Participatory Detection of Identity Theft on Mobile Social Platforms

R.VaraPrasad[1], Giridhar Babu[2], K.Abdul Azeez Khan[3]
G.Pullaiah College of Engineering and technology,Kurnool,Andhra Pradesh

Abstract- *Popularity of smart devices has led to increasing use of social networking services for various purposes. Despite its benefits, the ubiquity of social network services introduces vulnerabilities to malicious behavior such as Sybil attacks or identity theft. As the 5G Era leads to the convergence of social, wireless, and mobile networks by enabling synergistic interplay between these networks, it is possible to take advantage of mobile edge computing in the detection of compromised social profiles in mobile and online social network platforms. In this paper, we propose a framework for participatory detection of identity theft on social networking platforms which would exploit the computing power of the user equipment. The proposed framework empowers the connections of a user in a social platform to cooperate on the verification of a social profile. Through a proof-of-concept study, we show that the proposed framework can detect anomalous behavior in the social profile by having each connection work on a different feature subset without semantic analysis. Our numerical results show that if the initial matching threshold in a decision tree is set properly, compromised accounts can be identified by the mobile platforms of the connections without undergoing heavy central processing.*

Keywords- Mobile social networks, smart cities, user verification, machine learning

I. INTRODUCTION

Online and mobile social networking (MSN) services have been reported as effective tools to acquire and share useful data through citizens in smart cities [1, 2]. To integrate mobile social networks with the emerging smart city infrastuctures, acquiring tweets in real time, efficient storage and filtering mechanisms for the acquired data, semantic and sentiment analysis, as well as intuitive visualization are listed as major requirements [3].

Early efforts in social network research proposed using stochastic methods to model user behavior in online social platforms [4]. On the other hand, mobile social networks are prone to several security and privacy challenges [5]. With the convergence of wireless communications, mobile networks and social networking services, identification of genuine profiles and detection of anomalous user behavior on online and mobile social platforms has become a crucial challenge. Despite various benefits of mobile and online social networking services, user accounts are prone to malicious attacks concerning security, privacy and trust. Several types of defense mechanisms are available to distinguish malicious accounts from non-malicious ones [6, 7]. Besides the presence of Sybil-like malicious nodes that aim to control multiple nodes in a network, identity theft is the primary concern of all users who are participants of mobile social networks. Content-based identification of social network users is a common approach in all solutions against these types of attacks [8, 9]. Identity theft on mobile social networks is not only a threat that af-fects data acquisition in subsequent times but also crucial for them as they also offer on-the-fly financial transactions such as payments, fund transfers and so on [10].

In this paper, we propose a participatory solution to detect identity theft against a followed user in a mobile social network. To this end, we propose a conceptual framework that consists of multiple follower nodes, an aggregator node, and a followed node in a mobile social network. The aggregator node acts as the broker in the system as it starts the training process over the participant followers. We propose that the aggregator node aggregates multiple decisions that arrive from multiple decision trees each of which consists of a smaller subset of the followed node's feature set. By pulling real Twitter data and processing it to be applicable to simulation scenarios, we investigate the impact of training parameters on the success rate and responsiveness of the system. Our simulations show that by setting the initial matching threshold at the individual decision trees of the participants to 50%, high detection rates (close to > 90%) as well as efficient responsiveness to identity theft attempts can be achieved.

II. RELATED WORK AND MOTIVATION

Detection of compromised social networking accounts has re-cently attracted several researchers. To the best of our knowl-edge, the authors in [11] presented the first comprehensive platform, COMPA to detect compromised accounts on popular social network platforms. Statistical modeling and anomaly detection are two fundamental components of COMPA to form behavioral patterns on user accounts and detect anomalous pat-terns. COMPA aims to group similar messages and applies a threshold-based approach to tag a group of messages as anoma-lous cluster in the profile. The authors in [12] compare the solutions against identity spoofing in social media. In [8], the authors introduce a content and text similarity measure to de-tect compromised accounts on Twitter and Facebook. In [13], the authors aim to build social behavior profiles for Facebook users, and detect deviations from a typical social conduct of a user. In the same study, the authors report that eight behav-ioral components make up the social behavioral profile of a user. The authors in [14] introduce Success Outdegree Propor-tion, Reverse Pagerank, Recipient Clustering Coefficient and Legitimate Recipient Proportion criteria while limiting their focus to email accounts. In [15], the authors propose using deep learning methods (i.e. convolutional neural networks) to detect an anomaly on the account based on the uploaded pictures. As the interplay between mobile social networks and 5G communications is inevitable [16, 17], we seek to answer the following question: Is it viable to delegate less-complex computational decision tasks to a subset of the connections of a user and aggregate the individual decisions at another mobile node? In response to this question, the next section proposes a conceptual framework along with its proposed components and functions.

III. PARTICIPATORY DETECTION FRAMEWORK

As opposed to the existing centralized solutions, we propose to address identity theft on social platforms through commu-nity engagement. A minimalist illustration of the proposed framework is presented in Fig. 1. As seen in the figure, a user's identity is verified by the followers in a participatory manner. One follower acts as the aggregator and assumes the responsibility of initiating the training iterations, reconfiguring the matching thresholds as well as the feature weights, and aggregating the decision values collected from other users. The followers call API searches and work on JSON objects which are returned by the APIs (e.g. Twitter API). Every follower builds a decision tree of arbitrarily selected feature subset by using the weights provided by the aggregator node. Here, a fundamental assumption is that the aggregator node is a trusted entity. Indeed, the aggregation process in the aggre-gator node will result in additional battery use. However, since the aggregator node is a trusted entity, it has to be incentivized effectively at the time of its appointment.

As the decision trees are popular, powerful and easy to implement tools for knowledge discovery [18], we enforce decision trees to the participants; however considering the challenges of the mobile environment, we let each participant i select a small subset of the entire feature set provided by the aggregator, i.e. s_i F. Algorithm 1 lists the steps of the train-ing procedure from the aggregator's standpoint. The aggrega-tor maintains a number of parameters, namely the weight set for the features of the tweets of the followed user (= f$!_i$ g), the feature set (F), initial and current matching thresholds to be used in its aggregation function (T h_{init}; T $h_{current}$), cur-rent time window of interest for the iteration ($W_{current}$) and the iteration count. The training starts with the aggregator nodes' sharing the current window with the participants. Upon receiving the start window message accompanied with the feature weights, each participant picks its own small subset of features s_i F, and builds the decision tree accordingly. The decision value is sent to the aggregator node at the end of each iteration. Upon the receipt of each follower's decision, the aggregator forms the aggregated decision value (D_{aggr}) and calculates a new threshold value as a running average sum of the current threshold (T $h_{current}$) and the current ag-gregated decision value (D_{aggr}) as shown in (1). Thus, the current matching threshold has a higher weight on the next value of the matching threshold in comparison to the currently aggregated decision.

$$I\ h_{current} = \frac{itCount\ \ 1}{itCount} I\ h_{current} + \frac{1}{itCount} D_{aggr}\ (1)$$

Reconfiguration of the weight set is as follows: If the cur-rent decision is greater than the best decision, current weight set is assigned as the best weight set, otherwise the best weight set is maintained. In the next iteration, the aggregator sends the reconfigured version of the best weight set (^{0}) to the fol-lowers as shown below in (2). In the equation, the function randXY () returns a tuple fa; bg where a and b denote two indices in the feature weights set. The aggregator increases one of these weights and decreases the other weight by where

2 (0; 1]. At the end of the next iteration, if the aggregated matching decision value is greater than the best decision value, the reconfigured feature weights set ^{0} is set as the new best weight set, .

$$!_i^{0} = \begin{cases} & !_i;fig\ 6\ randXYf\ ()g \\ !_i & ;\quad i = a \qquad a; b\ = randXY\ () \\ & \wedge f \qquad g \end{cases} \quad (2)$$

IV. NUMERICAL RESULTS

4.1. Experimental settings

To evaluate the proposed framework we have set up a simulation environment by using 20 diverse and popular public Twitter accounts. We have used the code in [19] to pull the most recent 500 tweets of these accounts as of March 27, 2017 from Twitter API. The json files of the 20 accounts are available at [20]. In this proof of concept study, we present a randomly selected 5 Twitter accounts as the followed users (Followed-1...Followed-5) in our simulations.

In the preparation of the training data, we have used the following feature set for each tweet: F ={Geo_enabled, hashtagList, isF AV ed, isM arkedAsP ossiblySensitive, isP rotected, quoteList, retweets, symbolList, timezone, timeQuarterOf Day, urlList }. As some of these features (i.e.Geo_enabled, timezone, isProtected) are fixed and have the same value for all accounts, we have introduced a slight modi-fication as follows: An account's geo-enabled, and protected parameters are turned on and off at random intervals during the observation windows, and the tweets are assigned these modified values. similarly timezone of a user is changed at random intervals during the monitored timeline, and the time zone of the tweets during that interval is modified accordingly.

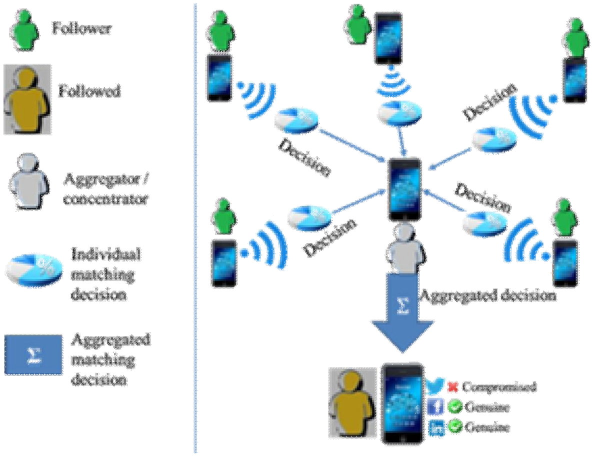
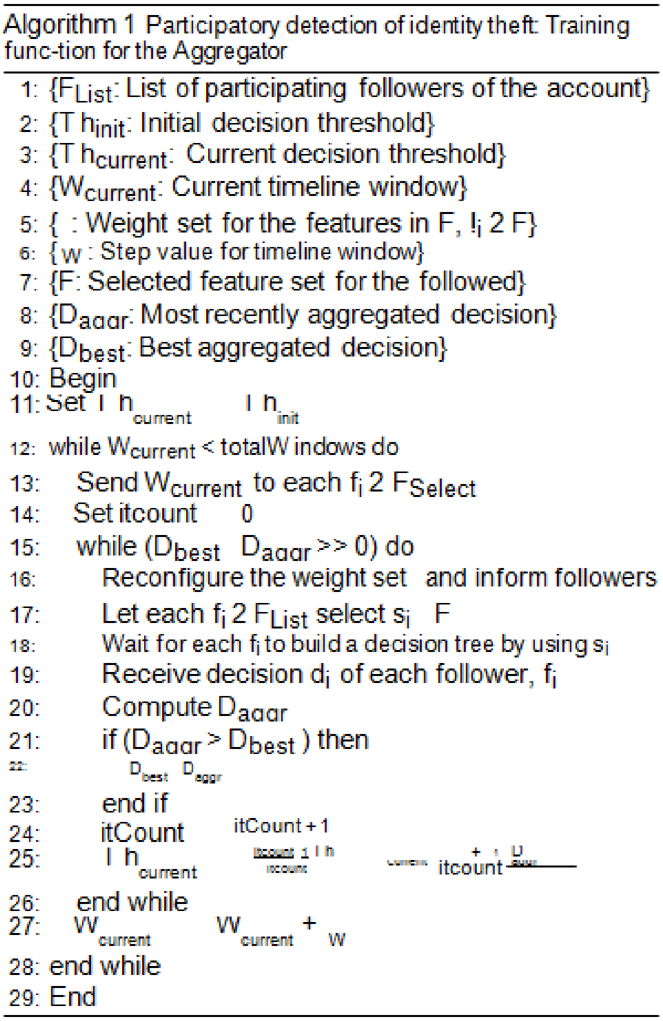


Fig. 1. Minimalist illustration of the proposed framework for participatory detection of social identity theft.

In our simulations, each mobile follower of a followed account in Fig. 1 is represented by a thread that generated a decision tree by selecting its own subset of features s_i out of the features set, F where 2 js_ij 5. In each iteration, we set the number of trees (i.e. followers) to 100. Due to the long training period, we set an upper bound of 10 on the number of iterations in each window size.

In order to investigate the impact of the initial match-ing threshold ($T\ h_{init}$), we set the initial threshold to five different values in the simulations as follows: $T\ h_{init}$ 2 f0:5; 0:6; 0:7; 0:8; 0:9g. For each followed user, the entire 500-tweet timeline is analyzed in windows of 100 tweets, incremented tweet by tweet for a total of 400 windows. A final decision is reached based on the decisions of all 400 windows.

At the end of the training, we inject Additive White Gaus-sian Noise (AWGN) to each timeline to simulate identity theft on the social platform. The followers run their decision trees at every 70 tweet windows which is continuously slided by 40 tweets. Thus, window-1 includes the tweets from tweet-0 to tweet-69 whereas window-2 includes the tweets from tweet-40 to tweet-109 and so on. The final values of the matching threshold $T\ h_{current}$ and the feature weight set obtained at the end of the training period are used by all followers.



Algorithm 1 Participatory detection of identity theft: Training func-tion for the Aggregator

```
{F_List: List of participating followers of the account}
{T h_init: Initial decision threshold}
{T h_current: Current decision threshold}
{W_current: Current timeline window}
{ : Weight set for the features in F, !_i 2 F}
{ w : Step value for timeline window}
{F: Selected feature set for the followed}
{D_aggr: Most recently aggregated decision}
{D_best: Best aggregated decision}
Begin
Set T h_current   T h_init
while W_current < totalW indows do
    Send W_current to each f_i 2 F_Select
    Set itcount    0
    while (D_best   D_aggr >> 0) do
        Reconfigure the weight set   and inform followers
        Let each f_i 2 F_List select s_i   F
        Wait for each f_i to build a decision tree by using s_i
        Receive decision d_i of each follower, f_i
        Compute D_aggr
        if (D_aggr > D_best ) then
            D_best   D_aggr
        end if
        itCount      itCount + 1
        T h_current      (itcount 1)/itcount T h_current + 1/itcount D_aggr
    end while
    W_current      W_current +   w
end while
End
```

4.2. Results

We investigate the impact of the initial matching threshold, $T\ h_{init}$ on the value of the final threshold, success ratio of detecting identity theft attempts, and the responsiveness of the framework. Final threshold is the final value obtained at the end of the training, and to be used by the decision trees of the mobile followers. The other performance metrics that we introduce are Identity Theft Detection Rate: (ITDR) and Responsiveness to Identity Theft Attempt (RITA) as formulated in (3)-(4). As seen in the equations, IT DR denotes the ratio of the truly detected noisy timeline windows over the entire noisy timeline whereas RIT A denotes the earliest window in the timeline where an anomaly has been detected. In a robust system, the former is aimed to be maximized whereas the latter is aimed to be minimized.

$$IT\,DR = {}^{P}{}_{w2}W \quad \underline{{}_{W}{}^{I}\,N_{w}} \qquad (3)$$

$$RIT\,A = \underset{w2W}{\mathrm{argminf}}T\,N_{W} = 1g \qquad (4)$$

Table 1. Simulation Settings

Variables	Values
Timeline length	latest 500 tweets
Number of participants (trees)	100
Number of Iterations	10
Initial Threshold(s)	0.5 | 0.6| 0.7| 0.8| 0.9
Size of Sliding Window	70
Sliding Window Step Size	40
Maximum Features per Tree	5
Total Features	11
Initial weights of the features	8f 2 F; f = 1
Identity theft	Simulated by AWGN

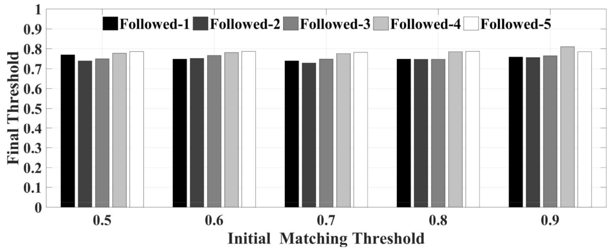


Fig. 2. Initial matching thresholds versus final thresholds (after training) for the five followed Twitter users

In the first step, we investigate the relation between the initial matching threshold and the final matching threshold that is to be used during the detection phase. We demonstrate the results for the five followed accounts as illustrated in Fig. 2. As seen in the figure, majority of the final thresholds fall between between [70%-80%] matching range with a few outliers. Tests were run using five initial thresholds, and all five led to final thresholds in this range. Regardless of the initial matching threshold, the final matching threshold takes its value in this range for the detection of identity theft on a followed account. Indeed, final feature weights are expected to be different.

In the second step, we test the IT DR performance of the proposed framework with respect to varying initial threshold values. As seen in fig. 3, when T hinit is set at 50%, the success rate leads to its highest value for all followed accounts whereas the IT DR performance under T hinit = 50% is followed by by T hinit = 70% and T hinit = 80%. In these three cases with all five of the above datasets, the noisy data was eventually flagged as invalid. Setting the initial matching threshold as high as 90% does not guarantee detection of identity theft on all followed users. As seen in the figures, the community was unable to detect the identity theft on the third followed user.

Lastly, we test the responsiveness to identity theft attempt (RIT A) of the proposed framework. Fig. 4 shows that in most cases, the training of the community (i.e. followers) can be fairly consistent regardless of the initial matching threshold. However, setting the initial threshold to an extremely high

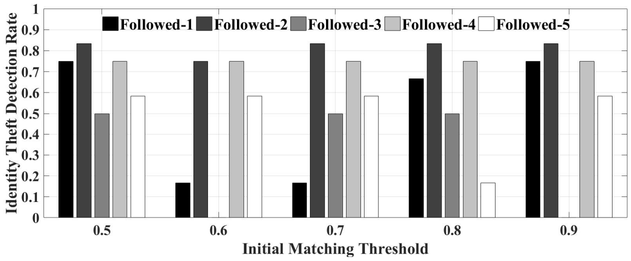


Fig. 3. Identity theft detection rate (IT DR) over 500 tweets when the followed users' identities were compromised

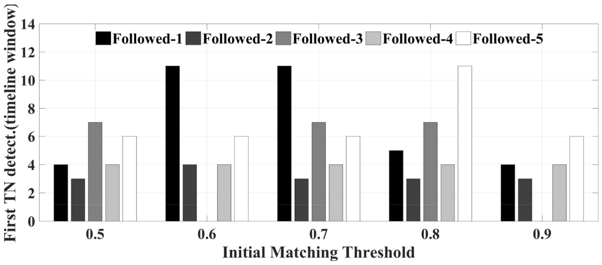


Fig. 4. Sensitivity of the proposed framework (i.e. RITA)

value ($T\,h_{init} = 0{:}9$) results in an exceptional situation where the follower community is unable to catch the noisy data on the timeline of the followed user (e.g. followed-3 in this example). Furthermore, setting the initial matching threshold to high values (e.g. $T\,h_{init} = 0{:}7$; $T\,h_{init} = 0{:}8$) in the training phase leads to extended detection time for some followed accounts (e.g. followed-1 can be detected at the 11th window. One reason of these outliers is interpreted as the random nature of the injected noise over the timeline in which the feature weights are generated leading to lower decision values for the original data, and thus a lower final threshold value is obtained.

V. CONCLUSION

In the 5G Era, digitally-formed social communities play an important role in the provision of smart services in a non-dedicated manner. We have introduced a participatory frame-work to detect identity theft on mobile social networks where each follower participates in the detection process of identity theft by running its own decision tree with the feature weights and matching thresholds that are dynamically reconfigured by an aggregator node. Our numerical results show that setting proper initial matching thresholds to be used by each partic-ipant follower would result in feasible performance in terms of responsiveness and success ratio. Furthermore, such initial setting has been shown to help the aggregator detect an anoma-lous condition as early as the second window of tweets over the entire timeline under study. We are currently investigating the impact of feature set, weights, participants and window size on the performance.

REFERENCES

[1] F. Delmastro, V. Arnaboldi, and M. Conti, “People-centric computing and communications in smart cities,” IEEE Communications Magazine, vol. 54, pp. 122–128, July 2016.

[2] L. Anthopoulos and P. Fitsilis, “Social networks in smart cities: Comparing evaluation models,” in IEEE First International Smart Cities Conference (ISC2), pp. 1–6, Oct 2015.

[3] M. Figueredo, J. Ribeiro, A. Emanuel, A. Cacho, H. Farias, J. Coelho, C. Prolo, and N. Cacho, “Using social network to support smart city initiatives,” in Pro-ceedings of the 21st Brazilian Symposium on Multimedia and the Web, WebMedia ’15, (New York, NY, USA), pp. 101–104, ACM, 2015.

[4] K. Lerman and T. Hogg, “Using stochastic models to describe and predict social dynamics of web users,” ACM Transactions on Intelligent Syst. Technol., vol. 3, pp. 62:1– 62:33, Sept. 2012.

[5] X. Liang, K. Zhang, X. Shen, and X. Lin, “Security and privacy in mobile social networks: challenges and solutions,” IEEE Wireless Communications, vol. 21, pp. 33–41, February 2014.

[6] Y. Zhou, D. W. Kim, J. Zhang, L. Liu, H. Jin, H. Jin, and T. Liu, “Proguard: Detecting malicious accounts in social-network-based online promotions,” IEEE Access, vol. 5, pp. 1990–1999, 2017.

[7] M. Al-Qurishi, M. Al-Rakhami, A. Alamri, M. Alrubaian, S. M. M. Rahman, and M. S. Hossain, “Sybil defense techniques in online social networks: A survey,” IEEE Access, vol. 5, pp. 1200–1219, 2017.

[8] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, “Towards detecting compromised accounts on social networks,” IEEE Transactions on Dependable and Secure Computing, vol. PP, no. 99, pp. 1–1, 2017.

[9] Y. Sha, Q. Liang, and K. Zheng, “Matching user accounts across social networks based on users message,” Procedia Computer Science, vol. 80, pp. 2423 – 2427, 2016.

[10] A. Beikverdi, I. Kim, and J. Song, “Centralized payment system using social networks account,” in 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, pp. 493–499, Dec 2014.

[11] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, “Compa: Detecting compromised accounts on social networks,” in Proceedings of the Network and Distributed System Security Symposium, San Diego, CA, 2012.

[12] D. Trang, F. Johansson, and M. Rosell, “Evaluating algorithms for detection of compromised social media user accounts,” in 2015 Second European Network Intelligence Conference, pp. 75–82, Sept 2015.

[13] R. Shahabadkar, M. K. B, and K. R. Shahabadkar, “Diagnosis of compromised accounts for online social performance profile network,” in 2016 International Conference on Wireless Communications, Signal Process-ing and Networking (WiSPNET), pp. 1552–1557, March 2016.

[14] X. Hu, B. Li, Y. Zhang, C. Zhou, and H. Ma, “Detecting compromised email accounts from the perspective of graph topology,” in Proceedings of the 11th International Conference on Future Internet Technologies, CFI ’16, (New York, NY, USA), pp. 76–82, ACM, 2016

[15] P. Chitrakar, C. Zhang, G. Warner, and X. Liao, “So-cial media image retrieval using distilled convolutional neural network for suspicious e-crime and terrorist ac-count detection,” in IEEE International Symposium on Multimedia (ISM), pp. 493–498, Dec 2016.

[16] G. Araniti, A. Orsino, L. Militano, L. Wang, and A. Iera, “Context-aware information diffusion for alerting messages in 5g mobile social networks,” IEEE Internet of Things Journal, vol. 4, pp. 427–436, April 2017.

[17] L. Jiang, H. Tian, Z. Xing, K. Wang, K. Zhang, S. Maharjan, S. Gjessing, and Y. Zhang, “Social-aware energy harvesting device-to-device communications in 5g networks,” IEEE Wireless Communications, vol. 23, pp. 20–27, August 2016.

[18] L. Rokach and O. Maimon, Data Mining with Decision Trees Theory and Applications. World Scientific, Dec 2007

[19] "C. Allison, Twitter Graphing Python." Online: https://github.com/ToferC/Twitter_graphing_python, 2016.

[20] "Analyzed accounts and tweets." Online: https://nextconlab.academy/TwitterAPISearch/jsons.zip, 2017.