The Sybil Attack - Theory and Practice

Kelong Cong
Delft University of Technology
k.cong@student.tudelft.nl

ABSTRACT

TODO

1. INTRODUCTION

Electronic commerce and online social networks are common phenomenons at the present time. They allow us to orchestrate many aspects of our lives in the comfort of our homes, behind the monitors of our devices. An online identity is often required to use such services, for examples we must create an account to Tweet[1] our friend, who must also have an account. In this scenario, users can choose to remain psudonymous if they are careful, where their real-life identity is uncorrelated with their online identity.

While creating pseudonyms is useful for protecting users' privacy, it also opens an alleyway for attackers. The Sybil attack, first described by Douceur[22], is an attack where an entity can assume multiple identities or Sybils, and then attack either another entity or undermine the whole system. For example, a malicious Twitter user can create many fake identities and have the fake identities follow his real identity, thus creating a false reputation. It is one of the most important attacks because it leads to a large number of consequences including but not limited to spreading false information, identity theift[8] and ballot stuffing[7]. Furthermore, to the best of our knowledge, there is no general solution for preventing the Sybil attack.

In this work, we survey various aspects of the Sybil attack. But in contrast with previous surveys, we include both the theoretical and practical aspects. First, we describe the Sybil attack in more detail and and illustrate its importance by looking at how researchers and black-hat hackers mounted the attack on real-world e-commerce and online social network systems in section 2. Since there is a large variety of Sybil attack defence mechanisms, from using trusted-third-party to exploiting the graph characteristics in online social networks, thus we classify these mechanisms by their "main idea" in section 3. Finally we present the related work and conclude in section 4 and section 5.

[1]A message sent using Twitter is a Tweet.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2016 ACM. ISBN 123-4567-24-567/08/06.

DOI: 10.475/123_4

2. THE SYBIL ATTACK

The Sybil attack is coined by Douceur[22] in 2002 in the context of peer-to-peer systems. In this section, we first introduce the Sybil attack using Douceur's original definition and outline the key (discouraging) theoretical results. Next, we review practical attacks in three types of systems (1) MANET (mobile ad-hoc networks) such as sensor networks, (2) reputation systems such as PageRank[58] but also include e-commerce systems such as eBay and (3) OSN (online social networks) such as Twitter and Facebook. We hope our review illuminates the alarming consequences of the Sybil attack.

2.1 Theoretical Results

Douceur defined the Sybil attack as forging multiple identities under the same entity[22]. An entity can be for example a physical user of the system and identities are how entities present themselves to the system. Thus a local entity has no direct knowledge of remote entities, only their identities. The forged identities do not necessarily follow the protocol specified by the underlying network, they may deviate arbitrarily from the protocol, i.e. they assume the characteristics of Byzantine fault[43]. We use these terms in the remainder of the survey.

The author modelled the system as a general distributed computing environment where there is no constraint on the topology, every node has limited computational resources and messages are guaranteed to be delivered. Under this model, the author proved that the Sybil attack is always possible without a central, trusted authority.

Cheng and Friedman proved an important result regarding the Sybil attack in reputation systems[11]. Reputation systems are commonly used in MANET, e-commerce and the internet in general, where entities are rewarded by their good behaviour and penalised otherwise. Google's PageRank[58] is an example of a reputation system, where a large number of links to a website makes it more reputable. Cheng and Friedman classified reputation systems into two categories,

1. symmetric reputation systems where the reputation score only depends on the network topology, popular reputation mechanisms such as PageRank[58] and EigenTrust[39] are examples of symmetric reputation systems, and

2. asymmetric reputation systems where there some nodes

are trusted and reputation scores are propagated through the trusted nodes, most OSN are examples of asymmetric reputation systems.

The authors formally proved that symmetric reputation systems are vulnerable to the Sybil attack. But in the asymmetric case, it is possible to construct a Sybil-proof reputation system.

2.2 The Sybil Attack in P2P FIle Sharing Networks

P2P (peer-to-peer) file sharing networks are distributed computer networks that are built for discovering and sharing files. BitTorrent[13] is likely the most popular P2P network at the time of writing. Due to their open and distributed nature, they are vulnerable to the Sybil attack.

2.2.1 *Denial of Service*

By exploiting vulnerabilities in the BitTorrent network, denial of service attack can be directed at any machine connected to the internet, not just machines in the network[72]. The main idea is to report the victim as the tracker (a server that coordinates the peers). El Defrawy, Gjoka and Markopoulou created a small scale proof-of-concept attack. Using only one machine, they could generate enough traffic to cripple small organisations and home users. The authors suggested that if Sybils are created to perform the same attack aimed at a single victim, then it could easily throttle links with much higher bandwidth[23].

2.2.2 *Index Poisoning*

P2P networks often implement a DHT (distributed hash table). The DHT in BitTorrent is called Mainline-DHT, based on Kademlia[53]. Keys are the infohashes (file identifiers) and values are the metadata of the files, these are distributed across all the participating peers. Every node stores a routing table and requests are routed iteratively to the node responsible for a particular key[48]. The goal of index poisoning is to corrupt routing table so that honest peers fail to find the values they want. It can be mounted by injecting Sybils into the DHT that do not follow the protocol. Wang and Kangasharju created honeypots in the BitTorrent network and detected as many as 300,000 Sybils[88]. Similar attacks are possible in other P2P networks such as Overnet[47].

2.3 The Sybil Attack in Online Social Networks

OSN (online social networks) are vulnerable to the Sybil attack even when most of them use a central, trusted authority such as Facebook. In OSN, users create profiles and form relationships with friends. In contrast with real world relationships, it is much easier to create relationships in OSN even with strangers. In 2008, Sophos conducted an experiment where they created a Facebook profile and send friend requests to 200 random users, and 41% of the users accepted the friend request[75]. A report by Facebook at the end of 2011 stated 5-6% of their accounts are fake[59]. Combining with the ability to create new identities with very little cost, it is possible to perform many types of attacks which we outline below.

2.3.1 *Identity Theft*

Authors of [8] created two attacks - profile cloning and cross-site profile cloning, targeting five social network sites including Facebook and LinkedIn. The iCloner system was created to automate these attacks.

In profile cloning, iCloner uses publicly available information to automatically create clones of the victim's profiles, effectively creating Sybils. iCloner then sent friend requests from the cloned profile to the friends of the victim. The fact that the victim may have many friends that they do not contact very often, e.g. friend from primary school living in another country, makes this attack highly effective. The authors found that the acceptance rate for cloned profiles was over 60%. Much higher than the acceptance rate of 30% for fictitious profiles. Once the friendship is established, it is possible to extract private information that is not available publicly and perform identity theft.

The idea of cross-site profile cloning is similar, except the cloned profile is created on another social network site that the victim does not yet use. Once the cloned profile is created, iCloner attempts to identify friends of the victim and begins sending friend requests. Similarly, 56% of the friend requests were accepted.

A more recent study created SbN (Socialbot Network) targeting Facebook[9]. Each socialbot is a Sybil created by the attack, it controls a forged profile and minic human behaviour to avoid detection. The attacker is the botmaster who coordinates the socialbots to achieve a common objective such as infiltrating the target OSN by creating friend relationships with real users. The authors found that infiltration success rate was as high as 80% and the FIS[77] (Facebook Immune System) was not sufficient to prevent the attack. Once the relationships are established, the botmaster can command the socialbots to start gathering private information which can then be used for identity theft.

These examples demonstrate that the carelessness of users and the ability to create Sybils makes OSN vulnerable to identity theft. Moreover, identity theft is only an entry point. Once trust relationships are established, the attacker can perform many other types of attacks such as spamming, phishing or astroturfing to gain advantage.

2.3.2 *Astroturf*

Astroturfing is an act of creating grassroot movement that are in reality carried out by a single entity, effectively spreading misinformation to legitimate users. It relies on the ability to create Sybils in the underlying social network. This type of attack is especially effective in social networks such as Twitter where a lot of the social interaction such as sending messages happen in the public.

In the 2010 Massachusetts senate race, Mustafaraj and Metaxas found evidence that Republican campaingners created fake Twitter accounts and used them to send spam. The spam caused Google real-time search results to tip in their favour thus causing a spread of misinformation[56]. Ratkiewicz et el. suggest that this type of attack can be mounted cheaply and may have a larger influence than traditional adversiting[61].

The Truthy system[61] is a web service that perform real-time analysis of Twitter to detect political astroturfing. In the 2010 U.S. midterm election, the authors found accounts which generated a lot of retweets but no original tweets. More importantly, they uncovered a network of bot accounts that injected thousands of tweets to smear the Democratic candidate.

In 2012, Wang et el. investigated two of the largeset crowdturfing[2] platforms in China that brings together buyers and sellers - Zhubajie and Sandaha. One of their services is perform astroturfing on Weibo (The Chinese Twitter). The authors found that the 5364 sellers collectively own 14151 Weibo accounts and the top 1% of the sellers own over 100 accounts. Furthermore, the business is growing and more than $4 million have been spent on these two platforms over five years[87].

2.3.3 Spam

[78] [27] [92]

2.4 The Sybil Attack in Reputation Systems

Reputation systems cultivate collaborative behaviour by allowing entities to trust each other based on community feedback, usually in the form of a reputation score. Entities decide whom to trust based on the reputation scores, thus entities are also incentivised to behave honestly. Reputation systems are found in many context. In e-commerce, namely eBay, researchers found that the merchant's reputation "is a statistically and economically significant determinant of auction prices"[34], and "buyers are willing to pay 8.1% more" for goods sold buy a reputable merchant[64]. The file sharing peer-to-peer network BitTorrent uses tit-for-tat as an ephemeral reputation system to encourage peers to upload in exchange for better download speeds[12]. The aforementioned PageRank[58] is also a reputation system, used for ranking reputable websites higher in Google's search results. Some OSN can be considered as reputation systems too, Twitter accounts with a lot of followers may be more reputable than accounts with fewer followers.

Unfortunately, reputation systems are also vulnerable to the Sybil attack. Worryingly, there appears to an industry built around it, and their products are easily accessible in the clearnet. In this section, we describe practical attacks on reputation systems.

2.4.1 Self-promoting

In self-promotion, the goal of the attacker is to illegitimately raise its own reputation. A common way to perform self-promition is to create Sybils and have them create positive reputation for the attacker's main identity.

Dini and Spagnolo studied the economics of buying reputation on eBay. The authors discovered many cheap items (around €0.7) for sell are simply there to boost feedback. For example, one of the item is titled "Apple Cranberry Crisp Recipe + 100% Positive Feedback". The authors successfully boosted their feedback by purchasing such items. But they made an unsuccessful attempted to place a bid on their own good with a fake account[21].

De Cristofaro et el. performed an emperical study on Facebook page promotion using like farms[17]. Some of the farms such as `SocialFormulae.com` are clearly operated by bots and the operator does not attempt to hide it, others such as `BlostLikes.com` tries to minic human users. The authors purchased the "1000 likes" service on their empty Facebook pages. In under a month, many empty pages have accumulated almost 1000 likes as promised by the like farms. The authors empty accounts were not terminated. Only a small number of the liker's account were terminated.

SEOClerks and MyCheapJobs are also evidences of marketplaces for self-promotion. Some of the top services include "1 million Twitter followers" at $849, "1000+ Instagram followers" at $10 and so on. The revenues of those two marketplaces are estimated to be at $1.3 million and $116 thousand, respectively[24]. Although the authors did not investigate the properties of the fake followers, there is little doubt that many of accounts used in these services are Sybils.

[2]Crowdsourced astroturfing.

2.4.2 Slandering

The goal of a slandering attack is to illegitimately produce negative feedback to undermine the reputation of the target. It is easy to imagine the improvement in effectiveness when using multiple Sybils. From the best of our knowledge, there are no published studies on real-world slandering. But research has shown having a negative feedback may harm the target's ability to do business[5].

2.4.3 Whitewashing

In whitewashing, attackers abuse the reputation system for temporary gain and then escape the consequences by joining the reputation system under a new identity to shed their bad reputation. Clearly, whitewashing is only possible when the Sybil attack is possible. Again, there are no studies on whitewashing in the real-world. But many have suggested that it is feasible attack[33, 52].

2.4.4 Denial of Service

The denial of service attack highly depends on the structure of the reputation system.

2.5 The Sybil Attack in MANET

2.6 TODO

a test bed for sybil attacks[35]
Quantifying Sybil attack[50]

3. DEFENCES

In this section we categorise various defence techniques against the sybil-attack in reputation systems.

3.1 Trusted Third Party

One of the earliest and best known reputation system is eBay[63]. The buyers and sellers rely on a trusted third party, in this case eBay, to gather and distribute feedbacks after every transaction. Even when there are no incentives to provide feedback, Resnick and Zeckhauser observed that feedback was provided more than half of the time[63], making eBay one of the most well-known online marketplaces.

In general, trusted third parties manage the issurance and verification of identities. Thus they can apply a fee on the peer for creating a new identity[62] or rate-limit the creation of new identities[22], making sybil-attacks more difficult. Furthermore, trusted third parties often have the ability to manipulate the identities. For example they could punish the attackers by disabling all of their identity when caught, making the sybil-attack much riskier especially when identities are costly.

Trusted third party is likely the most widely used technique in practice. Marketplaces such as Amazon or eBay, online forums such as Stackoverflow or Reddit, all use a form of trusted third party.

Unfortunately, a trusted third party is often a single point of failure. Moreover, being a centralised system, it is difficult to scale up to suit increasing user demands. In the remainder of this section, we focus on distributed techniques for preventing the sybil-attack.

Credence 06[86] - uses central authority to sign key

3.2 Costly Identity Creation

3.2.1 IP Address

3.2.2 Low reputation for new users

Feldman 04[25] - adaptive stranger, low score on entry

3.3 Indirect Information

EigenTrust[39] - doesn't prevent sybils, suggests to add cost in ID creation R2Trust[81] - credibility, tackles colluders, time decay factor

3.4 Graph Techniques

OSN (online social networks) such as Facebook can be viewed as a graph, where the nodes on the graph are identities created by users and the edges represent trust relationships.

Theory[70] Gal-Oz et al. [26] communities are collection of knots, sybils can form a knot? Regret[65, 66] - information from multiple dimensions Guha 04[28] - no mention of sybil attacks or attacks in general

3.4.1 Flow Based

BarterCast[54] SybilRes[18]

3.4.2 Topology

SybilGuard[95] SybilLimit[94] SybilInfer[16] SybilShield[71] - assuming sybils have bad connectedness SumUp[82] GateKeeper[83] - based on SumUp Social-network[84] - community detection

Distributed Sparse Cut Monitoring[42]

Other systems are built on top: ReDS[2] suggests to use sybilimit or sybilinfer SybilProof-DHT[45]

3.5 Reputation Transfer

Trust-transfer[68]

3.6 Self Registration

P-GRID 01[1] Self-registration[20] - distributed registration based on IP address

3.7 Cryptography Based Techniques?

Secure-Overlay[49] - ID crypto and SSS Privacy-preserving[67] - blockchain? Proof-of-stake[19] SybilConf[79]

3.8 Content Driven

[10]

3.9 Other

Parental control[80] - uses parents to "observe" find suspects, only for detection, requires a sybil-proof reputation scheme DSybil[96] - recommendation system, need historical data Symon[38] - pair peers together, likelihood for both to be sybils is low, the pair monitor each other to prevent attacks XRep 02[15] IP check, and checks digest, uses existing P2P systems like Gnutella

3.10 Unsorted?

Beth and PGP limits Sybil attack to some extent by using social graphs Beth 94[6] PGP (Zimmermann) 95[100]

Yu 00[93] Lee 03[44] - uses flooding, might not be scalable, only talks about DoS Marti 04[51] ARA 05[31] - no mention of sybil, prevents freeriding, prevents short-term abuse because reputation increases gradually FuzzyTrust Song 05[74] - uses fuzzy logic P2PRep/Fuzzy 06[4] - also fuzzy, does not prevent generation of false rumors Xiong 05[90] - no mention of sybil, but tries to mitigate false information PowerTrust 06[99] - uses "power nodes" (from power-law), no mention of sybil, some defence against colluders

Histos and Sopras[97], doesn't really have structure? Beta[36] Gupta et al.[30]

PeerTrust[89] - DHT, used P-GRID source code, has credibility rating

PerContRep[91]

3.11 Does not handle Sybil-attack?

TrustMe[73] is a reputation that focuses on anonymity, no mention of sybil attack

H-Trust[98] does not mention sybil

Coner et al.[14] assumes clients cannot perform sybil attack

TrustGuard 05[76] - assumes it is built on secure overlay networks (sybil-proof networks)

Scrivener 05[57] - assumes ID cannot be created and discarded

4. RELATED WORK

Reputation Surveys: [52] [37] ? [33] [41] [69] ? [32]

Sybil Surveys: [46] [55] [60] [29] [40] Sok[3] but also some contribution

Other: [85]

5. SUMMARY

6. REFERENCES

[1] K. Aberer. P-Grid: A self-organizing access structure for P2P information systems. In *International Conference on Cooperative Information Systems*, pages 179–194. Springer, 2001.

[2] R. Akavipat, M. N. Al-Ameen, A. Kapadia, Z. Rahman, R. Schlegel, and M. Wright. ReDS: A framework for reputation-enhanced DHTs. *IEEE Transactions on Parallel and Distributed Systems*, 25(2):321–331, 2014.

[3] L. Alvisi, A. Clement, A. Epasto, S. Lattanzi, and A. Panconesi. Sok: The evolution of sybil defense via social networks. In *Security and Privacy (SP), 2013 IEEE Symposium on*, pages 382–396. IEEE, 2013.

[4] R. Aringhieri, E. Damiani, D. Vimercati, S. De Capitani, S. Paraboschi, and P. Samarati. Fuzzy techniques for trust and reputation management in anonymous peer-to-peer systems. *Journal of the American Society for Information Science and Technology*, 57(4):528–537, 2006.

[5] S. Ba and P. A. Pavlou. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS quarterly*, pages 243–268, 2002.

[6] T. Beth, M. Borcherding, and B. Klein. Valuation of trust in open networks. In *European Symposium on Research in Computer Security*, pages 1–18. Springer, 1994.

[7] R. Bhattacharjee and A. Goel. Avoiding ballot stuffing in ebay-like reputation systems. In *Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems*, pages 133–137. ACM, 2005.

[8] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda. All your contacts are belong to us: automated identity theft attacks on social networks. In *Proceedings of the 18th international conference on World wide web*, pages 551–560. ACM, 2009.

[9] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu. The socialbot network: when bots socialize for fame and money. In *Proceedings of the 27th Annual Computer Security Applications Conference*, pages 93–102. ACM, 2011.

[10] K. Chatterjee, L. de Alfaro, and I. Pye. Robust content-driven reputation. In *Proceedings of the 1st ACM workshop on Workshop on AISec*, pages 33–42. ACM, 2008.

[11] A. Cheng and E. Friedman. Sybilproof reputation mechanisms. In *Proceedings of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems*, pages 128–132. ACM, 2005.

[12] B. Cohen. Incentives build robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer systems*, volume 6, pages 68–72, 2003.

[13] B. Cohen. Bep 3: The bittorrent protocol specification, Jan 2008.

[14] W. Conner, A. Iyengar, T. Mikalsen, I. Rouvellou, and K. Nahrstedt. A trust management framework for service-oriented environments. In *Proceedings of the 18th international conference on World wide web*, pages 891–900. ACM, 2009.

[15] E. Damiani, D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante. A reputation-based approach for choosing reliable resources in peer-to-peer networks. In *Proceedings of the 9th ACM conference on Computer and communications security*, pages 207–216. ACM, 2002.

[16] G. Danezis and P. Mittal. SybilInfer: Detecting Sybil Nodes using Social Networks. In *NDSS*. San Diego, CA, 2009.

[17] E. De Cristofaro, A. Friedman, G. Jourjon, M. A. Kaafar, and M. Z. Shafiq. Paying for likes?: Understanding facebook like fraud using honeypots. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, pages 129–136. ACM, 2014.

[18] R. Delaviz, N. Andrade, J. A. Pouwelse, and D. H. Epema. SybilRes: A sybil-resilient flow-based decentralized reputation mechanism. In *Distributed Computing Systems (ICDCS), 2012 IEEE 32nd International Conference on*, pages 203–213. IEEE, 2012.

[19] R. Dennis and G. Owenson. Rep on the Roll: A Peer to Peer Reputation System Based on a Rolling Blockchain. 2016.

[20] J. Dinger and H. Hartenstein. Defending the sybil attack in p2p networks: Taxonomy, challenges, and a proposal for self-registration. In *First International Conference on Availability, Reliability and Security (ARES'06)*, pages 8–pp. IEEE, 2006.

[21] F. Dini and G. Spagnolo. Buying reputation on eBay: Do recent changes help? *International Journal of Electronic Business*, 7(6):581–598, 2009.

[22] J. R. Douceur. The sybil attack. In *International Workshop on Peer-to-Peer Systems*, pages 251–260. Springer, 2002.

[23] K. El Defrawy, M. Gjoka, and A. Markopoulou. BotTorrent: Misusing BitTorrent to Launch DDoS Attacks. *SRUTI*, 7:1–6, 2007.

[24] S. Farooqi, M. Ikram, G. Irfan, E. De Cristofaro, A. Friedman, G. Jourjon, M. A. Kaafar, M. Z. Shafiq, and F. Zaffar. Characterizing Seller-Driven Black-Hat Marketplaces. *arXiv preprint arXiv:1505.01637*, 2015.

[25] M. Feldman, K. Lai, I. Stoica, and J. Chuang. Robust incentive techniques for peer-to-peer networks. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 102–111. ACM, 2004.

[26] N. Gal-Oz, E. Gudes, and D. Hendler. A robust and knot-aware trust-based reputation model. In *IFIP International Conference on Trust Management*, pages 167–182. Springer, 2008.

[27] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 35–47. ACM, 2010.

[28] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web*, pages 403–412. ACM, 2004.

[29] R. Gunturu. Survey of Sybil attacks in social networks. *arXiv preprint arXiv:1504.05522*, 2015.

[30] M. Gupta, P. Judge, and M. Ammar. A reputation system for peer-to-peer networks. In *Proceedings of the 13th international workshop on Network and operating systems support for digital audio and video*, pages 144–152. ACM, 2003.

[31] M. Ham and G. Agha. ARA: A robust audit to prevent free-riding in P2P networks. In *Fifth IEEE International Conference on Peer-to-Peer Computing (P2P'05)*, pages 125–132. IEEE, 2005.

[32] F. Hendrikx, K. Bubendorfer, and R. Chard. Reputation systems: A survey and taxonomy. *Journal of Parallel and Distributed Computing*, 75:184–197, 2015.

[33] K. Hoffman, D. Zage, and C. Nita-Rotaru. A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys (CSUR)*, 42(1):1, 2009.

[34] D. Houser and J. Wooders. Reputation in auctions: Theory, and evidence from eBay. *Journal of Economics & Management Strategy*, 15(2):353–369, 2006.

[35] A. A. Irissappane, S. Jiang, and J. Zhang. Towards a comprehensive testbed to evaluate the robustness of reputation systems against unfair rating attack. In *UMAP Workshops*, volume 12, 2012.

[36] A. Jøsang and R. Ismail. The beta reputation system. In *Proceedings of the 15th bled electronic commerce conference*, volume 5, pages 2502–2511, 2002.

[37] A. Jøsang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision support systems*, 43(2):618–644, 2007.

[38] B. Jyothi and J. Dharanipragada. Symon: Defending large structured p2p systems against sybil attack. In *2009 IEEE Ninth International Conference on Peer-to-Peer Computing*, pages 21–30. IEEE, 2009.

[39] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th international conference on World Wide Web*, pages 640–651. ACM, 2003.

[40] D. Koll, J. Li, J. Stein, and X. Fu. On the state of OSN-based Sybil defenses. In *Networking Conference, 2014 IFIP*, pages 1–9. IEEE, 2014.

[41] E. Koutrouli and A. Tsalgatidou. Taxonomy of attacks and defense mechanisms in P2P reputation systems-Lessons for reputation system designers. *Computer Science Review*, 6(2):47–70, 2012.

[42] A. Kurve and G. Kesidis. Sybil detection via distributed sparse cut monitoring. In *2011 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2011.

[43] L. Lamport, R. Shostak, and M. Pease. The byzantine generals problem. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 4(3):382–401, 1982.

[44] S. Lee, R. Sherwood, and B. Bhattacharjee. Cooperative peer groups in NICE. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 2, pages 1272–1282. IEEE, 2003.

[45] C. Lesniewski-Lass and M. F. Kaashoek. Whanau: A sybil-proof distributed hash table. NSDI, 2010.

[46] B. N. Levine, C. Shields, and N. B. Margolin. A survey of solutions to the sybil attack. *University of Massachusetts Amherst, Amherst, MA*, 7, 2006.

[47] J. Liang, N. Naoumov, and K. W. Ross. The Index Poisoning Attack in P2P File Sharing Systems. In *INFOCOM*, pages 1–12. Citeseer, 2006.

[48] A. Loewenstern and A. Norberg. Bep 5: Dht protocol, Jan 2008.

[49] E. K. Lua. Securing peer-to-peer overlay networks from sybil attack. In *Communications and Information Technologies, 2007. ISCIT'07. International Symposium on*, pages 1213–1218. IEEE, 2007.

[50] N. B. Margolin and B. N. Levine. Quantifying resistance to the sybil attack. In *International Conference on Financial Cryptography and Data Security*, pages 1–15. Springer, 2008.

[51] S. Marti and H. Garcia-Molina. Limited reputation sharing in P2P systems. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 91–101. ACM, 2004.

[52] S. Marti and H. Garcia-Molina. Taxonomy of trust: Categorizing P2P reputation systems. *Computer Networks*, 50(4):472–484, 2006.

[53] P. Maymounkov and D. Mazieres. Kademlia: A peer-to-peer information system based on the xor metric. In *International Workshop on Peer-to-Peer Systems*, pages 53–65. Springer, 2002.

[54] M. Meulpolder, J. A. Pouwelse, D. H. Epema, and H. J. Sips. Bartercast: A practical approach to prevent lazy freeriding in p2p networks. In *Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on*, pages 1–8. IEEE, 2009.

[55] A. Mohaisen and J. Kim. The Sybil attacks and defenses: a survey. *arXiv preprint arXiv:1312.6349*, 2013.

[56] E. Mustafaraj and P. T. Metaxas. From obscurity to prominence in minutes: Political speech and real-time search. 2010.

[57] A. Nandi, T.-W. J. Ngan, A. Singh, P. Druschel, and D. S. Wallach. Scrivener: Providing incentives in cooperative content distribution systems. In *Proceedings of the ACM/IFIP/USENIX 2005 International Conference on Middleware*, pages 270–291. Springer-Verlag New York, Inc., 2005.

[58] L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: bringing order to the web. 1999.

[59] E. Protalinski. Facebook: 5-6% of accounts are fake, 2012. Accessed 31-10-2016.

[60] G. Rakesh, S. Rangaswamy, V. Hegde, and G. Shoba. A survey of techniques to defend against sybil attacks in social networks. *International Journal of Advanced Research in Computer and Communication Engineering*, 3(5), 2014.

[61] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer. Truthy: mapping the spread of astroturf in microblog streams. In *Proceedings of the 20th international conference companion on World wide web*, pages 249–252. ACM, 2011.

[62] P. Resnick et al. The social cost of cheap pseudonyms. *Journal of Economics & Management Strategy*, 10(2):173–199, 2001.

[63] P. Resnick and R. Zeckhauser. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. *The Economics of the Internet and E-commerce*, 11(2):23–25, 2002.

[64] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on eBay: A controlled experiment. *Experimental economics*, 9(2):79–101, 2006.

[65] J. Sabater and C. Sierra. REGRET: reputation in gregarious societies. In *Proceedings of the fifth international conference on Autonomous agents*, pages 194–195. ACM, 2001.

[66] J. Sabater and C. Sierra. Social regret, a reputation model based on social relations. *ACM SIGecom Exchanges*, 3(1):44–56, 2002.

[67] A. Schaub, R. Bazin, O. Hasan, and L. Brunie. A trustless privacy-preserving reputation system. In *IFIP International Information Security and Privacy Conference*, pages 398–411. Springer, 2016.

[68] J.-M. Seigneur, A. Gray, and C. D. Jensen. Trust transfer: Encouraging self-recommendations without sybil attack. In *International Conference on Trust*

Management, pages 321–337. Springer, 2005.
[69] C. Selvaraj and S. Anand. A survey on security issues of reputation management systems for peer-to-peer networks. *Computer Science Review*, 6(4):145–160, 2012.
[70] S. Seuken and D. C. Parkes. On the Sybil-proofness of accounting mechanisms. 2011.
[71] L. Shi, S. Yu, W. Lou, and Y. T. Hou. Sybilshield: An agent-aided social network-based sybil defense among multiple communities. In *INFOCOM, 2013 Proceedings IEEE*, pages 1034–1042. IEEE, 2013.
[72] K. C. Sia. DDoS vulnerability analysis of BitTorrent protocol. *UCLA: Techical Report*, 2006.
[73] A. Singh and L. Liu. TrustMe: anonymous management of trust relationships in decentralized P2P systems. In *Peer-to-Peer Computing, 2003.(P2P 2003). Proceedings. Third International Conference on*, pages 142–149. IEEE, 2003.
[74] S. Song, K. Hwang, R. Zhou, and Y.-K. Kwok. Trusted P2P transactions with fuzzy reputation aggregation. *IEEE Internet computing*, 9(6):24–34, 2005.
[75] Sophos. Sophos facebook id probe shows 41% of users happy to reveal all to potential identity thieves, 2007. Accessed 30-10-2016.
[76] M. Srivatsa, L. Xiong, and L. Liu. TrustGuard: countering vulnerabilities in reputation management for decentralized overlay networks. In *Proceedings of the 14th international conference on World Wide Web*, pages 422–431. ACM, 2005.
[77] T. Stein, E. Chen, and K. Mangla. Facebook immune system. In *Proceedings of the 4th Workshop on Social Network Systems*, page 8. ACM, 2011.
[78] G. Stringhini, C. Kruegel, and G. Vigna. Detecting spammers on social networks. In *Proceedings of the 26th Annual Computer Security Applications Conference*, pages 1–9. ACM, 2010.
[79] F. Tegeler and X. Fu. SybilConf: computational puzzles for confining sybil attacks. In *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pages 1–2. IEEE, 2010.
[80] A. Tehale, A. Sadafule, S. Shirsat, R. Jadhav, S. Umbarje, and S. Shingade. Parental Control algorithm for Sybil detection in distributed P2P networks. *International Journal of Scientific and Research Publications*, 2(5), 2012.
[81] C. Tian and B. Yang. H^2 Trust, a reputation and risk based trust management framework for large-scale, fully decentralized overlay networks. *Future Generation Computer Systems*, 27(8):1135–1141, 2011.
[82] D. N. Tran, B. Min, J. Li, and L. Subramanian. Sybil-Resilient Online Content Voting. In *NSDI*, volume 9, pages 15–28, 2009.
[83] N. Tran, J. Li, L. Subramanian, and S. S. Chow. Optimal sybil-resilient node admission control. In *INFOCOM, 2011 Proceedings IEEE*, pages 3218–3226. IEEE, 2011.
[84] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove. An analysis of social network-based sybil defenses. *ACM SIGCOMM Computer Communication Review*, 40(4):363–374, 2010.
[85] D. S. Wallach. A survey of peer-to-peer security issues. In *Software Security-Theories and Systems*, pages 42–57. Springer, 2003.
[86] K. Walsh and E. G. Sirer. Experience with an object reputation system for peer-to-peer filesharing. In *USENIX NSDI*, volume 6, 2006.
[87] G. Wang, C. Wilson, X. Zhao, Y. Zhu, M. Mohanlal, H. Zheng, and B. Y. Zhao. Serf and turf: crowdturfing for fun and profit. In *Proceedings of the 21st international conference on World Wide Web*, pages 679–688. ACM, 2012.
[88] L. Wang and J. Kangasharju. Real-world sybil attacks in BitTorrent mainline DHT. In *Global Communications Conference (GLOBECOM), 2012 IEEE*, pages 826–832. IEEE, 2012.
[89] L. Xiong and L. Liu. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE transactions on Knowledge and Data Engineering*, 16(7):843–857, 2004.
[90] L. Xiong, L. Liu, and M. Ahamad. Countering feedback sparsity and manipulation in reputation systems. In *Collaborative Computing: Networking, Applications and Worksharing, 2007. CollaborateCom 2007. International Conference on*, pages 203–212. IEEE, 2007.
[91] Z. Yan, Y. Chen, and Y. Shen. PerContRep: a practical reputation system for pervasive content services. *The Journal of Supercomputing*, 70(3):1051–1074, 2014.
[92] C. Yang, R. Harkreader, J. Zhang, S. Shin, and G. Gu. Analyzing spammers' social networks for fun and profit: a case study of cyber criminal ecosystem on twitter. In *Proceedings of the 21st international conference on World Wide Web*, pages 71–80. ACM, 2012.
[93] B. Yu and M. P. Singh. A social mechanism of reputation management in electronic communities. In *International Workshop on Cooperative Information Agents*, pages 154–165. Springer, 2000.
[94] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao. Sybillimit: A near-optimal social network defense against sybil attacks. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 3–17. IEEE, 2008.
[95] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. Sybilguard: defending against sybil attacks via social networks. In *ACM SIGCOMM Computer Communication Review*, volume 36, pages 267–278. ACM, 2006.
[96] H. Yu, C. Shi, M. Kaminsky, P. B. Gibbons, and F. Xiao. Dsybil: Optimal sybil-resistance for recommendation systems. In *2009 30th IEEE Symposium on Security and Privacy*, pages 283–298. IEEE, 2009.
[97] G. Zacharia, A. Moukas, and P. Maes. Collaborative reputation mechanisms for electronic marketplaces. *Decision Support Systems*, 29(4):371–388, 2000.
[98] H. Zhao and X. Li. H-trust: A group trust management system for peer-to-peer desktop grid. *Journal of Computer Science and Technology*, 24(5):833–843, 2009.

[99] R. Zhou and K. Hwang. Powertrust: A robust and scalable reputation system for trusted peer-to-peer computing. *IEEE Transactions on parallel and distributed systems*, 18(4):460–473, 2007.

[100] P. R. Zimmermann. *The official PGP user's guide*. MIT press, 1995.