International Journal of Engineering Research and Modern Education (IJERME)
ISSN (Online): 2455 - 4200
(www.rdmodernresearch.com) Volume I, Issue I, 2016

VOTE TRUST: A NOVEL APPROACH TO FILTER USER BEHAVIOUR AND SYBILS IN OSN

K. P. Sarona* & M. Rajakumar**

* PG Scholar, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

** Associate Professor & Head, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

Abstract:

Most of the Online Social Networks (OSN) restrict users of certain age groups (for example the people below 18 years are not allowed to signup). Another main issue with Online Social Networks (OSNs) are they suffer from the creation of fake accounts that introduce fake product reviews, malware and spam. Existing defenses focus on using the social graph structure to isolate fakes. However, our work shows that Sybils could befriend a large number of real users, invalidating the assumption behind social-graph-based detection. In this paper, we proposed a novel concept which will block all Sybil communities as well as malicious and inappropriate messages in social networking communities. Here we use a modified version of Vote Trust algorithm, a scalable defense system that further leverages user-level activities along with a content filtering concept. VoteTrust models the friend invitation interactions among users as a directed, signed graph, and uses two key mechanisms to detect Sybils over the graph: a voting-based Sybil detection to find Sybils that users vote to reject, and a Sybil community detection to find other colluding Sybils around identified Sybils. We also deploy this algorithm along with a content filtering mechanism in our OSN application so as to show the effectiveness of this new algorithm. The content filtering mechanism uses two distinct algorithms for text content filtering and media content filtering. Text content filtering is based on natural language processing and the images and other media type is filtered based on well known skin detection algorithms.

Index Terms: Online Social Network, Sybil Attack, Sybil Detection & Spam

1. Introduction:

Online social networks are not only a way to keep in touch, but a way of life. Several features of online social networks are common to each of the more than 300 social networking sites currently in existence. The most basic feature is the ability to create and share a personal profile. This profile page typically includes a photo, some basic personal information (name, age, sex, and location) and extra space for listing your favorite bands, books, TV shows, movies, hobbies and Web sites.

Sybil attacks are one of the most prevalent and practical attacks against distributed systems. In this attack, a malicious user creates multiple fake identities, known as Sybil's, to unfairly increase their power and influence within a target community. Distributed systems are ill-equipped to defend against this attack, since determining a tight mapping between real users and online identities is an open problem. To date, researchers have demonstrated the efficacy of Sybil attacks against P2P systems, anonymous communication networks, and sensor networks.

One fundamental issue in today On-line Social Networks (OSNs) is to give users the ability to control the messages posted on their own private space to avoid that unwanted content is displayed. Up to now OSNs provide little support to this requirement. To fill the gap, in this paper, we propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a

flexible rule-based system, that allows users to customize the filtering criteria to be applied to their walls, and a Machine Learning based soft classifier automatically labeling messages in support of content-based filtering. The feature that makes this project unique is the ability to sense what is good and what is bad. The system provides a powerful rule layer exploiting a flexible language to specify Filtering Rules (FRs), by which users can state what content should not be displayed on their walls. In addition, the system provides the support for user-defined Black Lists (BLs) and blocks Adult images and Words from user wall.

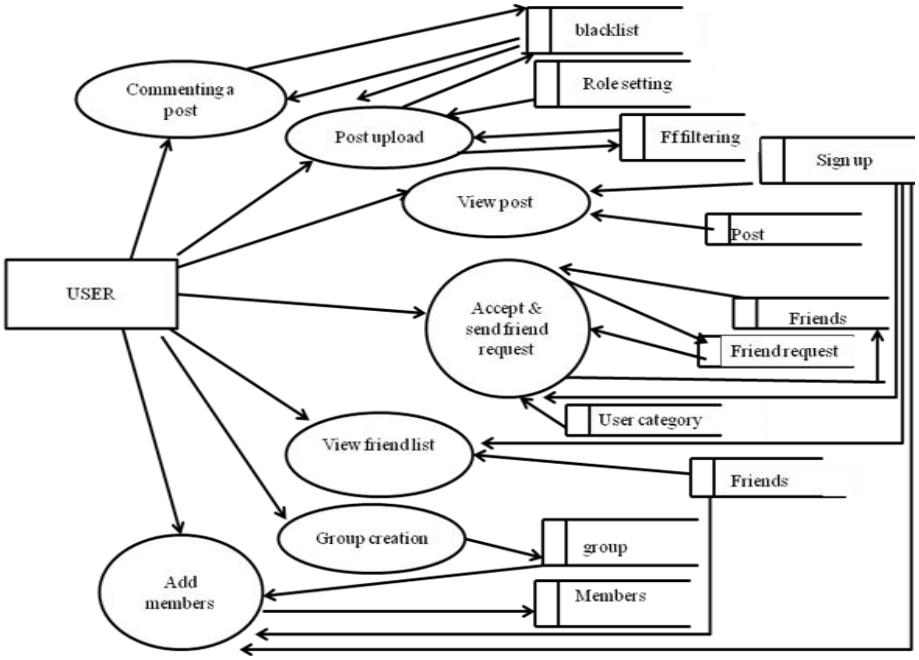


Figure 1: Data Flow Diagram

The reminder of this paper is organized as follows. Section II, describes the Related Works. Section III, describes the Proposed Work. Section IV, describes the Experimental Evaluation and Results. Section V summarizes the Conclusion and Future Enhancement and Section VI References.

2. Literature Survey:

Recently, there has been a great effort in defending Sybils (e.g., spammers) in OSNs. Some of them are analysed as below:

The features of OSN are attractive to adversaries who have incentives to distort the popularity and value of the resource. Generally, adversaries can launch Sybil attack or Identity Clone attack to achieve their malicious purposes. Different works are there to analyse this two type of attacks. Sybil attacks focus on creating multiple online user identities (Sybil identities) and try to achieve malicious results through these identities. In an Identity Clone attack (also called Profile Cloning attack) an adversary first creates similar or even identical profiles to impersonate victims in an OSN system. He then distorts the reputation and the value of a resource through the network involving faked profiles.

Sybil attacks and Identity Clone attacks look somehow similar in appearance since these similar attack patterns could confuse administrators of OSN systems. These works analyse and compare Sybil attacks and Identity Clone attacks in OSN systems based on their characteristics. Comparison between these attacks and characteristics of a Sybil attack and an Identity Clone attack based on the pre-requirements, network topologies and impacts of the attacks.

When considering the network topology, in the presence of Sybil attack a social network graph can be conceptually divided into two parts: one consisting of all genuine identities and the other consisting of all Sybil identities. The link connecting a genuine node to a Sybil node is called an attack edge. In the network of an Identity Clone attack, there are two types of nodes: genuine nodes and cloned nodes. For example, in Figure 2.1, a large solid sphere is the target node while a large hollow sphere is the cloned identity created by an adversary and impersonating the target. The small solid squares are the other genuine nodes and the hollow squares are the rest of the cloned nodes created by the adversary to impersonate the targets friends or potential friends.

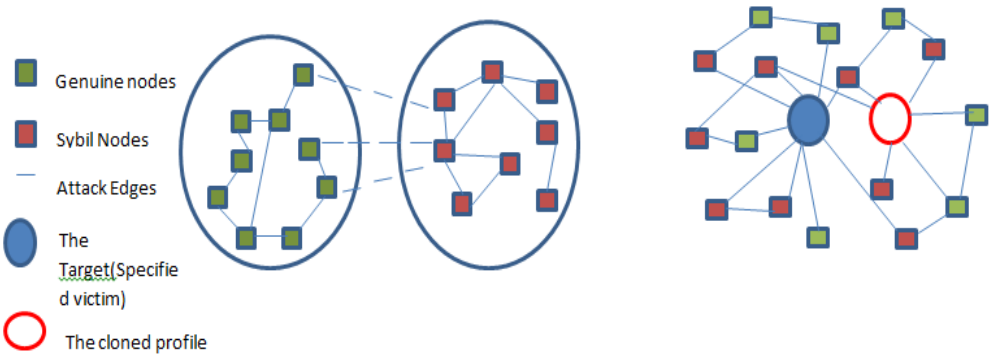


Figure 2: A network of a Sybil attack and identity clone attack

Now in attack impacts of Sybil attack and identity clone attack, Sybil attack can be used to affect the popularity, reputation, value and other characteristics of resources in OSN systems by using Sybil nodes. An adversary can boost invaluable resources and resource providers who have bad reputations. He can also downgrade valuable resources and reputable resource providers. In addition, the adversary can launch spam attacks by requesting the Sybil identities to propagate malicious messages to their neighbour nodes.

Similar to Sybil attacks, Identity Clone attacks can affect the popularity, reputation, value of resources in OSNs using fake profiles. Additionally, such attacks can also influence the choices made by victim's friends using the trust built in friendships. Now considering the Pre-Requirements of attacks, to launch a Sybil attack on an OSN system, an adversary needs to create multiple identities. Most OSN systems require a user to choose a username or input his e-mail address during the account registration process. This result indicates that adversary needs to have multiple unique usernames or a huge number of e-mail addresses for launching the attack. Thus, the adversary also needs to compromise this restriction in the registration process, in order to create many identities automatically. To launch an Identity Clone attack, an adversary also requires a number of unique usernames and e-mail addresses for creating identities in an OSN system. Additionally, the knowledge of a victim (identity that is cloned) is another pre-requirement in an Identity Clone attack. One of the limitations of this study is that the work does not contain any method to detect or defend against Sybil's.

Some decentralized protocols OSNs leverage the social graph structure to defend against Sybil's. This section discusses Sybil guard a novel decentralized protocol that limits the corruptive influence of Sybil attacks. Its design is based on a unique insight regarding social networks, where identities are nodes in the graph and (undirected) edges are human-established trust relations (e.g., friend relations). The edges connecting the honest region (i.e., the region containing all the honest nodes) and the Sybil region (i.e., the region containing all the Sybil identities created by malicious

users) are called attack edges. This protocol ensures that the number of attack edges is independent of the number of Sybil identities, and is limited by the number of trust relation pairs between malicious users and honest users.

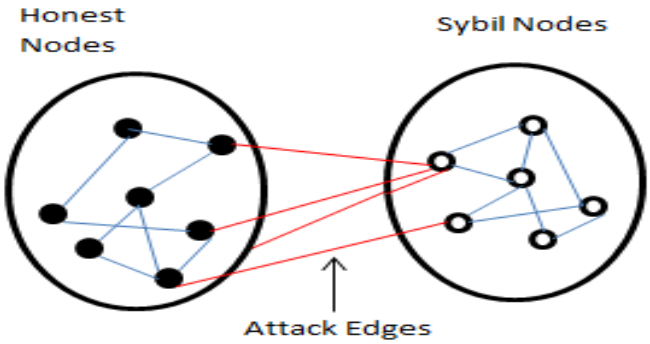


Figure 3: The social network with honest nodes and Sybil nodes

If malicious users create too many Sybil identities, the graph becomes strange in the sense that it has a small quotient cut ie.,: a small set of edges (the attack edges) whose removal disconnects a large number of nodes (all the Sybil identities) from the rest of the graph. Directly searching for such cuts is not practical. Instead, relies on a special kind of verifiable random walk in the graph and intersections between such walks. These walks are designed so that the small quotient cut between the Sybil region and the honest region can be used against the malicious users, to bound the number of Sybil identities that they can create.

The basic idea behind is that an honest node (called the verifier) decides whether or not to accept another node (called the suspect).The verifier only accepts a suspect whose random route intersects with the verifiers random route. Accepting S means that V is willing to receive service from and provide service to S. Ideally, the defense system should guarantee that V accepts only honest nodes.

3. Proposed Work:

The proposed system is based on content filtering system and a modified version of vote trust which can block text messages and even multimedia messages from entering into user's wall and message box. The system implements a text filtering system which is based on a knowledge database and natural language processing. The media filtering system implements a skin detection algorithm and the media content is blocked based on the occurrence of skin colored pixels. The blocked message will appear in the private message box, where a user can verify and block the messages if needed. The system implements a simplified version of vote trust algorithm to find Sybil groups and thus making social network safe and free of sybils.

To propose a system allowing OSN users to have a direct control on the messages posted on their walls. This is achieved through a flexible rule-based system that allows users to customize the filtering criteria to be applied to their walls, and a Machine Learning-based soft classifier automatically labeling messages in support of content-based filtering.

- ✓ High restriction to post vulgar messages on other's wall.
- ✓ Efficient content filtering system.
- ✓ Safe for any age groups.
- ✓ Sybil free.
- ✓ User's privacy is still preserved.

4. Experimental Analysis and Results:

Implementation is the process of translating design specification in to source code. The primary goal of implementation is to write source code and internal implementation. So that conformance of code to its specification can be easily verified, So that debugging, testing and modification are eased. The source is developed with clarity, simplicity and elegance.

The coding is done in a modular fashion giving such importance even to the minute detail so, when hardware and storage procedures are changed or now data is added, rewriting of application programs is not necessary. To adapt or perfect use must determine new requirements, redesign generate code and test exiting software/hardware. Traditionally such task when they are applied to an existing program has been called maintenance.

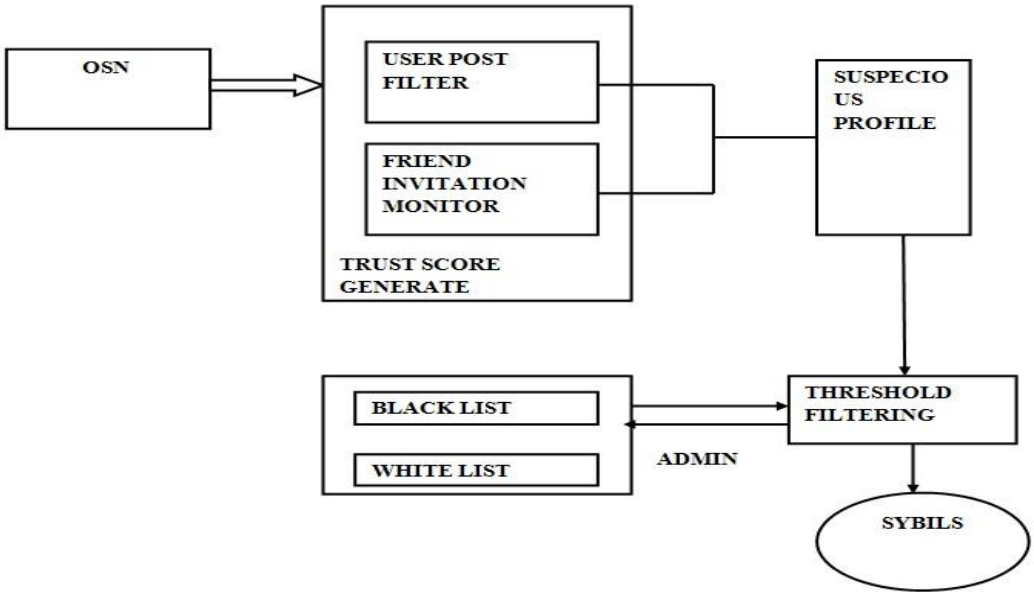
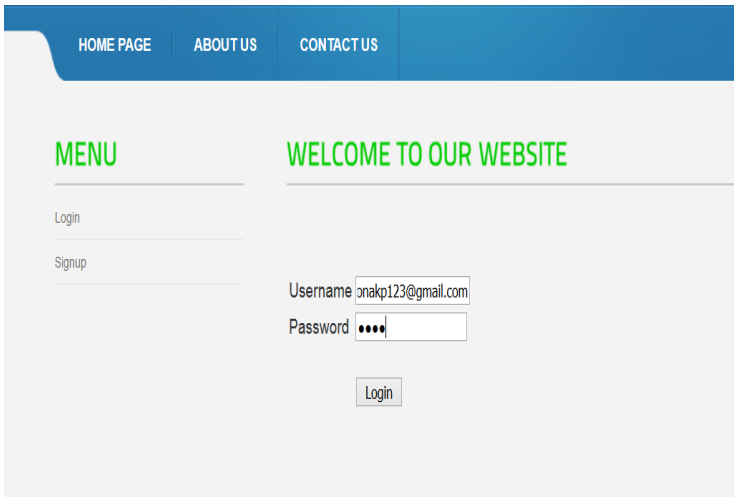


Figure 4: System Architecture

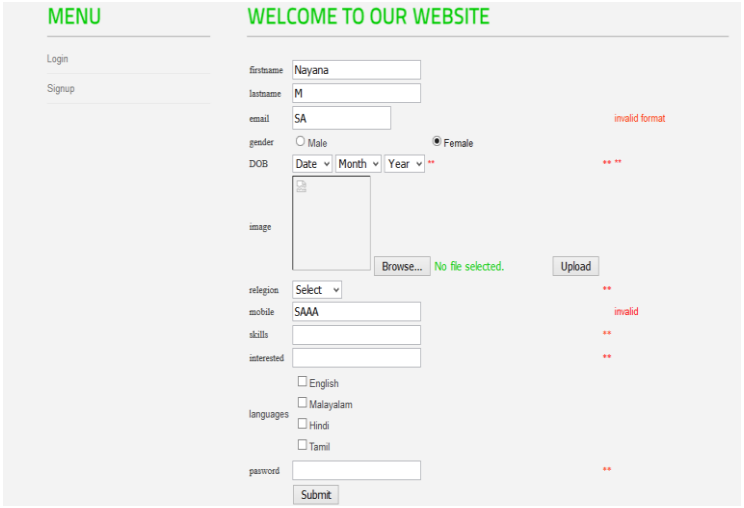
The system architecture for the proposed system including user behaviour analyzer starting from the Online Social Network (OSN), which consists of different users. Vote Trust uses behaviour analyzer logic in addition to the friend invitation graph among the users in OSN. Vote Trust models the request/confirm interactions among users as a friend invitation graph. The goal of Vote Credence is to take input as the friend invitation graph and outputs the classification of any user that can be modeled as a node in the graph. User behaviour analyser module uses the message filtering mechanism. After analysing the profiles, on the basis of the graph and user behaviour analyser some suspicious profile has been identified.

Now the role of the admin is to use a particular threshold value and select some trusted users as seeds to seed the trust. The vote propagation logic is to propagate the votes through the trusted seeds. Global vote aggregation is also performed by admin. After vote aggregation, the Sybil detection is performed by calculating the global acceptance rate and determines whether this value for each of the user is above or below the threshold value used by the admin. All the users that have the global acceptance rate below the threshold value are identified as sibyls. After identifying this Sybils, Vote Trust system deletes the Sybil account.

Login:



User Sign Up:



Post Messages:

5. Conclusion and Future Enhancement:

In this paper we have presented a system to filter undesired messages from OSN walls. The system exploits a soft classifier to enforce customizable content-dependent Filtering rules. This work is the first step of a wider project. The early encouraging results we have obtained on the classification procedure prompt us to continue with

other work that will aim to improve the quality of classification. In particular, future plans contemplate a deeper investigation on two interdependent tasks. The first concerns the extraction and/ or selection of contextual features that have been shown to have a high discriminative power. The second task involves the learning phase. Since the underlying domain is dynamically changing, the collection of pre classified data may not be representative in the longer term. The present batch learning strategy, based on the preliminary collection of the entire set of labeled data from experts, allowed an accurate experimental evaluation but needs to be evolved to include new operational requirements.

6. References:

1. N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in Proc. of NSDI, 2009.
2. B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based sybil defenses," in Proc. of SIGCOMM, 2010.
3. J. Xue, Z. Yang, X. Yang, X. Wang, L. Chen, and Y. Dai, "Vote trust Leveraging friend invitation graph to defend against social network sybils," in Proc. of INFOCOM, 2013.
4. J. Jiang, C. Wilson, X. Wang, P. Huang, W. Sha, Y. Dai, and B. Y. Zhao, "Understanding latent interactions in online social networks," in Proc. of IMC, 2010.
5. L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: automated identity theft attacks on social networks," in Proc. of WWW, 2009.
6. Y. Boshmaf, D. Logothetisy, G. Siganosz, J. L. Jorge Ler´ıax, M. Ripeanu, and K. Beznosov, "´Integro: Leveraging victim prediction for robust fake account detection in osns," in Proc. of NDSS, 2015.
7. Z. Gyongyi, H. Garcia-molina, and J. Pedersen, "Combating web spam with trust rank," in VLDB. Morgan Kaufmann, 2004, pp. 576–587.
8. Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," in nsdi, 2012.
9. L. G. Valiant, "A bridging model for parallel computation," Commun. ACM, vol. 33, no. 8, pp. 103–111, Aug. 1990.
10. A. Cheng and E. Friedman, "Sybil proof Reputation mechanisms," in Proc. Of P2PECON, 2005.