WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

SYBILBELIEF: A SEMI-SUPERVISED LEARNING APPROACH FOR STRUCTURE BASED SYBIL DETECTION

[a] K.Banuroopa, [b]N.Manoj bala

[a]Assistant Professor, Department of Computer Science, Karpagam Academy of Higher Education
[b]Department of Computer Science, Karpagam Academy of Higher Education

A B S T R A C T

Sybil attacks are a fundamental threat to the security of distributed system. There has been a growing interest in leveraging social network to mitigate Sybil attacks. We introduce Sybil belief a semi supervised learning framework to detect Sybil nodes. Sybil Belief takes a social network of the nodes in the system, a small set of known benign nodes, and, optionally, a small set of known Sybil's as input. We show that Sybil Belief is able to accurately identify Sybil nodes with low false positive rates and low false negative rates. Sybil Belief is resilient to noise in our prior knowledge about known benign and Sybil nodes. Sybil accounts in online social networks are used for criminal activities such as spreading spam or malware stealing other users' private information and manipulating web search results. Sybil defenses require users to present trusted identities issued by certification authorities. However, such approaches violate the open nature that underlies the success of these distributed systems.

Keywords: Sybil attacks, Loop belief propagation, distributed systems

1. Introduction

Sybil attacks, where a single entity emulates the behavior of multiple users, form a fundamental threat to the security of distributed systems .Example systems include peer-to peer networks, email, reputation systems, and online social networks. For instance, in 2012 it was reported that 83 million out of 900 million Face book accounts are Sybils. Sybil accounts in online social networks are used for criminal activities such as spreading spam or malware, stealing other users' private information , and manipulating web search results via "+1" or "like" clicks. Traditionally, Sybil defenses require users to present trusted identities issued by certification authorities. However, such approaches violate the open nature that underlies the success of these distributed systems. Recently, there has been a growing interest in leveraging social networks to mitigate Sybil attacks. These schemes are based on the observation that, although an attacker can create arbitrary Sybil users and social connections among themselves, he or she can only establish a limited number of social connections to benign users. As a result, Sybil users tend to form a community structure among them, which enables a large number of Sybil users to integrate into the system. Note that it is crucial to obtain social connections that represent trust relationships between users, otherwise the structure-based Sybil detection mechanisms have limited detection accuracy.

2. Methodology

Sybil Belief a semi-supervised learning framework, to perform both Sybil classification and Sybil ranking. Sybil Belief overcomes a number of drawbacks of previous work. We extensively evaluate the impact of various factors including parameter settings in Sybil Belief, the number of labels, and label noise on the performance of Sybil Belief using synthetic social networks. Sybil ranking mechanism achieves reasonably good performance. In their experimental evaluation, the authors found that using simple local community detection had equivalent results to using the state-of-art Sybil detection approaches. Several approaches have been proposed to propagate trust scores or reputation scores in file-sharing networks or auction platforms. The users or nodes involved in our projects are Sender, Intermediate and Receiver. In order to send file, the sender has to find out the list of nodes which are connected with the sender. From that available list he can choose receiver. Then the sender has to analyze the performance of each and every node which is connected with the sender. The performance analysis list will return the priority based result so that sender can choose the intermediate to send the file. The Intermediate will receive the file from sender then it will analyze the performance so that it can send data to another intermediate or receiver. In the receiver side, the receiver has to select the file path to receive the file from sender or intermediate. Then the receiver can view the file received file.

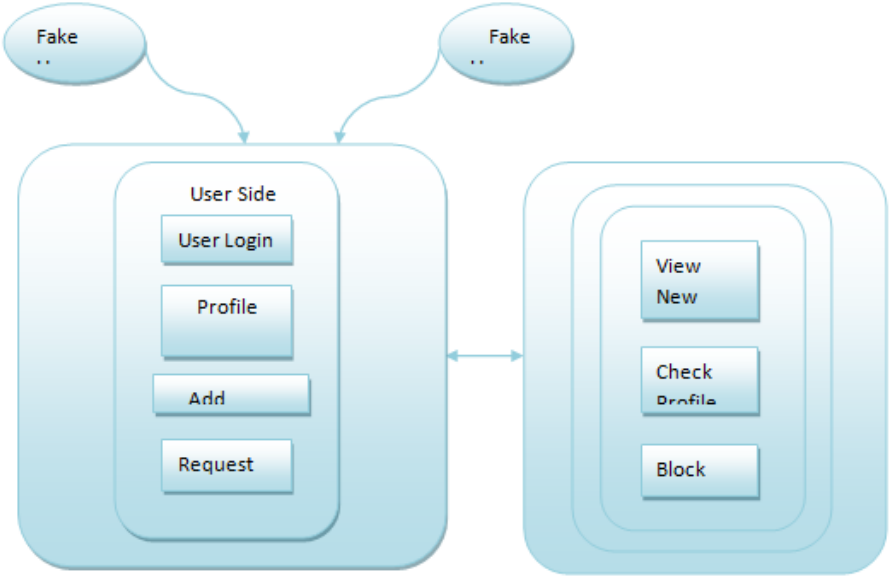

Fig-1 System architecture

A. Loopy Belief Propagation

Belief propagation algorithm exists for several types of graphical models Bayesian network and markov random field in particular. We describe here the variant that operates on a factor graph. Sybil Belief does not use random walks, and relies instead on the Markov Random Fields and Loopy Belief Propagation. Sybil Belief is able to incorporate information about known benign and known Sybil nodes. Our experimental results show that Sybil Belief performs an order of magnitude better than Sybil Limit and Sybil Infer. Moreover, Sybil Belief is scalable to large scale social networks, unlike above mechanisms. The Markov Random Fields (MRF) has many applications in electrical engineering and computer science such as computer vision and natural language processing. However, the application of MRFs to the security and privacy area is rather limited. Belief propagation algorithm exists for several types of graphical models bayesian network and markov random field, in particular. We describe here the variant that operates on a factor graph. A factor graph is a bipartite graph containing nodes corresponding to variables V and factors U, with edges between variables and the factors in which they appear. We can write the joint mass function. where $\mathbf{x}_u$ is the vector of neighboring variable nodes to the factor node u. Any Bayesian network or Markov random field can be represented as a factor graph. The algorithm works by passing real valued functions called *messages* along the edges between the nodes.

The name of the algorithm is clear from the previous formula: the complete marginalization is reduced to a sum of products of simpler terms than the ones appearing in the full joint distribution. In a typical run, each message will be updated iteratively from the previous value of the neighboring messages. Different scheduling can be used for updating the

messages. In the case where the graphical model is a tree, an optimal scheduling allows to reach convergence after computing each messages only once (see next sub-section). When the factor graph has cycles, such an optimal scheduling does not exist, and a typical choice is to update all messages simultaneously at each iteration. Upon convergence (if convergence happened), the estimated marginal distribution of each node is proportional to the product of all messages from adjoining factor.

4. Conclusion

In this paper, we propose Sybil Belief, a semi-supervised learning framework, to detect Sybil nodes in distributed systems. Sybil Belief takes social networks among the nodes in the system, a small set of known benign nodes, and, optionally, a small set of known Sybil nodes as input, and then Sybil Belief propagates the label information from the known benign and/or Sybil nodes to the remaining ones in the system. We extensively evaluate the influence of various factors including parameter settings in the Sybil Belief, the number of labels, and label noises on the performance of Sybil Belief. Moreover, we compare Sybil Belief with state-of-the-art Sybil classification and ranking approaches on real-world socialnetwork topologies. Our results demonstrate that Sybil Belief performs orders of magnitude better than previous Sybil classification mechanisms and significantly better than previous Sybil ranking mechanisms. Furthermore, Sybil Belief is more resilient to noise in our prior knowledge about known benign nodes and known Sybils. Interesting avenues for future work include evaluating Sybil-Belief and previous approaches with datasets containing real Sybils and applying our Sybil Belief framework to other security and privacy problems such as graph based Botnet detection, reputation systems ,and private information inference .

References

[1] J. R. Douceur, “The Sybil attack,” in IPTPS, 2002.

[2] Malicious/fake accounts in Face book, http://www.cnn.com/2012/08/02/tech/social-media/Face book-fake-accounts/index.html.

[3] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, “Design and evaluation of a real-time url spam filtering service,” in IEEE S & P, 2011.

[4] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, “All your contacts are belong to us: Automated identity theft attacks on social networks,” in WWW, 2009.

[5] P. L. Fong, “Preventing Sybil attacks by privilege attenuation: A design principle for social network systems,” in IEEE S & P, 2011. [6] Google Explores +1 Button To Influence Search Results, http://www.tekgoblin.com/2011/08/29/google-explores-1-button-toinfluence- search-results/. [6] Google Explores +1 Button To Influence Search Results, “http://www.tekgoblin.com/2011/08/29/google-explores-1-button-toinfluence-search-results/.”

[7] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, “An analysis of social network-based Sybil defenses,” in SIGCOMM, 2010.

[8] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman., “SybilGuard: Defending against Sybil attacks via social networks,” in SIGCOMM, 2006. [9] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, “SybilLimit: A nearoptimal social network defense against Sybil attacks,” in IEEE S & P, 2008.

[9] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, “SybilLimit: A nearoptimal social network defense against Sybil attacks,” in IEEE S & P, 2008.

[10] G. Danezis and P. Mittal, “Sybil Infer: Detecting Sybil nodes using social networks,” in NDSS, 2009.

[11] Praneesh, M., and R. Annamalai Saravanan. "Deep Stack Neural Networks Based Learning Model for Fault Detection and Classification in Sensor Data." *Deep Learning and Edge Computing Solutions for High Performance Computing*. Springer, Cham, 2021. 101-110.