Detecting Pickpocket Suspects from Large - Scale Public Transit Records

N.SRINIVASA RAO [1], KUDUPUDI VAISHNAVI[2].

[1] Assistant Professor, DEPT OF MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh

Email:- naagaasrinu@gmail.com

[2]PG Student of MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh

Email:- kudupudivaishnavi9@gmail.com

ABSTRACT

Massive data collected by automated fare collection (AFC) systems provide opportunities for studying both personal traveling behaviors and collective mobility patterns in urban areas. Existing studies on AFC data have primarily focused on identifying passengers' movement patterns. However, we creatively leveraged such data for identifying pickpocket suspects. Stopping pickpockets in the public transit system has been crucial for improving passenger satisfaction and public safety. Nonetheless, in practice, it is challenging to discern thieves from regular passengers. In this paper, we developed a suspect detection and surveillance system, which can identify pickpocket suspects based on their daily transit records. Specifically, we first extracted a number of useful features from each passenger's daily activities in the transit system. Then, we took a two-step approach that exploits the strengths of
unsupervised outlier detection and supervised classification models to identify thieves, who typically exhibit abnormal traveling behaviors. Experimental results demonstrated the effectiveness of our method. We also developed a prototype system for potential uses by security personnel.

I.INTRODUCTION

Public transit passengers can easily become distracted in crowded environments, where they are often rushing from one location to another. Having their focus drift from their belongings, they often become common targets of pickpockets [1, 2]. During the first 9 months of 2014, it was reported that 350 pickpockets were apprehended in the subway system and 490 on buses in Beijing.1 Many other big cities around the world, such as Barcelona, Rome, and Paris, also suffer from pickpocket problems.2 Indeed, it is challenging to detect theft activities committed by cunning thieves who know how to escape without being discovered. It is critical to provide a smart surveillance and tracking tool for transit system security personnel.

With rapid advances in information technology and infrastructure, transactional records collected by automated fare collection (AFC) systems are now available for understanding passengers' mobility patterns and urban dynamics [3, 4, 5, 6, 7]. Most existing studies focus on identifying regular, collective mobility patterns, such as commute flows and transit networks. Our study is the first to focus on identifying thieves based on AFC data. It is possible to detect thieves using AFC records because behavioral differences logged in the mobility footprints may be used to separate suspects from regular passengers. Examples of such behaviors include traveling for an extended length of time, making unnecessary transfers, and taking regular routes with random stops. Designing an intelligent system that automatically extracts specific, identified behavioral features and dynamically detects and tracks pickpocket suspects has become a possibility.

Detecting thieves based on AFC records is not a simple outlier detection problem. Fig. 1 shows the difference between a known thief and an outlier. We can see a number of trajectories between hot regions A and B. By careful examination, we see that most passengers move from one region to another using a near-optimal configuration (e.g., shortest time/distance, or a minimal number of transfers). However, a passenger (a known suspect) who took the path A -> C-> D->! B looks suspicious because there is no need to make transfers at C and D in order to reach B. Based on the above observation, passengers who exhibit such abnormal behaviors will be selected for further examination. In contrast, another passenger who travels from E to B is an outlier, since few passengers take the same path. However, this passenger is likely just a regular passenger who originates from a less crowded area. Detecting thieves is challenging also because not every trip made by a regular passenger looks normal. Regular commuters may occasionally make trips to visit friends or places of interest, and such trips may look suspicious by how much they deviate from regular passenger behaviors.

Adding to this complex landscape, a large number of AFC records are being collected from millions of passengers, when only a tiny fraction of passengers are actual pickpockets. Pinpointing such a small group of people within such a large scale dataset is analogous to searching for a needle in the haystack. Meanwhile, we need to effectively transform our knowledge based on model development into a decision support system. Such a system needs to provide real-time decision recommendations to guide security personnel to perform their work more efficiently.

II. EXISTING SYSTEM

- The System of existing literature focuses on finding patterns in passenger activity records. Such knowledge can be useful in a variety of applications, and plays a vital role in effectively finding and satisfying passenger needs. Examples include assessing the performance of the transit network, identifying and optimizing problematic or flawed bus routes, improving the accuracy of passenger flow forecasted between two regions, and making service adjustments that accommodate variations in ridership on different days. In particular, [4] estimated the crowdedness of various stations in the transportation network using AFC data. [9] measured the variability of transit behaviors on different days of the week.

- Existing studies that detect anomalies in urban sensing data can be divided into two categories: those based on locations, and those on trajectories. Along the line of location-based anomaly detection, [15] presented a framework that learned the context of different functional regions in a city, which provided the basis of our feature extraction approach.

- In addition, [16] attempted to discover casual relationships among spatiotemporal outliers. [17] mined representative terms from social media posts when location relevant events happened in the city, such as accidents or protests. [18] discovered black-hole or volcano patterns in human mobility data in a city, which could quickly identify gathering events, such as football matches or concerts. Detection of such anomalies can help send alerts, and provide input for intelligent decision support, such as smoothing the traffic flow.

Disadvantages

- There is no Smart Card Based Travelling due to Only Manual Passengers Activity Patterns and Manual Transit Records.
- Since thieves involved in victim reported events were not caught, the system could not identify them according to the no-checkout rule. Instead, the system manually labeled thieves according to their travel behaviors. Specifically, the system first identified all passengers on the vehicle during the same period of time, and then visualized their trajectories to ascertain whether their travel patterns were typical.

III. PROPOSED SYSTEM

❖ In the proposed system, the system adopted a comprehensive approach to the pickpocket detection problem. The overall framework of our solution is illustrated in this system. The system first partitioned the city area into regions with functional categories. Then, the mobility characteristics of passengers were extracted from transit records dynamically over time.

❖ A core component of the system was a two-step passenger classification process, the first step being regular passenger filtering, and the second step being suspect detection. Finally, system user feedback information, such as newly confirmed thieves, was entered as ground truth for future model training based on a utility function that strikes a tradeoff between effectiveness (i.e., performance) and relevance (i.e., recency). A more detailed description of this system may be found in this system.

❖ The contribution of our study can be summarized as follows. Firstly, we identified a number of features that may be extracted from AFC records and are potentially useful for distinguishing thieves from regular passengers.

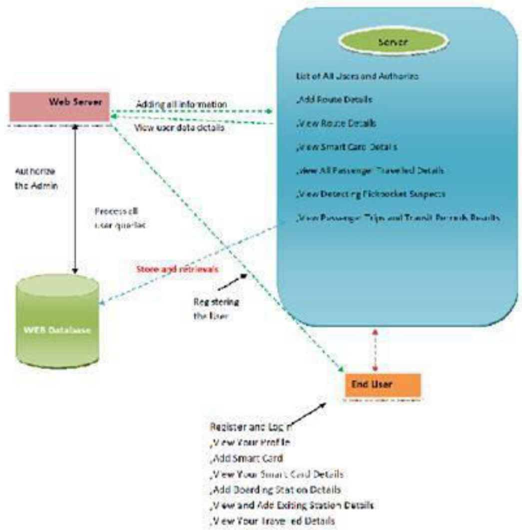
❖ Secondly, a two-step approach was proposed to make the suspect detection problem practical in a large-scale data environment where the positive and negative samples are extremely imbalanced.

❖ Thirdly, our dynamic filtering enhancement significantly reduced the everyday computation costs and maintained superior accuracy. Most importantly, a real system for the end user was designed and tested using real-world, large-scale data. As an applied data science study, our solution is the first to address an important social issue identifying pickpockets by using big data. The significance of this work has been recognized by a featured article in The Economist.

Advantages

- Effective techniques for Abnormal Traveling Behavior Detection.
- Smart Card Based Transit Record Maintenance.

IV. ARCHITECTURE DIAGRAM



V.I MPLEMENTATION

- Server

  In this module, the Web Server has to login by using valid user name and password. After login successful he can do some operations such as List of All Users and Authorize, Add Route Details, View Route Details, View Smart Card Details ,view All Passenger Travelled Details ,View Detecting Pickpocket Suspects ,View Passenger Trips and Transit Records Results
- User

  In this module, there are n numbers of users are present. User should register before doing some operations. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like View Your Profile, Add Smart Card, View Your Smart Card Details, Add Boarding Station Details, View and Add Exiting Station Details, View Your Travelled Details.

REFERENCES

[1] G. R. Newman and M. M. McNally, "Identity theft literature review," United States Department of Justice, Report 210459, July 2005.

[2] M. Felson and R. V. Clarke, "Opportunity makes the thief: Practical theory for crime prevention," Policing and Reducing Crime Unit: Police Research Series, Report 98, 1998.

[3] T. L. C. da Silva, J. A. F. de Macˆedo, and M. A. Casanova, "Discovering frequent mobility patterns on moving object data," in MobiGIS, 2014, pp. 60–67.

[4] I. Ceapa, C. Smith, and L. Capra, "Avoiding the crowds: understanding tube station congestion patterns from trip data," in UrbComp, 2012, pp. 134–141.

[5] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders travel patterns," Transportation Research Part C, vol. 36, pp. 1–12, 2013.

[6] K. Zheng, Y. Zheng, N. J. Yuan, S. Shang, and X. Zhou, "Online discovery of gathering patterns over trajectories," IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 8, pp. 1974–1988, 2014.

[7] Y. Liu, C. Liu, J. Yuan, L. Duan, Y. Fu, H. Xiong, S. Xu, and J. Wu, "Intelligent bus routing with heterogeneous human mobility patterns," Knowledge and Information

Systems, 2016, forthcoming, DOI: 10.1007/s10115-016-
0948-6.
[8] B. Du, C. Liu, W. Zhou, Z. Hou, and H. Xiong, "Catch
me if you can: Detecting pickpocket suspects from
large-scale transit records," in KDD, 2016, pp. 87–96.
[9] C. Morency, M. Tr´epanier, and B. Agard, "Analysing
the variability of transit users behaviour with smart
card data," in ITSC, 2006, pp. 44–49.
[10] M. J. Sung, "Analysis of travel patterns of the elderly
using transit smart card data," in TRB, no. 11-2357,
2011.
[11] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barab´asi,
"Understanding individual human mobility patterns,"
Nature, vol. 453, no. 7196, pp. 779–782, 2008.
[12] C. Song, Z. Qu, N. Blumm, and A.-L. Barab´asi, "Limits
of predictability in human mobility," Science, vol. 327,
no. 5968, pp. 1018–1021, 2010.
[13] C. Chen, D. Zhang, Z.-H. Zhou, N. Li, T. Atmaca, and
S. Li, "B-planner: Night bus route planning using largescale
taxi gps traces," in PerCom, 2013, pp. 225–233.
[14] W. Luo, H. Tan, L. Chen, and L. M. Ni, "Finding
time period-based most frequent path in big trajectory
data," in SIGMOD, 2013, pp. 713–724.
[15] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of
different functions in a city using human mobility and
pois," in KDD, 2012, pp. 186–194.