



ANALYTICAL MAPPING OF REGISTERED CRIMINAL ACTIVITIES IN VILNIUS CITY

Giedrė Beconytė¹, Agnė Eismontaitė², Denis Romanovas³

*Centre for Cartography, Vilnius University,
M. K. Čiurlionio g. 21, LT-03101 Vilnius, Lithuania
E-mail: ¹giedre.beconyte@gf.vu.lt (corresponding author)*

Received 03 October 2012; accepted 12 December 2012

Abstract. The paper describes the method and the basic results of research into data on criminal activities in Vilnius city. Approximately 100000 incidents registered by police in both 2010 and 2011 have been located and geocoded using their street address information. Analysis of territorial distribution of the incidents in general and of the most common types (assaults, robberies and thefts, motor vehicle thefts and minor offences) in two years reveals that there exist higher concentration areas for all types of crimes. Over one year such areas grew but generally retained their shape and location. The density of incidents is generally dependent on population density, but also tends to concentrate around some shopping centres and entertainment areas. Kernel density spatial analysis method produces visually expressive results and should be applied for fast visual analysis and comparison of data.

Keywords: crime, delinquency, geographic data, density map, location quotient, assault, theft, city.

1. Introduction

Distribution and concentration of crimes and delinquency have always been important in understanding city life. Geographic approach has been successfully applied since early 20th century in the U.S. when first (non-digital) crime distribution maps were made. New research methods have been developed (Boba 2005; Bruce 2008) and more intensely applied in order to support or to disprove theories about differentiation of criminal activities between city districts, such as social disorganization theory explaining street crime levels by characteristics of neighbourhood (Zhang, Peterson 2007). More intense research into geography of crimes began in the second half of the 20th century as computers and GIS technology allowed processing large volumes of geographic data and efficient visualizations of the results (Maltz *et al.* 2000). Traditionally, such research is still more popular in the U.S. where stronger differentiation between city zones is observed and processes of succession are much faster than in Central or Eastern Europe.

The main purpose of the research on crime and delinquency distribution conducted in Vilnius University in 2009–2012 is to reveal the spatial pattern of overall distribution in Vilnius City and distribution of different types of offences. We expected to observe a spatial trend of change of crime rates in 2011. We did not have initial hypothesis about the location and number of highest concentration areas, though it had been anticipated that they would match neither the pattern with crime rates spatially dispersing from the city centre outward nor the opposite.

2. Data and data processing

Data on criminal activities in Vilnius City in 2010 and 2011 were obtained from the registry of Vilnius County Police headquarters. The dataset of 2010 did not include Grigiškės. The incident record typically consists of street address, number of injured/fatalities, date and time when information was submitted to police offices or by 112 (common emergency telephone number). The address information had not been geocoded and used for spatial analysis.

The incidents were initially classified into types but different sets of type values had been used in 2010 and 2011. The authors grouped several types of incidents into four large types: assaults, burglaries and thefts, motor vehicle thefts and minor offences. These four types together cover about 58.8 percent of all analysed incidents of each year (Fig. 1). Due to their different character (thus different factors that influence the spatial pattern) each of the four types was analysed separately.

Geographic coordinates of the registered incidents were determined using street address information. Vilnius city address database containing 48674 address points was used as reference dataset.

Because number of registered incidents was very large, the geocoding process had to be automated. Initially we attempted to employ ArcGIS *Geocoding* tool but lack of proper address locators and a poor documentation made it inefficient. Then *Google Geocoding API* was tested, but it could not be used due to its limitations of

use: *Google Geocoding API* may only be used in conjunction with a *Google map* and query limit of 2500 location requests per day exists. In order to avoid such complications a custom geocoding program was written in Python.

The original addresses where incidents had been registered were already split into components: city name, street name and optionally house number, all stored in separate table columns. Such structure facilitated identification of the address components. Street names was the most problematic part of the address data due to misspellings and inconsistency of spelling of the compound names, particularly personal names that in different records would or would not include the title, first or second given name or abbreviation. They all had to be transformed into the same form.

In order to produce better address matching results, reference dataset was altered in this way: city type was separated from its name (same was done with street name) and city name was reduced to its root. Reference dataset was also populated by two new fields: city name soundex field and street name soundex field. That was necessary to assure that small misspelling error in address did not result in a false match.

Then every address record was processed as follows:

1. Address information was standardized. If address city name included city type, it was separated from the city name string and was standardized, e.g., “m.”, “miest.”, “miestas”, “mieste” (different forms and abbreviations of “town”) into “miestas” (nominative of “town”). The same was done with the street names. Address strings were converted into lowercase.
2. Soundex representations of city name and street name were calculated.
3. The reference dataset was queried for address soundex representation and all matches were returned.
4. The best matching address out of the returned list was computed using Levenshtein algorithm for measuring the amount of difference between the

strings (Levenshtein 1965) and corresponding co-ordinates were returned.

5. If registered incident address contained house number and the latter was found in the reference dataset, the coordinates of the address point were returned. If registered incident address contained house number, but such number was not in the reference dataset, linear interpolation was made between two presumably closest addresses on the same street. If registered incident address did not contain house number, location of a random house of the matched city street was returned.

About 99.4 percent of the 97912 and 117417 incidents that occurred correspondingly in 2010 and 2011 were successfully located. 84.7 percent of incidents were matched to address points using exact incident address point co-ordinates, 5.0 percent of incidents were located approximately by interpolating nearest address point co-ordinates along the street where incident had occurred and 9.7 percent of incidents were located by randomly choosing incident's street address. Only 1186 incidents could not be even approximately located.

Four main reasons why the incidents could not be located are:

1. Too many misspells in the address field (e.g., “Vilniaus m. UKMR”, “Vilniaus m. Rodūnioskelias”, “Vilniaus m. Ghelvonų”).
2. Other locator than address was used (e.g., “Vilniaus m. centras”, “Vilniaus m. s.b. Vyturys”).
3. Vilnius city address database did not contain address points for the incident street.
4. Matching algorithm resulted in a wrong match. The false matches were identified manually by scanning the address matching log.

Some records have been located manually. Mainly unclassified incidents with completely incorrect location information could not be matched.

After additional filtering (some located incidents were outside Vilnius City), the total of 97812 incidents of 2010 and 116997 incidents of 2011 were used in further calculations.

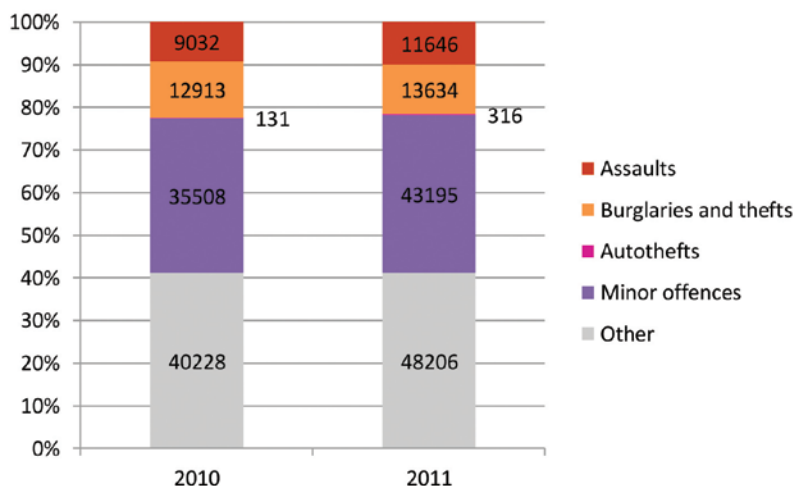


Fig. 1. Incidents by type in 2010 and 2011 in all districts of Vilnius City

3. Methods of calculation

The probability of density of incidents has been estimated using spatial kernel density method based on the quadratic kernel function (Silverman 1986: 76) and a corresponding ArcGIS tool that calculates the density of point features in a neighbourhood around each cell of an output raster (Chainey, Ractliffe 2005; Gibin *et al.* 2007). Whereas in a simple point density calculation points that fall within the search area are summed, and divided by the search area size to get density value for each cell, kernel density also evaluates influence of the occurrences registered in the neighbourhood. We used 1500 meter neighbourhood radius and the 60×60 meter cell size for the output raster dataset for Vilnius city and correspondingly 400 meter neighbourhood radius (with exception of small districts of Naujamiestis and the Old Town where 200 meter radius was used for better precision) and the 10×10 meter cell size for individual districts. The points have been weighted depending on the total number of incidents that were registered at the same address. The output raster datasets of 2010 (Fig. 2) and 2011 (Fig. 3) were produced and subtracted to identify the changes in the crime and delinquency landscape of Vilnius in one year.

Location quotient technique (Brantingham, P. L., Brantingham, P. J. 1997; Harries 1999) has been applied to determine the relative weight of four major types of incidents. Location quotient was calculated as an index for comparing a district's share of a particular type of

criminal activity with the share of that same activity at the city level. It allows evaluating the deviation of impact of a particular type of criminal activity in a district. We applied formula:

$$LQ_{ij} = \frac{C_{ij} / \sum_{j=1...21} C_{ij}}{\sum_{i=1...4} C_{ij} / \sum_{i=1...4, j=1...21} C_{ij}}, \quad (1)$$

where C_{ij} is a rate of criminal activity i in district j .

Location quotients were estimated for assaults, burglaries and thefts, motor vehicle thefts and minor offences separately.

4. Incident maps

The initial density maps (Figs 1, 2) show distribution of probability of criminal activity for all Vilnius City based on incident data of correspondingly 2010 and 2011. After many experiments, relative density values have been grouped into 8 classes (1 – low, 8 – high) using natural (statistical) breaks method in order to achieve best visual expression. It can easily be seen that crime rates grew in 2011 but the concentration areas retained their general shape and structure.

The same classification method was used for the larger scale density maps of four major types of crimes in the central part of Vilnius. Neither total density nor

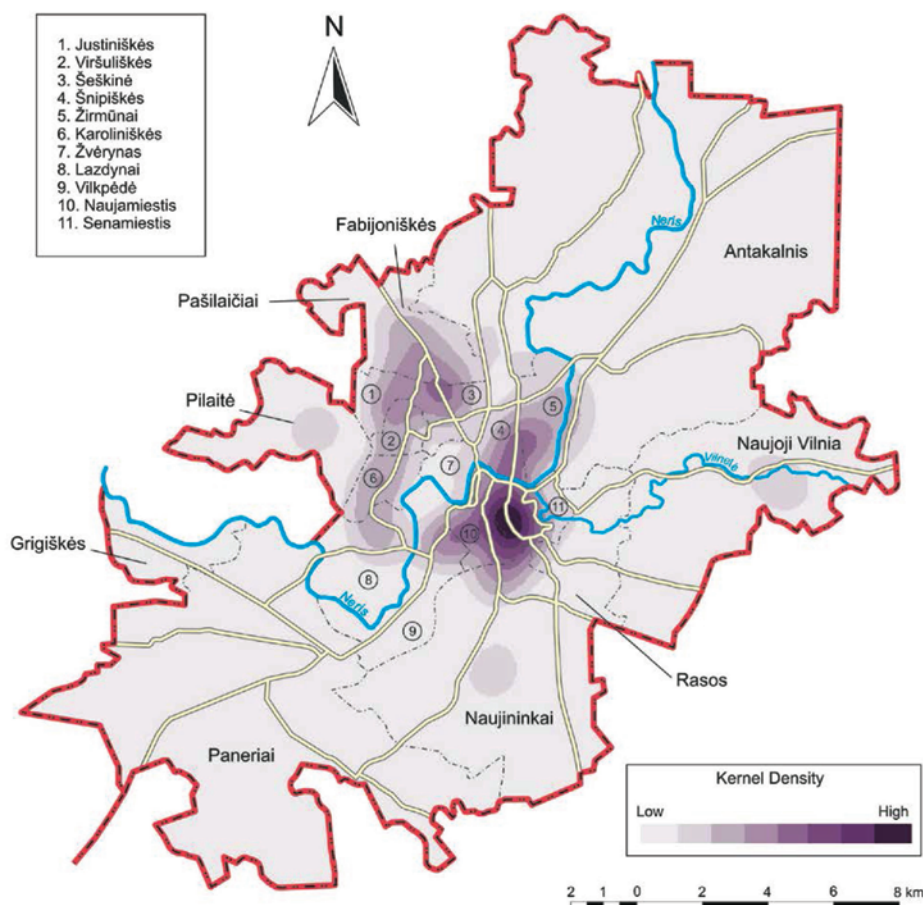


Fig. 2. Probability density of incidents by type in 2010 in all districts of Vilnius City

density by each type of crimes has any characteristic distribution in peripheral districts and the relative density value for those districts is “low”, that can be explained by much lower population density outside the central part of the city.

There are two large concentration areas: the major one of the Old Town and secondary one of the flat block residential districts with the centre in Pašilaičiai. These areas roughly match the highest population density areas of Vilnius city. The central part of Vilnius including those two areas was analysed in more detail and by different types of criminal activity. The results are shown on four larger scale maps (Figs. 3–7).

The patterns of distribution of four different types of criminal activities (Figs. 4–7) reveal their dependence on the population density as general configuration of each crime concentration areas is always similar to the area of high population density and to the general incident density areas. Distribution of the most common types of incidents, namely minor offences (such as hooliganism and other public nuisances, Fig. 5) and burglaries and thefts (Fig. 7) is very similar to the general pattern and to each other. However, assaults (Fig. 4) tend to concentrate in several smaller areas different from highest population density areas and motor vehicle thefts (Fig. 5) have strongly different distribution pattern with highest concentration area in newer residential districts of Fabijoniškės and Pašilaičiai that have relatively lower crime rates.

Location quotient choropleth maps (Fig. 8) show prevalence of particular type of criminal activity in each district independently from the total number of registered incidents. Thus, assaults strongly dominate in crime landscape of Paneriai and Naujoji Vilnia, robberies and thefts in Verkiai and motor vehicle thefts in Pašilaičiai, Fabijoniškės and Grigiškės.

5. Conclusions

Density analysis can be applied for analysis of criminal activities and produces results that are visually very expressive. However, if volumes of initial data are large and not completely consistent, pre-processing such as geocoding and generalisation may be time consuming.

The authors have not made any assumptions about dependency between crime rates and specific population groups. Indeed, the spatial pattern of crime rates in Vilnius does not seem to be related with specific communities but rather concentrates around specific public areas, such as shopping and entertainment centres.

The spatial pattern of crime rates in 2010–2011 showed stability with significant concentration in the Old Town and some immediate neighbourhoods and not so strong concentration in Central-Western residential districts. Overall incident rate is roughly proportional to population density but patterns of different major types of crimes deviate from Vilnius population density pattern in different ways.

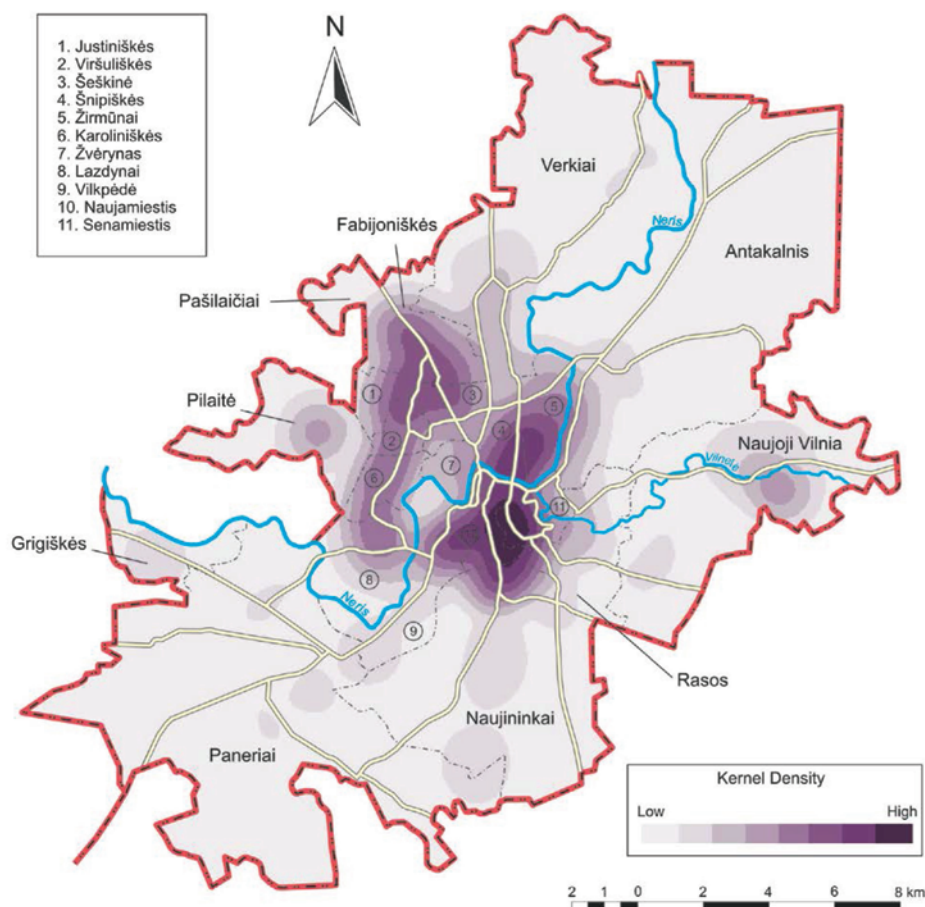


Fig. 3. Probability density of incidents by type in 2011 in all districts of Vilnius City

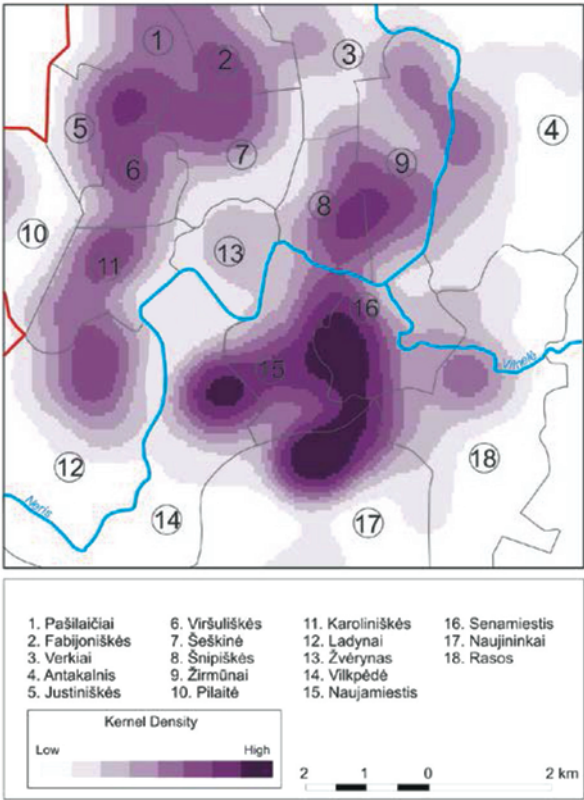


Fig. 4. Density of assaults in the central part of Vilnius City in 2011

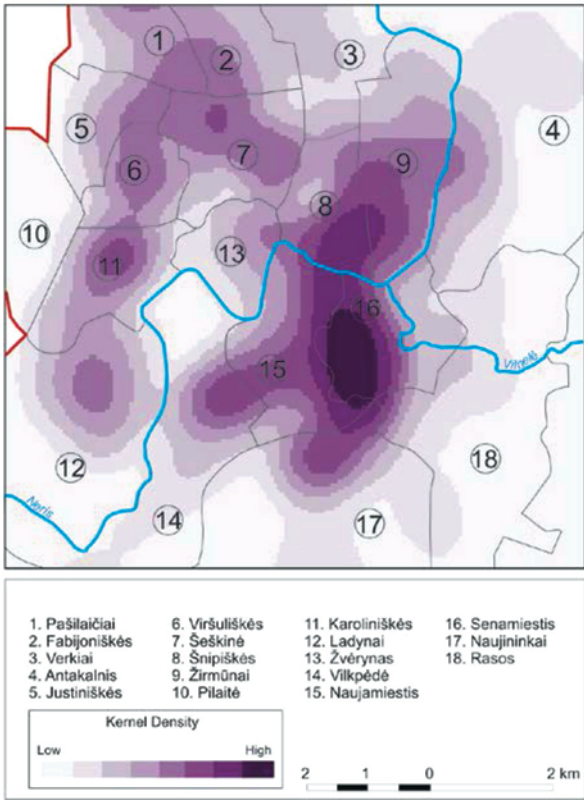


Fig. 5. Density of burglaries and thefts in the central part of Vilnius City in 2011

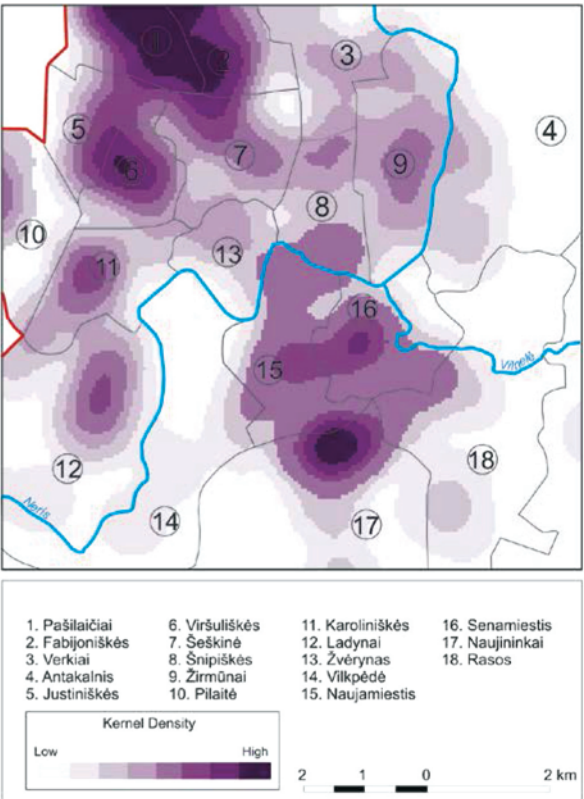


Fig. 6. Density of motor vehicle thefts in the central part of Vilnius City in 2011

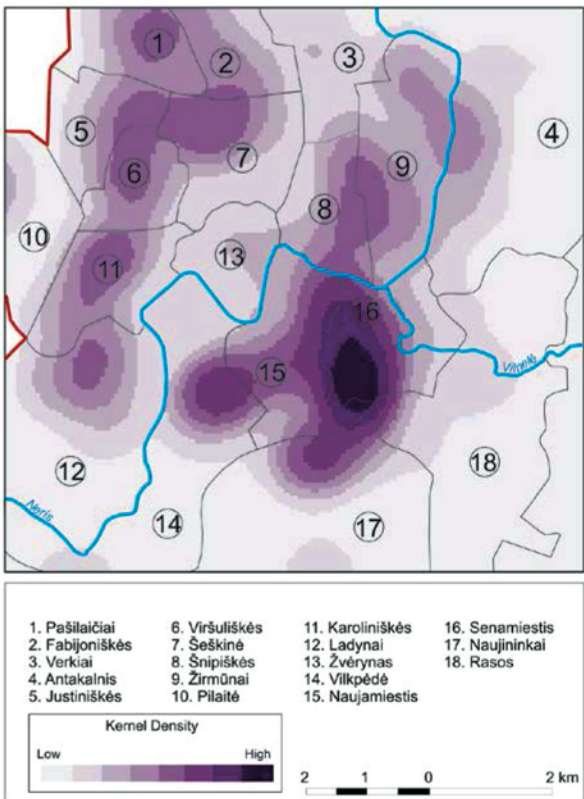


Fig. 7. Density of minor offences in the central part of Vilnius City in 2011

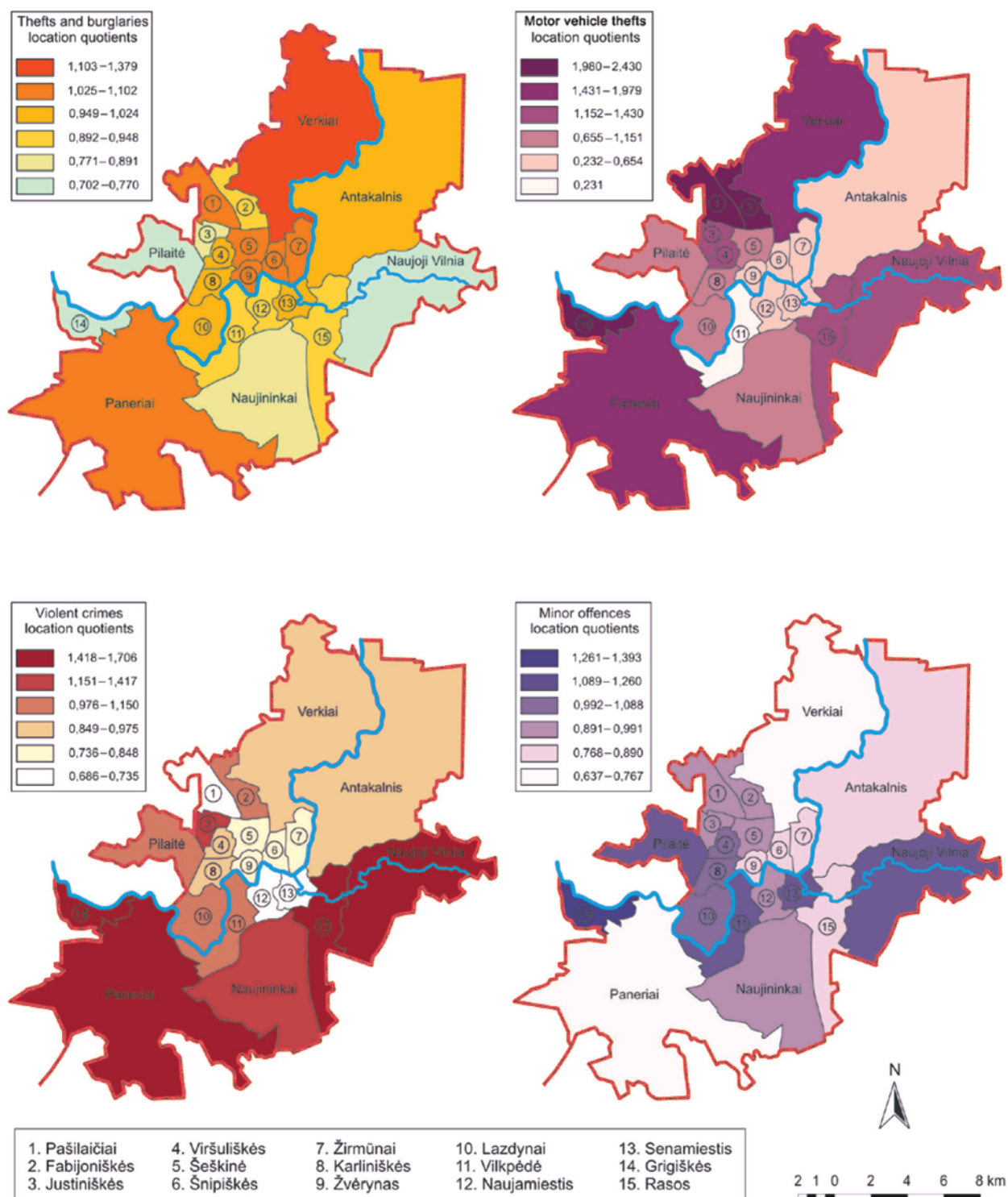


Fig. 8. Location quotients for four major types of incidents in all districts of Vilnius City in 2011

Density analysis can be efficiently complemented by other methods of spatial analysis. In our research location quotient was estimated to determine structural differences of criminal activity in different districts of Vilnius. Burglaries and thefts relatively prevail in rather dissimilar districts of Šnipiškės, Verkių and Žvėrynas. Assaults tend to prevail in the areas most distant from the city centre (Rasos, Grigiškės, Naujoji Vilnia) with exception of Pašilaičiai residential district. Motor vehicle thefts concentrate almost solely around Pašilaičiai.

The geocoded point data and visualizations of registered incidents of 2010 and 2011 are available at Lithuanian Spatial Information Portal www.geoportal.lt.

References

- Boba, R. 2005. *Crime Analysis and Crime Mapping*. California: Sage Publications. 320.
- Brantingham, P. L.; Brantingham, P. J. 1997. Mapping crime for analytic purposes: location quotients, counts, and rates, in

- D. Weisburd, J. T. McEwen (Eds.). *Crime Mapping and Crime Prevention*. New York: Criminal Justice Press, 263–288.
- Bruce, C. V. 2008. *Exploring Crime Analysis: Readings on Essential Skills*. Kansas: International Association of Crime Analysts, 342–363.
- Chainey, S.; Ractliffe, J. 2005. *GIS and Crime Mapping*. Chichester: John Wiley and Sons. 422.
- Gibin, M.; Longley, P.; Atkinson, P. 2007. Kernel density estimation and percent volume contours in general practice catchment area analysis in urban areas, in *Proceedings of Geographical Information Science Research Conference (GIS-RUK)*, 11–13 April 2007, NUI Maynooth, Ireland, 270–276 [online], [cited 26 November 2012]. Available from Internet: <http://ncg.nuim.ie/gisruk/materials/proceedings/>
- Harries, K. 1999. *Mapping Crime: Principle and Practice*. Washington DC: U.S. Department of Justice, Office of Justice Program, National Institute of Justice, Crime Mapping Research Center. 206.
- Levenshtein, V. 1965. Binary codes capable of correcting deletions, insertions, and reversals, *Doklady Akademii Nauk SSSR* 163(4): 845–848.
- Maltz, M. D.; Gordon, A. C.; Friedman, W. 2000. *Mapping Crime in Its Community Setting: Event Geography Analysis*. New York: Springer-Verlag. 197.
- Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman & Hall. 22.
- Zhang, H.; Peterson, M. P. 2007. A spatial analysis of neighbourhood crime in Omaha, Nebraska using alternative measures of crime rates, *Internet Journal of Criminology* 31 [online], [cited 26 November 2012]. Available from Internet: <http://www.internetjournalofcriminology.com/Zhang%20Peterson%20-%20A%20SPATIAL%20ANALYSIS%20OF%20NEIGHBOURHOOD%20CRIME.pdf>
- Giedrė BECONYTĖ**. Professor, Vilnius University, M. K. Čiurlionio g. 21, LT-03101 Vilnius, Lithuania. Ph +370 640 16583, Fax: +370 5 2398296, e-mail: giedre.beconyte@gf.vu.lt.
- A member of the Commission of Theoretical Cartography at the International Cartographic Association, the Executive secretary of Lithuanian cartographic society, has published more than 40 papers addressing information engineering methods of cartography, actively participated in the projects on national thematic cartography and in the process of developing the infrastructure of Lithuanian spatial data.
- Research interests: thematic cartography, system analysis and engineering, graphic design and cartography.
- Agnė EISMONTAITĖ**. Master of Science in Cartography at the JSC “HNIT-Baltic”, S. Konarskio g. 28A, LT-03127 Vilnius, Lithuania. Ph + 370 5 215 0575, Fax +370 5 215 0576, e-mail: agne.eismontaite@gmail.com.
- A graduate with Magna Cum Laude from Vilnius University (2012) actively participating in activities of Lithuanian cartographic society.
- Research interests: social geography, thematic cartography, GIS.
- Denis ROMANOVAS**. MSc student at Vilnius Gediminas Technical University, GIS programmer at the National Centre of Remote Sensing and Geoinformatics “GIS-Centras”, Sėlių g. 66, LT-08109 Vilnius, Lithuania. Ph +370 5 2724 741, Fax +370 5 3737 723, e-mail: d.romanovas@gis-centras.lt.
- Research interests: GIS analysis and application programming.