Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 17 (2013) 861 – 869

Procedia
Computer Science

Information Technology and Quantitative Management (ITQM2013)

Detecting Sybil Nodes in Anonymous Communication Systems

Wenyu Zang[a,b, c,d], Peng Zhang[b,c], Xiao Wang[a,b,c,d], Jinqiao Shi[b,c] , Li Guo[b,c]*

[a]Institute of Information Engineering, CAS, Beijing, China
[b]*Institute of Computing Technology, CAS, Beijing, China*
[c]*National Engineering Laboratory for Information Security Technologies, Beijing,China*
[d]*Graduate University, CAS, Beijing, China*

Abstract

As one of the most fundamental problems in open-membership systems, Sybil attack has attracted wide attentions from both industrial and academic fields. Many anonymous communication and censorship-resistant systems have incorporated Sybil defense in their designs based on social networks. This paper proposed two-class undirected mixed membership stochastic blockmodels to discover adversary's Sybil identities in the network. Different from existing algorithms, our model discriminates Sybil nodes by simulating the generative process of social networks. And we gave a matrix B to describe the interaction probability of Sybil and honest nodes, which can be used to observe contaminated situation of the network. The effectiveness of the algorithm is validated by experiments over both simulated and real-world social networks. The method can be also used to further improve the accessibility of anonymous and anti-censorship systems.

*Keywords:*Sybil attack, anonymous communication, Accessibility, probability graph model;

1. Introduction

Sybil attack [1] is one of the most challenging and fundamental problems in open-membership systems. By creating multiple Sybil identities, the adversary can "suppress" the honest users in a wide scope of tasks, such as voting schemes for email spam [2], data replication in DHT [3]. For example, in user-generated content sites such as Youku, Youtube and Zhihu, the attacker can register many identities and collude to manipulate the

* Corresponding author: Tel: +86-10-82546739; fax: +86-10-82546701
E-mail address: zangwenyu@nelmail.iie.ac.cn.

doi:10.1016/j.procs.2013.05.110

content ranking result which is based on users' voting. Other kinds of Sybil attacks have also been observed in peer-to-peer systems like Maze [4] [5] and eMule [6].

Recently, Sybil attack, due to its significant impact on system accessibility, has aroused wide attention in the design of anonymous and censorship-resistant systems. Research work can be classified into the following three aspects: (i) Traditional Sybil attacks and defense. By compromising multiple proxies in the anonymous communication infrastructure, the attacker can perform selective DoS attack [7], misleading routing attack [8] [9] or even information harvest attack [10] [11]. Defense methods of these attacks are highly related to the attacks, ranging from DoS probing [12] and collusion detection to client authorization [13], etc. (ii) Sybil defense in peer-to-peer anonymous communication systems. In these systems, user relies on each other to relay his traffic to the destination. In order to prevent a malicious user from obtaining the entire view of the system, a user is allowed to connect to only a few neighbors. And the neighbor relationship between two users can be established if they know each other in the real world [10] [14] [15]. (iii) Sybil defense in centralized anonymous communication systems. One key challenge in the design of these systems is to distributing proxies to each user while avoiding attackers from enumerating them with many Sybil identities. Recently, social relationship is adopted into the design of proxy distribution strategies with the purpose of limiting adversary's ability of creating multiple Sybils [16] [17] [18]. For example, in Psiphon [16], new users can make a registration and learn proxy addresses only through the invitation by an existing registered user. We can leverage the social relationships in these designs to further discover potential Sybils and mitigate Sybil attacks.

It can be observed that, although the adversary can create arbitrary Sybil identities in the system (by connecting to existing Sybils or through the invitation of other Sybils), it cannot establish many social connections to honest nodes. Based on the same observation, a variety of Sybil defense schemes [19] [20] [21] have been proposed to distinguish Sybil nodes from honest ones. These schemes take the social graph and a known honest node as input. They can be viewed as graph partitioning algorithms that divide nodes in a graph into two parts: the Sybil region and the honest region. This paper assumes that, both the honest region and Sybil region in the network is fast mixing. We leverage the idea in community detection field to detect the Sybil nodes. We used a probability graph model named Two-class undirected mixed member stochastic Blockmodels to detect the Sybil nodes through simulate the generative process of a social network. In our model every nodes are belonging to both Sybil and honest group with different probabilities. Nodes in different group (Sybil and honest nodes) have different interaction probabilities; we used a matrix B to describe these probabilities. Finally our model used observed edges between all nodes to infer the probability of Sybil for each node and interaction probabilities matrix B.

2. Two-class undirected mixed member stochastic Blockmodels Algorithm in Sybil discovery

In this section we describe our model named two-class undirected mixed member stochastic Blockmodels. In our model, we assume that the observed network can be described as a graph G=(N,Y), where $Y_{pq} \epsilon \{0,1\}$ represents whether there is an edge between node p and q. Our goal is to calculate the probability of Sybil for each node and estimate the probability of having a link between honest and Sybil nodes.

2.1. Problem statement and notation illustration

As show in figure 1(a) we assume that although an attacker can create arbitrary Sybil identities in social networks, he or she cannot establish an arbitrarily large number of social connections to non-Sybil nodes. As a result, Sybil nodes tend to be poorly connected to the rest of the network, compared to the non-Sybil nodes. Sybil defence schemes leverage this observation to identify Sybils. We detect Sybil nodes from the limited capacity of Sybils to establish social links.

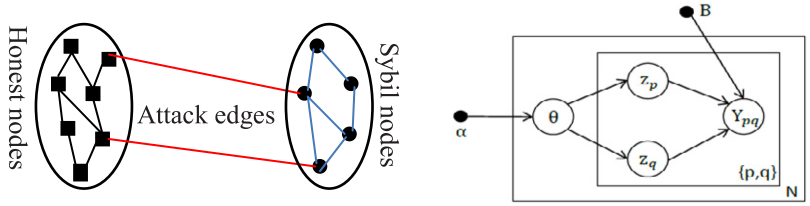

Fig. 1. (a) Sybil attack based on social network; (b) Graphical model representation of Probabilistic Block-model.

The notations used in the following passages and its illustrations are showed in table 1.

Table 1. Notation and illustration

Notation	Illustration
α	Beta prior denote global distribution of Sybil and honest node
B	Bernoulli rate matrix denote interaction probabilities
θ	Per node probability of becoming a Sybil
z_p	Whether a node is a Sybil or not in one sampling
Y_{pq}	Whether there is an edge between node p and node q
γ	Variational prior of parameter θ
φ	Variational prior of parameter z
ρ	Sparsity parameter denote the sparsity of edges

2.2. Two-class undirected mixed member stochastic Blockmodels

Two-class undirected mixed member stochastic Blockmodels(we will call it Blockmodels for short in the following passages) is a generative probability graph model for a network graph. Different from other models, we do not discriminate which nodes are Sybil nodes directly; instead we simulate the network generative process. We assume there are two latent groups exist, and the observed network is generated according to distributions of group-membership for each node and a matrix of group-group interaction strength. The distribution of each node is specified by latent simplicial vectors. Each node is associated with a randomly drawn vector θ_i for node i, where θ_{ik} denotes the probability of node i belonging to group k. That is each node is belonging to both two groups with different probabilities. The probabilities of interactions between different groups are defined by a matrix of Bernoulli rates $B_{(K*K)}$, where B(g,h) represents the probability of having a link between a node from group g and group h.

The complete generative process for a graph $G = (N, Y)$ is as follows:

- For each node $i \in N$
  --Draw a 2 dimensional mixed membership vector $\theta_i \sim Beta(\alpha)$
- For each pair of nodes $(p, q) \in N \times N$
  --Draw membership indicator $z_p \sim Bornoulli(\theta_p)$
  --Draw membership indicator $z_q \sim Bornoulli(\theta_q)$
  --Sample the value of their interaction $Y(p, q) \sim Bornoulli(z_p B z_q)$

The Blockmodels is represented as a probabilistic graphical model in Figure 1(b). As the figure makes clear, there are three levels to the Blockmodels representation. The parameters α and B are global level parameters, assumed to be sampled once in the process of generating a network. The variables θ are node-level variables, sampled once per node. Finally, the variables $\mathrm{z_p}, \mathrm{z_q}$ and $\mathrm{Y_{pq}}$are edge-level variables and are sampled once for each edge in the network graph.

3. Inference and Parameter Estimation

We have described the Blockmodels and illustrated its generative process in the above. In this section, we turn our attentions to procedures for inference and parameter estimation under Blockmodels. As a result, we will firstly calculate per-node probability of Sybil and per-pair roles through posterior inference and then obtain Beta parameter α and Bernoulli rate matrix B through parameter estimate.

3.1. Posterior Inference

The posterior inference problem is to compute the posterior distribution of the latent variables given a collection of observations. According to probabilistic graphical model show in figure 1, we can obtain the joint probability of the edges Y and latent variables in Probabilistic Block-model as show in formulation (1)

$$\mathrm{P}\big(\mathrm{Y}, \theta_{1:N}, z_p, z_q \big| \alpha, B\big) = \prod_{pq} P(Y(p,q)\big|z_p, z_q, B)P(z_p\big|\theta_p)P(z_q\big|\theta_q) \prod_N(\theta_n|\alpha) \tag{1}$$

The posterior probability can be obtained by computing the integral for latent variables in function (1).

$$\mathrm{P}(Y \mid \alpha, B) = \int_\theta \Sigma(\prod_{pq} P(Y(p,q)\big|z_p, z_q, B)P(z_p\big|\theta_p)P(z_q\big|\theta_q) \prod_N(\theta_n|\alpha))d_\theta$$

As this formulation is too complicated, a closed form solution does not exist. We used variational EM to give an approximate estimate of the parameters. We begin by bounding the log of the marginal probability of the data with Jensen's inequality,

$$\log(\mathrm{Y}|\alpha, \beta) \geq E_q\big[logp\big(Y, \theta, z_p, z_q \big| \alpha, \beta\big)\big] - E_q[logq(\theta, z_p, z_q)] \tag{2}$$

And then we assumed that latent variables q depends on a set of free parameters (as show in figure 2). We specify q as the mean-field fully-factorized family.

$$\mathrm{q}\big(\theta, z_p, z_q \big| \gamma, \varphi_p, \varphi_q\big) = \prod_p q_1(\theta_p|\gamma_p) \prod_{p,q} (q_2(z_p|\varphi_p)q_2(z_q|\varphi_q))$$

where $\mathrm{q_1}$ is a Beta, $\mathrm{q_2}$ is a Bornoulli, $\{\gamma_\mathrm{p}, \varphi_\mathrm{p}, \varphi_\mathrm{q}\}$ are the set of free variational parameters that are optimized to tighten the bound.

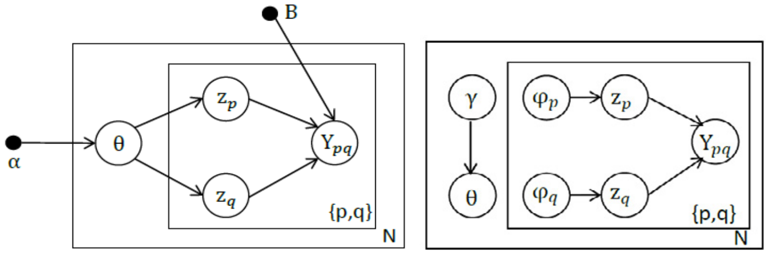

Fig. 2. (a) Graphical model representation of Probabilistic Block-model; (b) Graphical model representation of the variational distribution used to approximate the posterior in Probabilistic Block-model.

By minimizing the KL divergence between q and the true posterior, we can tighten the bound with respect to the variational parameters. And then we get the updates for variational parameters:

$$\varphi_p \propto e^{E_q[log\theta_{pg}]} \prod_h (B(g,h)^{Y(p,g)} \cdot (1 - B(g,h))^{1-Y(p,g)})^{\varphi_{p,h}}$$

$$\varphi_q \propto e^{E_q[log\theta_{ph}]} \prod_g \left(B(g,h)^{Y(p,g)} \cdot \left(1 - B(g,h)\right)^{1-Y(p,g)}\right)^{\varphi_{p,g}}$$

for g,h = 1,2. Where $E_q[\log\theta_k] = \psi(\alpha_k) - \psi(\sum_k \alpha_k)$, and $\psi(\alpha_k)$ is derivation of log-gamma function. The update for the variational Beta parameters$\gamma_{p,k}$ is

$$\gamma_{p,k} = \alpha_k + \sum_q \varphi_{p,k} + \sum_q \varphi_{q,k} \qquad \text{for all node p} = 1,2, \dots \text{n and k} = 1,2.$$

3.2. Parameter estimate

We used a virational expectation-maximization (EM) algorithm to calculate hyper-parameters $\{\alpha, \mathrm{B}\}$ of the model. Maximize the lower bound in formulation (2) we can obtain a local optimum solution of $\{\alpha, \mathrm{B}\}$.

A closed form solution is not exit; we use Newton-Raphson method to achieve an approximate solution.

$$\alpha_k^{n+1} = \alpha_k^n + L(\alpha_k^n) \,/\, L'(\alpha_k^n)$$

$$L'(\alpha_k^n) = N\big(\psi(\textstyle\sum_k \alpha_k) - \psi(\alpha_k)\big) + \sum_p (\psi\big(\gamma_{p,k}\big) - \psi(\textstyle\sum_k \gamma_{p,k}))$$

The approximate MLE of B is

$$\mathrm{B(g,h)} = \frac{\sum_{p,q} \mathrm{Y(p,q)} \cdot \varphi_{pg} \cdot \varphi_{qh}}{(1-\rho) \cdot \sum_{p,q} \cdot \varphi_{pg} \cdot \varphi_{qh}}$$

Finally, the approximate MLE of the sparsity parameter ρ is

$$\rho = \frac{\sum_{p,q}(1 - \mathrm{Y(p,q)})(\sum_{g,h} \varphi_{pg}\varphi_{qh})}{\sum_{p,q}\sum_{g,h} \varphi_{pg}\varphi_{qh}}$$

4. Experiments and Results

We presented a study of both simulated data and real-world data to validate our algorithms. The simulated data experiment is showed in part 4.1 and real-world data experiment is showed in part 4.2.

4.1. Experiments on simulated data

We generate the simulated data in the following steps:

(1) Build a graph $\mathrm{G}_1$of N_1nodes and M_1 edges, with Barab´asi–Albert graphs of N = 128 and m_0 = 8.

(2) Make a copy and graph G_1 as G_2 (with nodes $\mathrm{N}_2 = \mathrm{N}_1$ and edges $\mathrm{M}_2 = \mathrm{M}_1$)

(3) Select random edges $\mathrm{e}_1 = (\mathrm{n}_{1,i}, \mathrm{n}_{1,j})$ and $\mathrm{e}_2 = (\mathrm{n}_{2,i}, \mathrm{n}_{2,j})$ from G_1 and G_2 respectively; Delete e_1 and e_2; Add edges $(\mathrm{n}_{1,\mathrm{i}}\mathrm{n}_{2,\mathrm{i}})$ and $(\mathrm{n}_{1,\mathrm{j}}, \mathrm{n}_{2,\mathrm{j}})$ to the Graph

(4) Repeat (3) for 7 times to get G.

The simulated data shows in figure 3 (a).

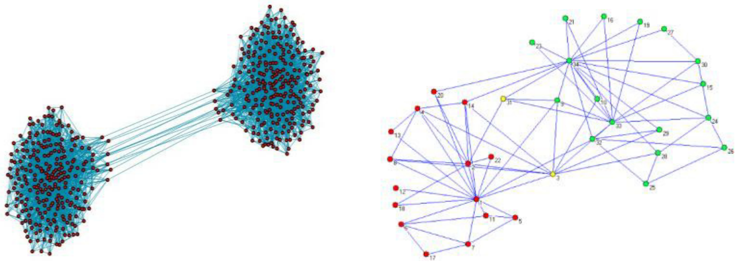

Fig. 3. (a) The synthetic network used in section 4.1 for discovering Sybil nod; (b) Node connection situation in karate dataset.

The Classification Accuracy of our algorithm on simulated data is 100%. We obtain the interaction probability matrix which shows in figure 4(a). We can see that B(1,1)=B(1,2)=0.14, which further verification our algorithm.

4.2. Experiments on real-world data

We used a real-world dataset (karate) to verify our algorithm. Karate dataset is social network of friendships between 34 members of a karate club at a US university in the 1970s which is widely used to testing algorithms. We assume one group in the karate dataset is honest users and another group is Sybil nodes. Through our algorithm we find some interesting phenomenons.

As figure 3 (b) shows, all 34 nodes can be divided into two groups. 15 nodes (left) are discovered to be Sybil nodes and two nodes (node 3 and node 31, which have equal probability to be honest node and Sybil node) are most suspicious node. We should pay more intentions to node 3 and node 31, these two nodes either the most dangerous Sybil node which earning trusts from honest users or infected honest users.

We also can obtain an interaction probability matrix which shows in figure 4(b). Honest nodes tend to connected with honest user and Sybil node tend to connected with Sybil node.

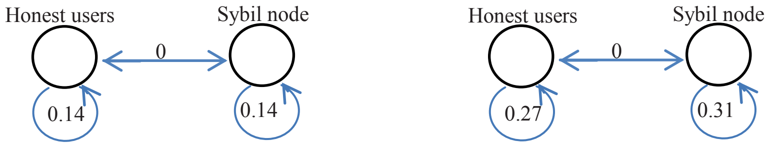

Fig. 4. (a) Estimated blockmodel in simulated data, B; (b) Estimated blockmodel in karate data, B.

Through this model, we can not only parting the Sybil nodes from honest users but also discover the network contaminated situation. If the value of parameter B(1,2)=B(2,1) is high, we can say that this network is invaded by invaders, otherwise if the value of parameter B(1,2)=B(2,1) is nearly to zero, we can say that the network is safety.

5. Related work

Avoiding Sybil attack is one of the most fundamental issues in open-membership systems. In practice, Sybil attacks have a significant influence on the accessibility of the anonymous communication and anti-censorship infrastructures. According to the different Sybil defence methods adopted in these infrastructures, research into this area can be divided to two classes: traditional Sybil attack/defence and social network-based Sybil defence. As for the social network-based Sybil defence, a variety of Sybil discovery algorithms have been proposed to isolate Sybil nodes from others. The accessibility of these infrastructures can be improved by applying these algorithms to mitigate Sybil attack.

Traditionally, the adversary can compromise multiple proxies in the anonymous communication infrastructure to perform attacks like DoS [7], misleading routing [8] [9] or ever information harvest [10] [11]. Defense methods of these attacks are highly specified to the implementation of these attacks. For example, [7] proposed selective DoS attack over Tor [22] where the adversary controls multiple relays in the Tor network. By collusion with each other, these relays can identify connections through them and selectively reject connections they cannot compromise. As a result, the chance for the adversary to compromise user anonymity is significantly increased. To evade selective DoS attack in Tor, [12] gave an algorithm to detect such attacks. In [11], the adversary adopted the idea of Sybil attack and deploys a few malicious middle relays in the Tor network. By observing incoming connections to these relays, they found many secret Tor bridges [23] which is essential to Tor's censorship-resistant design. [13] shows how the impact of this attack can be reduced by authorization of connections.

Recently, social network-based design is cooperated into the design of anonymous communication infrastructures with the purpose of avoiding Sybil attacks. In peer-to-peer anonymous communication systems, a user can only connect to a few neighbors that they already know in the real world [10] [14] [15]. Every user servers as a client as well as a proxy. User's traffic to the destination is transferred through a path in the network in a hop-by-hop fashion. Adopting social relations in peer-to-peer anonymous communication design can significantly reduce adversary's ability to crawl arbitrary honest users by joining the network multiple times or creating multiple Sybil identities [10]. For example, Freenet [14] adopts the peer-to-peer design and transfers users' data through recursive routing algorithm for the purpose of limiting users' knowledge about other nodes. It also introduced darknet mode in version 0.7, where connection between two peers is allowed only if they have prior out-of-band agreement, presumably based on mutual trust in the real life. [10] presented the concept of MCON as a communication system that hides the identities of its members from both insider and outsider attackers. In order to achieve this goal, MCON adopts peer-to-peer design and get bootstrapped based on social relationships between users. And it grows by having existing members invite new nodes based on social relationships. Social network plays an important role in the design of peer-to-peer anonymous communication systems.

In the centralized anonymous communication systems, proxies are disseminated to users by a trusted distribution center who has a full list of all online proxies in the system. The distribution center has two conflicting goals: to disseminate resources widely and to prevent adversaries from enumerating them with many Sybil identities. In order to limit attacker's ability to creating multiple valid Sybils, social relationship is incorporated into the design of proxy distribution strategies [16] [17] [18]. For example, in the well-known web-based proxy Psiphon [16], new users can make a registration and learn proxy address only through the invitation by an existing registered user. The state-of-art distribution for Tor bridges [23] rBridge is proposed in [18]. rBridge also employs an introduction-based mechanism to invite new users for the purpose of resisting Sybil attacks.

Incorporating social network in the design of anonymous communication systems also presents us an opportunity to discover Sybil identities based on the connections between nodes. Variety of social network-based Sybil defense algorithms are presented in the past few years [19] [20] [21]. They all aim to distinguish Sybil identities from honest ones. Most of them are based on a primary assumption that compared to the honest nodes, Sybil nodes tend to be poorly connected to the rest of network. For example, SybilGuard [19] assumes the honest part of the network (based on social relationships in the reality) is fast mixing. It also makes the assumption that, the adversary may create many Sybils but relatively few connections between these Sybils and honest nodes. Most recently, [24] shows the possibility of using community detection algorithms for Sybil defense.

6. Conclusion and Future Work

In this paper we proposed a generative model to solve the problem in Sybil discovery. Different from existing models, we do not discriminate which nodes are Sybil nodes directly; instead we simulate the network generation process to parting Sybil nodes from honest nodes. Furthermore, we give a matrix B to describe the interaction probability between Sybil nodes and honest nodes. The algorithm shows good performance on both synthetic and real-world data. In future work, we will try to obtain real connection relations in anonymous communication systems, such as the Free net, to further validate two-class undirected mixed member stochastic Blockmodels .

Acknowledgements

This work is supported by National Natural Science Foundation of China (Grant No.61100174, 6127250 0,61003167),National High Technology Research and Development Program of China, 863 Program (GrantNo.2011AA010701, 2011AA01A103),National Key Technology R&D Program (Grant No.2012BAH37B04) and the “Strategic Priority Research Program” of Chinese Academy of Science(Grant XDA06030200).

References

[1] J. Douceur, The sybil attack, in: P. Druschel, F. Kaashoek, A. Rowstron (Eds.), Peer-to-Peer Systems, Vol. 2429 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2002, pp. 251–260.
[2] Vipul’s razor. URL http://razor.sourceforge.net/
[3] M. Castro, P. Druschel, A. Ganesh, A. Rowstron, D. S. Wallach, Secure routing for structured peer-to-peer overlay networks, SIGOPS Oper. Syst. Rev. 36 (SI) (2002) 299–314.
[4] Q. Lian, Z. Zhang, M. Yang, B. Zhao, Y. Dai, X. Li, An empirical study of collusion behavior in the maze p2p file-sharing system, in: Distributed Computing Systems, 2007. ICDCS ’07. 27th International Conference on, 2007, p. 56.
[5] M. Yang, Z. Zhang, X. Li, Y. Dai, An empirical study of free-riding behavior in the maze p2p file-sharing system, in: M. Castro, R. Renesse (Eds.), Peer-to-Peer Systems IV, Vol. 3640 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2005, pp. 182–192.
[6] emule-project.net. URL http://www.emule-project.net/
[7] N. Borisov, G. Danezis, P. Mittal, P. Tabriz, Denial of service or denial of security?, in: Proceedings of the 14th ACM conference on Computer and communications security, CCS ’07, ACM, New York, NY, USA, 2007, pp. 92–102.
[8] M. K. Reiter, A. D. Rubin, Crowds: anonymity for web transactions, ACM Trans. Inf. Syst. Secur. 1 (1) (1998) 66–92.
[9] M. Rennhard, B. Plattner, Introducing morphmix: peer-to-peer based anonymous internet usage with collusion detection, in: Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society, WPES ’02, ACM, New York, NY, USA, 2002, pp. 91–102.
[10] E. Vasserman, R. Jansen, J. Tyra, N. Hopper, Y. Kim, Membership-concealing overlay networks, in: Proceedings of the 16th ACM conference on Computer and communications security, ACM, New York, USA, 2009, pp. 390–399.
[11] Z. Ling, J. Luo, W. Yu, M. Yang, X. Fu, Extensive analysis and large-scale empirical evaluation of tor bridge discovery, in: INFOCOM, 2012 Proceedings IEEE, 2012, pp. 2381 –2389.
[12] N. Danner, D. Krizanc, M. Liberatore, Detecting denial of service attacks in tor, in: R. Dingledine, P. Golle (Eds.), Financial Cryptography and Data Security, Vol. 5628 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 273–284.
[13] R. Smits, D. Jain, S. Pidcock, I. Goldberg, U. Hengartner, Bridgespa: improving tor bridges with single packet authorization, in: Proceedings of the 10th annual ACM workshop on Privacy in the electronic society, WPES ’11, ACM, New York, NY, USA, 2011, pp. 93–102. doi:10.1145/2046556.2046569. URL http://doi.acm.org/10.1145/2046556.2046569
[14] I. Clarke, O. Sandberg, B. Wiley, T. Hong, Freenet: A distributed anonymous information storage and retrieval system, in: H. Federrath (Ed.), Designing Privacy Enhancing Technologies, Vol. 2009 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2001, pp. 46–66.
[15] Y. Sovran, A. Libonati, J. Li, Pass it on: social networks stymie censors, in: Proceedings of the 7th international conference on Peer-topeer systems, IPTPS’08, USENIX Association, Berkeley, CA, USA, 2008, pp. 3–3.
[16] Psiphon design overview 1.0 (2009). URL http://psiphon.ca/documents/Psiphon Design Overview 1 0.pdf
[17] D. McCoy, J. A. Morales, K. Levchenko, Proximax: measurement-driven proxy dissemination (short paper), in: Proceedings of the 15thinternational conference on Financial Cryptography and Data Security, FC’11, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 260–267.
[18] Q. Wang, Z. Lin, N. Borisov, N. Hopper, rbridge: User reputation based tor bridge distribution with privacy preservation, NDSS’13

(toappear), 2013.
[19] H. Yu, M. Kaminsky, P. B. Gibbons, A. D. Flaxman, Sybilguard: defending against sybil attacks via social networks, IEEE/ACM Trans.Netw. 16 (3) (2008) 576–589.
[20] G. Danezis, P. Mittal, SybilInfer: Detecting Sybil Nodes using Social Networks, in: NDSS, 2009.
[21] H. Yu, P. B. Gibbons, M. Kaminsky, F. Xiao, Sybillimit: a near-optimal social network defense against sybil attacks, IEEE/ACM Trans.Netw. 18 (3) (2010) 885–898.
[22] R. Dingledine, N. Mathewson, P. Syverson, Tor: the second-generation onion router, in: Proceedings of the 13th conference on USENIXSecurity Symposium - Volume 13, USENIX Association, 2004, pp. 21–21.
[23] R. Dingledine, N. Mathewson, Design of a blocking-resistant anonymity system, Tech. rep. (2006).
[24] B. Viswanath, A. Post, K. P. Gummadi, A. Mislove, An analysis of social network-based sybil defenses, in: Proceedings of the ACM SIGCOMM 2010 conference, SIGCOMM ’10, ACM, New York, NY, USA, 2010, pp. 363–374.
[25] E. Airoldi,D.Blei,S.Fienberg,E.Xing. Mixed Membership Stochastic Blockmodels. JMLR,2008,pp.1981-2014
[26]D. Blei, A. Ng, M. Jordan. Latent Dirichlet Allocation. JMLR, 2003, pp.993-1022
[27]P. Zhang, J. Li, P. Wang, B. Gao,X. Zhu, L. Guo. Enabling Fast Prediction for Ensemble Models on Data Streams. KDD,2011
[28]P. Zhang, B. Gao, P. Liu, Y. Shi, L. Guo. A Fromework for Application-Driven Classification of Data Streams. Neurocomputing, 2012
[29]X. Zhu, P. Zhang, X. Lin, Y. Shi, Active Learning from Stream Data Using Optimal Weight Classifier Ensemble. IEEE Transactions on System, Man, Cybernetics, Part B, Vol. 40 (6), 2010, pp.1607- 1621.