ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940www.ijsem.org
Vol 14, Issuse.4 Dec 2023

CATCH ME IF YOU CAN: DETECTING PICKPOCKET SUSPECTS FROM LARGE-SCALE TRANSIT RECORDS

L. PRIYANKA[1], B. TIRU RANGA NAYAKA[2], BANDA SAI RAKSHITHA[3], BATTU VAMSHI[4], BINGI MANISH KUMAR YADAV[5]

ABSTRAT:

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.

Keywords: AFC, theft activity, pickpocket, abnormal activity.

I INTRODUCTION

1) 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

1Assistant Professor, Dept of CSE, MALLA REDDY INSTITUTE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS),Dhulapally,Secundrabad, Hyderabad, Telangana, India.
2,3,4,5 UG Students, Dept of CSE, MALLA REDDY INSTITUTE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS),Dhulapally,Secundrabad, Hyderabad, Telangana, India.

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. In this paper, we adopted a comprehensive approach to the pickpocket detection problem.

RELATED STUDY

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.

PROPOSED SYSTEM

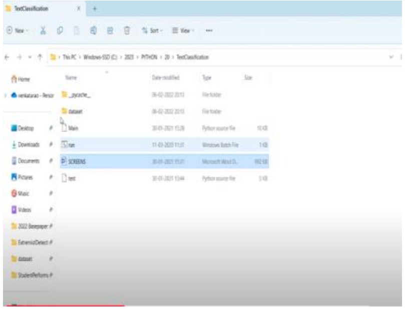
In the proposed system, the system adopted a comprehensive approach to the pickpocket detection problem. The overall framework of our solution is illustrated in this system. The system first partitioned the city area into regions with functional categories. Then, the mobility characteristics of passengers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

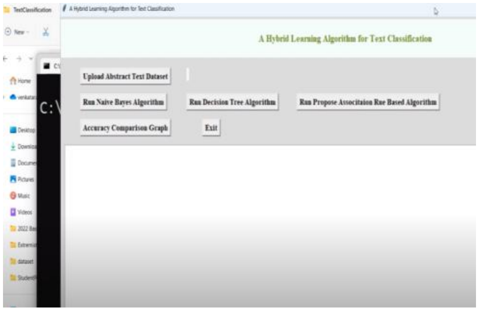
ISSN2454-9940www.ijsem.org
Vol 14, Issuse.4 Dec 2023

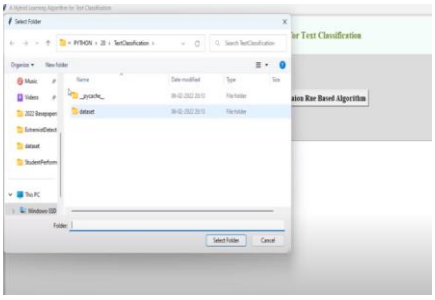
can be summarized as follows. Firstly, we identified a number of features that may be extracted from AFC records and are potentially useful for distinguishing thieves from regular passengers. Secondly, a two-step approach was proposed to make the suspect detection problem practical in a large-scale data 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 realworld, 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.

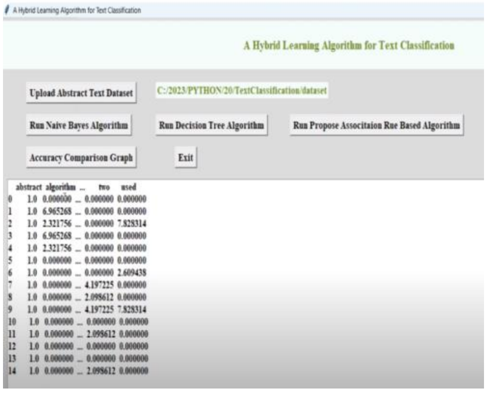
WORKING METHODOLOGY

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.

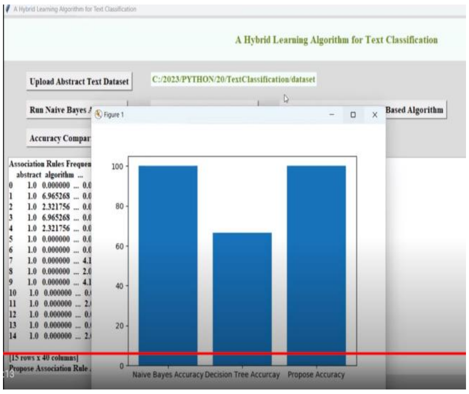



TextClassification
A Hybrid Learning Algorithm for Text Classification
New
A Hybrid Learning Algorithm for Text Classification
Upload Abstract Text Dataset
Run Naive Bayes Algorithm
Run Decision Tree Algorithm
Run Propose Associtaion Rue Based Algorithm
Accuracy Comparison Graph
Exit
Home
Desktop
Downloads
Documents
Pictures
Music
Videos




A Hybrid Learning Algorithm for Text Classification
A Hybrid Learning Algorithm for Text Classification
Upload Abstract Text Dataset
C:/2023/PYTHON/20/TextClassification/dataset
Run Naive Bayes Algorithm
Run Decision Tree Algorithm
Run Propose Associtaion Rue Based Algorithm
Accuracy Comparison Graph
Exit

```
   abstract  algorithm  ...       two      used
0       1.0   0.000000  ...  0.000000  0.000000
1       1.0   6.965268  ...  0.000000  0.000000
2       1.0   2.321756  ...  0.000000  7.828314
3       1.0   6.965268  ...  0.000000  0.000000
4       1.0   2.321756  ...  0.000000  0.000000
5       1.0   0.000000  ...  0.000000  0.000000
6       1.0   0.000000  ...  0.000000  2.609438
7       1.0   0.000000  ...  4.197225  0.000000
8       1.0   0.000000  ...  2.098612  0.000000
9       1.0   0.000000  ...  4.197225  7.828314
10      1.0   0.000000  ...  0.000000  0.000000
11      1.0   0.000000  ...  2.098612  0.000000
12      1.0   0.000000  ...  0.000000  0.000000
13      1.0   0.000000  ...  0.000000  0.000000
14      1.0   0.000000  ...  2.098612  0.000000
```


A Hybrid Learning Algorithm for Text Classification
A Hybrid Learning Algorithm for Text Classification
Upload Abstract Text Dataset
C:/2023/PYTHON/20/TextClassification/dataset
Run Naive Bayes A
Based Algorithm
Accuracy Compar
Figure 1
100
80
60
40
20
0
Naive Bayes Accuracy
Decision Tree Accurcay
Propose Accuracy

```
Association Rules Frequen
   abstract  algorithm  ...
0       1.0   0.000000  ...  0.0
1       1.0   6.965268  ...  0.0
2       1.0   2.321756  ...  0.0
3       1.0   6.965268  ...  0.0
4       1.0   2.321756  ...  0.0
5       1.0   0.000000  ...  0.0
6       1.0   0.000000  ...  0.0
7       1.0   0.000000  ...  4.1
8       1.0   0.000000  ...  2.0
9       1.0   0.000000  ...  4.1
10      1.0   0.000000  ...  0.
11      1.0   0.000000  ...  2.
12      1.0   0.000000  ...  0.
13      1.0   0.000000  ...  0.
14      1.0   0.000000  ...  2.

[15 rows x 40 columns]
Propose Association Rule
```

CONCLUSION

In this paper, we developed a suspect detection and tracking system by mining large-scale transit records. The system assists in identifying pickpocket suspects' and enables active surveillance in high-risk areas. Specifically, we first constructed a feature representation for profiling passengers. Then, we established a novel two-step framework to distinguish regular passengers from pickpocket suspects. Finally, we leveraged real-world datasets from multiple sources for model training and validation, and implemented a prototype system for end users. Experimental results on real-world data showed the effectiveness of our proposed approach.

REFERANCES

[1] Awad W.A, (2012), ―Machine Learning Algorithms In Web Page Classification‖, International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 5,

[2] Apoorva M. &Anupam S. (2017), ―From Machine Learning to Deep Learning: Trends and Challenges‖ CSI Communications Pp 10-11

[3] AlexyBhowmick&Shyamanta M. Hazarika 2016, ―Machine Learning for E-mail Spam Filtering: Review, Techniques and Trends‖ arXiv:1606.01042v1 [cs.LG] 3 Jun 2016

[4] Asha T, Shravanthi U.M, Nagashree N, & Monika M (2013), ― Building Machine Learning Algorithms on Hadoop for Bigdata‖ International Journal of Engineering and Technology Volume 3 No. 2

[5] Arthur V. Ratz, 11 Mar 2018, ―Naïve Bayesian Anti-Spam Filter Using Node.JS, JavaScript And Ajax Requests‖ under The Code Project Open License (CPOL)

[6] Bart V. L. (2017), Machine Learning:A Revolution in Risk Management and Compliance?‖, the capco institute journal of financial transformation

[7] Davidson T, Danawarmsley, Micheal Macy & Ingmar Webar, 2017, ‗Automated hate speech detection and the problem of offensive language', proceedi.ngs of the eleventh international association for the advancement of artificial intelligence (AAAI) conference on web and social media (ICWSM), www.aaai.org.

[8] M. Durairaj and A. AlaguKarthikeyan, (2017), ―Efficient Hybrid Machine Learning Algorithm for text Classification,‖ International Journal on Recent and Innovation Trends in Computing and Communication, Vol.5(5), 680-688, 2017. ISSN:2321-8169. [UGC approved journal list No. 49222, IF: 5.75

[9] Georgios K. Pitsilis, HeriRamampiaro and HelgeLangseth, 2018 ―Detecting Offensive Language in Tweets Using Deep Learning‖ arXiv:1801.04433v1 [cs.CL] 13 Jan 2018

[10] Gröndahl, T., Pajola, L., Juuti, M., Conti, M., &Asokan, N. (2018). All You Need Is "Love": Evading Hate SpeechDetection. In Proceedings of the 11th ACM Workshop on Artificial Intelligence and Security (pp. 2-12). New York: ACM. https://doi.org/10.1145/3270101.3270103.