Security Journal
https://doi.org/10.1057/s41284-019-00226-1

ORIGINAL ARTICLE

Crime prevention of bus pickpocketing in Beijing, China: does air quality affect crime?

Ning Ding[1,2] · Yiming Zhai[1]

Abstract
In recent years, along with the development of the urban public transport system, bus pickpocketing crimes have garnered increasing attention. In this paper, a complete and clear structure including finding patterns based on data, applying criminological theory to explain patterns, and using empirical investigation method to verify theoretical explanation are presented. We found that temperature and season are not clearly correlated with bus pickpocketing. The AQI and $PM_{2.5}$ indices, however, demonstrated significant correlations with daily bus pickpocketing incidents: the worse the air quality, the more bus pickpocketing occurred. Then two empirical investigations were carried out to verified that crime pattern theory and rational choice theory can be used to explain the impact of air quality on bus pickpocketing crime. Furthermore, we utilized the SVM method to predict daily bus pickpocketing crime risk with an accuracy rate of 81%. The results of this paper can provide early warnings of urban bus pickpocketing and help police reduce crimes.

Keywords Bus pickpocketing · Air quality · Crime prediction · Empirical analysis · Support vector machine

Introduction

With the development of the urban public transport system, buses have become an indispensable part of people's daily activities. Buses are also the site of extensive pickpocketing. Bus pickpockets are thieves who steal items (often wallets, passports, and other valuables) from people's clothing and bags while they are on the bus. Bus stations and routes are high-traffic areas, passengers are mobile, and they tend to

✉ Yiming Zhai
yiming_zhai@126.com

1 School of Criminal Investigation and Forensic Science, People's Public Security University of China, Beijing 100038, China

2 Public Security Behavioral Science Lab, People's Public Security University of China, Beijing 100038, China

Published online: 05 December 2019

have little awareness of security; thus, pickpockets can escape easily following a crime, and most victims fear retaliation if confronted. These issues have made controlling bus pickpocketing difficult, although the crime seriously affects passengers' personal safety and property.

Studies on bus pickpocketing have focused primarily on various modes of public transport (bus, subway, aircraft, and ship) and types of crime (violent crime, pickpocketing, graffiti, destruction of public property, etc.) (Ye and Wu 2014). Research has also examined the impact of environmental factors on bus station crime rates (Loukaitou-Sideris 1999; Loukaitou-Sideris et al. 2001; Pearlstein and Wachs 1982) and the difference between static and non-static (i.e., moving) crime on public transport (Newton 2004). Newton (2004) found that crime may occur on moving vehicles on the public transport system, making perpetrators difficult to locate. Public transport areas involve a mixture of static and non-static events.

Considerable criminology research has explored the relationship between environmental factors and crime. Quetelet (1842) pointed out that crime patterns change with the seasons. Peng et al. (2011) found that robbery and theft decreased in extreme weather conditions, while burglary increased on days with more hours of daylight. Kuo and Sullivan (2001) discovered that chances of property and violent crime decline with a greater proportion of greenery around an apartment.

In recent years, most cities in China, including Beijing, have been plagued by serious issues in air quality degradation because of an increase in private cars and industrial pollution emissions (Matus et al. 2012; Tie et al. 2006; Wang et al. 2014). *Berkeley Earth* noted that approximately 1.6 million people die of diseases related to fine particulate matter ($PM_{2.5}$) annually in China, accounting for nearly 17% of all deaths. Around 4000 people die of air pollution every day in China. Rohde and Muller (2015) analyzed data from nearly 1500 air quality monitoring stations in China and discovered that 92% of Chinese people breathed unhealthy air for at least 120 h between April 5, 2014 and August 5, 2015. Pollutants can directly affect lung functioning, increasing the body's susceptibility to hypoxia. The World Health Organization deemed air pollution carcinogenic in 2013. Several studies have examined the spatial distribution of air pollution, including air quality in certain areas of China (Lee et al. 2002; Niu et al. 2015; Wang et al. 2014; Zhang et al. 2015) along with the distribution of the air quality index (AQI) in China as a whole (Pu et al. 2017). However, little research has explored the potential correlation between bus pickpocketing crime and air quality.

In this paper, several environmental factors, such as the AQI and $PM_{2.5}$ indices, season, and temperature, are tested to determine their potential correlations with the number of bus pickpocketing. The support vector machine (SVM) method is used to predict the daily risk of bus pickpocketing. The data and methods used in this paper are introduced in "Data collection" section. "The correlation between environment factors and crime" section evaluates the correlation between environmental factors and the daily number of bus pickpocketing incidents using the Pearson correlation coefficient. In "Temporal pattern of bus pickpocketing crime" section, the temporal distribution of bus pickpocketing is analyzed by the *K*-means cluster method. "Empirical investigations to verify the analysis" section uses two empirical investigations to verify the above data analysis and applications of criminological theory.

The SVM prediction method is utilized in "The SVM method for bus pickpocketing crime prediction" section and conclusions are presented in "Conclusions" section.

Data collection

The analysis is based on daily bus pickpocketing numbers provided by Beijing Municipal Public Safety Bureau from 2014 for 357 days, and the source of this data has been used in several papers (Peng et al. 2011; Chen et al. 2013; Chen and Lu 2017; Hu et al. 2017). Crime pattern theory suggests that criminal activities are most likely to occur when the potential offenders and victims occupy the same space and time (Santos 2016). For bus pickpocketing incidents, the time window (business hours: 7 a.m. to 11 p.m.) and place (buses) are more concentrated, providing favorable conditions in which to identify potential patterns.

AQI and $PM_{2.5}$ are two commonly used indicators of air quality in China. The AQI is a number used by government agencies to communicate to the public the current or forecasted level of air pollution. $PM_{2.5}$ refers to atmospheric particulate matter (PM) with diameter of less than 2.5 micrometers, approximately 3% the diameter of a human hair. Table 1 includes the individual air quality index (IAQI), calculated based on HJ633-2012 from China's Ministry of Environmental Protection. AQI represents the maximum of all individual air quality indices. In Table 2, AQI is divided into different levels, each of which exerts a different degree of impact on human health. China's government website does not provide historical air quality data; thus, data for AQI, PM2.5, and temperature were obtained from tianqihoubao.com (Chen et al. 2017) and rp5 (http://rp5.ru/) (Hu et al. 2017). Data on these sites came from daily weather data published on the official website. $PM_{2.5}$ data were also available directly from the official website of the U.S. Embassy in China, *stateair.net* (Xing et al. 2016). Two different sets of $PM_{2.5}$ data from China and the U.S. Embassy in China employed different calculation methods, but both reflected air quality to some extent; therefore, $aPM_{2.5}$ and $bPM_{2.5}$ were used to distinguish data from China and the U.S. embassy.

Table 1 Individual air quality index and corresponding concentration limits of pollutant (HJ633-2012)

IAQI	Contaminant project concentration limits					
	SO_2 (μg/m^3)	NO_2 (μg/m^3)	PM_{10} (μg/m^3)	CO (μg/m^3)	O_3 (μg/m^3)	$PM_{2.5}$ (μg/m^3)
0	0	0	0	0	0	0
50	50	40	50	2	100	35
100	150	80	150	4	160	75
150	475	180	250	14	215	115
200	800	280	350	24	265	150
300	1600	565	420	36	800	250
400	2100	750	500	48	1000	350
500	2620	940	600	60	1200	500

Table 2 The range of AQI and relevant information

AQI	Air quality level	Air quality description	Color	The influence on health
0–50	I	Excellent	Green	Normal activity
51–100	II	Good	Yellow	Sensitive people should reduce outdoor activities
101–150	III	Slightly polluted	Orange	Part of healthy people may manifest symptoms
151–200	IV	Moderately polluted	Red	Healthy people may manifest symptoms
201–300	V	Heavy polluted	Purple	The disease symptoms of cardiovascular and respiratory systems may aggravate
> 300	VI	Seriously polluted	Maroon	Healthy people also will be obviously discomfort

The description is derived from People's Republic of China National Environmental Protection Standards (HJ633-2012)

Methods and results

The correlation between environment factors and crime

In this paper, the Pearson correlation coefficient was calculated to determine the correlation between four environmental factors (AQI, $PM_{2.5}$, season, and temperature) and daily bus pickpocketing data using SPSS 19. Table 3 shows that the three parameters (AQI, $aPM_{2.5}$, and $bPM_{2.5}$) were significantly correlated with the number of daily bus pickpocketing incidents. The Pearson correlation coefficient was positive, implying that the worse the air quality, the higher the risk of bus pickpocketing.

This phenomenon can be explained from two aspects: (1) the impact of a decline in visibility, and (2) psychological impact. When the air quality is poor, the number of fine particles in the air increases, thus reducing visibility. In low-visibility environments, the probability of dangerous driving increases. All drivers, including those driving buses, may reduce their speed to improve their response time in case of an emergency. As the average speed of city buses declines, passengers stay in them for longer. These conditions afford bus pickpockets more time to commit multiple crimes in one place, improving the efficiency of criminal activity and increasing the

Table 3 The correlation between environmental factors and the daily number of bus pickpocket

Project	Correlation	
	Pearson correlation	Significant (bilateral)
AQI	0.119*	0.024
$aPM_{2.5}$	0.116*	0.028
$bPM_{2.5}$	0.127*	0.016
Season	− 0.068	0.202
Temperature	− 0.040	0.456

*Significant correlation at 0.05 (bilateral) number of bus pickpocketing crimes. In addition, low visibility allows pickpockets to escape unnoticed. Even when passengers realize they have been victims, it is difficult to recognize perpetrators and provide clues to the police, which improves the income-risk ratio for bus pickpockets.

In terms of psychological factors, bus pickpockets' decision-making process can be explained by rational choice (RC) theory, proposed by Cornish and Clarke (1987). RC theory states that people have reasoning ability and consider means, purpose, and cost-effectiveness to make rational choices. Assuming that crime in general is a purposive action to meet needs of offenders related to money, status, sexuality, and pleasure, fulfilling such needs requires individuals to make decisions. Offenders are also constrained by their personal abilities and information availability along with other factors. Under an RC theory framework, bus pickpockets will consider favorable or unfavorable conditions such as air quality, timing of the crime, the social environment, and other features when rationally calculating the costs and benefits of engaging in criminal activity. For example, if an offender chooses to go out on hazy days to identify potential pickpocketing targets, the risk of being caught and the possible health consequences of poor air quality must be considered. Therefore, when air quality is poor, going outside can pose direct threats to human health. Bus pickpockets would likely not refrain from pickpocketing simply because of health risks. For passengers, however, poor weather can adversely affect their mood and make them less vigilant.

Regarding criminal profit, offenders may find it easier to pickpocket under hazy conditions compared to good weather and they have the lower likelihood of being captured. The potential threat that poor weather poses to pickpockets' health is far less than potential profits, hence the decision is made to venture out on hazy days to find targets. For another example, bus pickpockets tend to commit crimes during rush hours (8 a.m. and 6 p.m.) due to the greater opportunity which is caused by higher traffic. Regarding criminal profit, the proceeds of crime during these time windows are much higher than other times while the risk of being captured remains the same. Similarly, when considering the social environment, if strengthening police enforcement during a certain time, the risk of bus pickpockets being captured increases substantially compared to at other times. In this case, assuming no significant changes in profit, bus pickpockets may opt not to commit crimes during those certain times. As shown, RC theory is helpful in examining criminals' decision-making process and ultimate choices.

A few studies have revealed a strong correlation between temperature and crime (Barnett and Adger 2007; Brunsdon et al. 2009; Mares 2013; Ranson 2014). Horrocks and Menclova (2011) found temperature to have a strong impact on violent crime as well as property crime. A classic explanation for how temperature influences crime is Cohen and Felson's (1979) routine activity (RA) theory. In RA theory, excluding inclement weather that prohibits travel, high temperatures may increase people's outdoor activities and social interactions. High temperatures also increase the probability of a criminal identifying a suitable target and present greater opportunity for crime. Presumably, if people are more likely to go outside in warmer weather, then the number of people taking the bus should increase and may result in higher risks of bus pickpocketing. RA theory summarizes conditions for committing crimes thusly: (1) a potential criminal has the capacity to commit crimes; (2) the criminal identifies a suitable target or victim; and (3) no authority figure or guardian is present to protect the victim. Like temperature, other environmental factors such as air quality and seasonality affect the availability of a suitable target.

However, SPSS results in this study demonstrate that the Pearson coefficients of temperature and seasonality exceed 0.05; hence, as the second condition in RA theory, these factors did not appear to exert a considerable impact on bus pickpocketers, unlike air quality. In Fig. 1, we can also see that the temperature and air quality curves are very different within a year. The main reason is that temperature is largely determined by the season, but the periodicity of the bus pocketing does not strictly match the seasonal changes.

Temporal pattern of bus pickpocketing crime

To forecast and prevent bus pickpocketing using correlation analysis, it is first necessary to know more about the regular patterns of the crime. This paper selected a dataset including the number of daily bus pickpocketing in 2014 to summarize the temporal pattern of bus pickpocketing incidents in a day and a year. Cluster analysis is commonly used for data mining and temporal-spatial analysis of crime (Murray and Estivill-Castro 1998; Murray et al. 2001). Owing to the advantages of simplicity, accuracy, and ease of operation, *K*-means cluster analysis is widely used in various fields including criminology. For example, Nath (2006) used *K*-means clustering to analyze crime patterns. This paper uses this method to analyze the temporal distribution of bus pickpocketing crime (day and year) and its steps can found in Fig. 2.

Figures 3 and 4 indicates that bus pickpocketing had a low incidence in June, July, November, and February within a year; May, October, January, August, September, and November were transition periods; and bus pickpocketing peaked in March and April. In a day, 8 p.m.–11 p.m. and 12 p.m.–1 p.m. were low-incidence periods for bus pickpocketing; 9 a.m.–11 a.m., 1 p.m.–5 p.m., and 7 p.m.–8 p.m. were moderate-incidence periods; and 7 a.m.–9 a.m. and 6 p.m.–7 p.m. were high-incidence

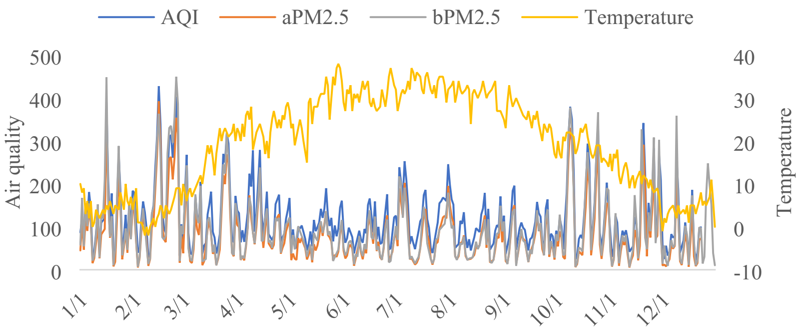
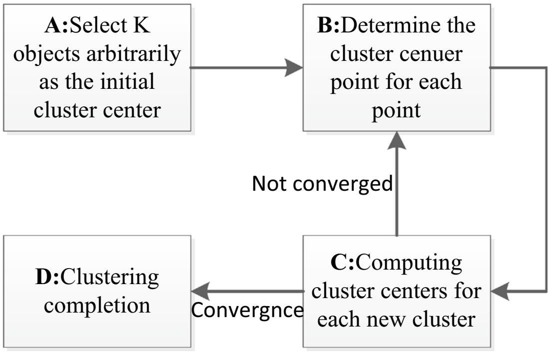


Fig. 1 Changes in environmental factors within a year

Fig. 2 The *K*-means steps to analyze the bus pickpocketing



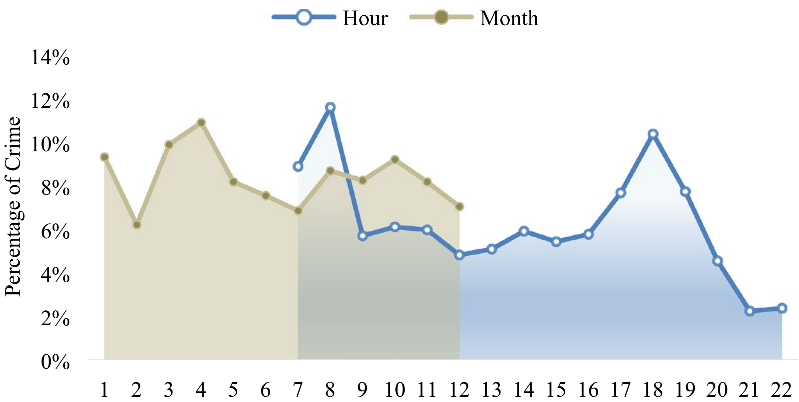


Fig. 3 The trend of bus pickpocketing crime in 1 day and 1 year

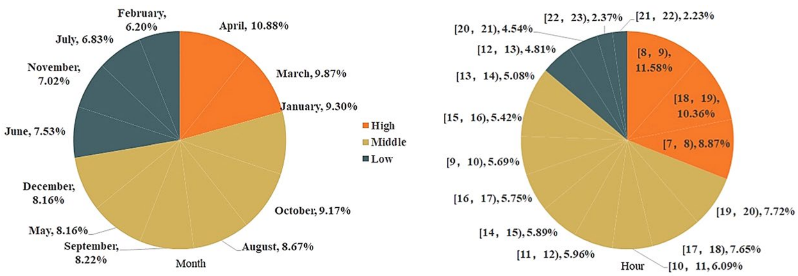


Fig. 4 Temporal clustering analysis of bus pickpocketing crime in 1 day and 1 year

periods. Hence, March and April annually and 7 a.m.–9 a.m. and 6 p.m.–7 p.m. daily represent key times for preventing bus pickpocketing.

Empirical investigations to verify the analysis

In the previous section, we have introduced a series of theories including crime pattern theory, rational choice (RC) theory and routine activity (RA) theory to explain the patterns of bus pickpocketing and the impact of environmental factors on it. What's more, it also can be seen from Fig. 3 that the two peaks of pickpocketing coincide with people's commuting time, so the pattern of criminal behavior is closely related to the pattern of human social activities. Obviously, air quality affects or changes people's social activities, which affects crime. In the haze weather, people who chose bicycles or walking may prefer to take the bus. The number of people taking buses will be more than in non-haze weather, and the number of bus pickpocketing crimes will increase.

In order to verify the above data analysis and theoretical applications, or to answer our questions: Why does air quality affect bus pickpocketing crime? We did some empirical investigations (EI) to explore.

In EI 1, in order to verify whether more people choose the bus to commute in the haze weather, we counted the number of people getting on and off the bus during the same time in different weather conditions. In order to calculate the quantity more conveniently, we used the road monitoring system of the public security department. We got three people who did not know the purpose of the experiment to randomly select the cameras of the three bus stations and let them calculate the number of people getting on and off the bus at 8:00–8:30 for 4 days. In order to control the variables, the dates we selected were all working days, and the weather was sunny and in four different air levels. Divide 8:00–8:30 into three sections (10 min each) and calculate the average number of people getting on and off the bus. It can be seen from Table 4 that in severe haze weather, more people would choose to travel by bus, which is consistent with the purpose of our investigations.

In EI 2, we designed a questionnaire to investigate the probability of people choosing bicycles, buses, taxis and walking in short trips, as well as the impact of rainy and snowy weather, haze weather and cold weather on people's travel. These three abnormal weathers we have chosen can be arranged into three levers according to their negative impact on travel: (1) haze, (2) cold, and (3) rainy and snowy. We received a total of 149 valid questionnaires and the results were shown in Tables 5, 6 and 7. Firstly, it can be seen from Table 5 that 44% of people will choose a bus as their transportation in the haze. What's more,

Table 4 Average number of people getting on and off the bus at different place and different air quality

Time	AQI	Lever	Camera 1	Camera 2	Camera 3
2018-06-14	155	IV	40	48	52
2018-06-15	111	III	27	38	46
2018-06-29	80	II	24	36	41
2018-07-05	48	I	23	48	38

Table 5 Questionnaire survey on means of transportation in haze days

Means of transportation	Subtotal	Proportion
Walking	28	18.79%
Bus	65	43.62%
Sharing bikes	13	8.72%
Taxi	43	28.86%
Effective number	149	

Table 6 Questionnaire survey on means of transportation in cold days

Means of transportation	Subtotal	Proportion
Bus	54	36.24%
Sharing bikes	14	9.4%
Walking	25	16.78%
Taxi	56	37.58%
Effective number	149	

Table 7 Questionnaire survey on means of transportation in rainy and snowy days

Means of transportation	Subtotal	Proportion
Taxi	76	51.01%
Sharing bikes	2	1.34%
Walking	27	18.12%
Bus	44	29.53%
Effective number	149	

haze weather is the only factor that makes more people choose buses rather than rainy and snowy weather or cold weather. Secondly, at the bad weather lever 1, the haze weather does not have a very urgent impact on human physiology, so people choose the bus, a faster and cheap transportation; When the human body feels cold, in order to get out of the current environment as soon as possible, the number of people who choose a faster transportation taxi is almost the same as the number of people taking the bus; The impact of rain and snow on people's travel is the most urgent, so half of them choose the fastest transportation taxi. The EI 2 is consistent with our investigations purpose too.

From the above two empirical investigations, we can conclude that the haze weather makes more people choose the bus as their transportation, so people who take the bus will be more than usual, and the bus will be more crowded. It will increase the chance of bus pickpockets to commit crimes and the number of pickpocketing crimes will be more. Therefore, the above two empirical investigations have verified that crime pattern theory and rational choice theory can be used to explain the impact of air quality on bus pickpocketing crime.

The SVM method for bus pickpocketing crime prediction

The SVM method is a supervised learning algorithm proposed by Vapnik and Chervonenkis in 1963. It is helpful in dealing with small samples and nonlinear and high-dimensional pattern recognition problems. Compared with an artificial neural network, SVM is simpler and has greater generalizability; hence, it has been widely used in recent years for data mining and crime hotspot analysis (Chang et al. 2005; Kianmehr and Alhajj 2006, 2008; Sathyadevan and Gangadharan 2014). SVM can also achieve more desirable classification prediction based on target characteristics. The experiment was carried out using scikit-learn (downloaded from https://scikit-learn.org/).

Figure 5 shows the overall condition of 357 days of bus pickpocketing in 2014(Jan 1-Dec 23). At first, we wanted to predict the number of the crime per day by regression, however, it had little practical significance considering that the number of crimes was only in the range of 0 to 13. Therefore, we just divided crime risk into two categories: high and low and chose the 6.5 (half of 13) as a boundary. We also found high-risk crimes account for nearly 20% of the total, which seems to be in some ways consistent with the 80/20 Rule in criminology. In our work, "+ 1" represented a high risk of bus pickpocketing and "– 1" represented a low risk. The characteristic quantity was

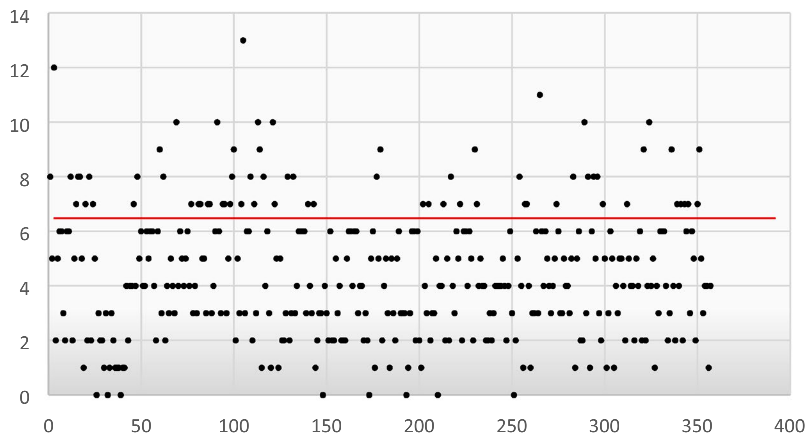


Fig. 5 The scatter plot of daily bus pickpocket crime number in 2014

Table 8 SVM and Naive Bayes prediction results

	Accuracy	Precision	Recall	F1-score
SVM	0.81	0.68	0.81	0.71
Naive Bayes	0.79	0.76	0.79	0.77

composed of AQI, $aPM_{2.5}$, $bPM_{2.5}$ and rest-day. Then, we chose 70% of 357 samples from 2014 as the training set; the remaining 30% samples served as a test set. To find the most accurate prediction method, we also use the naive Bayesian method in addition to the SVM method.

Table 8 shows the SVM method and Naive Bayes method prediction results. Although the support vector machine has higher accuracy and recall, its precision and F1-score are not the better. Part of the reason is that the number of bus pickpocketing crimes in Beijing is not more. If in some places where the number of bus pickpocketing is more, the model can be improved from two aspects. First, the crime level can be divided into three levels or five levels according to the actual situation, and the prevention and control measures will be more flexible; the second is that the impact of the small value on the prediction accuracy will be correspondingly reduced.

In practical work, once police find that the prediction of bus pickpocketing crimes continues to be at high risks in the future, they could organize special crime crackdown activities to maintain social stability. Furthermore, when the government prepares to organize large-scale activities, the public security department can suggest whether to strengthen the intensity of social management based on the prediction of crimes in the coming period.

Conclusions

This paper demonstrates a strong correlation between $PM_{2.5}$, AQI and the number of bus pickpocketing crime using crime statistics and observed climate records collected in Beijing, China. Results indicate that the worse the air quality, the more bus pickpocketing occurred. Although a few studies have shown strong relationships between temperature and crime, temperature and season demonstrated no obvious correlation with bus pickpocketing. After analyzing the relationship between temporal distribution and the number of bus pickpocketing, this paper revealed that every March and April (annually) and 7 a.m.–9 a.m. and 6 p.m.–7 p.m. (daily) were key times to improve bus pickpocketing prevention. Then two empirical investigations were carried out to verified that crime pattern theory and rational choice theory can be used to explain the impact of air quality on bus pickpocketing crime. Furthermore, we utilized the SVM method to predict daily bus pickpocketing crime risk with an accuracy rate of 81%.

In order to answer the questions: Why does air quality affect the bus pickpocketing crime? A complete and clear structure including finding patterns based on data, applying criminological theory to explain patterns, and using empirical investigation method to verify theoretical explanation is presented in our paper. In bus pickpocketing crimes, some of environment factors(air quality in this paper) affect the choices of offenders and victims (rational choice theory), and it gives more chances to offenders to commit crimes (crime pattern theory). In this way, more bus pickpocketing crimes occur.

If the police department can accurately predict the risk of a crime, it can bring a variety of benefits. First, rationally arrange the intensity of police work. When the risk is high in the future, arrange more police forces in advance; When the risk is low, few police is enough.This way can effectively improve the combat effectiveness of the police. Second, to control the criminal situation. As recent crime trends continue to rise and future criminal risks are increasing, appropriate special strikes can be launched to prevent a crime from seriously affecting people's lives. Third, serving the national large-scale activity plan. social security is one of the important factors determining whether a city can hold large-scale international activities. The future public security situation can provide the following information: whether the address of the event is reasonable, the intensity of social governance before the activity and the content of the crime in the emergency plan.

Though the information and data available for analysis were limited, this paper still provided insights regarding the correlation between air quality and bus pickpocketing. The results provide implications for crime prevention and analysis for police. Considering Beijing's great social security environment and special position in China, the total amount of bus pickpocketing crimes is relatively small. Collecting more crime data from different cities (e.g., Henan, Guizhou) in China and continuing to predict crime in other regions will be a focus in our future work.

Acknowledgements This work is supported by the National Natural Science Foundation of China (71904194) and National Key R&D Program of China (No. 2018YFC0822800).

Appendix

Principles of SVM

When using SVM to classify two types of data, training samples are divided into "1" and "− 1" categories that are then trained to get a hyperplane that is maximally distant from the positive samples and negative samples. Test samples are plotted in a high-dimensional space to distinguish whether they are positive or negative according to an optimal separating hyperplane. This study employed a training set:

$$\{(x_i, y_i)[x_i \in R^n, y_i \in R, i = 1, 2, 3 \dots,]\} \tag{1}$$

where $x_i \epsilon R^n$ represents the eigenvector, and $y_i \epsilon \{+1, -1\}$ is the category mark. In this paper, "+ 1" indicates a high crime rate, whereas "− 1" is low. Most real-world problems are not linearly separable; SVM uses the kernel technique to automatically realize nonlinear mapping on a feature space. SVM uses the kernel function $\Phi(x)$ to map the training set data x_i to a high-dimensional linear feature space to find an optimal hyperplane:

$$\Phi(x) \cdot \omega + b = 0 \tag{2}$$

where $\omega \in R^n$ and $x \in R^n$, b is the offset. Therefore, two samples can be separated properly with the largest distance between the two sample types. The discriminant function is

$$y(x) = \text{sign}[\omega \cdot \Phi(x) + b]. \tag{3}$$

The problem of an optimal regression hyperplane can be transformed into a programming function as follows:

$$\min\left\{\frac{1}{2}||\omega||^2\right\} + C\sum\nolimits_{i=1}^{n}[C(\xi_i) + C(\xi^*)]. \tag{4}$$

The constraint of the function is

$$\begin{cases} y_i - (\omega \cdot x_i) - b \le \varepsilon + \xi_i \\ (\omega \cdot x_i) + b - y_i \le \varepsilon + \xi_i^* \end{cases} \quad i = (1, 2, 3, \ldots, n), \tag{5}$$

where $\xi_i \geq 0$ and $\xi_i^* \geq 0$ are each slack variables, $c(\xi_i)$ is the loss function, and C is the penalty term constant. By using the Lagrangian multiplier to solve the dual form of the quadratic programming problem with linear constraints, we get:

$$\max\left\{L_{\mathrm{D}} = \sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} a_i a_j y_i y_j \Phi(x_i) \cdot \Phi(x_j) = \sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} a_i a_j y_i y_j K(x_i, x_j)\right\} \tag{6}$$

with the following restrictions:

$$0 \le a_i \le C, \quad \sum_{i=1}^{l} a_i y_i = 0, \tag{7}$$

where $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, called a kernel function with the discriminant function of

$$y(x) = \text{sign}\left[\omega^* \cdot \Phi(x) + b^*\right] = \text{sign}\left[\sum_{x_i \in SV} a_i^* y_i K(x_i, x) + b^*\right], \tag{8}$$

where ω^*, b^*, a_i^* are the optimal solutions.

There are several common kernel functions of the SVM method:

(1) Linear kernel function: $Kx_i, x = x \cdot x_i$
(2) Polynomial kernel function: $Kx_i, x = [(x \cdot x_i) + 1]^q, q = 1, 2, \ldots, n$
(3) Radial basis function (RBF):$Kx_i, x = \exp\left(-\frac{|x - x_i|^2}{2\sigma^2}\right)$
(4) Cauchy kernel function: $Kx_i, x = \tanh(v(x \cdot x_i) + C)$

The RBF kernel is most often used in the above function because of its good learning ability. It is an ideal classification function for any type of sample (low-dimensional, high-dimensional, small, large, etc.). This study used the RBF kernel as well.

Principles of Naive Bayes

The probability that the quantity level of the number of bus pickpocketing on a certain day belongs to category c is:

$$P(c|d) \propto P(c) \prod_{1 \le k \le n_d} P(t_k|c), \tag{9}$$

where $P(c)$ is the prior probability that the number of bus pickpocketing on a certain day belongs to category c, $(t_1, t_2, t_1, \ldots, t_{n_d})$ is the characteristic attribute of day d, and n_d represents the total amount of all feature attributes of d. After knowing that the number of bus pickpocketing on a certain day belongs to the prior probability of category c_i, it is necessary to find the most likely category for a certain day d. For Naive Bayes, it is the category of the estimate of maximum a posteriori (MAP). i.e. the formula (2) (where $c \in C$):

$$c_{\mathrm{map}} = \arg\max \hat{P}(c|d) = \arg\max \hat{P}(c) \prod_{1 \le k \le n_d} \hat{P}(t_k|c). \tag{10}$$

Calculating the product of the conditional probability using Eq. (2) may cause the lower bound of the floating point number to overflow, so introduce the logarithm to get the formula (3) (where $c \in C$):

$$c_{\mathrm{map}} = \arg\max \hat{P}(c|d) = \arg\max \left[\log \hat{P}(c) + \sum_{1 \le k \le n_d} \log \hat{P}(t_k|c) \right]. \tag{11}$$

References

Barnett, J., and W.N. Adger. 2007. Climate change, human security and violent conflict. *Political Geography* 26 (6): 639–655.

Brunsdon, C., J. Corcoran, G. Higgs, and A. Ware. 2009. The influence of weather on local geographical patterns of police calls for service. *Environment and Planning B: Planning and Design* 36 (5): 906–926.

Chang, W., D. Zeng, and H. Chen. 2005. Prospective spatio-temporal data analysis for security informatics. In *Proceedings of the 2005 IEEE intelligent transportation systems, 2005*, 1120–1124. IEEE. http://ieeexplore.ieee.org/abstract/document/1520208/.

Chen, H., Y. Lin, Q. Su, and L. Cheng. 2017. Spatial variation of multiple air pollutants and their potential contributions to all-cause, respiratory, and cardiovascular mortality across China in 2015–2016. *Atmospheric Environment* 168: 23–35. https://doi.org/10.1016/j.atmosenv.2017.09.006.

Chen, P., and Y. Lu. 2017. Exploring co-offending networks by considering geographic background: An investigation of electric bicycle thefts in Beijing. *The Professional Geographer*. https://doi.org/10.1080/00330124.2017.1325753.

Chen, P., H. Yuan, and D. Li. 2013. Space-time analysis of burglary in Beijing. *Security Journal* 26 (1): 1–15.

Cohen, L.E., and M. Felson. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review* 44: 588–608.

Cornish, D.B., and R.V. Clarke. 1987. Understanding crime displacement: An application of rational choice theory. *Criminology* 25 (4): 933–948.

Horrocks, J., and A.K. Menclova. 2011. The effects of weather on crime. *New Zealand Economic Papers* 45 (3): 231–254.

Hu, X., P. Chen, H. Huang, T. Sun, and D. Li. 2017. Contrasting impacts of heat stress on violent and nonviolent robbery in Beijing, China. *Natural Hazards* 87 (2): 961–972. https://doi.org/10.1007/s11069-017-2804-8.

Kianmehr, K. and R. Alhajj. 2006. Crime hot-spots prediction using support vector machine. In *IEEE international conference on computer systems and applications, 2006*, 952–959. IEEE. http://ieeexplore.ieee.org/abstract/document/1618468/.

Kianmehr, K., and R. Alhajj. 2008. Effectiveness of support vector machine for crime hot-spots prediction. *Applied Artificial Intelligence* 22 (5): 433–458.

Kuo, F.E., and W.C. Sullivan. 2001. Environment and crime in the inner city: Does vegetation reduce crime? *Environment and Behavior* 33 (3): 343–367.

Lee, Y.C., G. Calori, P. Hills, and G.R. Carmichael. 2002. Ozone episodes in urban Hong Kong 1994–1999. *Atmospheric Environment* 36 (12): 1957–1968.

Loukaitou-Sideris, A. 1999. Hot spots of bus stop crime: The importance of environmental attributes. *Journal of the American Planning Association* 65 (4): 395–411.

Loukaitou-Sideris, A., R. Liggett, H. Iseki, and W. Thurlow. 2001. Measuring the effects of built environment on bus stop crime. *Environment and Planning B: Planning and Design* 28 (2): 255–280.

Mares, D. 2013. Climate change and levels of violence in socially disadvantaged neighborhood groups. *Journal of Urban Health* 90 (4): 768–783.

Matus, K., K.-M. Nam, N.E. Selin, L.N. Lamsal, J.M. Reilly, and S. Paltsev. 2012. Health damages from air pollution in China. *Global Environmental Change* 22 (1): 55–66.

Murray, A.T., and V. Estivill-Castro. 1998. Cluster discovery techniques for exploratory spatial data analysis. *International Journal of Geographical Information Science* 12 (5): 431–443.

Murray, A.T., I. McGuffog, J.S. Western, and P. Mullins. 2001. Exploratory spatial data analysis techniques for examining urban crime: Implications for evaluating treatment. *British Journal of Criminology* 41 (2): 309–329.

Nath, S.V. 2006. Crime pattern detection using data mining. In *WI-IAT 2006 workshops. 2006 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology workshops, 2006*, 41–44. IEEE. http://ieeexplore.ieee.org/abstract/document/4053200/.

Newton, A.D. 2004. Crime on public transport: 'Static' and 'non-static' (moving) crime events. *Western Criminology Review* 5 (3): 25–42.

Niu, H., W. Hu, W. Pian, J. Fan, and J. Wang. 2015. Evolution of atmospheric aerosol particles during a pollution accumulation process: A case paper. *World Journal of Engineering* 12 (1): 51–60. https://doi.org/10.1260/1708-5284.12.1.51.

Pearlstein, A., and M. Wachs. 1982. Crime in public transit systems: An environmental design perspective. *Transportation* 11 (3): 277–297.

Peng, C., S. Xueming, Y. Hongyong, and L. Dengsheng. 2011. Assessing temporal and weather influences on property crime in Beijing, China. *Crime, Law and Social Change* 55 (1): 1–13.

Pu, H., K. Luo, P. Wang, S. Wang, and S. Kang. 2017. Spatial variation of air quality index and urban driving factors linkages: Evidence from Chinese cities. *Environmental Science and Pollution Research* 24 (5): 4457–4468.

Quetelet, A. 1842. *A treatise on man*. Franklin. http://ocw.abuad.edu.ng/courses/literature/21l-017-the-art-of-the-probable-literature-and-probability-spring-2008/readings/quetelet_exce.pdf.

Ranson, M. 2014. Crime, weather, and climate change. *Journal of Environmental Economics and Management* 67 (3): 274–302.

Rohde, R.A., and R.A. Muller. 2015. Air pollution in China: Mapping of concentrations and sources. *PLoS ONE* 10 (8): e0135749.

Santos, R.B. 2016. *Crime analysis with crime mapping*. Thousand Oaks, CA: Sage. https://books.google.com/books?hl=zh-CN&lr=&id=G1O0DAAAQBAJ&oi=fnd&pg=PP1&dq=Crime+Analysis+with+Crime+Mapping&ots=TiDwofR4pt&sig=kH7VFrPLxPZQeYZIwakyCiBnJFE.

Sathyadevan, S., S. Gangadharan, et al. 2014. Crime analysis and prediction using data mining. In *2014 first international conference on networks & soft computing (ICNSC)*, 406–412. IEEE. http://ieeexplore.ieee.org/abstract/document/6906719/.

Tie, X., G.P. Brasseur, C. Zhao, C. Granier, S. Massie, Y. Qin, et al. 2006. Chemical characterization of air pollution in Eastern China and the Eastern United States. *Atmospheric Environment* 40 (14): 2607–2625.

Wang, Y., Q. Ying, J. Hu, and H. Zhang. 2014. Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013–2014. *Environment International* 73: 413–422.

Xing, Y.-F., Y.-H. Xu, M.-H. Shi, and Y.-X. Lian. 2016. The impact of PM2.5 on the human respiratory system. *Journal of Thoracic Disease* 8 (1): 69.

Ye, W.J., and S. Wu. 2014. Identifying crime patterns of bus pickpocketing using weighted spatial-temporal association rules mining. *Journal of Geo-Information Science* 04: 537–544.

Zhang, F., L. Wang, J. Yang, M. Chen, Z. Wei, and J. Su. 2015. The characteristics of air pollution episodes in autumn over the southern Hebei, China. *World Journal of Engineering* 12 (3): 221–236. https://doi.org/10.1260/1708-5284.12.3.221.