Deep Neural Network-based Relationship Identification Framework to Discriminate Fake Profile Over Social Media

Suneet Joshi[1], Deepak Singh Tomar[2]
Department of Computer Science
Maulana Azad National Institute of Technology
Bhopal, India

***Abstract*—Involvement of social media like personal, business and political propaganda activities, attracts anti-social activities and has also increased. Anti-social elements get a wider platform to spread negativity after hiding their identity behind fake and false profiles. In this paper, an analytical and methodological user identification framework is developed to significantly binds implicit and explicit link relationship over the end-users graphical perspective. Identify malicious user, its communal information and sockpuppet node. Apart from that, this work provides the concept of the deep neural network approach over the graphical and linguistic perspective of end-user to classify as malicious, fake and genuine. This concept also helps identify the trade-off between the similarity of nodes attributes and the density of connections to classifying identical profile as sockpuppet over social media.**

***Keywords*—Social media; anomaly detection; malicious activity; spam account; fake account; sockpuppet; deep neural network**

I. Introduction

Social media has entered our lives in many areas, among 7.5 billion people globally; 3.1 billion are active on social media. Many activities, such as communication, entertainment, political campaigning, and shopping, are carried out on social media platforms [1]. As a result of this, huge data generated spontaneously on social media platforms continuously emerge. The spread and popularity of social media have attracted the attention of antisocial elements. These people, unfortunately, use social media to scam or cyberbullying activity through a fake account.

Ungenuine user-profiles opened by users for mischievous purposes in social networks such as Facebook, Twitter and LinkedIn are called fake accounts. Fake accounts are usually opened for lack of trust, fear or hiding from anyone, protecting oneself from the potential loss of important news and accessing information by hiding. Apart from this, fake accounts are also opened in celebrities' names to gather followers, run ad campaigns, run negative campaigns about a brand, or get personal information and profile information of users. The credibility and global expansion of social media can infer that fake accounts opened using individuals or companies' names can pose a major problem.

Domenico et al. [2] state that false profiles on social networks are those that do not comply with the terms and conditions established by the platform, they do not belong to real people, they do not belong to the person they indicate, and they pretend to be real profiles existing. They also indicate fake, manual or artisan profiles (created by people) and those generated and manipulated manually and automatically (bots or robots). They mention that there are different types of "tasks" of a fake profile: stalker, cyberbullying, gamers, spammers, pornography, digital reputation, media manipulation, cybercrime.

There are different categories of fake profiles, generated for different purposes. Some of them (gamers or stalkers) may be harmless. Still, others have a clear intention of causing damage or seeking financial gain for themselves, insults, extortion, threats, scams and worst Cases, corruption and grooming of minors.

Recently, researchers applying classification approach to detect fake account over social media. But due to a lack of graphical and linguistic implicit information [3], [4] for end node, the performance of this research does not get significant results. On the other hand, linguistic pattern and geocommunal information of end-user are crucial characteristics to identify the pattern of the end-user.

However, graphical communal characteristics depend upon the implicit and explicit link relationship. The explicit link relationship easily extracted from the graphical structure. Whereas, extraction of the implicit link relationship is a challenging task. Mining of linguistics and behavioural pattern of user-generated content such as, like, dislike, follow, comment and share lead to extract implicit graphical structure.

In this paper, an analytical and methodological user identification framework is developed to significantly binds implicit and explicit link relationship over the end-users graphical perspective. Identify malicious user, its communal information and sockpuppet node. Apart from that, this work provides the concept of the deep neural network approach over the graphical and linguistic perspective of end-user to classify as malicious, fake and genuine. This concept also helps identify the trade-off between the similarity of nodes attributes and the density of connections for Influence maximization.

The organization of the paper is as follows. In the second part, the relevant literature is given, and the social media analysis and fake account detection programs are briefly mentioned. In the third part, the algorithms we developed and used are mentioned. While the evaluation results are mentioned in the fourth section, results and suggestions are given in the last

section.

II. Related Work

Social networking has become an increasingly important application in recent years, because of its unique ability to enable social contact over the internet for geographically dispersed users. A social network can be represented as a graph, in which nodes represent users, and links represent the connections between users.

The purpose of the literature survey is to gain and understand the diverse and dynamic nature of social media data for feature extraction to extract Misuse of Fake Profiles for Review Spam On Social Media [1-7], Detection of fake review spreading community [8,10].

Along with that total eight articles (published in 2016 to 2019) presented in this paper are summarized in Table 1 that contains six columns. The main task of the articles is illustrated in the second column. Column third illustrates method used. Column fourth illustrate method and algorithm used for account verification in different application. Whereas sixth column describes the name of data sets and its source that has been used for evaluating different methodology.

Cresci et al. [5] developed a behaviour model inspired by biological DNA in detecting spambots in social networks in another bot research. By changing the genetic algorithm's different parameters, it was determined how advanced bots escaped from detection techniques, in another Galindo et al. [6] examined political bots in the General Elections. The accounts considered in the study using three different data sets are grouped as bot or human. To classify the data set, features such as the age and location of the relevant Twitter accounts, the length of the user step, the sickness per tweet, and the time between two retweets were used. AdaBoost, logistic regression, support vector machine and naive Bayes have been tested as the classification algorithm. Logistic regression worked the best among them. Based on the data in Chasma, author can say that bots retweet slightly less than real accounts. Also, bots include more external URLs in their tweets than original accounts.

Ruiz et al. [7] claim that when detecting bots on Twitter, follower friends' ratio will not always give us correct results. They think that bots can unfollow accounts that do not automatically follow back. Instead, the text in the tweets of bot accounts is more uniform than the actual accounts. They use text entropy to measure similarity. It also deals with the methods used to access Twitter to detect bot and human accounts. For example, most human accounts use the web or mobile application, while the bots have stated that they use other applications such as API, they also stated that human accounts have a more complex timing behaviour than bots and cyborgs. In this study, they use multiple classification methods as bots or human accounts in the Twitter social network. The process of updating has been carried out. By applying feature extraction techniques to the data set, it has been prepared for the dilution process.

III. Proposed Work

A graphical, linguistics and social theory based relationship identification (RIF) framework is developed to identify malicious end-user over social media, as shown in Fig. 1. This framework amalgamates linguistics, temporal and contextual ethics of user-generated content with profile and graphical information.

The RIF framework extract feature vector to delineate user behaviours and similarity index over social media. Classifying identical profile concerning to similar user via Jaccard coefficient over linguistics pattern of tweets and provide linguistics, temporal and contextual meaning to develop a mathematical model for classifying identical profile as sockpuppet over social media.

A. Data Extraction

RIF framework analyze and extract user pattern from user-generated content, profile and graphical information of social media user. This approach encapsulates social media mining concepts, theories, with the concept of natural language processing to extract the communal intersection of user-generated and profile content from social media.

B. User Feature Vector

RIF framework examine and correlate user profile (u_f), generated data (c_f) with graphical perspective (g_f) of social media data as .

$$\rho = \{u_f \bowtie c_f \bowtie g_f\} \tag{1}$$

The taxonomy of user feature includes profile, content, and graph-based feature, as shown in Fig. 2. Whereas in this work profile-based feature comprises validation of profile information such as suspicious user profile is verified or not, profile age, profile cover, and picture as

$$u_f = \begin{cases} verified & \text{if } v_f = y \\ age & \text{if } not\ immediate \\ cover & \text{if } not\ default \\ picture & \text{if } not\ default \end{cases} \tag{2}$$

However, content-based features include temporal, contextual validation of user-generated data, grammatical quality, and emotional context of surfing nature as shown in Fig. 3.

Temporal taxonomy comprises time interval between tweets(t_g), retweets (rt_g)and its frequency(t_f,rt_f). Contextual content includes term and document frequency of user tweets, Whereas linguistics feature reflects the standard of language script and sensitivity incorporate susceptibility of the user while tweets.

$$t_f = \begin{cases} t_g & = time \\ rt_g & = time \\ t_f & = number \\ rt_f & = number \end{cases} \tag{3}$$

However, graph-based features include validation of structural and relational nature of end-user such as number of friends, follower, friend distribution, etc.

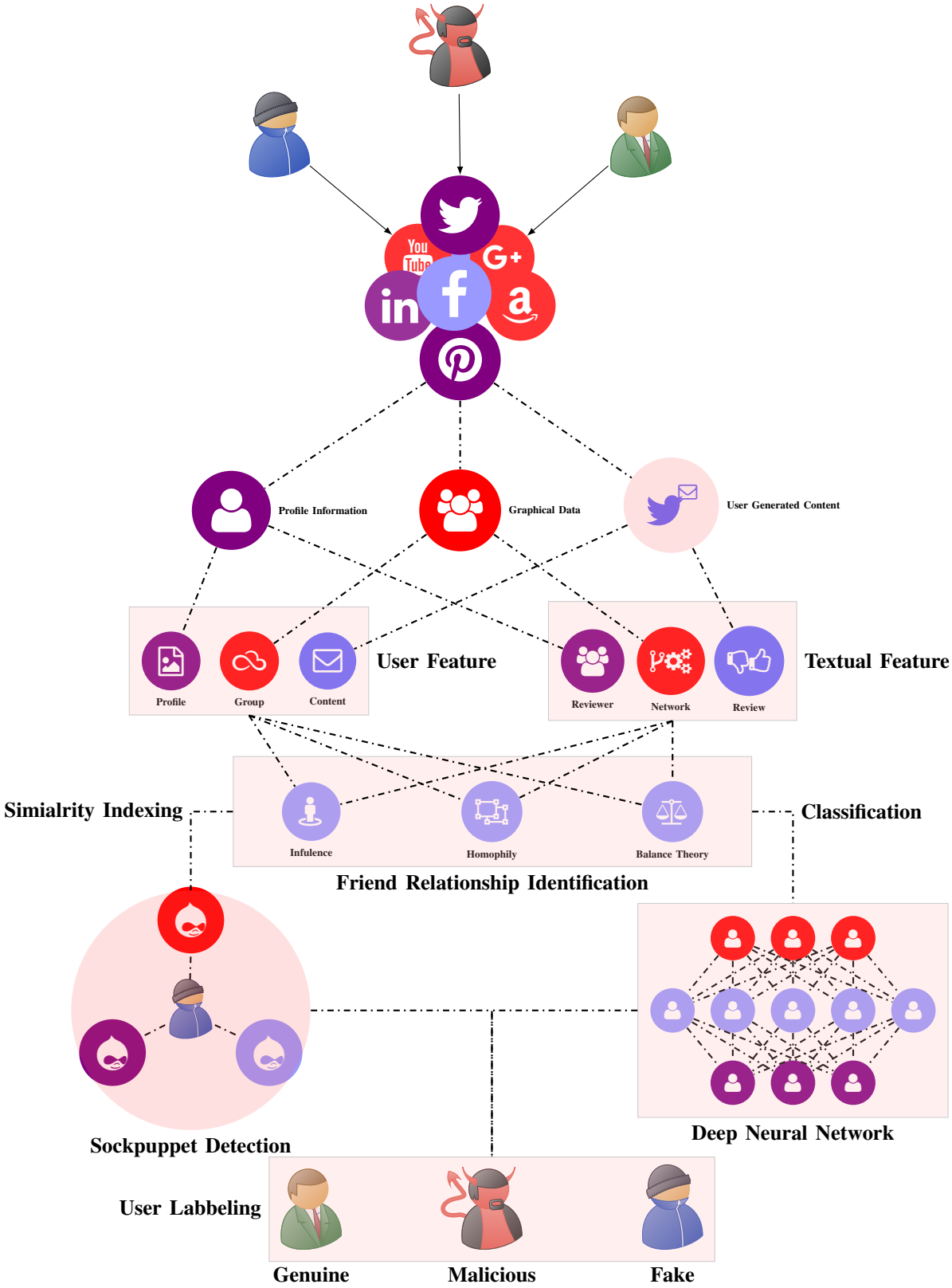

Fig. 1. Proposed Relationship Identification (RIF) Framework for Fake User Identification.

TABLE I. ARTICLE SUMMARY:FAKE ACCOUNT DETECTION

R	Task	Approach	Algo	Data Set	Research Gap
[8]	Cross-Platform Fake account identification	Friend Relationship-Based User Identification	Graph Based	Crude Data Set	Multiple Dimensions Profile Information
[9]	Anomalous Compromised Account Detection	Statistical Anomaly Detection Techniques	Graph Based	Twitter Dataset	Discriminate weightage of features
[10]	Sybil Attacks detection Via Fake profile	Deep-Regression	Graph Based	USA Election Tweeter data set	Handling noisy and malicious data
[11]	Sybil Attacks detection Via Fake profile	Pairing-based Cryptography	Graph Based	Twitter and YouTube dataset	Handling optimized defensive features
[12]	Mining Fake Account	Social Media Mining	Machine Learning	Deceptive Accounts Dataset	Detection of identity deception
[13]	Contextual long short-term memory architecture to detect bots	Deep Neural Network	Machine Learning	Cresci and Collaborators Dataset	Scrutinize social media Conversation in different contexts
[14]	Location labeling for Spam Account detection	Similarity based Social Media Mining	Machine Learning	Twitter API Dataset	Handle dynamic information
[15]	Detection of malicious profiles	Petri net structure analyzes	GB-Machine Learning	Crude dataset	Optimization of irrelevant features

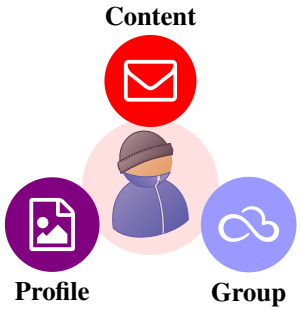

Fig. 2. User Feature over Social Media.

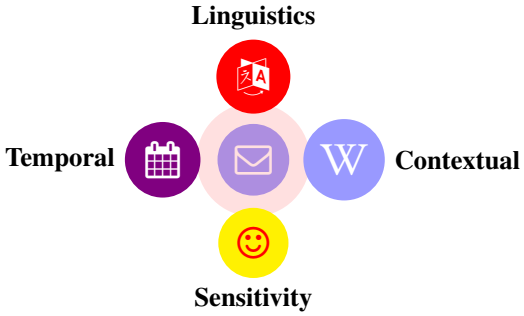

Fig. 3. Content Based Feature over Social Media.

C. Textual Feature

Textual feature of social media user generated content are classified into three class behalf of content, reviewer profile and network dimension as shown in Fig. 4. RIF framework examine and correlate following Review, Reviewer and Network centric feature.

(a) **Reviewer Centric Feature [16], [17], [18]**
- Number of reviews
- Number of Shared/helpful votes
- Time interval between reviews
- Percentage of positive and negative reviews
- Ratio of verified purchase
- Verified stay flag
- Rating deviation
- Review length

(b) **Review Centric Feature[19], [20], [21], [22], [23], [17], [18], [24], [25]**
- Content Similarity Score (Nearly Duplicates)
- Percentage of Pronouns/ Nouns/ Adjective / Verbs
- Lexical Validity
- Lexical Diversity
- Content Diversity
- Syntactical Diversity
- Active and Passive Voice
- Picture and Links
- Emotivenss
- Content Relevancy
- Sentiment Score
- Linguistic inquiry and Word Count
- Product Information Matching
- First Review Flag

(c) **Network Centric Feature (NCF) [26], [27]**
- IP address
- GPS Information
- Timestamp
- Traffic Patterns (IP density)
- Device Information

D. Relationship Identification

After identifying profile and textual feature of end-user as seed profile , relationship identification employed balance theory to extracts hidden relationships of other similar profile with seed profile as implicit link relationship. For instance, consider g(v,r_e) as a social media graph having 11 users nodes and 9 relationship edges, as shown in Fig. 5. Then

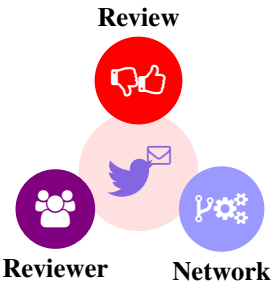

Fig. 4. Textual Feature over Social Media Post.

after applying the balance theory of SMM, two hidden implicit relationships are extracted over graph g(v,r_e), as shown in Fig. 6 by the red line.

After extracting the secret relationship, nodes are hierarchically differentiated according to their implicit status derived through the status theory. After applying the status theory, node colour over the clique are changed. The Degree of the brightness of node color has shown its hidden implicit statuses over the clique, as shown in Fig. 7.

Simultaneously, the graph transmit effects as explicit characteristics extracted through Influence, Homophily, and Confounding correlation theory. Higher status communal node changes the belongingness of its lower status node into their respective community through the Influence theory. Whereas, homophily builds the belongingness of similar characteristics node over the same community. However, any online forum creates an environment to make individuals similar, as confounding.

After extracting implicit information from social media through social theory, NCF generates vertex degree vector and reachability matrix, as shown in equation 6.16 and 6.17.

$$n_d^v = \left\{n_d^1, n_d^2, n_d^3, ...n_d^m\right\} \forall\, m \leq n-1 \tag{4}$$

Where,n_d^v is represent node degree vector and n_d^i is the number of node having degree i in desire clique structure. Whereas, $node_{rm}$ represent node rechability square matrix having n*n dimension and r_{v_i,v_j} is the modular distance between node v_i and v_j

$$node_{rm} = \left[r_{v_i,v_j}\right]_{n*n} \tag{5}$$

After extracting node feature vector and matrix , multiplication of vertex degree vector and node reachability matrix return $A_{i,j}$ as the highest influence node. Simultaneously, the K-means algorithm builds the community of similar nodes with a similarity index of the Jaccard coefficient over the initial point $A_{i,j}$.

IV. Environmental Setup and Result Analysis

The comparative analysis is present interesting and useful facts regarding the state-of-the-art of malicious account classification technique. For performance evaluation of DNN based RIF framework with basic stand-alone classifiers such as Random Forest (RF), Bagging Classifier, J48 Classifier, Random Tree, and Logistic Regression has been carried out over two different interaction and structural anomalies social media data set, namely Crude and Cresci Collaborators (CCDS) data set. Crude dataset [10] has 6824 profile data(Fake+ Genuine),59153788 tweets, 4899493 followers, 16236669 Likes, 67976 listed count 1367 URL Shared. Simultaneously, CCDS [11] has 3474 genuine accounts, 8377522 genuine tweets, 991 fake account, and 1610176 fake tweets.

Performance evaluation of Random Forest (RF) for malicious account classification with and without user feature and social theory is described in Table I.

The RF algorithm acquires 67.09%, 66.98%, 68.12% and 80.21% 78.45%, 81.78% precision with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table II and Fig. 8(a). The RF algorithm's performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The RF algorithm acquires 1.53%, 1.36%,3.09% and 33.02%, 30.10%, 35.62% improvement over the precision with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 8(b).

Whereas, RF algorithm acquires 67.34%, 66.14%, 68.92% and 78.41% 74.24%, 79.12% recall with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table II and Fig. 8(c).The RF algorithm's performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The RF algorithm acquires 3.09%, 1.26%,5.51% and 41.15%, 33.65%, 35.62% improvement over the recall with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 8(d).

Simultaneously, RF algorithm acquires 67.84%, 65.91%, 69.46% and 78.9% 76.14%, 79.98% F1-Score with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table II and Fig. 8(e).The RF algorithm's performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The RF algorithm acquires 4.19%, 1.23%, 6.68% and 38.37%, 33.53%,40.27% improvement over the F1-Score with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 8(f).

However, RF algorithm acquires 92.94%, 92.1%, 93.45% and 95.78%, 94.56%, 96.2% Accuracy with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table II and Fig. 8(g).The RF algorithm's performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The RF algorithm acquires 1.41%, .49%, 1.96% and 5.46%, 4.12%,5.92% improvement over the Accuracy with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 8(h).

The Bagging algorithm acquires 66.59%, 65.19%, 67.82% and 75.22%, 74.61%, 76.15% precision with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table III and Fig. 9(a). The RF algorithm's performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Bagging acquires 3.43%, 1.26%, 5.34% and 42.25%, 41.09%, 44.01% improvement over the precision with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 9(b).

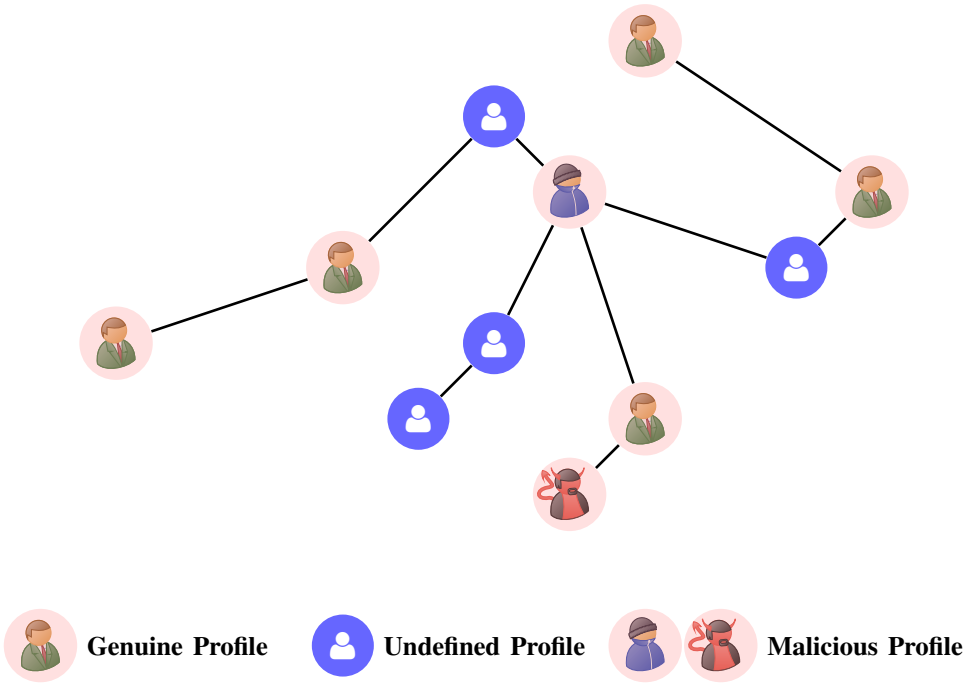

Fig. 5. Structure of user Community.

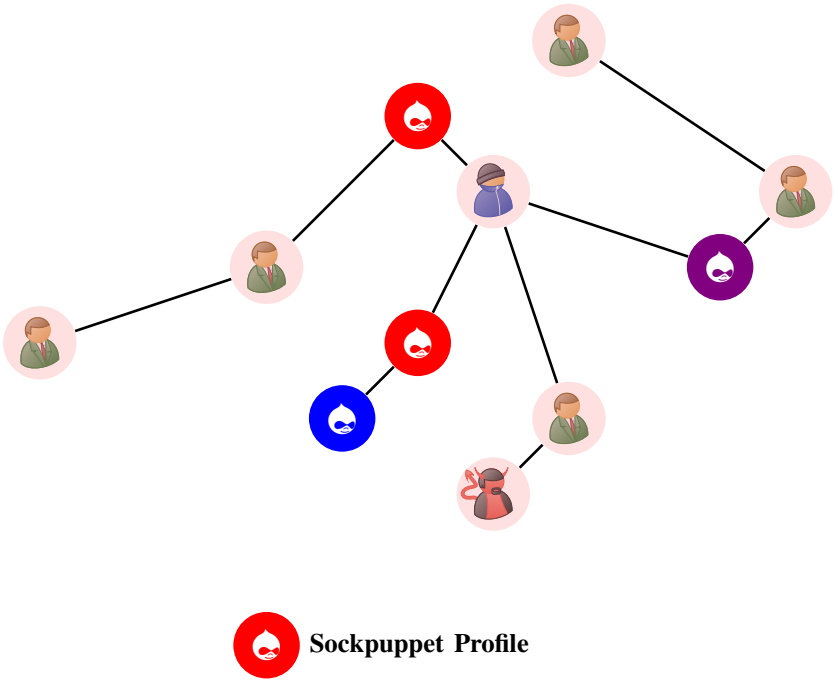

Fig. 6. Identify Status of Profile Via Balance Theory.

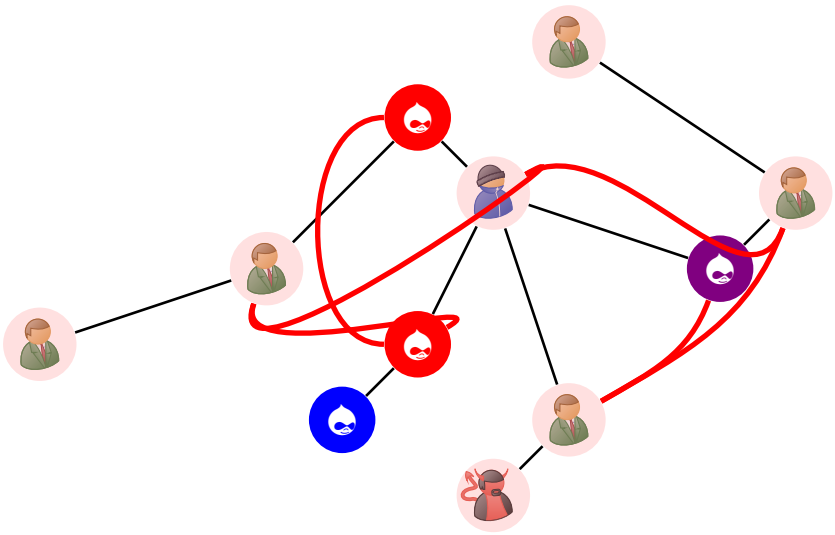

Fig. 7. Implicit Link Structuring via Influence and Homophily.

TABLE II. MALICIOUS ACCOUNT CLASSIFICATION THROUGH RANDOM FOREST

Data Set	Evaluation Parameter	Standalone	User Feature	Social Theory	Fusion
Crude	Precision	66.08	67.09	66.98	68.12
	Recall	65.32	67.34	66.14	68.92
	F1-Score	65.11	67.84	65.91	69.46
	Accuracy	91.65	92.94	92.10	93.45
CCSD	Precision	60.30	80.21	78.45	81.78
	Recall	55.55	78.41	74.24	79.12
	F1-Score	57.02	78.90	76.14	79.98
	Accuracy	90.82	95.78	94.56	96.20

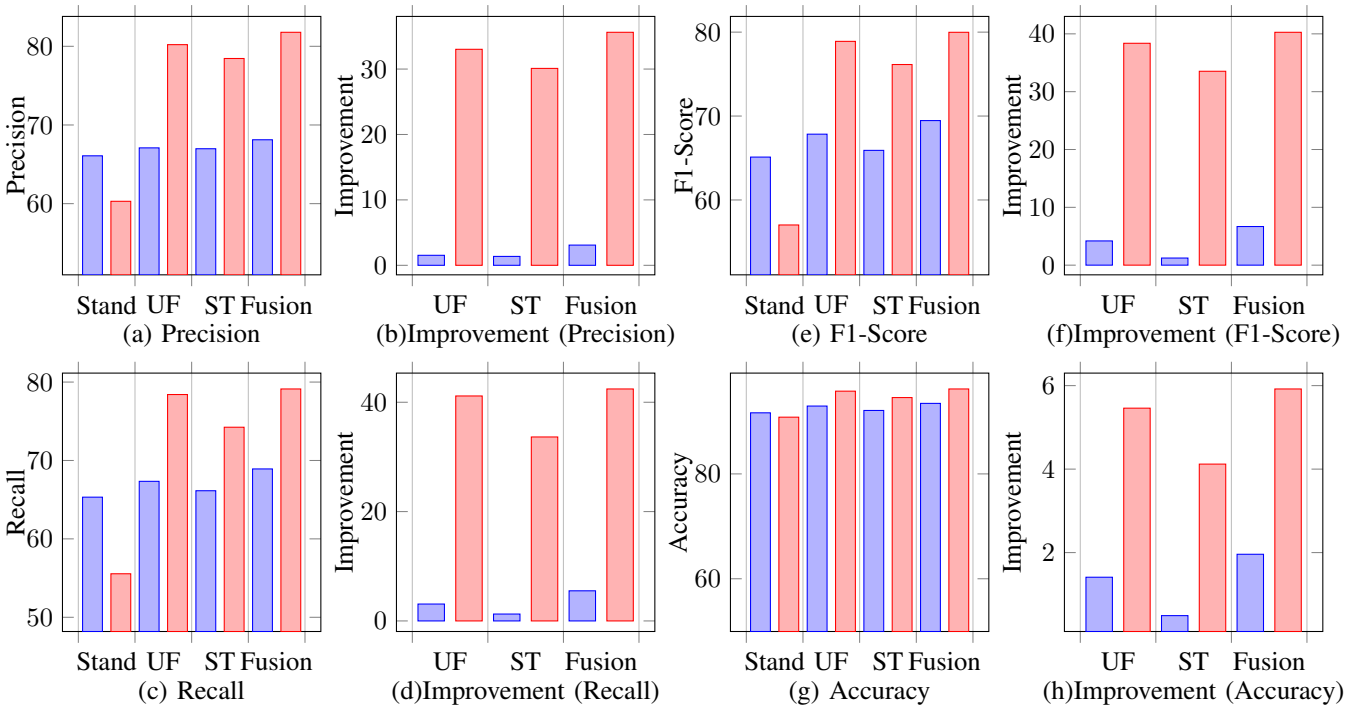

Fig. 8. Performance Evaluation of Malicious Account Classification Through Random Forest.

TABLE III. MALICIOUS ACCOUNT CLASSIFICATION THROUGH BAGGING

Data Set	Evaluation Parameter	Standalone	User Feature	Social Theory	Fusion
Crude	Precision	64.38	66.59	65.19	67.82
	Recall	63.04	66.69	64.85	67.98
	F1-Score	55.36	65.24	64.12	68.72
	Accuracy	90.94	92.74	90.42	93.14
CCSD	Precision	52.88	75.22	74.61	76.15
	Recall	48.84	73.16	70.58	73.89
	F1-Score	49.83	73.65	71.25	75.28
	Accuracy	89.44	95.55	92.18	95.98

Whereas, Bagging acquires 66.69%, 64.85%, 67.98% and 73.16%, 70.58%, 73.89% recall with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table III and Fig. 9(c).The Bagging performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The RF algorithm acquires 5.79%, 2.87%, 7.84% and 49.80%, 44.51%, 51.29% improvement over the recall with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 9(d).

Simultaneously, Bagging acquires 65.24%, 64.12%, 68.72% and 73.65%, 71.25%, 75.28% F1-Score with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table III and Fig. 9(e). The Bagging performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Bagging acquires 17.85%, 15.82%, 24.13% and 47.80%, 42.99%, 51.07% improvement over the F1-Score with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 9(f).

However, Bagging acquires 92.74%, 91.42%, 93.14% and 95.55%, 92.18%, 95.98% Accuracy with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in table 3 and figure 9(g).The Bagging performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Bagging acquires 1.98%, .53%, 2.42% and 6.83%, 3.06%, 7.31% improvement over the Accuracy with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 9(h).

The J48 algorithm acquires 63.52%, 62.78%, 64.15% and 70.6%, 64.52%, 72.82% precision with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table IV and Fig. 10(a). The J48 performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Bagging acquires 4.75%, 3.53%, 5.79% and 52.19%, 39.08%, 56.97% improvement over the precision with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 10(b).

Whereas,J48 acquires 62.02%, 59.69%, 62.84% and 69.23%, 65.82%, 71.56% recall with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in table 4 and figure 10(c).The J48 performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The J48 acquires 5.14%, 1.19%, 6.53% and 64.52%, 56.42%, 70.06% improvement over the recall with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 10(d).

Simultaneously, J48 acquires 61.73%, 58.36%, 62.84% and 69.19%, 66.58%, 70.69% F1-Score with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table IV and Fig. 10(e).The J48 performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The J48 acquires 19.68%, 13.14%, 21.13% and 58.73%, 52.74%, 62.17% improvement over the F1-Score with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 8(f).

However, J48 acquires 91.87%, 91.25%, 93.14% and 93.95%, 92.56%, 94.64% Accuracy with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table IV and Fig. 8(g).The J48 performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The J48 acquires 1.77%, 1.09%, 3.18% and 6.50%, 4.92%, 7.28% improvement over the Accuracy with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 10(h).

The Random tree algorithm acquires 64.67%, 64.08%, 65.84% and 72.76%, 70.28%, 73.58% precision with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table V and Fig. 11(a). The J48 performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Random tree acquires 1.57%, 0.64%, 0.41% and 41.04%, 36.23%, 42.62% improvement over the precision with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 11(b).

Whereas, Random tree acquires 65.13%, 64.56%, 66.18% and 50.86%, 49.86%, 52.69% recall with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table V and Fig. 11(c). The Random tree performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Random tree acquires 4.59%, 3.68%, 6.28% and 7.48%, 5.37%, 11.35% improvement over the recall with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 11(d).

Simultaneously, Random tree acquires 63.78%, 62.86%, 64.27% and 54.15%, 52.85%, 55.28% F1-Score with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table V and Fig.

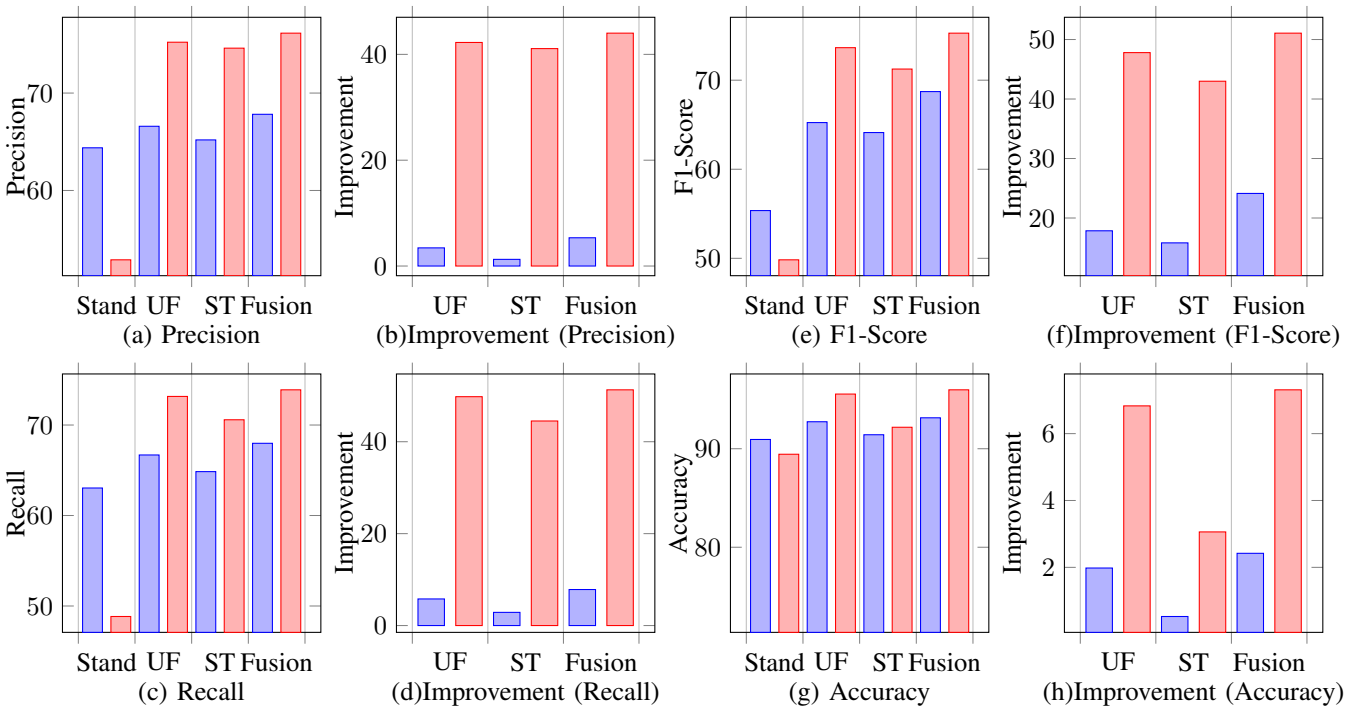

Fig. 9. Performance Evaluation of Malicious Account Classification Through Bagging.

TABLE IV. MALICIOUS ACCOUNT CLASSIFICATION THROUGH J48

Data Set	Evaluation Parameter	Standalone	User Feature	Social Theory	Fusion
Crude	Precision	60.64	63.52	62.78	64.15
	Recall	58.99	62.02	59.69	62.84
	F1-Score	51.58	61.73	58.36	62.48
	Accuracy	90.27	91.87	88.25	92.14
CCSD	Precision	46.39	70.6	64.52	72.82
	Recall	42.08	69.23	65.82	71.56
	F1-Score	43.59	69.19	66.58	70.69
	Accuracy	88.22	93.95	92.56	94.64

TABLE V. MALICIOUS ACCOUNT CLASSIFICATION THROUGH RANDOM TREE

Data Set	Evaluation Parameter	Standalone	User Feature	Social Theory	Fusion
Crude	Precision	63.67	64.67	64.08	65.84
	Recall	62.27	65.13	64.56	66.18
	F1-Score	54.61	63.78	62.86	64.27
	Accuracy	90.65	91.54	91.05	92.47
CCSD	Precision	51.59	72.76	70.28	73.58
	Recall	47.32	50.86	49.86	52.69
	F1-Score	48.36	54.15	52.85	55.28
	Accuracy	89.06	94.84	92.42	95.58

TABLE VI. MALICIOUS ACCOUNT CLASSIFICATION THROUGH LOGISTIC REGRESSION

Data Set	Evaluation Parameter	Standalone	User Feature	Social Theory	Fusion
Crude	Precision	52.97	60.92	59.85	62.56
	Recall	53.42	61.54	60.21	63.08
	F1-Score	53.13	61.11	60.48	62.46
	Accuracy	88.31	90.17	88.23	91.47
CCSD	Precision	57.21	63.18	61.56	64.58
	Recall	46.03	92.54	89.95	94.12
	F1-Score	56.44	92.72	90.56	93.86
	Accuracy	76.75	84.67	83.41	85.98

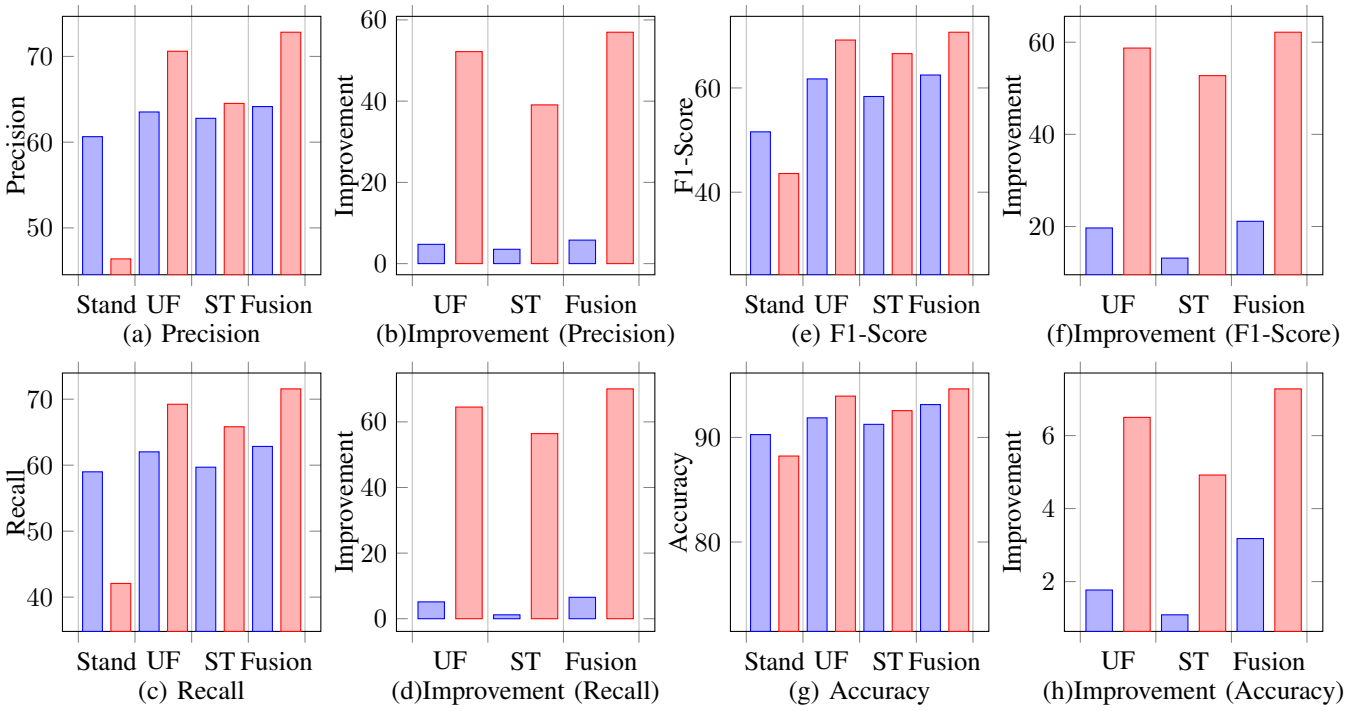

Fig. 10. Performance Evaluation of Malicious Account Classification Through J48.

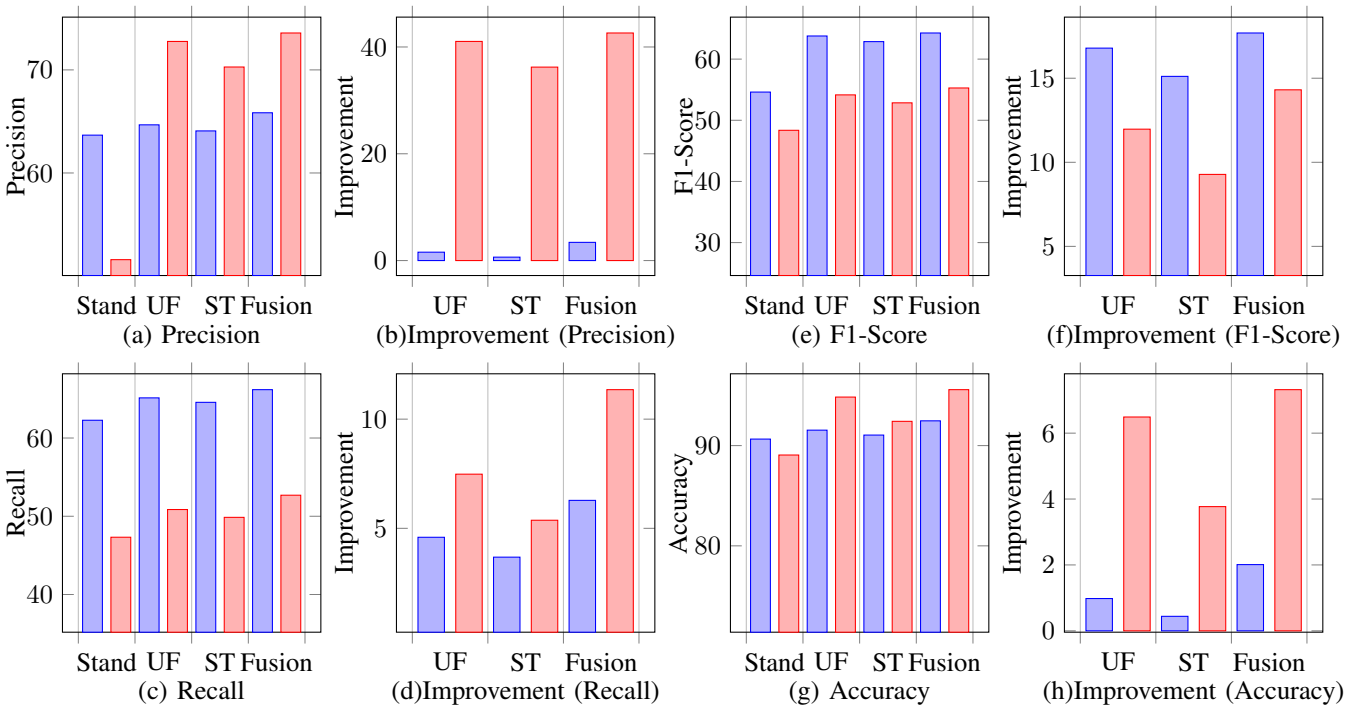

Fig. 11. Performance Evaluation of Malicious Account Classification Through Random Tree.

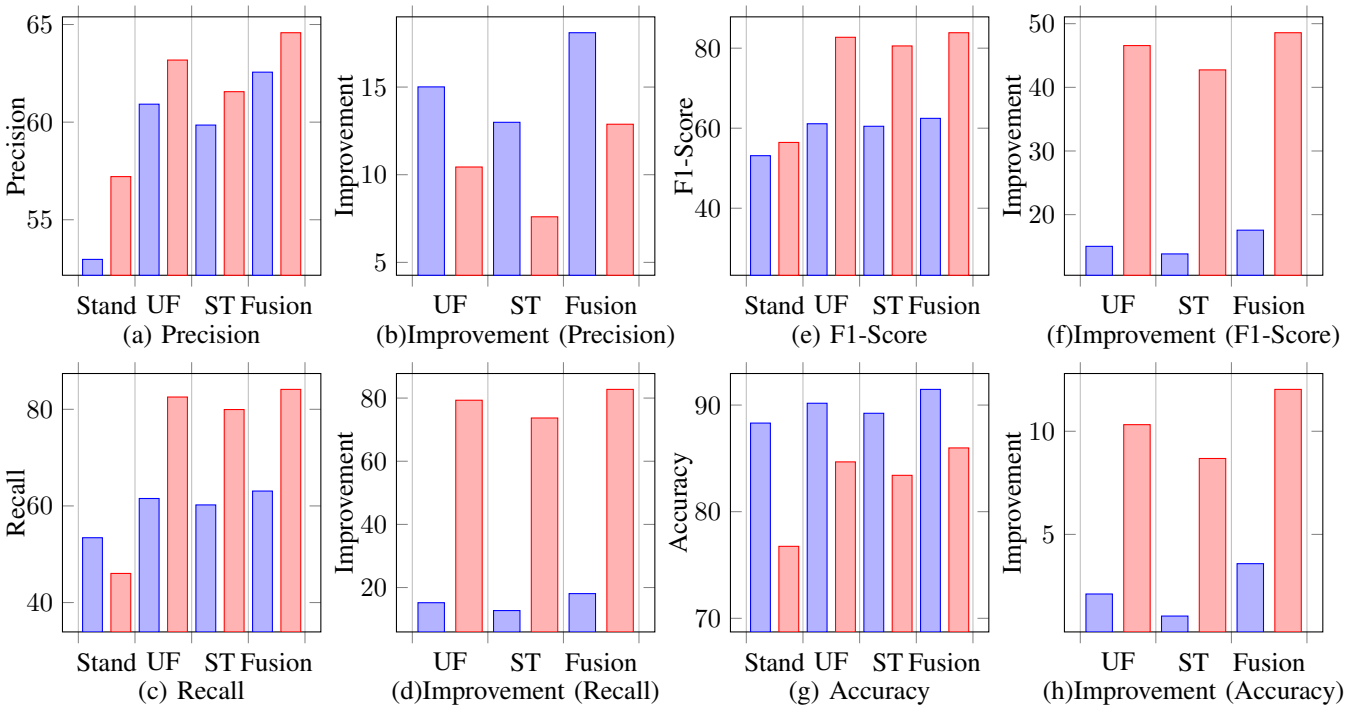

Fig. 12. Performance Evaluation of Malicious Account Classification Through Logistic Regression.

TABLE VII. MALICIOUS ACCOUNT CLASSIFICATION THROUGH PROPOSED WORK

Data Set	Evaluation Parameter	Random Forest	Bagging	J48	Random Tree	Logistic Regression	Proposed Work
Crude	Precision	68.12	67.82	64.15	65.84	62.56	75.89
	Recall	68.92	67.98	62.84	66.18	63.08	76.42
	F1-Score	69.46	68.72	62.48	64.27	62.46	77.52
	Accuracy	93.45	93.14	93.14	92.47	91.47	95.89
CCSD	Precision	81.78	76.15	72.82	73.58	64.58	82.49
	Recall	79.12	73.89	71.56	52.69	84.12	87.76
	F1-Score	79.98	75.28	70.69	55.28	83.86	86.19
	Accuracy	96.2	95.98	94.64	95.58	85.98	98.54

11(e). The Random tree performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Random tree acquires 16.79%, 15.11%,17.69% and 11.97%, 9.28%, 14.31% improvement over the F1-Score with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 11(f).

However, Random tree acquires 72.76%, 70.28%, 73.78% and 94.84%, 92.42%, 95.58% Accuracy with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table V and Fig. 11(g). The Random tree performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Random tree acquires .98%, .44%, 2.01% and 6.49%, 3.77%, 7.32% improvement over the Accuracy with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 11(h).

The Logistic algorithm acquires 60.92%, 59.85%, 62.56% and 63.18%, 61.56%, 64.58% precision with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table VI and Fig. 12(a). The Logistic performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Logistic acquires 15.01%, 12.99%, 18.10% and 10.44%, 7.60%, 12.88% improvement over the precision with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 12(b).

Whereas, Logistic regression algorithm acquires 61.54%, 60.21%, 63.08% and 82.54%, 79.95%, 84.12% recall with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table VI and Fig. 12(c). The Logistic performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The J48 acquires 15.20%, 12.71%, 18.08% and 79.32%, 73.69%, 82.75% improvement over the recall with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 12(d).

Simultaneously, Logistic regression algorithm acquires 61.11%, 60.48%, 62.46% and 82.72% 80.56%, 83.86% F1-Score with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table VI and Fig. 12(e). The Logistic performance is significantly boosted up after rectifying network information by user fea-

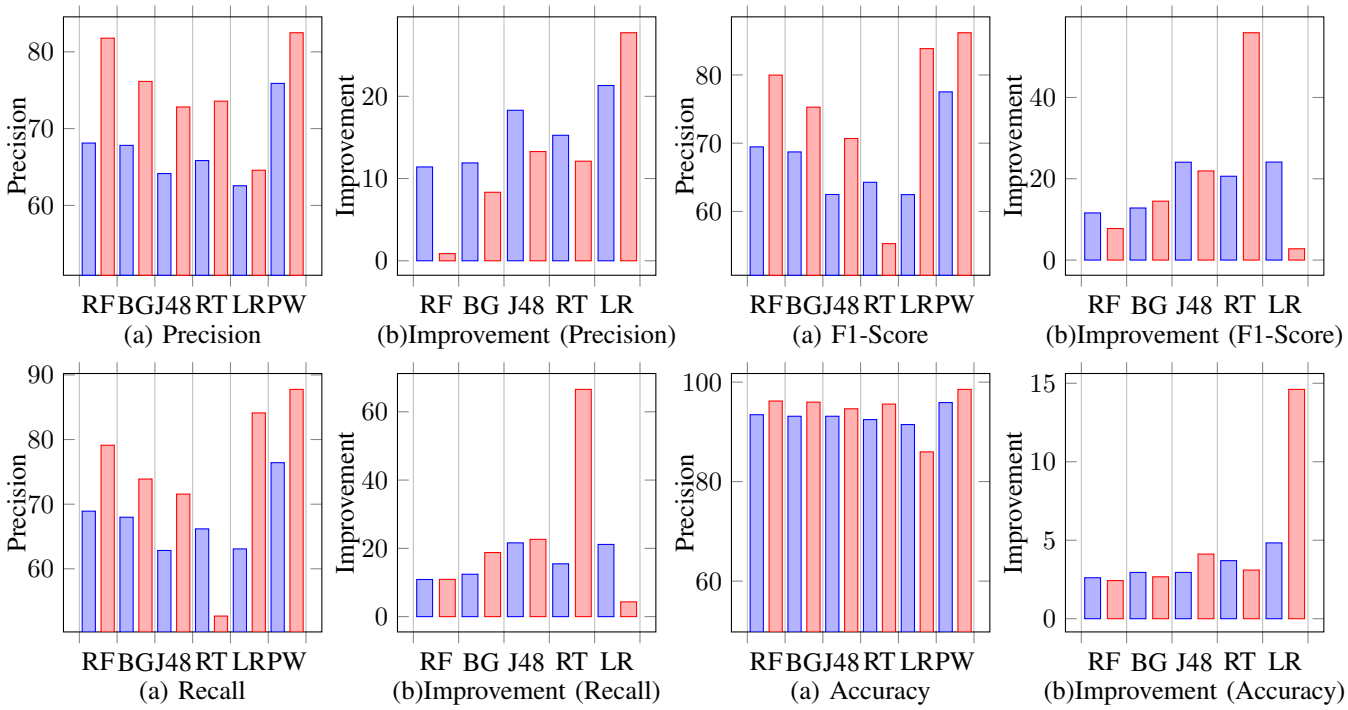

Fig. 13. Performance Evaluation of Malicious Account Classification Through Proposed Work.

ture, social theory, and fusion of both. The Logistic acquires 15.02%, 13.83%, 17.56% and 46.56%, 42.74%, 48.58% improvement over the F1-Score with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 12(f).

However, Logistic regression algorithm acquires 90.17%, 89.23%, 91.47% and 84.67% ,83.41%, 85.98% Accuracy with user feature, social theory, and fusion of both respectively over Crude and CCDS data set as shown in Table VI and Fig. 12(g).The Logistic performance is significantly boosted up after rectifying network information by user feature, social theory, and fusion of both. The Logistic acquires 2.11%, 1.04%, 3.58% and 10.32%, 8.68%, 12.03% improvement over the Accuracy with user feature, social theory, and fusion of both over Crude and CCDS data set, as shown in Fig. 12(h).

Whereas Proposed work acquire 75.89% , 82.49% precision, 76.42% , 87.76% recall, 77.52%, 86.19% F1-Score, and 95.89%, 98.54% Accuracy respectively over Crude and CCDS data set as shown in Table VII and Fig. 13. However its gain 11.41% - 21.31% and 0.87% - 27.73% improvement in precision, 10.88% - 21.61% and 10.92% - 66.56% improvement in recall, 11.60% - 24.11% and 2.78% - 55.92% improvement in F1-Score, and 2.61% - 4.83% and 2.43% - 14.61% improvement in Accuracy over Crude and CCDS data set, as shown in Fig. 13.

V. Conclusion

Online Social Network (OSN) is a network hub where people with similar interests or real world relationships interact. As the popularity of OSN is increasing, the security and privacy issues related to it are also rising. Fake and Clone profiles are creating dangerous security problems to social network users. Cloning of user profiles is one serious threat, where already existing userâ€™s details are stolen to create duplicate profiles and then it is misused for damaging the identity of original profile owner. They can even launch threats like phishing, stalking, spamming, etc. Fake profile is the creation of profile in the name of a person or a company which does not really exist in social media, to carry out malicious activities. In this paper graphical, linguistics and social theory based relationship identification (RIF) framework is developed to identify malicious end-user over social media.This framework amalgamates linguistics, temporal and contextual ethics of user-generated content with profile and graphical information. The RIF framework extract feature vector to delineate user behaviors and similarity index over social media. Classifying identical profile concerning to similar user via Jaccard coefficient over linguistics pattern of tweets and provide linguistics, temporal and contextual meaning to develop a mathematical model for classifying identical profile as sockpuppet over social media. RIF framework achieve maximum 82.49% precision, 87.76% recall, 86.19% F1-Score, and 98.54% Accuracy. However its gain maximum 27.73% improvement in precision, 66.56% improvement in recall, 55.92% improvement in F1-Score, and 14.61% improvement in Accuracy.

References

[1] N. K. Singh, D. S. Tomar, and A. K. Sangaiah, “Sentiment analysis: a review and comparative analysis over social media,” *Journal of Ambient Intelligence and Humanized Computing*, May 2018.

[2] G. D. Domenico, J. Sit, A. Ishizaka, and D. Nunan, “Fake news, social media and marketing: A systematic review,” *Journal of Business Research*, vol. 124, pp. 329–341, 2021.

[3] N. Singh and D. Tomar, “Comprehensive analysis of scope of negation for sentiment analysis over social media,” *Journal of Theoretical and Applied Information Technology*, vol. 97, pp. 1704–1719, 03 2019.

[4] N. K. Singh and D. S. Tomar, "Feature fusion for negation scope detection in sentiment analysis: Comprehensive analysis over social media," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, 2019.

[5] S. Cresci, M. Petrocchi, A. Spognardi, and S. Tognazzi, "On the capability of evolved spambots to evade detection via genetic engineering," *Online Social Networks and Media*, vol. 9, pp. 1–16, 2019.

[6] J. Pastor-Galindo, M. Zago, P. Nespoli, S. López Bernal, A. Huertas Celdrán, M. Gil Pérez, J. A. Ruipérez-Valiente, G. Martínez Pérez, and F. Gómez Mármol, "Twitter social bots: The 2019 spanish general election data," *Data in Brief*, vol. 32, p. 106047, 2020.

[7] J. Rodríguez-Ruiz, J. I. Mata-Sánchez, R. Monroy, O. Loyola-González, and A. López-Cuevas, "A one-class classification approach for bot detection on twitter," *Computers and Security*, vol. 91, p. 101715, 2020.

[8] X. Zhou, X. Liang, H. Zhang, and Y. Ma, "Cross-platform identification of anonymous identical users in multiple social media networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, pp. 411–424, Feb 2016.

[9] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towards detecting compromised accounts on social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, pp. 447–460, July 2017.

[10] M. Al-Qurishi, M. Alrubaian, S. M. M. Rahman, A. Alamri, and M. M. Hassan, "A prediction system of sybil attack in social network using deep-regression model," *Future Generation Computer Systems*, vol. 87, pp. 743 – 753, 2018.

[11] M. Al-Qurishi, S. M. M. Rahman, M. S. Hossain, A. Almogren, M. Alrubaian, A. Alamri, M. Al-Rakhami, and B. Gupta, "An efficient key agreement protocol for sybil-precaution in online social networks," *Future Generation Computer Systems*, vol. 84, pp. 139 – 148, 2018.

[12] E. Van Der Walt and J. Eloff, "Using machine learning to detect fake identities: Bots vs humans," *IEEE Access*, vol. 6, pp. 6540–6549, 2018.

[13] S. Kudugunta and E. Ferrara, "Deep neural networks for bot detection," *Information Sciences*, vol. 467, pp. 312 – 322, 2018.

[14] M. Celik and A. S. Dokuz, "Discovering socially similar users in social media datasets based on their socially important locations," *Information Processing and Management*, vol. 54, no. 6, pp. 1154 – 1168, 2018.

[15] S. R. Sahoo and B. Gupta, "Hybrid approach for detection of malicious profiles in twitter," *Computers and Electrical Engineering*, vol. 76, pp. 65 – 81, 2019.

[16] Y. Liu and B. Pang, "A unified framework for detecting author spamicity by modeling review deviation," *Expert Systems with Applications*, vol. 112, pp. 148 – 155, 2018.

[17] R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Information Processing and Management*, vol. 56, no. 4, pp. 1234 – 1244, 2019.

[18] Y. Liu, B. Pang, and X. Wang, "Opinion spam detection by incorporating multimodal embedded representation into a probabilistic review graph," *Neurocomputing*, vol. 366, pp. 276 – 283, 2019.

[19] J. K. Rout, A. K. Dash, and N. K. Ray, "A framework for fake review detection: Issues and challenges," in *2018 International Conference on Information Technology (ICIT)*, pp. 7–10, 2018.

[20] L. Li, B. Qin, W. Ren, and T. Liu, "Document representation and feature combination for deceptive spam review detection," *Neurocomputing*, vol. 254, pp. 33 – 41, 2017. Recent Advances in Semantic Computing and Personalization.

[21] W. Liu, J. He, S. Han, F. Cai, Z. Yang, and N. Zhu, "A method for the detection of fake reviews based on temporal features of reviews and comments," *IEEE Engineering Management Review*, vol. 47, pp. 67–79, Fourthquarter 2019.

[22] M. Petrescu, K. Oâ€™Leary, D. Goldring, and S. B. Mrad, "Incentivized reviews: Promising the moon for a few stars," *Journal of Retailing and Consumer Services*, vol. 41, pp. 288 – 295, 2018.

[23] E. F. Cardoso, R. M. Silva, and T. A. Almeida, "Towards automatic filtering of fake reviews," *Neurocomputing*, vol. 309, pp. 106 – 116, 2018.

[24] L. You, Q. Peng, Z. Xiong, D. He, M. Qiu, and X. Zhang, "Integrating aspect analysis and local outlier factor for intelligent review spam detection," *Future Generation Computer Systems*, vol. 102, pp. 163 – 172, 2020.

[25] S. Noekhah, N. binti Salim, and N. H. Zakaria, "Opinion spam detection: Using multi-iterative graph-based model," *Information Processing & Management*, vol. 57, no. 1, p. 102140, 2020.

[26] S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi, "Netspam: A network-based spam detection framework for reviews in online social media," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 1585–1595, July 2017.

[27] J. K. Rout, A. Dalmia, K. R. Choo, S. Bakshi, and S. K. Jena, "Revisiting semi-supervised learning for online deceptive review detection," *IEEE Access*, vol. 5, pp. 1319–1327, 2017.