support of computer methods. Rossmo (2000) also obtained hit score percentages of 15% or less in almost all tested series except in two cases. All analyses were performed in R 3.6.3 software, while graphs were plotted in the ggplot2 package (Wickham, 2009).

Data

Thanks to cooperation with Police Headquarters, data was obtained from the *KSIP (Krajowy System*

Informacyjny Policji – National Police Information System), which contained data on 10634 burglaries from 2011. From among them, the incidents for which offenders were assigned (had been found) were selected. Data obtained from the *KSIP* allow for obtaining a large amount of material, but on the other hand using them is associated with potential inaccuracies. This is due to the fact that they are not created on the basis of court judgements. For the purposes of the current study, two or more events that have a common perpetrator or perpetrators are considered to be a series. Since it is possible that offenders only cooperated on some events and not on others, each offender was treated as a separate series and with anchor points attributed to only that one offender. This may negatively affect the assessment of accuracy, as a group of offenders may have operated in the region of the place of residence of one offender (the decision maker, the person in charge of the group) – in such a situation, profiling may accurately predict the place of residence of that one offender, but incorrectly that of the others. During the research, there was no access to case files, so there was no way to determine who in a given group was the ‘main’ perpetrator. As the problem of the possibility of the composition of the groups changing over time could not be resolved, it was considered justified to treat each offender as a separate series.

However, if more than one location was assigned to a perpetrator, then the best result was selected by testing each of the geographic profiling methods. This was dictated by practical considerations – if the perpetrator was detected thanks to the first (more probable) location, further search would not make sense, and including the remaining locations in the analysis would be groundless and would only distort the results.

The final analysis did not include perpetrators who lived outside the search area, which was based on the fact that since the aim of the research is to compare geographical profiling methods and the perpetrator resides outside the search area, any method used will be inaccurate. Thus, taking such events into account will only distort the results; moreover, as indicated by Paulsen’s (2007) research, as well as research on the Polish population of perpetrators (Górski, 2018b), it is possible to predict whether the perpetrator is a stable perpetrator. This means that it is possible to discard perpetrators who do not live within the search area even before they are detected, and geographic profiling should be used in other cases.

Finally, the series were divided into three groups based on series length. The first group included series consisting of two events, the second group included three or four events, and the last group consisted of

series of five or more events. These groups contained 15, 8, and 15 series of events, respectively, for a total of 38 series of thefts. Only 15 detected series consisted of five or more events. Such a small sample size is not unheard of in this type of study; as Mburu and Helbich (2014) pointed out – often such analyses are done with even smaller samples. Snook et al. (2002) and Snook, Taylor and Bennell (2003) conducted research with a sample of ten series in both cases, while Rossmo (2000) conducted research on a sample of fifteen series of events.

Analyses

A total of 11 methods supporting geographical profiling were included in the analyses: seven centrophographic methods and four based on probability grids. An exact list of the studied methods together with their brief description can be found in Table 1. The area of search for the perpetrator was considered to be the area of their activity enlarged by about 20%. The two described methods for assessing the accuracy of the geographic profile were included in the study: the distance of the most probable point to the offender’s place of residence (error distance) and the more probable percentage of the area from the offender’s place of residence (hit score percentage).

Since probability grid-based methods were included in the analysis, it was necessary, as a first step, to calculate the parameters of the distance decay function. In this analysis, the distribution of the distances travelled was considered for all serial offender events and not only for stable offenders or those residing within the search area, as such a limitation was without meaning for this analysis. The analysis concerned 416 events, and after discarding the so-called outliers, there remained 381 – and it was on their basis that the parameters of three distributions of the distance travelled (exponential, normal and lognormal) were calculated. The distributions are shown together with the kernel density estimation (calculated on the basis of the same sample) in Figure 1.

In the next step, geographical profiles were created for all analysed series and an analysis of their accuracy was carried out using the two described methods. The results were divided into three groups depending on the series length. For such assessments of accuracy, three measures were calculated: mean, median¹ and

¹ The reader may at times feel lost, as both the mean and median appear twice. Once as centrophographic methods of creating a geographic profile, and then again as statistics of the obtained results. For the sake of clarity, the first situation is de-

Table 1
Geographic profiling methods with description

Measure name	Description
Centographic methods	
First	Distance to the first known crime in series
Last	Distance to the last known crime in series
Median	Distance to median of coordinates of crime locations
Mean	Distance to mean of coordinates of crime locations
Minimum distance	Distance to point of minimum sum of distances to the crime locations (also known as spatial median)
Harmonic	Distance to harmonic mean of coordinates
Circle center	Distance to center of the circle which diameter is segment between two crime locations furthest away from each other
Distance decay functions based on which probability surface was calculated	
Exponential	Exponential distribution fitted to distribution of journey to crime distances
Normal	Normal distribution fitted to distribution of journey to crime distances
Lognormal	Lognormal distribution fitted to distribution of journey to crime distances
Kernel	Kernel distribution function of journey to crime distances

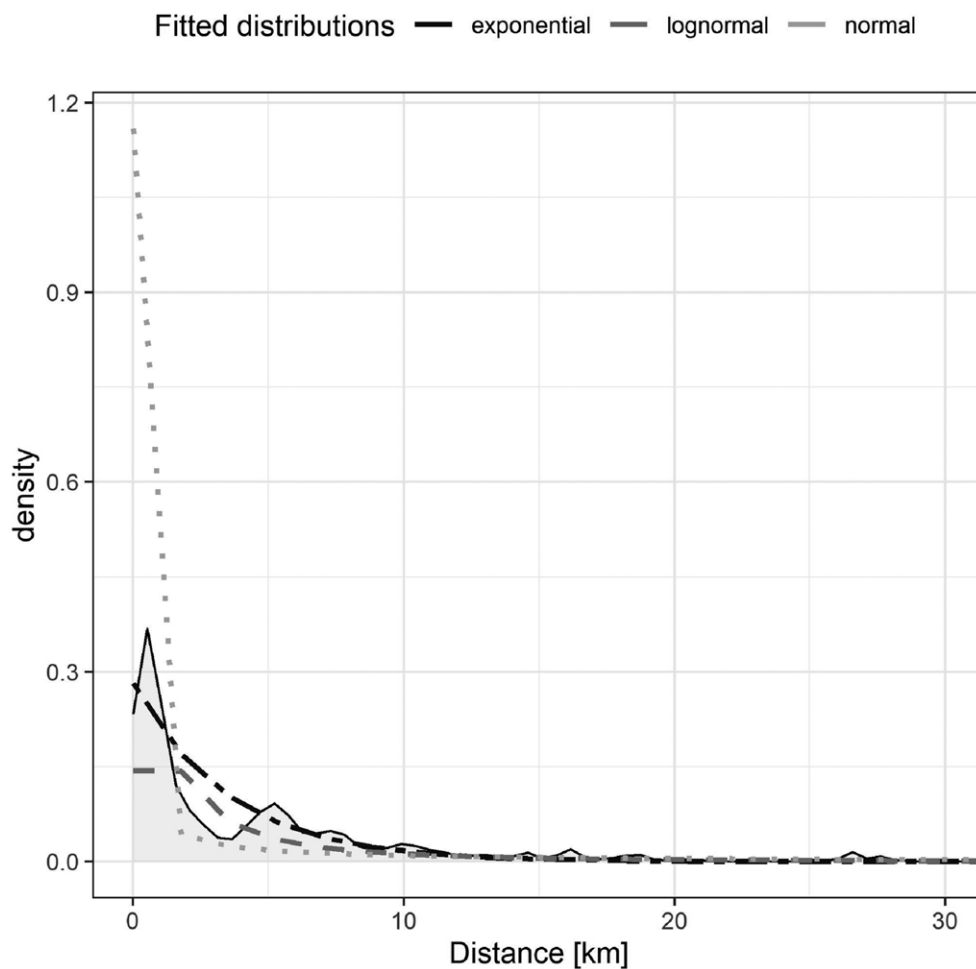


Fig. 1. Kernel density function and fitted distributions.

Table 2
Geographic profiling measures and two methods of counting their accuracy

	Hit score percentage [%]								
	Mean			Median			Standard deviation		
Series Length	2	3–4	5 and more	2	3–4	5 and more	2	3–4	5 and more
First	51.77	32.45	32.59	54.71	23.52	20.05	36.66	32.98	34.89
Last	40.72	18.79	32.02	36.59	16.35	20.27	34.83	16.13	26.99
Median	49.98	67.95	46.78	44.13	70.26	44.68	31.38	29.41	30.72
Mean	55.47	32.04	26.68	60.21	19.41	23.8	34.57	30.81	19.22
Minimum distance	49.05	27.18	23.92	45.53	24.62	26.94	33.07	31.67	21.97
Harmonic	55.44	32.05	26.68	60.26	19.44	24.03	34.62	30.82	19.22
Circle.center	55.47	49.9	43.67	60.21	33.64	41.12	34.57	35.25	33.3
Exponential	38.19	29.05	20.32	49.29	30.58	22	30.06	27.1	19.92
Normal	52.19	67.91	48.16	42	71.22	43.11	35.17	31.94	27.94
Lognormal	41.25	42.82	22.49	33.27	44.37	20.71	27.76	22.82	23.23
Kernel	54.07	46.56	36.48	51.25	36.09	33.8	30.86	35.9	31.29
	Error distance [km]								
	Mean			Median			Standard deviation		
Series length	2	3–4	5 and more	2	3–4	5 and more	2	3–4	5 and more
First	3.45	2.6	1.47	4.33	0.9	0.86	3.06	3.06	1.67
Last	2.05	1.48	3.54	0.75	1.05	1.86	2.33	1.73	4.58
Median	3.27	4.75	4.94	3.41	3.42	3.07	2.6	4.4	5.11
Mean	1.7	1.57	1.41	1.86	0.53	0.94	1.17	1.73	1.27
Minimum distance	2.42	1.84	1.43	0.92	0.72	0.44	2.58	2.27	1.81
Harmonic	1.68	1.57	1.41	1.87	0.53	0.94	1.16	1.73	1.27
Circle.center	1.7	2	2.23	1.86	1.35	0.82	1.17	2.03	2.44
Exponential	2.69	1.94	1.43	0.64	0.9	0.44	3.25	2.31	1.89
Normal	2.13	2.35	2.38	2.15	2.11	2.57	1.17	1.5	1.32
Lognormal	2.82	1.95	1.63	1.15	0.94	0.58	3.16	2.23	1.94
Kernel	2.62	1.94	1.89	2.68	0.6	1.05	1.86	2.17	1.84

standard deviation. The median seems to be the most important out of the mentioned measures, as it indicates how large a percentage of the area would need to be checked in order to find 50% of the perpetrators. The results are presented in Table 2. The best values in each column are shown in bold, i.e. the profiling method that was optimal for the given length of the series

scribed as the mean/median of the coordinates and the second as the mean/median of the results. The reader may at times feel lost, as both the mean and median appear twice. Once as centrographic methods of creating a geographic profile, and then again as statistics of the obtained results. For the sake of clarity, the first situation is described as the mean/median of the coordinates and the second as the mean/median of the results.

when the given measure of accuracy was applied. It should be noted that for both methods (hit score percentage and error distance) of evaluating profile accuracy, the lower the value, the more accurate the profile was.

The results of both methods of evaluating accuracy were then presented in graphs: in Figure 2, the results describing accuracy measured as hit score percentage, and in Figure 3, accuracy measured as error distance.

To check the relationship between accuracy and series length, linear regression coefficients were calculated for each geographic profiling method. In the case of assessing the accuracy using the hit score percentage, all directional coefficients were negative except for the median of the coordinates. In the case of

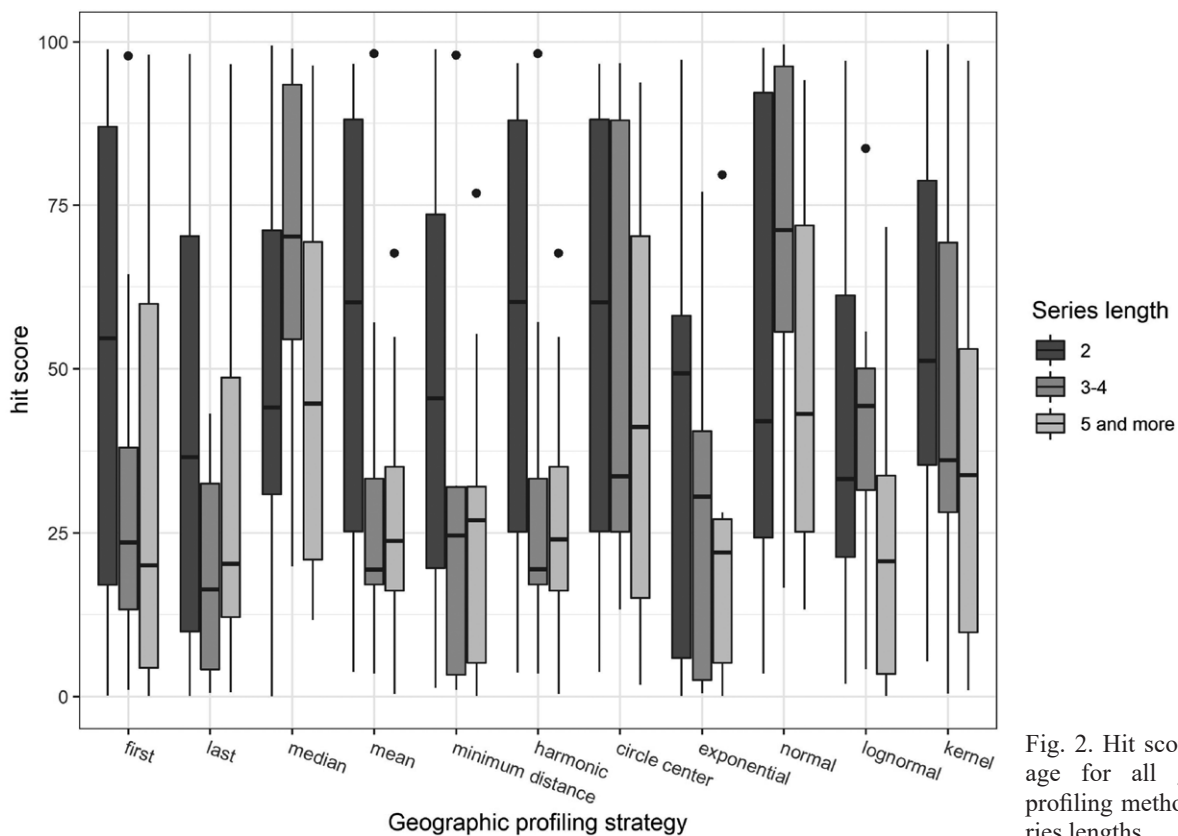


Fig. 2. Hit score percentage for all geographic profiling methods and series lengths.

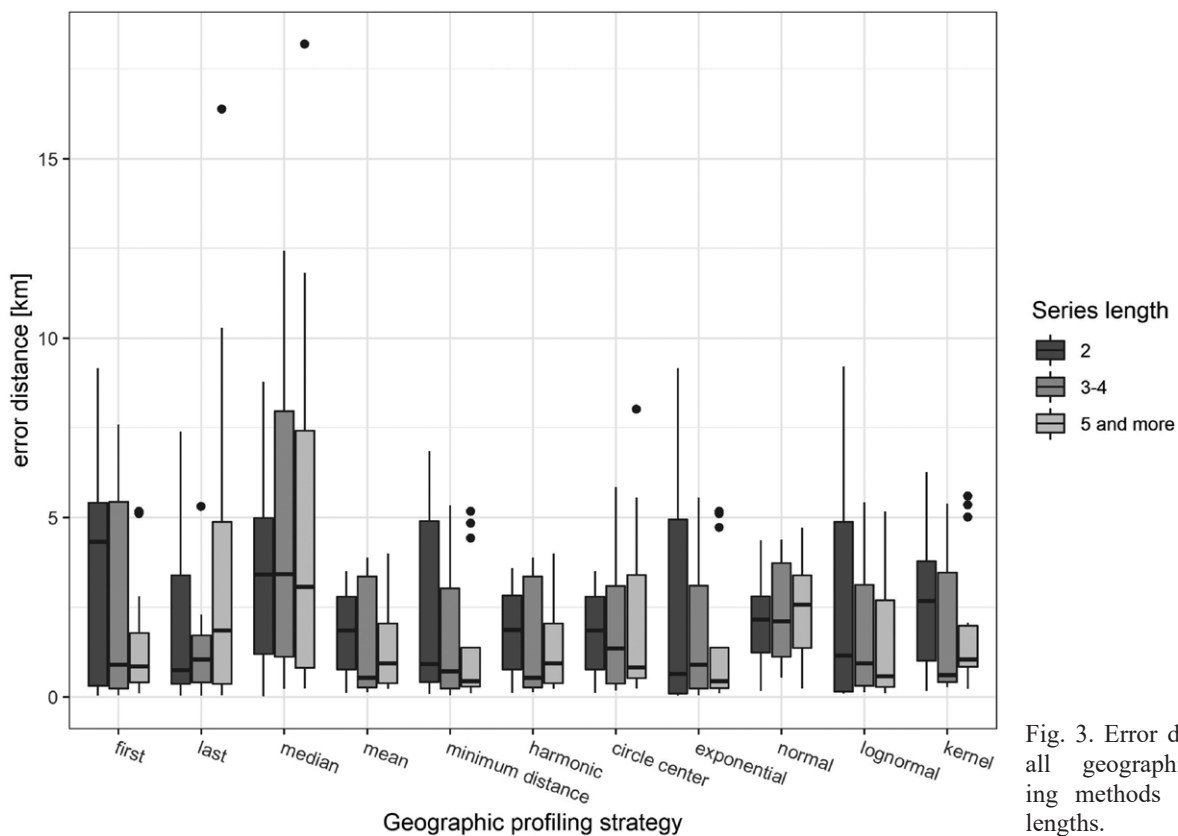


Fig. 3. Error distance for all geographic profiling methods and series lengths.

error distance, negative directional coefficient values were obtained for distance from: the first event in the series, the mean of the coordinates, the minimum distance point, and the harmonic mean. However, in the case of methods based on probability grids, negative values of the directional coefficient were obtained for exponential distribution, lognormal distribution and density estimator.

To check whether differences are statistically significant for measurements taken for series longer than five events, the Friedman test was performed. This is a non-parametric test for repeated measurements (Field, Miles, Field, 2017). Since different profiling methods were tested on the same series, it is more adequate than a test for independent samples. The analysis performed on the hit score percentage variable showed no significant differences (Chi-squared = 15.78, $p = 0.11$), while the comparison for error distance turned out to be statistically significant (Chi-squared = 31.5, $p < 0.001$). The *post-hoc* test showed that there were significant differences between the median of the coordinates and most of the other tested methods (these were the mean of the coordinates, the distance from the centre of the circle, the harmonic mean of the coordinates, the exponential distribution, and the lognormal distribution).

Since the studied sample consisted mostly of short series and the hit score percentage values were higher than the predicted 15%, one more analysis was performed for the probability planes calculated on the basis of the exponential and lognormal functions. The analysis aimed to predict for what series length such accuracy will be obtained. It was decided to approximate the relationship between series length and accuracy calculated as hit score percentage using an exponential function. These two methods of creating geographic profiles were chosen because they gave some of the best results and because, for these methods, both ways of assessing profile accuracy indicated an improvement in prediction accuracy with series length. The approximate relationship is described by the equation:

$$\text{hit score percentage} = a \cdot e^{(-b \cdot \text{length of series})}$$

Then, after calculating the parameters a and b of such an exponential function, it was checked for which length of series the predicted hit score percentage would be below 15%. This resulted in a minimum of ten events for the exponential function and a minimum of thirteen events for the lognormal function. Note that the applied model assumes that the accuracy

will tend towards zero. However, such an assumption does not have to be fulfilled.

Discussion

Hypothesis 1 was partially confirmed. For the hit score percentage measure of accuracy, almost all tested methods showed improvement with the length of the series. This is due to the negative directional coefficients in the regression equations and is also evident in Figure 2. The results were not as clear for error distance. This may suggest two things: it is possible that this measure is less sensitive to change, which is consistent with the arguments presented in the theoretical section, namely that hit score percentage is a better method for evaluating geographic profiles. The second reason may result from the properties of the tested methods of creating geographic profiles. It is possible that the methods that showed progress with the length of the series – using both ways of evaluating accuracy – are the methods that show the best improvement.

Hypothesis 2 was also partially confirmed. The probability grid calculated on the basis of the exponential distribution had the best mean accuracy in terms of the hit score percentage for series consisting of 5 or more events. The median of the results was also the best for this method when accuracy was measured using the distance error method for series with a length of two and five or more events.

Hypothesis 3 was not confirmed. The obtained data were far from the hit score percentage value of 15%. The closest to this value were the results obtained for the profiling method based on the distance from the last event in the series, for series of 3–4 events. This was the smallest group, consisting of just 8 series, so it is likely that this result was due to chance.

The results obtained are ambiguous. On the one hand the obtained accuracy values were far from expected; the accuracy of the tested methods was lower than in the mentioned studies. On the other hand, there are elements of the research that give hope for the possibility of using geographic profiling in Poland. Firstly, almost every tested method showed improvement with increasing series length. This is evident both from the negative values of linear regression coefficients and the values presented in Table 2. Secondly, it should be remembered that the data in the *KSIP* database may potentially contain errors, which means that it is possible that with more complete data, better accuracy could be obtained.

In this context, the changes in prediction accuracy based on the first event in the series are interesting. If

the data were fully reliable, then this accuracy should not change with the length of the series of events. However, this is not the case; an improvement in the predictions is visible, which suggests that the shorter series are in fact shorter portions of series that in their entirety encompass more events. This, of course, conflicts with the recommendations of Rossmo (2000), who emphasised the need to have complete information about the events that make up a series. However, one must also take into account the arguments of Snook et al. (2005a), that in reality there is never such complete and certain information, and therefore a geographical profile should also be useful in situations of incomplete data. Thus, it is possible that the accuracy calculated in this way better reflects real situations where geographic profiling would be used.

In addition, it is worth noting the predicted series lengths at which a hit score percentage of the order of 15% should be expected. This value is much higher than the threshold of five events proposed by Rossmo (2000). In the two tested methods, for the *hit score percentage* to be at most 15%, the series should be a minimum of ten and thirteen events in length. Thus, it appears that it is necessary for the series to be much longer even assuming that the hit score percentage will approach zero with the length of the series. The method by which this value was calculated is obviously an attempt to extrapolate the data, but this and the other results show how important a factor series length is, and the five events proposed by Rossmo (2000) should be regarded as an absolute minimum. The results are also consistent with the research by Snook et al. (2005b) – similarly to their analysis, in the Polish population of perpetrators, for shorter series, more complicated methods based on probability grids did not have an advantage over simple centographic methods. It is only with longer series that such methods allow us to obtain better results.

It is worth repeating the performed analyses on more reliable data sets, as well as on longer series. Such analyses should make it possible to indicate the best methods for geographic profiling, and to determine more precisely for which series lengths accurate predictions can be obtained. Research on other types of events is of course also needed. The analyses conducted so far suggest that it is possible to create geographic profiles on the Polish population of perpetrators. The tested methods were relatively accurate; however, the obtained results were worse than expected, and the reasons for that can be sought in the incompleteness of the data. Therefore, based on the results obtained it is recommended to be extremely cautious when creating geographical profiles, pay special attention to

the length of the series and make sure that the data on which the predictions are based are complete.

References

1. Canter, D., Hammond, L. (2007). Prioritizing burglars: comparing the effectiveness of geographical profiling methods. *Police Practice and Research*, 8(4), 371–384. <https://doi.org/10.1080/15614260701615086>.
2. Canter, D., Hammond, L., Youngs, D., Juszczak, P. (2012). The efficacy of ideographic models for geographical offender profiling. *Journal of Quantitative Criminology*, 29(3), 423–446. <https://doi.org/10.1007/s10940-012-9186-6>.
3. Canter, D., Larkin, P. (1993). The environmental range of serial rapists. *Journal of Environmental Psychology*, 13(1), 63–69. doi:10.1016/s0272-4944(05)80215-4.
4. Chainey, S. (2008). *Crime mapping case studies: practice and research*. Chichester: John Wiley & Sons.
5. Chainey, S., Ratcliffe, J. (2008). *GIS and crime mapping*. Chichester: Wiley.
6. Field, A. P., Miles, J., Field, Z. (2013). *Discovering statistics using R*. SAGE Publications Ltd.
7. Goldschneider, M. (2012). Przestępczość na warszawskiej Pradze z perspektywy współczesnych teorii ekologicznych w kryminologii. *Archiwum Kryminologii*, 34, 207–252.
8. Górski, M. (2018a). The distance covered by perpetrators and their behaviour at the scene of the incident in the case of burglaries and robberies, *Problems of Forensic Sciences*, 113, 27–50.
9. Górski, M. (2018b). Buffer zone and circle theory on the example of burglaries and robberies, *Problems of Forensic Sciences*, 113, 51–66.
10. Levine, N. (2013). Chapter 13: Journey-to-crime estimation. (In) N. Levine, *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations* (v 4.0). Houston: Ned Levine Associates, Washington: National Institute of Justice. <http://www.nij.gov/CrimeStat>.
11. Mburu, L., Helbich, M. (2014). Evaluating the accuracy and effectiveness of criminal geographic profiling methods: The case of Dandora, Kenya. *The Professional Geographer*, 67(1), 110–120. <https://doi.org/10.1080/00330124.2014.886921>.
12. Mordwa, S. (2019). Profilowanie geograficzne, czyli wykorzystanie analiz przestrzennych do wykrywania sprawców przestępstw. *Space, Society, Economy*, 30, 51–75. <https://doi.org/10.18778/1733-3180.30.03>.
13. Paulsen, D. (2006). Human versus machine: a comparison of the accuracy of geographic profiling methods. *Journal of Investigative Psychology and Offender Profiling*, 3(2), 77–89. <https://doi.org/10.1002/jip.46>.

14. Paulsen, D. (2007). Improving geographic profiling through commuter/marauder prediction. *Police Practice and Research*, 8(4), 347–357. doi:10.1080/15614260701615045.
15. Rich, T., Shively, M. (2004). *A methodology for evaluating geographic profiling software*. Cambridge, MA: Abt Associates.
16. Rossmo, D. K. (2005a). Geographic heuristics or shortcuts to failure?: response to Snook et al. *Applied Cognitive Psychology*, 19(5), 651–654. <https://doi.org/10.1002/acp.1144>.
17. Rossmo, D. K. (2005b). An evaluation of NIJ's evaluation methodology for geographic profiling software, 9.03.2005. Retrieved June 30, 2021 from: <http://www.txstate.edu/gii/geographic-profiling/publications/contentParagraph/01/document/Response%20to%20NIJ%20GP%20Evaluation%20Methodology.doc>.
18. Snook, B., Canter, D., Bennell, C. (2002). Predicting the home location of serial offenders: a preliminary comparison of the accuracy of human judges with a geographic profiling system. *Behavioral Sciences & the Law*, 20(1–2), 109–118. <https://doi.org/10.1002/bsl.474>.
19. Snook, B., Taylor, P. J., Bennell, C. (2003). Geographic profiling: the fast, frugal, and accurate way. *Applied Cognitive Psychology*, 18(1), 105–121. <https://doi.org/10.1002/acp.956>.
20. Snook, B., Taylor, P. J., Bennell, C. (2005a). Shortcuts to geographic profiling success: a reply to Rossmo (2005). *Applied Cognitive Psychology*, 19(5), 655–661. <https://doi.org/10.1002/acp.1142>.
21. Snook, B., Zito, M., Bennell, C., Taylor, P. J. (2005b). On the complexity and accuracy of geographic profiling strategies. *Journal of Quantitative Criminology*, 21(1), 1–26. <https://doi.org/10.1007/s10940-004-1785-4>.
22. Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Corresponding author

Michał Górski
College of Inter-Faculty Individual Studies in
Mathematics and Natural Sciences, University of
Warsaw
ul. Stefana Banacha 2C
PL 02-097 Warszawa
e-mail: mgorski@student.uw.edu.pl

TRAFNOŚĆ METOD PROFILOWANIA GEOGRAFICZNEGO NA PRZYKŁADZIE WŁAMAŃ NA TERENIE WARSZAWY

Profilowanie geograficzne ostatnimi laty cieszy się coraz większym zainteresowaniem, pojawiają się także nowe publikacje polskojęzyczne poruszające ten temat. Zazwyczaj obejmują one problemy związane z profilowaniem geograficznym, takie jak rozkład odległości pokonanej przez sprawców (np. Goldschneider, 2012; Górski, 2018a), lub są to tylko teoretyczne publikacje (np. Mordwa, 2019). Nie ma jednak badań sprawdzających trafność profilowania geograficznego na polskiej próbie. Jak wskazują Mburu i Helbich (2014), takie badania nie istnieją poza wysoko rozwiniętymi krajami zachodnimi. Autorzy wymieniają USA, Kanadę, Wielką Brytanię i jedno badanie w Niemczech, natomiast sami wykonali swoje analizy na populacji kenijskiej. Zatem w Europie kontynentalnej brakuje tego typu analiz, a aktualne badanie ma na celu wypełnić tę lukę.

Profilowanie geograficzne to zbiór metod mających na celu przewidzenie miejsca przebywania sprawcy serii przestępstw na podstawie lokalizacji miejsc zdarzeń (Rossmo, 2000; Canter, Hammond, Youngs, Juszczak, 2012). Są to metody optymalizacji wskazujące najbardziej prawdopodobny obszar, co w konsekwencji ma przyspieszyć proces znajdowania sprawcy. Zastosowanie profilowania po pierwsze pozwala obniżyć koszty poszukiwań poprzez minimalizację obszaru (bądź liczby osób typowanych), jaki trzeba sprawdzić do momentu znalezienia faktycznego sprawcy (Chainey, 2008). Po drugie umożliwia znalezienie sprawcy zanim dopuści się on kolejnego przestępstwa, co jest wyjątkowo istotne w przypadku przestępstw ciężkich (Chainey, Ratcliffe, 2008).

Teorie będące podstawą profilowania geograficznego zostały opisane wielokrotnie w polskojęzycznej literaturze (np. Mordwa, 2019), także na łamach czasopisma „Problems of Forensic Sciences” (Górski, 2018a). Natomiast konieczne jest opisanie działania programów do tworzenia profili geograficznych (metody ilościowe profilowania geograficznego) oraz metod oceny trafności przewidywań tworzonych przez takie narzędzia.

Oprogramowanie do tworzenia profili geograficznych

Aby przybliżyć działanie oprogramowania do tworzenia profili geograficznych, trzeba opisać koncepcję funkcji oporu przestrzeni. Jest to koncepcja wywodząca się z analiz poruszania się ludności w różnych (niezwiązanych z przestępczością) celach (Levine, 2013). Funkcja

oporu przestrzeni to rozkład odległości, jakie osoby pokonują od miejsca zamieszkania do danego celu (Levine, 2013). W kontekście przestępczości wprowadzono jednak dwie modyfikacje. Po pierwsze – miejsce zamieszkania uogólniono jako punkt zakotwiczenia sprawy (Chainey, Ratcliffe, 2008), czyli lokalizację, z której sprawca dokonuje swoich przestępstw (niekoniecznie miejsce zamieszkania; Rossmo, 2000; Chainey, 2008). Po drugie – wprowadzono koncepcję strefy buforowej, czyli obszaru w bezpośrednim sąsiedztwie miejsca zamieszkania sprawcy, w którym prawdopodobieństwo dokonania przestępstwa jest mniejsze (Canter, Larkin, 1993).

Programy do tworzenia profili geograficznych działają w następujący sposób: cały obszar poszukiwań jest podzielony na małe fragmenty ułożone w szachownicę; dla każdego punktu z takiej siatki wyliczane jest prawdopodobieństwo znalezienia sprawcy na podstawie odległości do lokalizacji miejsc zdarzeń oraz funkcji oporu przestrzeni (Chainey, Ratcliffe, 2008). W związku z tym na podstawie odległości danego punktu do każdej z lokalizacji zdarzeń można wyliczyć prawdopodobieństwo pokonania przez sprawcę odległości z danego miejsca. To w konsekwencji przekłada się na prawdopodobieństwo, że tam znajduje się punkt zakotwiczenia sprawy. Ponieważ takie wyliczenia dotyczą całej siatki, powstaje dzięki temu płaszczyzna prawdopodobieństwa, a punkty o największych wartościach wskazują obszary o najwyższym prawdopodobieństwie znalezienia sprawcy (Chainey, Ratcliffe, 2008). Metody tego typu będą dalej nazywane opartymi na siatce bądź płaszczyźnie prawdopodobieństwa (Mburu, Helbich, 2014).

Niektóre dużo prostsze metody, zwane centrograficznymi, polegają na wskazaniu jednego punktu, od którego należy rozpocząć poszukiwania (Mburu, Helbich, 2014). Te punkty to na przykład średnia, mediana przestrzenna lub pierwsze albo ostatnie znane zdarzenie w serii. Aby ocenić, jakie metody są trafniejsze, trzeba najpierw omówić różne metody oceny trafności profili.

Ocena trafności profilu geograficznego

W literaturze przyjęto dwie podstawowe metody oceny trafności metod profilowania geograficznego. Jedna polega na ustaleniu odległości punktu o największym prawdopodobieństwie do faktycznego miejsca zamieszkania sprawcy (*error distance*; Rich, Shively, 2004). Umożliwia ona porównanie metod polegających na

tworzeniu płaszczyzny prawdopodobieństwa z metodami opierającymi się na wskazaniu tylko jednego punktu jako miejsca rozpoczęcia poszukiwań. Ponieważ jednak przy zastosowaniu metod opartych na funkcji oporu przestrzeni prawdopodobieństwo nie spada tak samo w każdym kierunku, to miara *error distance* nie pozwala brać tego pod uwagę (Rossmo, 2005b). Możliwe jest nawet, że dwa (lub więcej) obszary zostaną wskazane jako przewidywane miejsca zakotwiczenia sprawcy. W takiej sytuacji jeśli sprawca znajduje się w nieznacznie mniej prawdopodobnym obszarze, to intuicyjnie profil można by ocenić jako trafny (Rossmo, 2005b). Wbrew tej intuicji wartość *error distance* może być duża, a w konsekwencji profil oceniony jako nietrafny. Przykład takiej sytuacji został zaprezentowany przez Rossmo (2005b).

Między innymi po to, by rozwiązać ten problem, stosuje się drugą metodę oceny trafności profili, tzw. *hit score percentage*. Polega ona na podzieleniu powierzchni obszaru, jaki należy sprawdzić (od najbardziej prawdopodobnego punktu aż do znalezienia faktycznego sprawcy), przez pole powierzchni całego obszaru poszukiwań (Rich, Shively, 2004; Rossmo, 2000, 2005b). Takie podejście rozwiązuje opisany wcześniej problem, ale też lepiej określa stopień, w jakim profil pozwolił ograniczyć początkowy obszar poszukiwań. Np. *hit score percentage* wynoszący 5% oznacza, że jeśli poszukiwano by sprawcy zgodnie z prawdopodobieństwem, należałoby sprawdzić 5% całego obszaru poszukiwań do czasu znalezienia właściwej osoby. Na ten moment nie ma porozumienia w tej kwestii: istnieją zwolennicy wszystkich wymienionych podejść, ale także innych metod będących powiązaniem powyższych (Rich, Shively, 2004; Paulsen, 2006).

Minimalna długość serii zdarzeń

Rossmo (2000) stawia warunek minimum pięciu powiązanych zdarzeń, aby stworzone przewidywania były trafne i by można było wykorzystać profil geograficzny. Wartość tych pięciu zdarzeń opiera na wynikach symulacji komputerowych. Z taką granicą nie zgadzają się Snook, Taylor i Bennell (2005a), wskazując na słabości zastosowanej przez Rossmo metody i twierdząc, że jest to niewystarczające uzasadnienie ustalenia granicy minimum pięciu zdarzeń. Co więcej, autorzy argumentują, że przy przyjętej przez Rossmo (2000) metodologii prawdziwa minimalna długość serii powinna wynosić 17 zdarzeń. Zwracają także uwagę na to, że ustawienie tak wysokiej granicy może powodować, że profilowanie geograficzne będzie przydatne w bardzo niewielu sytuacjach.

Badania nad trafnością profilowania geograficznego

Nie ma w Polsce badań dotyczących trafności metod algorytmicznych profilowania geograficznego. Wynika to najprawdopodobniej z ograniczonej dostępności zarówno oprogramowania, jak i braku odpowiednich danych do przeprowadzenia takich analiz. Spośród zagranicznych badań warto zwrócić uwagę na pracę Paulsena (2006) oraz Snooka, Zito, Bennella i Taylora (2005b). W pierwszej z wymienionych autor porównywał trafność stosowanych metod z oceną wcześniej przeszkolonych osób i ich predykcji bez wsparcia programów komputerowych. Dane w tych badaniach składały się z 25 serii różnego typu zdarzeń, natomiast jeśli wziąć pod uwagę przewidywania uczestników badania, to faktyczni sprawcy znajdowali się w najbardziej prawdopodobnych 5% przestrzeni rzadziej niż w analogicznym obszarze wyliczonym na podstawie algorytmów (8% trafności wskazań uczestników, 48% dla rozkładu wykładniczego; Paulsen, 2006). Ta przewaga oprogramowania zniknęła, kiedy wykonano taką analizę dla top 15% powierzchni – w tych warunkach wskazania uczestników były porównywalne z niektórymi rezultatami wykorzystania programów komputerowych.

Snook i in. (2005b) w swoich badaniach wykazali, że metody centrograficzne dla krótszych serii mogą dawać takie same (lub lepsze) rezultaty jak metody oparte na siatkach prawdopodobieństwa. Autorzy sprawdzali trafność różnych metod w zależności od długości serii. W większości badanych warunków najtrafniejszą metodą było wyznaczenie środka okręgu, którego średnicą jest odcinek łączący dwa najodleglejsze od siebie nawzajem zdarzenia. Dopiero dla serii długości dziesięciu zdarzeń trafniejsze były przewidywania na podstawie punktu, którego suma odległości do wszystkich zdarzeń jest najmniejsza, natomiast dla serii długości dziesięciu zdarzeń najtrafniejsze były przewidywania na podstawie liniowej funkcji oporu przestrzeni (Snook i in., 2005b). Zatem dopiero od serii długości dziesięciu zdarzeń optymalne było wykorzystanie metod wyliczających całą płaszczyznę prawdopodobieństwa. Wcześniej inne, prostsze metody były optymalne. Aby zastosować większość metod centrograficznych, nie potrzeba oprogramowania, wystarczy mapa, ołówek i linijka. Zatem jeśli te metody są wystarczające, to potrzeba stosowania oprogramowania jest wątpliwa. W teście przeprowadzonym na próbie serii kradzieży z włamaniem Canter i Hammond (2007) wykazali, że w zależności od stosowanej metody w top 5% płaszczyzny prawdopodobieństwa znajduje się od 72% do 46% sprawców. Jest to znacznie więcej niż uzyskał Paulsen (2006), ale w przeciwieństwie do wyników uzyskanych przez Snooka i in. (2005b) posłużenie się środkiem okręgu okazało się najmniej trafną metodą.

Warto zwrócić uwagę, że w badaniach Paulsena (2006) zastosowanie siatki prawdopodobieństwa wyliczonej na podstawie rozkładu wykładniczego było jedną z najlepszych metod. Przy wykorzystaniu tej metody najczęściej sprawców znajdowało się w top 5% profilu geograficznego, natomiast w top 15% profilu lepsze wyniki uzyskały tylko przewidywania z wykorzystaniem programów Rigel i Dragnet. Także w badaniach Cantera i Hammond (2007) funkcja wykładnicza dawała jedne z najlepszych wyników.

Aktualne badania

Głównym celem aktualnych badań było ustalenie, jak sprawdzają się w polskiej populacji sprawców różne metody algorytmiczne profilowania geograficznego w zależności od długości analizowanej serii zdarzeń. Jak już wcześniej wspomniano, tego typu badania są rzadkie, a w Polsce nie miały dotąd miejsca. W związku z powyższym postawiono następujące hipotezy badawcze:

Hipoteza 1. Większość badanych metod profilowania geograficznego będzie dawała lepsze wyniki trafności wraz ze zwiększaniem długości serii.

Hipoteza 2. Rozkład wykładniczy będzie najlepszą metodą przewidywania miejsca zamieszkania sprawcy.

Hipoteza 3. Dla najlepszych metod profilowania geograficznego średni *hit score percentage* będzie mniejszy niż 15%.

Funkcja wykładnicza była przewidywana jako dająca najlepsze wyniki ze względu na badania Cantera i Hammond (2007), którzy analizowali podobną grupę sprawców, oraz na podstawie wyników Paulsena (2006), w których okazała się jedną z najlepszych metod. Granica 15% została wybrana także ze względu na badania Paulsena (2006), do których porównanie umożliwi ocenę, czy testowane metody dają lepsze wyniki trafności niż te, w których wyniki uzyskuje się bez wsparcia metod komputerowych. Także Rossmo (2000) uzyskiwał *hit score percentage* 15% lub mniej w prawie wszystkich badanych seriach poza dwoma przypadkami. Wszystkie analizy wykonano w programie R 3.6.3, natomiast wykresy w pakiecie ggplot2 (Wickham, 2009).

Dane

Dzięki współpracy z Komendą Główną Policji pozyskano dane z systemu KSIP, które zawierały 10 634 kradzieży z włamaniem z 2011 roku. Spośród nich wybrano te zdarzenia, do których byli przypisani sprawcy. Dane pozyskane z KSIP pozwalają na uzyskanie dużej ilości materiału, ale z drugiej strony posługiwanie się nimi wiąże się z potencjalnymi niedokładnościami. Wynika

to z faktu, że nie są one tworzone w oparciu o wyroki sądowe. Na potrzeby aktualnych badań za serię uznawano dwa lub więcej zdarzeń mających wspólnego sprawcę bądź sprawców. Ponieważ możliwe jest, że przestępcy współpracowali tylko przy niektórych zdarzeniach, a przy innych nie, każdego sprawcę traktowano jako oddzielną serię z punktami zakotwiczenia przypisanymi tylko temu jednemu sprawcy. Może to negatywnie wpływać na ocenę trafności, gdyż grupa sprawców mogła działać w obrębie miejsca zamieszkania jednego z nich (osoby decyzyjnej, kierującej daną grupą) – w takiej sytuacji profilowanie może trafnie przewidywać miejsce zamieszkania tego jednego sprawcy, ale błędnie pozostałych. W ramach badań nie było dostępu do akt spraw, nie było więc możliwości ustalenia, kto w danej grupie byłby „głównym” sprawcą. Ponieważ nie można było rozwiązać problemu możliwości zmiany składu grup wraz z biegiem czasu, uznano, że traktowanie każdego sprawcy jako oddzielnej serii jest uzasadnione.

Jeśli jednak jakiemuś sprawcy przypisana była więcej niż jedna lokalizacja, to testując każdą z metod profilowania geograficznego, wybierano najlepszy wynik. Było to podyktowane względami praktycznymi – jeśli sprawca zostałby wykryty dzięki pierwszej (bardziej prawdopodobnej) lokalizacji, to dalsze poszukiwania nie miałyby sensu, a dołączenie pozostałych miejsc w analizie jest bezpodstawne i tylko zaburzyłoby wyniki.

W ostatecznej analizie nie uwzględniono sprawców, którzy mieszkali poza obszarem poszukiwań, co wynikało z tego, że skoro celem badań jest porównanie metod profilowania geograficznego, a sprawca znajduje się poza obszarem poszukiwań, to każda stosowana metoda będzie nietrafna. Uwzględnienie więc takich zdarzeń będzie tylko zaburzać wyniki, co więcej, jak wskazują badania Paulsena (2007), a także badania na polskiej populacji sprawców (Górski, 2018b), możliwe jest przewidywanie, czy sprawca jest tzw. sprawcą stabilnym. Oznacza to, że możliwe jest odrzucenie sprawców niemieszkających w obrębie obszaru poszukiwań jeszcze przed ich wykryciem, a w pozostałych przypadkach należałoby stosować profilowanie geograficzne.

Ostatecznie podzielono serie na trzy grupy na podstawie długości serii. Do pierwszej z nich należały serie składające się z dwóch zdarzeń, do drugiej – z trzech lub czterech, a ostatnia składała się z serii pięciu lub więcej zdarzeń. Do tych grup należały odpowiednio 15, 8 i 15 serii zdarzeń, w sumie 38 serii kradzieży. Zaledwie 15 wykrytych serii składa się z pięciu lub więcej zdarzeń. Tak mała liczebność próby nie jest niespotykana w tego typu badaniach; jak wskazali Mburu i Helbich (2014) – często takie analizy są wykonywane na jeszcze mniejszych próbach. Snook i in. (2002) oraz Snook i in. (2003) przeprowadzili badania na próbach dziesięciu serii w obu przypadkach, natomiast Rossmo (2000) – na próbie piętnastu serii zdarzeń.

Analizy

W analizach uwzględniono w sumie 11 metod wspomagających profilowanie geograficzne: siedem metod centrograficznych oraz cztery oparte na siatkach prawdopodobieństwa. Dokładną listę badanych metod wraz z ich krótkim opisem można znaleźć w tabeli 1. Za obszar poszukiwania sprawcy uznawano obszar jego działalności powiększony o około 20%. W badaniach uwzględniono dwie opisywane metody oceny trafności profilu geograficznego: odległość najbardziej prawdopodobnego punktu do miejsca zamieszkania sprawcy (*error distance*) oraz procent przestrzeni bardziej prawdopodobny od miejsca zamieszkania sprawcy (*hit score percentage*).

Ponieważ w analizie wzięto pod uwagę metody oparte na siatkach prawdopodobieństwa, konieczne było, aby w pierwszym kroku obliczyć parametry funkcji oporu przestrzeni. W tej analizie uwzględniono rozkład pokonanej odległości dla wszystkich zdarzeń sprawców seryjnych, a nie tylko sprawców stabilnych bądź znajdujących się w obrębie obszaru poszukiwań, gdyż takie ograniczenie było bez znaczenia dla tej analizy. Analiza dotyczyła 416 zdarzeń, a po odrzuceniu tzw. *outlierów* pozostało ich 381 – i to na ich podstawie wyliczono parametry trzech rozkładów pokonanej odległości (wykładniczego, normalnego oraz lognormalnego). Rozkłady zobrazowano wraz z jądrowym estymatorem gęstości (wyliczonym na podstawie tej samej próby) na wykresie 1.

W następnym kroku stworzono profile geograficzne dla wszystkich analizowanych serii oraz wykonano analizę ich trafności dwiema opisanymi metodami. Wyniki podzielono na trzy grupy w zależności od długości serii. Dla tak wykalkulowanych ocen trafności wyliczono trzy miary: średnią, medianę¹ oraz odchylenie standardowe. Najważniejszą z wymienionych wydaje się mediana, gdyż wskazuje ona, jak duży procent przestrzeni należałoby sprawdzić, aby znaleźć 50% sprawców. Wyniki zaprezentowano w tabeli 2. Pogrubione zostały najlepsze wartości w każdej kolumnie, czyli metoda profilowania, która przy danej długości serii była optymalna, gdy się wzięło pod uwagę daną miarę trafności. Należy zaznaczyć, że dla obu miar trafności profilu im mniejsza wartość, tym trafniejszy był profil.

Następnie wyniki obu miar trafności zaprezentowano na wykresach: na wykresie 2 wyniki opisujące trafność mierzoną jako *hit score percentage*, natomiast na wykresie 3 – trafność mierzoną jako *error distance*.

¹ Czytelnik może momentami czuć się zagubiony, gdyż zarówno średnia, jak i mediana pojawiają się dwukrotnie. Raz jako centrograficzne metody tworzenia profilu geograficznego, a potem ponownie – jako statystyki otrzymanych wyników. Dla jasności wywodu pierwsza sytuacja jest opisywana jako średnia/mediana współrzędnych, a druga jako średnia/mediana wyników.

Aby sprawdzić, jak wygląda zależność trafności od długości serii, wyliczono współczynniki regresji liniowej dla każdej metody tworzenia profilu geograficznego. W przypadku oceny trafności *hit score percentage* wszystkie współczynniki kierunkowe były ujemne poza medianą współrzędnych. W przypadku *error distance* ujemne wartości współczynnika kierunkowego uzyskano dla odległości od: pierwszego zdarzenia w serii, średniej współrzędnych, punktu minimalnej odległości, średniej harmonicznej. Natomiast w przypadku metod opartych na siatkach prawdopodobieństwa ujemne wartości współczynnika kierunkowego uzyskano dla rozkładu wykładniczego, rozkładu lognormalnego oraz estymatora gęstości.

Aby sprawdzić, czy różnice są statystycznie istotne dla pomiarów wykonanych dla serii dłuższych niż pięć zdarzeń, wykonano test Friedmana. Jest to test nieparametryczny dla powtórzonych pomiarów (Field, Miles, Field, 2017). Ponieważ różne metody profilowania testowano na takich samych seriach, jest on bardziej adekwatny niż test dla prób niezależnych. Analiza wykonana na zmiennej *hit score percentage* wykazała brak istotnych różnic (chi-kwadrat = 15,78, $p = 0,11$), natomiast porównanie dla *error distance* okazało się statystycznie istotne (chi-kwadrat = 31,5, $p < 0,001$). Test *post-hoc* wykazał, że istotne różnice są między medianą współrzędnych a większością pozostałych testowanych metod (były to: średnia współrzędnych, odległość od środka okręgu, średnia harmoniczna współrzędnych, rozkład wykładniczy, rozkład lognormalny).

Ponieważ badana próba składała się w większości z krótkich serii, a wartości *hit score percentage* były wyższe od przewidywanych 15%, wykonano jeszcze jedną analizę dla płaszczyzn prawdopodobieństwa wyliczanych na podstawie funkcji wykładniczej oraz lognormalnej. Analiza ma na celu przewidywanie, dla jakiej długości serii uzyskamy taką trafność. Postanowiono przybliżyć zależność między długością serii a trafnością liczoną jako *hit score percentage* za pomocą funkcji wykładniczej. Wybrano te dwie metody tworzenia profili geograficznych, ponieważ dawały one jedne z najlepszych wyników oraz ze względu na to, że dla tych metod oba sposoby oceny trafności profili wskazywały poprawę trafności przewidywań wraz z długością serii. Przybliżoną zależność opisuje równanie:

$$\text{hit score percentage} = a \cdot e^{(-b \cdot \text{długość serii})}$$

Następnie, po obliczeniu parametrów a i b takiej funkcji wykładniczej, sprawdzono, dla jakiej długości serii przewidywany *hit score percentage* będzie poniżej 15%. W efekcie dla funkcji wykładniczej uzyskano minimum dziesięć zdarzeń, natomiast dla lognormalnej – minimum trzynaście. Należy zwrócić uwagę, że wykorzystany model zakłada, że trafność będzie zbiegać do zera. Jednak założenie takie nie musi być spełnione.

Dyskusja

Hipoteza 1 została częściowo potwierdzona. Dla miary trafności *hit score percentage* prawie wszystkie testowane metody wykazywały poprawę wraz z długością trwania serii. Wynika to z ujemnych współczynników kierunkowych w równaniach regresji, jest również widoczne na wykresie 2. Wyniki nie były tak jednoznaczne dla *error distance*. Może to sugerować dwie rzeczy: możliwe jest, że ta miara jest mniej wrażliwa na zmiany, co jest spójne z argumentami zaprezentowanymi w części teoretycznej, mianowicie że *hit score percentage* jest lepszą metodą oceny profili geograficznych. Drugi powód może wynikać z właściwości testowanych metod tworzenia profili geograficznych. Możliwe jest, że metody, które wykazywały postęp wraz z długością serii przy obu sposobach oceny trafności, są metodami, które wykazują najlepszą poprawę.

Hipoteza 2 także została częściowo potwierdzona. Siatka prawdopodobieństwa wyliczona na podstawie rozkładu wykładniczego miała najlepszą średnią trafności *hit score percentage* dla serii składających się z 5 lub więcej zdarzeń. Mediana wyników była także najlepsza dla tej metody, kiedy trafność była mierzona metodą *distance error* dla serii długości dwóch oraz pięciu lub więcej zdarzeń.

Hipoteza 3 nie została potwierdzona. Uzyskane dane były dalekie od wartości *hit score percentage* wynoszącej 15%. Najbliższe tej wartości były wyniki uzyskane dla metody profilowania opartej na odległości od ostatniego zdarzenia w serii przy seriach długości 3–4 zdarzeń. Była to grupa najmniej liczna, składająca się z zaledwie 8 serii, jest zatem prawdopodobne, że wynik taki był efektem przypadku.

Uzyskane wyniki są niejednoznaczne. Z jednej strony uzyskano wartości trafności dalekie od oczekiwanych, a trafność testowanych metod była niższa niż w przytoczonych wcześniej badaniach. Z drugiej strony są elementy przeprowadzonych badań, które dają nadzieję na możliwość stosowania profilowania geograficznego w Polsce. Po pierwsze prawie każda testowana metoda wykazywała poprawę wraz ze wzrostem długości serii. Wynika to zarówno z ujemnych wartości współczynników regresji liniowych, jak i wartości zaprezentowanych w tabeli 2. Po drugie – należy pamiętać, że dane w bazie KSIP mogą potencjalnie zawierać błędy, co oznacza możliwość, że przy pełniejszych danych uzyskano by lepszą trafność.

W tym kontekście interesujące są zmiany trafności przewidywań na podstawie pierwszego zdarzenia w serii. Jeśli dane byłyby w pełni wiarygodne, to trafność ta nie powinna się zmieniać wraz z długością serii zdarzeń. Jednakże tak nie jest, widoczna jest poprawa przewidywań, co sugeruje, że w przypadku krótszych serii w rzeczywistości są to krótsze fragmenty serii, które w całości

obejmują więcej zdarzeń. Jest to w oczywistej sprzeczności z zaleceniami Rossmo (2000), który akcentował konieczność posiadania pełnej informacji o zdarzeniach wchodzących w skład danej serii. Trzeba jednak brać także pod uwagę argumenty Snooka i in. (2005a), że w rzeczywistości nigdy nie ma takiej pełnej i pewnej informacji, a zatem profil geograficzny powinien być przydatny także w sytuacjach niekompletnych danych. Możliwe więc, że trafność wyliczona w ten sposób bardziej odpowiada realnym sytuacjom, w których profilowanie geograficzne byłoby stosowane.

Dodatkowo należy zwrócić uwagę na przewidywane długości serii, przy których należy spodziewać się *hit score percentage* rzędu 15%. Jest to wartość dużo wyższa od proponowanej przez Rossmo (2000) granicy pięciu zdarzeń. W dwóch testowanych metodach aby *hit score percentage* wynosił najwyżej 15%, serie powinny być długości minimum dziesięciu i trzynastu zdarzeń. Zatem okazuje się, że konieczne jest, aby serie były dużo dłuższe nawet przy założeniu, że współczynnik *hit score percentage* będzie zbliżał się do zera wraz z długością serii. Metoda, przy pomocy której wyliczono te wartości, to oczywiście próba ekstrapolacji danych, jednak zarówno to, jak i pozostałe wyniki pokazują, jak ważnym czynnikiem jest długość serii, a proponowane przez Rossmo (2000) pięć zdarzeń należy traktować jako absolutne minimum. Wyniki są także spójne z badaniami Snooka i in. (2005b) – podobnie jak w przeprowadzonych tam analizach także w polskiej populacji sprawców dla krótszych serii bardziej skomplikowane metody, oparte na siatkach prawdopodobieństwa, nie dawały przewagi nad prostymi metodami centrograficznymi. Dopiero przy dłuższych seriach metody takie pozwalają uzyskać lepsze wyniki.

Warto powtórzyć przeprowadzone analizy na pewniejszych zbiorach danych, a także na dłuższych seriach. Takie analizy powinny pozwolić na wskazanie najlepszych metod do tworzenia profili geograficznych, a także na bardziej precyzyjne ustalenie, dla jakich długości serii trafne przewidywania mogą być uzyskane. Badania nad innego typu zdarzeniami oczywiście także są potrzebne. Dotychczas przeprowadzone analizy sugerują, że można tworzyć profile geograficzne na polskiej populacji sprawców. Testowane metody były relatywnie trafne, jednak uzyskane wyniki były gorsze od spodziewanych, a przyczyn tego upatruje się w niekompletności danych. Dlatego na podstawie uzyskanych wyników zaleca się wyjątkową ostrożność przy tworzeniu profili geograficznych, zwracanie szczególnej uwagi na długość serii oraz upewnienie się, że dane, na podstawie których tworzone są przewidywania, są kompletne.

