International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

COMPOSITE BEHAVIORAL MODELING FOR IDENTITY THEFT DETECTION IN ONLINE SOCIAL NETWORKS

Dasari Kusuma Prabha (MCA Scholar), B V Raju College, Vishnupur, Bhimavaram, West Godavari District, Andhra Pradesh, India, 534202.

Dr. V. Bhaskara Murthy, B V Raju College, Vishnupur, Bhimavaram, West Godavari District, Andhra Pradesh, India, 534202.

ABSTRACT- In this work, we aim at building a bridge from coarse behavioral data to an effective, quick-response, and robust behavioral model for online identity theft detection. We concentrate on this issue in online social networks (OSNs) where users usually have composite behavioral records, consisting of multidimensional low-quality data, e.g., offline check-ins and online user-generated content (UGC). As an insightful result, we validate that there is a complementary effect among different dimensions of records for modeling users' behavioral patterns. To deeply exploit such a complementary effect, we propose a *joint* (instead of *fused*) model to capture both online and offline features of a user's composite behavior. We evaluate the proposed joint model by comparing it with typical models and their fused model on two real-world datasets: Foursquare and Yelp. The experimental results show that our model outperforms the existing ones, with the area under the receiver operating characteristic curve (AUC) values 0.956 in Foursquare and 0.947 in Yelp, respectively. Particularly, the *recall* (true positive rate) can reach up to 65.3% in Foursquare and 72.2% in Yelp with the corresponding *disturbance rate* (false-positive rate) below 1%. It is worth mentioning that these performances can be achieved by examining only one composite behavior, which guarantees the low response latency of our method. This study would give the cybersecurity community new insights into whether and how real-time online identity authentication can be improved via modeling users' composite behavioral patterns.

1. INTRODUCTION

The rapid development of the Internet, more and more affairs, e.g., mailing health caring shopping booking hotels, and purchasing tickets, are handled online Meanwhile, the Internet also brings sundry potential risks of invasions, such as losing financial identity theft and privacy leakage Online accounts serve as the agents of users in the cyber world. Online identity theft is a typical online crime which is the deliberate use of another person's account usually as a method to gain a financial advantage or obtain credit and other benefits in another person's name. As a matter of fact, compromised accounts are usually the portals of most cybercrimes such as blackmail fraud and spam Thus, identity theft detection is essential to guarantee users' security in the cyber world. Traditional identity authentication methods are mostly based on access control schemes, e.g., passwords and tokens But users have some overheads in managing dedicated passwords or tokens. Accordingly, the biometric identification is delicately

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

introduced to start the era of password-free. However, some disadvantages make these access control schemes still incapable of being effective in real-time online services

1) They are not *nonintrusive*. Users have to spend extra time in the authentication.

2) They are not *continuous*. The defending system will fail to take further protection once the access control is broken. Behavior-based suspicious account detection [16], [18], [19] is a highly anticipated solution to pursue a nonintrusive and continuous identity authentication for online services. It depends on capturing users' suspicious behavior patterns to discriminate the suspicious accounts. The problem can be divided into two categories: fake/sybil account detection [20] and compromised account detection [21]. The fake/Sybil account's behaviors usually do not conform to the behavioral pattern of the majority. Meantime, the compromised account usually behaves in a pattern that does not conform to its previous one, even behaves like fake/sybil accounts. It can be solved by capturing *mutations* of users' behavioral patterns.Comparing with detecting compromised accounts, detecting fake/sybil accounts is relatively easy since the latter's behaviors are generally more detectable than the former's. It has been extensively studied and can be realized by various population-level approaches, e.g., clustering [22], [23], classification [5], [24]–[26] and statistical or empirical rules [8], [27], [28]. Thus, we *only* focus on the compromised account detection, commonly called *identity theft detection*, based on individual-level behavioral models.

2. EXISTING SYSTEM

introduced hand movement, orientation, and grasp (HMOG), a set of behavioral features to continuously authenticate smartphone users. Rajoub and Zwiggelaar used thermal imaging to monitor the periorbital region's thermal variations and test whether it can offer a discriminative signature for detecting deception. However, these biometric technologies usually require expensive hardware devices which makes it inconvenient and difficult to popularize. explored a multimodal deception detection approach that relied on a novel dataset of 149 multimodal recordings, and integrated multiple physiological, linguistic, and thermal features. These works indicated that users' behavior patterns can represent their identities. Many studies turn to utilize users' behavior patterns for identifications. Behavior-based methods were born at the right moment, which plays important roles in a wide range of tasks including preventing and detecting identity theft. Typically, behavior-based user identification includes two phases: user profiling and user identifying. User profiling is a process to characterize a user with his/her history behavioral data. Some works focus on statistical characteristics, such as the mean, variance, median, or frequency of a variable, to establish the user profile. Naini *et al.* [55] studied the task of identifying the users by matching the histograms of their data in the anonymous dataset with the histograms from the original dataset. But it mainly relied on experts' experience since different cases usually have different characteristics. proposed a behavior-based method to identify compromises of individual high-

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

profile accounts. However, it required high-profile accounts which were difficult to obtain. Other researchers discovered other features, such as tracing patterns, topic and spatial distributions, to describe user identity. conducted a study on online user behavior by collecting and analyzing user clickstreams of a well-known OSN. developed a topic model extending the LDA to identify the active users. presented a technique based on principal component analysis (PCA) that accurately modeled the "like" behavior of normal users in Facebook and identified significant deviations from it as anomalous behaviors. proposed an approach that involved the novel collection of online news stories and reports on the topic of identity theft. Lichman and Smyth [48] proposed MKDE model to accurately characterize and predict the spatial pattern of an individual's events. Tsikerdekis and Zeadally presented a detection method based on nonverbal behavior for identity deception, which can be applied to many types of social media. These methods above mainly concentrated on a specific dimension of the composite behavior and seldom thought about utilizing multidimensional behavior data. explored the complex interaction between social and geospatial behavior and demonstrated that social behavior can be predicted with high precision. It indicated that composite behavior features can identify one'sproposed a probabilistic generative model combining the use of spatiotemporal data and semantic information to predict user's behavior. presented POISED, a system that leverages the differences in propagation between benign and malicious messages on social networks to identify spam and other unwanted content. These studies implied that composite behavior features are possibly helpful for user identification.

Disadvantages

1) LDA model performs poorly in both datasets which may indicate its performance is strongly sensitive to the data quality.

2) CF-KDE and LDA model performs not well in Yelp dataset comparing to Foursquare dataset, but the fused model [17] observes a surprising reversion.

3) The joint model based on *relative anomalous score Sr* outperforms the model based on *logarithmic anomalous score Sl* .

4) The joint model (i.e., JOINT-SR, the joint model in the following content of the system all refer to the joint model based on *Sr*) is indeed superior to the fused model.

3. PROPOSED SYSTEM

In this article, we propose an approach to detect identity theft by using multidimensional behavioral records which are possibly insufficient in each dimension. According to such characteristics, we choose the online social network (OSN) as a typical scenario where most users' behaviors are coarsely recorded [39]. In the Internet era, users' behaviors are composited by offline behaviors, online behaviors, social behaviors, and perceptual/cognitive behaviors. The behavioral data can be collected in many applications, such as offline check-ins in location-based services (LBSs), online tips-posting in instant messaging services, and social relationship making in online social services. Accordingly, we design our method based on users' composite behaviors by these categories. In OSNs, user behavioral data

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

that can be used for online identity theft detection are often too low-quality or restricted to build qualified behavioral models due to the difficulty of data collection, the requirement of user privacy, and the fact that some users have a few several behavioral records. We devote ourselves to proving that a high-quality (effective, quick response, and robust) behavioral model can be obtained by integrally using multidimensional behavioral data, even though the data is extremely insufficient in each dimension.

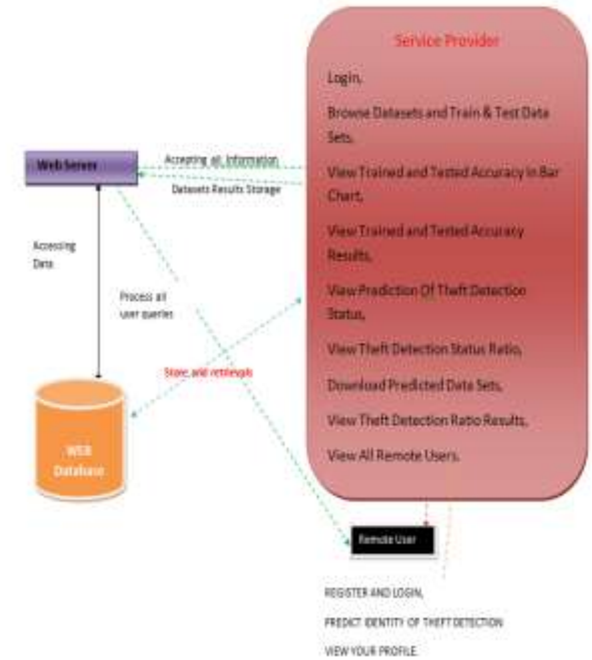
Advantages

1) We propose a joint model, CBM, to capture both online and offline features of a user's composite behavior to fully exploit coarse behavioral data.

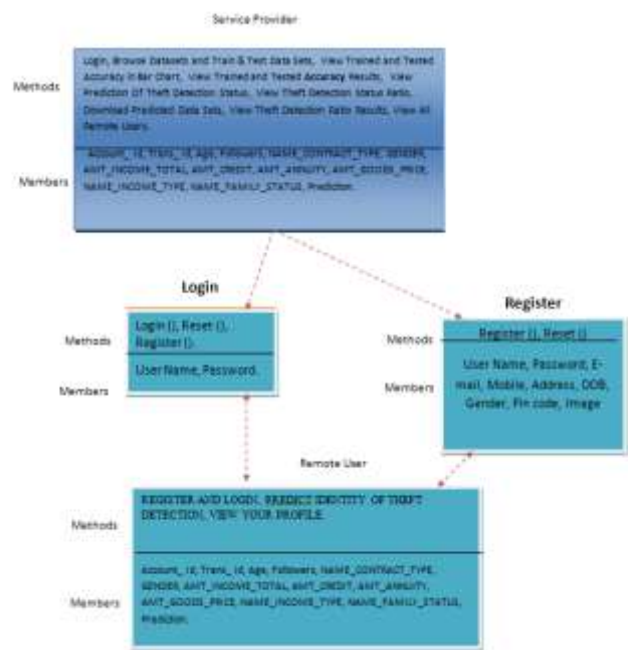
2) We devise a relative anomalous score Sr to measure the occurrence rate of each composite behavior for realizing real-time identity theft detection.

3) We perform experiments on two real-world datasets to demonstrate the effectiveness of CBM. The results show that our model outperforms the existing models and has the low response latency.

Architecture Diagram



➢ **Class Diagram :**



4. SYSTEM STUDY

4.1 FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are

ECONOMICAL FEASIBILITY
TECHNICAL FEASIBILITY
SOCIAL FEASIBILITY

5. SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

6. CONCLUSION

We investigate the feasibility of building a ladder from low-quality behavioral data to a high-performance behavioral model for user identification in OSNs. By deeply exploiting the complementary effect among OSN users' multidimensional behaviors, we propose a joint probabilistic generative model by integrating online and offline behaviors. When the designed joint model is applied to identity theft detection in OSNs, its comprehensive performance, in terms of the detection efficacy, response latency, and robustness, is validated by extensive evaluations on real-life OSN datasets. Particularly, the joint model significantly outperforms the existing fused model. Our behavior-based method mainly aims at detecting identity

7. REFERENCES

[1] J. Onaolapo, E. Mariconti, and G. Stringhini, "What happens after you are pwnd: Understanding the use of leaked Webmail credentials in the wild," in *Proc. Internet Meas. Conf.*, Nov. 2016, pp. 65–79.

[2] A. Mohan, "A medical domain collaborative anomaly detection framework for identifying medical identity theft," in *Proc. Int. Conf. Collaboration Technol. Syst. (CTS)*, May 2014, pp. 428–435.

[3] Y.-A. de Montjoye, L. Radaelli, V. K. Singh, and A. S. Pentland, "Unique in the shopping mall: On the reidentifiability of credit card metadata," *Science*, vol. 347, no. 6221, pp. 536–539, Jan. 2015.

[4] P. Hyman, "Cybercrime: It's serious, but exactly how serious?" *Commun. ACM*, vol. 56, no. 3, pp. 18–20, Mar. 2013.

[5] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: Automated identity theft attacks on social networks," in *Proc. 18th Int. Conf. World Wide Web (WWW)*, 2009, pp. 551–560.

[6] J. Lynch, "Identity theft in cyberspace: Crime control methods and their effectiveness in combating phishing attacks," *Berkeley Technol. Law J.*, vol. 20, no. 1, pp. 259–300, 2005.

[7] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towards detecting compromised accounts on social networks," *IEEE Trans. Dependable Secure Comput.*, vol. 14, no. 4, pp. 447–460, Jul. 2017.

[8] T. C. Pratt, K. Holtfreter, and M. D. Reisig, "Routine online activity and Internet fraud targeting: Extending the generality of routine activity theory," *J. Res. Crime Delinquency*, vol. 47, no. 3, pp. 267–296, Aug. 2010.

[9] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time URL spam filtering service," in *Proc. IEEE Symp. Secur. Privacy*, May 2011, pp. 447–462.

[10] H. Li *et al.*, "Bimodal distribution and co-bursting in review spam detection," in *Proc. 26th Int. Conf. World Wide Web*, Apr. 2017, pp. 1063–1072.

[11] A. M. Marshall and B. C. Tompsett, "Identity theft in an online world," *Comput. Law Secur. Rev.*, vol. 21, no. 2, pp. 128–137, Jan. 2005.

[12] B. Schneier, "Two-factor authentication: Too little, too late," *Commun. ACM*, vol. 48, no. 4, p. 136, Apr. 2005.

[13] M. V. Ruiz-Blondet, Z. Jin, and S. Laszlo, "CEREBRE: A novel method for very high accuracy event-related potential biometric identification," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1618–1629, Jul. 2016.

[14] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Biometric recognition in automated border control: A survey," *ACM Comput. Surv.*, vol. 49, no. 2, p. 24, 2016.

[15] B. A. Rajoub and R. Zwiggelaar, "Thermal facial analysis for deception detection," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 1015–1023, Jun. 2014.

[16] M. M. Waldrop, "How to hack the hackers: The human side of cybercrime," *Nature*, vol. 533, no. 7602, pp. 164–167, May 2016.

[17] C. Wang, B. Yang, J. Cui, and C. Wang, "Fusing behavioral projection models for identity theft detection in online social networks," *IEEE Trans. Comput. Social Syst.*, vol. 6, no. 4, pp. 637–648, Aug. 2019.

[18] C. Shen, Y. Li, Y. Chen, X. Guan, and R. A. Maxion, "Performance analysis of multi-motion sensor behavior for active smartphone authentication," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 48–62, Jan. 2018.

[19] C. Wang and H. Zhu, "Representing fine-grained co-occurrences for behavior-based fraud detection in online payment services," *IEEE Trans. Dependable Secure Comput.*, early access, May 4, 2020, doi: 10.1109/TDSC.2020.2991872.

[20] H. Zheng *et al.*, "Smoke screener or straight shooter: Detecting elite sybil attacks in user-review social networks," in *Proc. 25th Annu. Netw. Distrib. Syst. Secur. Symp. (NDSS)*, San Diego, CA, USA, Feb. 2018, pp. 259–300