Feature-based and Graph-based Approaches to Spam call detection problem

Dr. Ngoc C Le and comrades

School of Applied Mathematics and Informatics
Hanoi University of Science and Technology
1 Dai Co Viet, Hanoi

***Abstract*—Spam over Internet Telephony (SPIT) is becoming increasingly severe and has attracted significant attention from telecom vendors due to its massive harm to finance and users' experience. There has been an urgent call for mechanisms to prevent SPIT attackers from fraud and unpleasant activities. After a thorough studying of previous researchers' works, there are two impressive approaches namely feature-based and graph-based that inspire us to develop a more competent mechanisms. With feature-based, we shall establish a set of 11 features which there are momentous distinguish in distribution between SPITers and legitimate users. On the other hand, with graph-based, on the foundations of CallRank algorithm in [1], we conduct a modification in the way how we weight the edges linking two users. Implementing on the dataset crawled from iCaller - a spam call blocker application, we shall propose acceptable results as a justification to our methodologies. Some conclusions are supposed to be given in the last section.**

***Index Terms*—Spam over Internet Telephony (SPIT) · Voice over Internet Protocol (VoIP) · Feature-based · Graph-based · Supervised learning methods · ROC**

I. Introduction

In the years to come, along with the rapid development of Internet technology for voice communication, VoIP telephony has experienced a significant rise in the number of subcribers. However, this is interupted by Spam over Internet Telephony (SPIT), a form of abuse and fraud which is proportionally increasing worldwide. SPIT is one of the most serious threats in VoIP due to severe financial losses and mental irritations caused by it to telephony users. Consider session initiation protocol (SIP), the most widely adopted technology in signaling protocol, as an illustration, it is easy for attackers to initiate spam calls to thousands of users' IP addresses with the help of low cost and convenience in using SIP.

In light of the necessity for preventing spammers from abusing, on the foundation of the experience during SPAM defense, a great many of approaches to solve the problem have been proposed. They consist of noticable factors namely reputation-based [2] , call-frequency-based [3], dynamic backlisting, fingerprinting [4], clarifying suspicious calls by capchas [5] and the use of some machine leanrning algorithms like support vector machine [6], [7] and semi-supervised clustering [8]. After a thorough revision, we highly appreciated two effective approaches namely feature-based and graph-based to discover potential SPIT. By some modifications of previous works and the variety of machine learning algorithms, we shall design a SPIT detection system through feature-based and graph-based approaches.

Our critical contributions in this paper can be listed as follows

- We take into consideration 11 distinguished features that are carefully revised and proven to be worth-noting.
- We implement our dataset with 8 different machine learning algorithms, reveal fundamental factors for evaluation in order to figure out which algorithm is regarded as state-of-the-art.
- We give a deeper insight into graph-based approach by developing CallRank [1] with some modifications and provide evidence that our methodology is of great effectiveness.

The remaining of our paper is organized as follows. Section 2 is used for summarizing related works with regard to feature-based approach and graph-based approach. In section 3, we shall give details about data used for training, methodology and evaluation factors. This is followed in section 4 by our implementation and some assessments on the results in comparison to previous researchers as well as some drawbacks existing. Some conclusions are given in section 5.

II. Related Works

A. SPIT detection mechanisms

SPIT turns into an extreme danger for VoIP clients as a result of the downturn in voice call costs, contrasting with

current foundations, and the absence of a worldwide lawful and administrative system. It influences the private existence of the client and his correspondences and turns into a wellspring of commotion for them. That threat works as an inspiration for researchers to come up with many mechanisms to dectect SPIT. A thorough review enables us to sum up with two outstanding mechanisms namely Behaviour-based and CallRank.

1) Behaviour-based: In [9], Kusumoto et al. took into account several call patterns including unsuccessful call-making rate, average call duration, user relationships. After a process of computing those different patterns, Naive Bayes is put into activation to identify spammers. The collection of previously used criteria can be listed as follows.

- Rate of answered calls [10], [11]
- Rate of dismissed calls [7], [9], [10], [11]
- Number of call-making attempts [7], [8], [10], [11]
- Call duration [7], [8], [1], [9], [10], [11]

It is worth-noticing that call duration pattern is taken advantage by most researchers. Combining with common behaviours from SPIT attackers, we develop a set of criteria which has the details given in subsection 3.2.

With a view to determining whether a phone number is potential SPIT or not, we conducted our experiment with 8 supervised learning algorithms reviewed in subsection 2.2.

2) CallRank: One well-known way to approach the problem is that the authors focused on reputation-based on the principle of computing reputation score. The concept of using reputation score also lies in the works [12], [13], [14] who made effort to establish a trust chain between the caller and the callee. We highly emphasize CallRank [1] because the authors approached the problem quite effectively by graph-based method. The basic idea in [1] is to build social network relationships and global reputations for users. Based on that, we compute impact score of a client in the network. We expect that the legitimate users may have a higher impact score than SPIT attackers, therefore it is possible to detect spammers.

B. Machine Learning Algorithms

When it comes to perform difficult tasks in artificial intelligence, in comparison to traditional algorithms, machine learning algorithms are of greater effectiveness in aspect of accuracy, time performance. Our work puts priority on supervised learning. In supervised learning, on the grounds that the classes are well-predetermined, a machine evolves during a learning process and classifies according to a classification model. Let us describe some classification algorithms used in our framework. Our chosen algorithms are Logistic regression, k-nearest neighborhood, Naive Bayes, Decision tree, Random forest, Bagging, AdaBoost and XGBoost.

III. Methodology

A. Dataset - iCaller App

We crawl data from iCaller in the period of time from the year 2018 to 2021. iCaller is an application being developed by Grooo International Jsc. The purpose lying behind the development of this app is to prevent calls that are related to fraud, loan, deabt, advertising, real_estate and other irritational activities. Currently, it is working based on direct reports from users all over the world. To be more specific, supposing that user A has just received a bothering call from user B who advertised about their insurance benefits. After that call, A considers that B is a spammer, then A marks B's phone number as an advertising spammer in the application user interface easily. Since then, whenever B initiates a call to A, the app would push a warning about the identity of B on A's phone and ask if A wants to answer that phone call. In order to eliminate all limitations, we have an intention that is to apply machine learning algorithms in the system for automatical checking. Furthermore, our methodology is capable of increasing accuracy in classification because in reality, it is quite likely that the users mistakenly mark a caller as spammer. We also emphasize in clarifying users' identities that would help people who receive calls to decide whether to answer the call or not. What's more, iCaller does not have access to the telecommunications infrastructure of telecom vendors, which makes the previous SPIT methods helpless in practical use. However, we have a solution that utilizes a server to cross-collect among phone numbers in the network, that helps us collect and store the data needed for implementing the methods we will propose in this paper.

As I mentioned above, iCaller enables users to mark label for a phone number based on 6 types of spammers reported namely report_advertising, report_loan, report_debt, report_cheat, report_real_estate, report_other and in the case of confirming a legitimate user, report_not_spam is put at use (see figure 2). Let us reveal our dataset's description that is extracted from iCaller's database. For each call in respective to each record, we collect data about that call, type of users which are member, spammer or unlabelled. iCaller does not have access to the telecommunications infrastructure of telecom vendors, which makes the previous SPIT methods helpless in practical use. However, we have a solution that utilizes a server

	member_phone	phone	type	time	in_contact	duration
299058	8.456921e+10	84939619767	1	11/28/2020 6:47	0	49
689753	8.433392e+10	84989230242	3	1/1/2021 9:45	1	36
476633	8.489882e+10	8.429E+11	1	12/5/2020 11:00	0	8
650656	8.498912e+10	84898424194	1	1/3/2021 11:57	1	56
512472	8.498333e+10	84981596541	3	12/13/2020 9:48	1	0
450309	8.497851e+10	84963059108	1	12/18/2020 14:31	1	15
595627	8.492794e+10	84922627759	1	12/26/2020 17:51	1	28
288596	8.498764e+10	84982830201	2	11/30/2020 13:43	1	14
449939	8.434348e+10	84963087963	1	11/17/2020 5:40	1	7
805678	8.493353e+10	84939289988	2	1/2/2021 9:03	0	51

Fig. 1: Sample of dataset

	phone	from_contact	report_advertising	report_loan	report_debt	report_cheat	report_real_estate	report_other
36552	84975503535	1	0	0	0	0	4	1
93735	84900	1	1	1	0	3	5	2
93757	84965303997	1	0	0	0	0	0	7
93776	84394955554	0	0	2	2	0	1	2
93849	8418001090	1	1	0	1	0	0	1
...	...	...	...	...	...	...	...	...
914974	84931141036	0	3	0	0	0	0	0
924098	84877283378	0	1	0	2	0	0	0
928861	84932710282	1	0	0	2	1	0	0
928908	84932779367	1	0	2	2	2	2	0
1039300	84899909039	1	4	0	0	0	0	0

Fig. 2: Sample of dataset

to cross-collect among phone numbers in the network, that helps us collect and store the data needed for implementing the methods we will propose in this paper. Recall that A calls B. Such information can be explained as shown in table I.

B. Feature-based

1) Understanding features: We shall explain 5 criteria to detect SPIT researched by three biggest telecommunication vendors in Vietnam

- Frequency call: the number of calls initiated from a leased phone in a period of time. For example, minimally 200 outgoing calls per day from 8am to 6pm.
- The rate of calls which have short duration: For example, more than 80% of calls are less than 25 seconds in duration.
- The rate of calls with short period of time between calls: It is the rate of calls with short time interval between two consecutive calls on the total number of calls. For example, more than 50% of calls have less than 20 seconds between calls.
- Rate of calls to unrelated subscribers: The rate of calls to unrelated numbers (never called before, not in contact list) over the total number of outgoing calls. Example: 90% of numbers called are different, without repeating.
- Behavioral characteristics: The phone number is mainly used for outgoing call, not receiving and sending SMS.

Column	Value	Meaning
phone		phone number
from_contact	0	that phone number is new user
	1	that phone number is taken from a member's contact list
report_loan	0	not Loan Spam
	1	Loan Spam reported
report_advertising	0	not Advertising Spam
	1	Advertising Spam reported
report_debt	0	not Debt Spam
	1	Debt Spam reported
report_cheat	0	not Cheating Spam
	1	Cheating Spam reported
report_real_estate	0	not Real_estate Spam
	1	Real_estate Spam reported
report_other	0	no other type of spam
	1	other type of spam reported
time		when the call happened
in_contact	0	B's number is in A's contact list
	1	B's number is not in A's contact list
duration		duration of the call
type	1	successful received call of member
	2	successful initiated call of member
	3	unsuccessful received call of member
	4	unsuccessful initiated call of member

TABLE I: Extracted information about call and users

However, for specific areas, there needs to have some adjustments to suit the characteristics of that area. In combination with such criteria used by former researchers, we develop a set of criteria that are proven to be helpful after a conscientious review. Here, we shall show graphs to give more knowledge about our chosen features.

In comparison to other researches, our distribution of time by day experiences no outstanding difference between Spammers and legitimate users (see figure 3). As a result, we tried

to figure out another distribution of time and come up with the feature related to number of calls per hour.

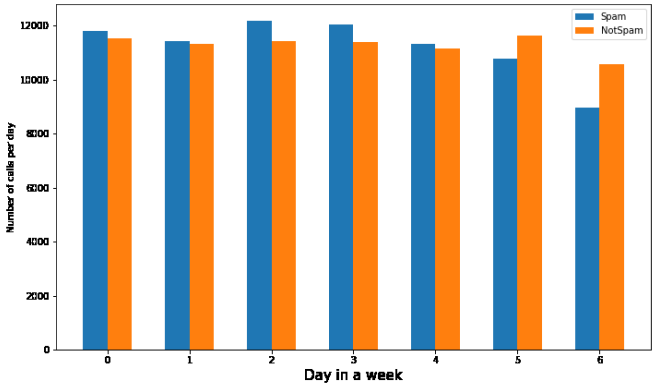
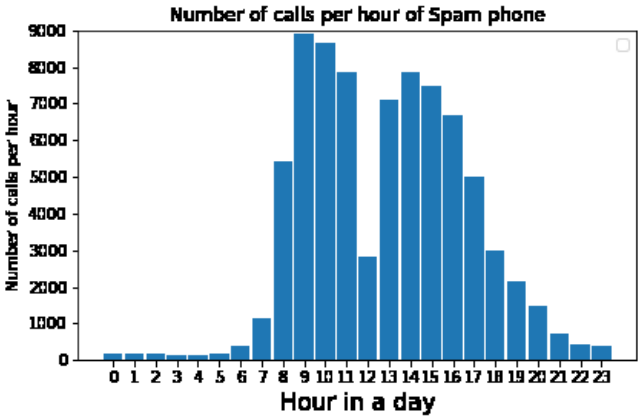
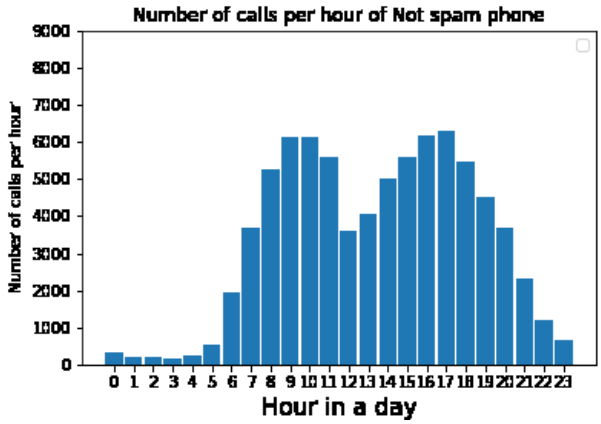


Fig. 3: Number of calls per day between Spam and notSpam



(a) Number of calls per hour of Spammers



(b) Number of calls per hour of legitimate users

Fig. 4: in_hour

Figure 4 reveals the noticable difference in distribution of "in_hour" feature that presents a higher number of calls made by spammers during work hours compared to members' one. In Vietnam, most companies have a schedule that is work hours start from around 7 a.m to 5 p.m. However, it can be observed that there is a significant discrepancy in the number of calls between spammers and members during 5 p.m - 6 p.m The reason for this is probably overtime policy of certain enterprises. Therefore, we decide to choose the range of "in_hour" feature from 7 a.m to 6 p.m.

We strongly believe that most spammers use their own phone numbers for work, in addition, the numbers that are stored the members' directory may have a relationship with the member and therefore are unlikely to be spammers , so we consider by assigning the binary variable [0,1], where 0 means that this number is not in the member's directory, 1 vice versa. However, we noticed that there were some SPIT numbers existing in the members' directory, we surmised that there was a case where the member did the action of adding that contact for the purpose of remembering spammers, or it is the relationship between seller - buyer. To make the observation more meaningful, we averaged over the total number of calls. We also observed that the rate of phone numbers existing in directory of legitimate members is markedly higher than those of spammers (see figure), which is a good criteria in the classification model.

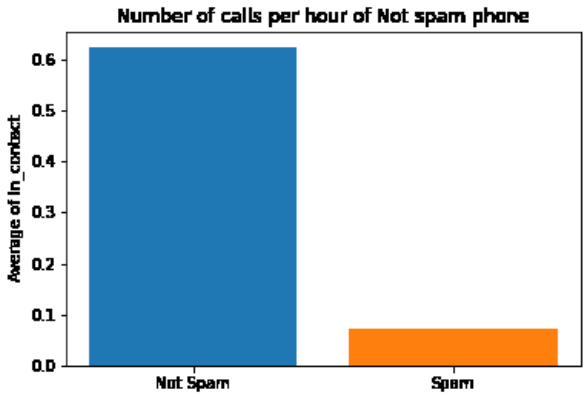


Fig. 5: Number of phones in directory of members and spammers

The set consists of 11 features is sampled in figure 6 and exhibited in table II.

call_to	call_in	call_to_miss	call_in_miss	duration_call_to	duration_call_in
2.708050	3.688879	2.708050	3.332205	6.364751	7.257708
0.693147	0.000000	0.000000	0.000000	2.846133	0.000000
6.819970	4.982439	7.075262	4.447482	10.374554	8.600068
3.367296	3.091042	3.135494	3.044522	6.695799	6.364751
3.238244	1.441367	2.852481	1.945910	7.252362	6.192523
3.688879	3.610918	3.401197	2.564949	7.471363	6.916715
0.000000	0.000000	0.693147	0.000000	0.000000	0.000000
1.032810	0.000000	1.742602	0.000000	0.580658	0.000000
0.000000	2.197225	0.693147	3.761200	0.000000	3.178054
3.076744	0.000000	2.629382	0.672943	7.912739	0.000000

(a)

avg_duration_call_to	avg_duration_call_in	avg_in_contact	in_hour	avg_success
3.077967	3.112840	0.512951	4.343805	0.376686
2.846039	0.000000	0.000000	0.693147	0.693147
2.798686	3.207819	0.051535	7.508758	0.358812
2.842579	2.717757	0.171272	3.931826	0.416394
3.570678	3.979767	0.032387	3.700606	0.428670
3.289278	3.091040	0.144250	4.394449	0.539715
0.000000	0.000000	0.000000	0.000000	0.000000
0.098667	0.000000	0.000000	0.112489	0.065802
0.000000	0.378436	0.000000	3.951244	0.000000
4.416036	0.000000	0.000000	3.568280	0.511333

(b)

Fig. 6: Sample dataset of 11 features

Feature	Meaning
call_to	number of calls initiated by a client
call_in	number of calls received by a client
call_to_miss	number of missed calls initiated by a client
call_in_miss	number of missed calls received by a client
duration_call_to	total amount of duration initiated by a client
duration_call_in	total amount of duration received by a client
avg_duration_call_to	average amount of duration initiated by a client
avg_duration_call_in	average amount of duration received by a client
avg_incontact	average number of phones stored in member's directory
in_hour	rate of calls made during work hour (7 a.m - 6 p.m)
avg_success	rate of successful calls

TABLE II: 11 features

2) Algorithms and Evaluation metrics: SPIT detection is a classification problem with a binary parameter where 0 means notspam and 1 means spam. To classify, we make use of 11 features above and 8 classification algorithms namely: Logistic regression, k-nearest neighborhood, Naive Bayes, Decision tree, Random forest, Bagging, AdaBoost and XGBoost. With XGBoost, the librabry availably intergrated in xgboost is used, while we use sklearn library for the other algorithms.

To estimate the effectiveness of the model, we use the k-fold cross-validation technique, which is often used to compare and select the best model for a problem. This technique is easy to understand, implement, and produces more reliable confidence intervals than other methods.

For evaluation process, we shall take into account accuracy, precision, recall and AUC. Within this paper, it is only possible for us to explain briefly about these evaluation factors. Accuracy assigns a measurement to how close it is to the accepted or true value, while precision means how close measurements of the same item are to each other. Precision is self-reliant from accuracy. That signals high precision may not lead to high accuracy, in some cases, even low accuracy. And, high accuracy with low precision is also possible to happen. As a

result, the more appreciated quality scientific observation is F1 score which is harmonic mean of precision and recall on the assumption that this value is not equal to 0. We also put emphasize on ROC - a curve that shows the performance of a classification model at all classification thresholds. AUC (Area under the curve) can also be used as an effective factor.

C. Graph-based

Necessary data is also crawled from the database of iCaller application. With graph-based, we aim at computing reputation score for each user, then the assessment to determine whether that user is a spammer or not is of simplicity. Our method consist of two fundamental steps as following

- First, we construct a graph representing a network between users where each user is represented by a node, and the calling relationship between users is illustrated by an arc. Each arc is weighted by the total duration between users.
- Then, with the help of Eigentrust, by calculating the centrality value as the reputation index, we come to a conclusion that the trusted-users group has a much higher score than the spam-users.

Let us explain why we build a mechanism around call duration. Our methodology is inspired by a simple perception that a legitimate client ordinarily has a large number of the calls that last for long spans. On the other hand, a spammer/salesperson are likely to try to reach as many individuals as would be prudent by making an enormous number of moderately short calls. A spammer may regularly receive no incoming call or a very small number of calls. The distinction in call patterns is that, for a spammer, the call pattern is, to a great extent, unidirectional while it is bidirectional for legitimate clients. We exploit this distinction in call patterns and use call duration to make considerable accreditations as confirmation of an understood degree of trust.

Given a graph $G = (V, A)$ which stands for a social network among phone users where V is vertex set, each v in V represents a member. Naturally, A is known as the set of arcs. Recall that $(u, v) \in A$ if and only if u has made a call to v. To illustrate, in figure 7, phone number 1091 once initiated calls to four other phone numbers and those calls were answered, so there exist arcs linking phone number 1091 to those four numbers. Denote w_{uv} is the weight of arc (u, v), then we come up with the normalized local trust rate which has formulation as follows

$$c_{uv} = \frac{d_{uv}}{\sum_j d_{uj}}$$

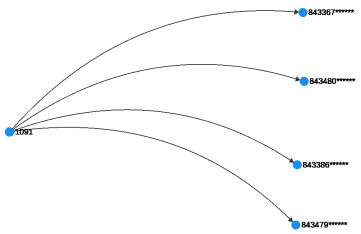


Fig. 7: Spammer and Legitimate Users network

with d_{uv} is the total duration of calls that user u has made to user v. c_{uv} plays the main role in our arc-weighting methodology. It is worth noticing that when it comes to a totally new user, he/she hasn't made a call, as a result, total duration of calls initiated by him/her is 0. In order to avoid the cases of dividing for zero, if there exists an user i has not made any call to any nearby users, we need an initial weight $c_{ij} = \frac{1}{|V|}$ where $|V|$ is the number of vertices.

After building a graph that represents the network among users, we use Eigentrust which is a centrality measure to evaluate the reputation of users on the social network. This is based on the assumption that the legitimate level of each user is determined by nearby users in the network.

Assume that we have m peers in the social network. We shall display some denotions contributing to Eigentrust algorithm.

Symbol	Description
t	Vector of reputation score
C	Local trust value
e	initial distribution

The initial distribution vector e has unit 1-norm and its component is defined by $e_i = \frac{1}{m}$. Hence, to calculate the reputation values of users denoted by vector $\overrightarrow{t}$, we compute the stationary distribution $\overrightarrow{t}$ by solving the following equation

$$\overrightarrow{t} = \left(C^T\right)^n * \overrightarrow{e}$$

with large n is the number of iterations. The algorithm is clearly illustrated by the following procedure:

IV. Experiments and Discussion

A. Feature-based

We use different machine learning algorithms to predict SPITters on a dataset of 4150 spammers and 4150 legitimate members. Then, we use the k-fold technique to compare and select the model with the best results. We divide the random data set into 10 parts and train over 10 times. At each attempt, we will choose 1 part as the validation data and the rest as train data. That helps us to evaluate differently and more accurately.

	classifiers_name	Accuracy	F1 Score	Recall	Precision	AUC
7	XGBClassifier	0.975904	0.976038	0.979167	0.972930	0.996980
4	RandomForest	0.975502	0.975610	0.977564	0.973663	0.996540
5	BaggingClassifier	0.975502	0.975551	0.975160	0.975942	0.995786
6	AdaBoostClassifier	0.969076	0.969089	0.967147	0.971038	0.993122
0	LogisticRegression	0.952209	0.952533	0.956731	0.948372	0.987313
1	KNN	0.961446	0.961446	0.959135	0.963768	0.985764
2	Naive_bayes	0.932932	0.929447	0.881410	0.983021	0.978825
3	DecisionTree	0.942972	0.942880	0.939103	0.946688	0.942981

Fig. 8: Result of 8 different algorithms

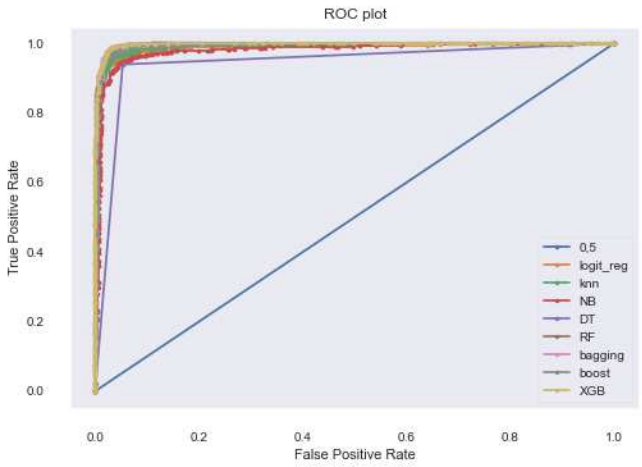


Fig. 9: ROC exhibition of 8 algorithms

It is easily observed in figure 8 and 9 that XGB is state-of-the-art, while Random Forest also helps to yeild quite exceptional result. It is natural that the algorithms using ensemble learning method always give higher results than the others. It is also reasonable that the algorithms Linear Regression, K-Nearest Neighborhood, Naive Bayes give worse results compared to other approaches. The reason for this is because these algorithms approach as a linear classification and it is not suitable with our dataset.

B. Graph-based

After computational process, the reputation score of users is given in figure figure 10

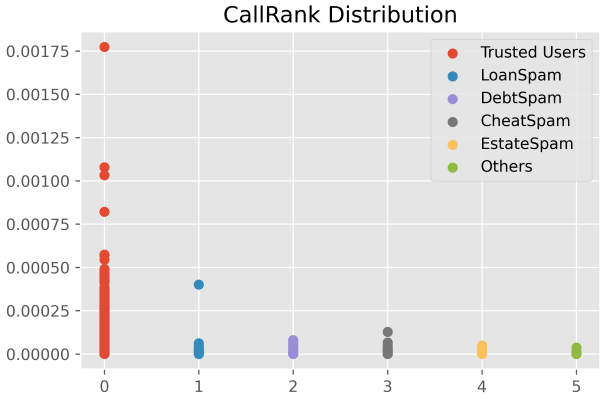


Fig. 10: CallRank distribution

An easy observation is that the reputation scores of legitimate users are much higher than those of spammer group. Table III shows the 95% confidence interval of each group.

Type	Bound
Trusted Users	18.0715 - 25.0715
LoanSpam	1.489 - 3.489
DebtSpam	0.3921 - 2.3921
CheatSpam	0.26 - 2.26
EstateSpam	0.1214 - 1.1214
Others	0.2967 - 2.2967

TABLE III: Confidence Interval

Although the results obtained are acceptable, this method has a noticable limitation. An outstanding criteria to detect SPITers is that most SPITers has a very low rate of $\frac{\text{call\_in}}{\text{call\_to}}$ where "call_in" is the number of received calls and "call_in" is the number of initiated calls. However, in the proposed graph-based approach, there is a shortage of nodes' orders which serves as another factor to weight each arc between two users. We intend to address this issue as part of our future work.

V. Conclusion

This work focuses on developing a mechanism to cope with SPIT attackers using two approaches: feature-based and graph-based. With feature-based, the dataset crawled from iCaller application is made use of to extract 11 useful features of SPIT attacks. Those criteria serve training process of eight

supervised classification algorithms (Logistic regression, k-nearest neighborhood, Naive Bayes, Decision tree, Random forest, Bagging, AdaBoost and XGBoost). A provisional study of those algorithms is conducted and the state-of-the-art algorithm, XGBoost, is revealed based on an evaluation of ROC curve. With graph-based, we apply some modification from CallRank algorithm and also yeild quite exceptional result. However, there is still a lack of information about numbers of calls from each node. Therefore, in future work, we have an intention to add a penalty function to display the rate between received calls and initiated calls at each node in order to increase our method's accuracy in general.

REFERENCES

[1] V. Balasubramaniyan, M. Ahamad, and H. Park, "Callrank: Combating spit using call duration, social networks and global reputation." 01 2007.

[2] P. Kolan and R. Dantu, "Socio-technical defense against voice spamming," *TAAS*, vol. 2, 03 2007.

[3] D. Shin, J. Ahn, and C. Shim, "Progressive multi gray-leveling: A voice spam protection algorithm," *Network, IEEE*, vol. 20, pp. 18 – 24, 10 2006.

[4] H. Yan, K. Sripanidkulchai, H. B. Zhang, Z.-Y. Shae, and D. Saha, "Incorporating active fingerprinting into spit prevention systems," 2006.

[5] R. Schlegel, S. Niccolini, S. Tartarelli, and M. Brunner, "Spam over internet telephony (spit) prevention framework," 01 2007, pp. 1 – 6.

[6] M. Nassar, R. State, and O. Festor, "Monitoring sip traffic using support vector machines," 01 2008.

[7] M. Nassar, O. Dabbebi, R. Badonnel, and O. Festor, "Risk management in voip infrastructures using support vector machines," 11 2010, pp. 48 – 55.

[8] Y.-S. Wu, S. Bagchi, N. Singh, and R. Wita, "Spam detection in voice-over-ip calls through semi-supervised clustering," 06 2009, pp. 307–316.

[9] T. Kusumoto, E. Chen, and M. Itoh, "Using call patterns to detect unwanted communication callers," 07 2009, pp. 64–70.

[10] R. J. Ben Chikha, T. Abbes, and A. Bouhoula, "A spit detection algorithm based on user's call behavior," in *2013 21st International Conference on Software, Telecommunications and Computer Networks - (SoftCOM 2013)*, 2013, pp. 1–5.

[11] R. jabeur ben chikha, T. Abbes, W. Chikha, and A. Bouhoula, "Behavior-based approach to detect spam over ip telephony attacks," *International Journal of Information Security*, vol. 15, 03 2015.

[12] Y. Rebahi and D. Sisalem, "Sip service providers and the spam problem," in *In 2nd Workshop on Securing Voice over IP*, 2005.

[13] P. Patankar, G. Nam, G. Kesidis, and C. Das, "Exploring anti-spam models in large scale voip systems," 07 2008, pp. 85–92.

[14] Y. Soupionis and D. Gritzalis, "Aspf: Adaptive anti-spit policy-based framework," 09 2011, pp. 153 – 160.

[15] B. Mathieu, S. Niccolini, and D. Sisalem, "Sdrs: A voice-over-ip spam detection and reaction system," *Security and Privacy, IEEE*, vol. 6, pp. 52 – 59, 01 2009.

[16] J. Quittek, S. Niccolini, S. Tartarelli, M. Stiemerling, M. Brunner, and T. Ewald, "Detecting spit calls by checking human communication patterns," 06 2007, pp. 1979–1984.