INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 699-703
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016
March 25, 2016

NEIGHBOUR SIMILARITY TRUST AGAINST SYBIL TACKS IN P2P E-COMMERCE

[1]Jennifer.F , [2]Kalaiselvi.S , [3]Ramadevi.K,
[1,2]Dept of Information and Technology, SKP Engineering college, Tiruvannamalai,
[3]Asst prof, Dept of Information and Technology, SKP Engg College, Tiruvannamalai.

Abstract:

Open-access distributed systems such as peer-to-peer systems are particularly vulnerable to *sybil attacks*, where a ma- licious user creates multiple fake identities (called *sybil nodes*). Without a trusted central authority that can tie identities to real human beings, defending against sybil attacks is quite challenging. Among the small number of decentralized approaches, our recent SybilGuard protocol leverages a key insight on social networks to bound the number of sybil nodes accepted. Despite its promising direction, SybilGuard can allow a large number of sybil nodes to be accepted. Furthermore, SybilGuard assumes that social networks are fast-mixing, which has never been confirmed in the real world. This paper presents the novel SybilLimit protocol that leverages the same insight as SybilGuard, but offers dramatically improved and near-optimal guarantees. The number of sybil nodes in our experiments for a million-node system. We further prove that SybilLimit's guarantee is a most a factor away from optimal when considering approaches based on fast-mixing social networks. Finally, based on three large-scale real-world social networks, we provide the first evidence that real-world social networks are indeed fast-mixing. This validates the fundamental assumption behind SybilLimit's and SybilGuard's approach.

Index Terms—Social networks, sybil attack, sybil identities, SybilGuard, SybilLimit.

1. INTRODUCTION

The honest users in a wide variety of collaborative tasks. Exampleswill be the adversary that helps it to accept most of the honest nodes. While this is true, SybilLimit is still needed to bound the number of sybil nodes accepted and also to prevent from growing beyond . Second, the benchmark set is itself a set with $o(1)$ fraction of sybil nodes. Thus, it may appear that an application can just as well use the nodes in directly and avoid the full SybilLimit protocol. However, the set is constructed randomly and may not contain some specific suspects that wants to verify.SYBIL attacks [1] refer to individual malicious users cre- ating multiple fake identities in *open-access* distributed systems (such as peer-to-peer systems). These open-access systems aim to provide service to any user who wants to use the service .observed in the real world [2] in the Maze peer-to-peer system. Researchers have also demonstrated [3] that it is surprisingly easy to launch sybil attacks in the widely used eMule system [4] When a malicious user's sybil nodes comprise a large fraction of the nodes in the system, that one user is able to "outvote".

2. THE SYBILGUARD APPROACH

Recently, we proposed SybilGuard [13], a new protocol for defending against sybil attacks without relying on a trusted central authority. SybilGuard leverages a key insight regarding *social networks* (Fig. 1). In a social network, the vertices (nodes) are identities in the distributed system and the (undi- rected) edges correspond to human-established trust relations in the real world. 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 sybil nodes per attack edge, yielding nearly200 times

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 699-703
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016
March 25, 2016

improvement over SybilGuard. Putting it another way, with SybilLimit, the adversary needs to establish nearly(the attack edges) whose removal disconnects a large number of nodes (all the sybil identities). On the other hand, "fast-mixing" [14] social networks do not tend to have such cuts. SybilGuard leverages the small quotient cut to limit the size of sybil attacks. has more than a handful of network positions, the attacker can fabricate arbitrary network coordinates.In reputation systems, colluding sybil nodes may artificially increase a (malicious) user's rating (e.g., in Ebay). Some sys- tems such as Credence [32] rely on a trusted central authority. There are existing distributed defenses [33], [34] to prevent such artificial rating increases. These defenses, how- ever, cannot bound the number of sybil nodes accepted, and in fact, all the sybil nodes can obtain the same rating as the ma- licious user. Sybil attacks and related problems have also SybilGuard is a completely decentralized protocol and en- ables any honest node (called the *verifier*) to decide whether or not to *accept* another node S means that is willing tasks with S. verifiers out of the honest nodes, where is some small con- stant close to 0. (The remaining nodes get degraded, not prov- able, protection.) Assuming fast-mixing social networks and as- suming the number of attack edges is , SybilGuard guarantees that any such verifier, with probability of at least δ (δ being a small constant close to 0), will accept at sybil nodes per attack edge and at least $(1 \quad \epsilon)n$ honest nodes.

2.1. SybilLimit: A Near-Optimal Protocol for Real-World Social

In this paper, we present a new protocol that leverages the same insight as SybilGuard but offers dramatically improved and near-optimal guarantees. We call the protocol *SybilLimit* because: 1) it limits the number of sybil nodes accepted; and it is near-optimal and thus pushes the approach to the limit. running an approximation algorithm [27] for minimal quotient cut will bound the number of sybil identities accepted within $O(\log n)$ per attack edge, where is the number of honest iden- tities. Also assuming global knowledge and further focusing on scenarios where only $o(n)$ honest identities are seeking to be accepted, SumUp [26] uses adaptive maximum flow on the so- cial network to bound the number of sybil identities (voters) ac- cepted per attack edge within $o(1)$.Similarly, the complete design of Ostra [24] and SybilInfer[25] also assume global knowledge about the social network. Even though both works [24], [25] allude to decentralized de- signs, none of them provides a complete design that is decen- tralized. Ostra does not provide guarantees that are provable. SybilInfer only proves that sybil nodes will increase the mixing time of the graph and thus affect the probability that a random walk starting from a region will end within that region. There is no result proven on *how much* the probability is affected. Sybil- Infer determines the probability via sampling, which by itself has unknown estimation error. As a result, SybilInfer is not able to prove an end-to-end guarantee on the number of sybil nodes accepted.

In contrast to all these above efforts, SybilLimit avoids the need for any global knowledge by using a decentralized secure random route technique. It provably bounds the number of sybil identities accepted per attack edge within $O(\log n)$ while ac- cepting nearly all honest nodes. The relationshipleverages the same insight as SybilGuard but offers dramatically improved and near-optimal guarantees. We call the protocol *SybilLimit* because: 1) it limits the number of sybil nodes accepted; and it is near-optimal and thus pushes the approach to the limit. running an approximation algorithm [27] for minimal quotient cut will bound the number of sybil identities accepted within $O(\log n)$ per attack edge, where is the number of honest iden- tities. Also assuming global knowledge and further focusing on scenarios where only $o(n)$ honest identities are seeking to be accepted, SumUp [26] uses adaptive maximum flow on the so- cial network to bound the number of sybil identities (voters) ac- cepted per attack edge within $o(1)$.Similarly,

the complete design of Ostra [24] and SybilInfer[25] also assume global knowledge about the social network. Even though both works [24], [25] allude to decentralized de- signs, none of them provides a complete design that is decen- tralized. Ostra does not provide guarantees that are provable. SybilInfer only proves that sybil nodes will increase the mixing time of the graph and thus affect the probability that a random walk starting from a region will end within that region. There is no result proven on *how much* the probability is affected. Sybil- Infer determines the probability via sampling, which by itself has unknown estimation error. As a result, SybilInfer is not able to prove an end-to-end guarantee on the number of sybil nodes accepted. In contrast to all these above efforts, SybilLimit avoids the need for any global knowledge by using a decentralized secure random route technique. It provably bounds the number of sybil identities accepted per attack edge within $O(\log n)$ while ac- cepting nearly all honest nodes. The relationship between Sybil- Guard and SybilLimit is discussed in more detail in Sections IV and V-C.

Finally, orthogonal to SybilLimit's goal of limiting the number of accepted sybil nodes, Ostra and SumUp further leverage feedback to modify the weight of the edges in the so- cial network dynamically. Neither of these two feedback-based heuristics offers a provable guarantee. Our recent work on DSybil [28] also uses feedback to defend against sybil attacks in the context of recommendation systems and provides strong provable end-to-end guarantees. In scenarios where feedback is available, we expect that combining these feedback-based techniques with SybilLimit can further strengthen the defense.

3. OTHER SYBIL DEFENSES

=Some researchers [29] proposed exploiting the *bootstrap tree* of DHTs. Here, the insight the number of accepted sybil nodes per attack edge within (see Table I). This is a factor reduction from SybilGuard's guarantee. In our experiments on the million-nods ynthetic social network used in [13], SybilLimit accepts on average around While its direction is promising, SybilGuard suffers from two major limitations. First, although the end guarantees of Sybil- Guard are stronger than previous decentralized approaches, theyare still rather weak in the ab osolute sense: Each attack edge allowssybil nodes to be accepted. permutatio where d is the degree of the node—as a *one-to-one mapping* from incoming edges to outgoing edges. A random route entering via edge will always exit via edge . This precomputed permutation, or *routing table*, serves to introduce *external correlation* across multiple random routes. Namely, once two random routes traverse the same directed edge, they will merge and stay merged (i.e., they *converge*). Furthermore, the outgoing edge uniquely deter- mines the incoming edge as well; thus the random routes can be *back-traced*. These two properties are key to SybilGuard's guarantees. As a side effect, such routing tables also introduce *internal correlation* within a single random route. Namely, if a random route visits the same node more than once, the exiting edges will be correlated. We showed [13] that such correlation tends to be negligible, and moreover, in theory it can be removed entirely using a more complex design. Thus, we ignore internal correlation from now on.Without internal correlation, the behavior of a single random route is exactly the same as a random walk. In connected and nonbipartite graphs, as the length of a random walk goes toward infinity, the distribution of the last node (or edge) traversed be- comes independent of the starting node of the walk. Intuitively, this means when the walk is sufficiently long, it "forgets" where it started. This final distribution of the last node (or edge) tra- versed is called the node (or edge) *stationary distribution* [14] of the graph. The edge stationary distribution (of any graph) is always a uniform distribution, while the node stationary distri- bution may not be. *Mixing*

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 699-703
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016
March 25, 2016

time [14] describes how fast we ap- proach the stationary distribution as the length of the walk increases. More precisely, mixing time is the walk length needed to achieve a certain *variation distance* [14], Δ, to the stationary distribution. Variation distance is a value in [0,1] that describes the "distance" between two distributions—see [14] for the pre- cise definition. A small variation distance means that the two distributions are similar. For a graph (family) with nodes, we say that it is *fast-mixing* if its mixing time is SybilLimit adopts a similar system model and attack model as SybilGuard [13].

3.1. Accepting Honest Nodes:

SybilGuard, each node performs a random route of length . A verifier only accepts a suspect S if S's random route intersects with 's. Theorem 2 tells us that 's random route will stay in the honest region with probability of at least for Theorem.further implies that with high probability, a long will include independent random nodes drawn from the node stationary distribution. It then follows from the generalized Birthday Paradox [42] that an honest suspect S will have a random route that intersects with random once for each verifier with reversed routing tables.

3.2. Performance Overheads:

While SybilLimit uses the same technique as SybilGuard to do random routes, the overhead in- curred is different because SybilLimit uses multiple instances of the protocol with a shorter route length. Interestingly, using instances of the random route protocol does not incurextra storage or communication overhead by itself. First, a node does not need to store routing tables since it can keep a single random seed.

3.3. Bounding the Number of Sybil Nodes Accepted

Intersect with's non-escaping random route, a sybil sus- pect's random route must traverse one of the attack edges. where there is only a single attack edge,

n	total number of nodes in the honest region
m	total number of edges in the honest region
g	total number of attack edges
r	number of random routes that each verifier and suspect performs
w	length of individual random routes (in SybilLimit)
l	length of individual random routes in SybilGuard
V	verifier node
S	suspect node

Fig.1. General Structure

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 699-703
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016
March 25, 2016

Thus, a node needs to send only messages instead of messages. SybilLimit inherits the idea from Sybil-Guard that an honest node should not have an excessive number of neighbors. This restriction helps bound the number of addi- tional attack edges the adversary gets when an honest node is compromised. If there are too many neighbors, SybilLimit will (internally)only use a subset of the node's edges while ignoring all others. This implies that rd will not be too large on average (e.g., 20). Finally, the total number of bits a node needs to send in the protocol is linear with the number of random routes times the length of the routes. Thus, the total number of bits sent in the d messages in SybilLimit is , as compared to in SybilGuard. The intersection condition requires that S's tails and tails must intersect (instance number is ignored when determining intersection), with S being registered at the intersecting tail. In contrast, SybilGuard has an intersection condition on nodes (in- stead of on edges or tails). For the balance condition, main- tains counters corresponding to its tails. Every accepted sus- pect increments the "load" of some tail. The balance condition from the set of non-escaping tails from honest suspects. The reason is that random routes are back-traceable, and starting from a non-escaping tail, one can always trace back to the starting node of the random route, encountering only honest nodes. This means that an honest suspect will never need average node degree being 10, an average node using Sybil- Guard needs to send 400 KBs of data every few days.

CONCLUSION

All these random routes need to be performed only one time (until the social network changes) and the relevant information will be recorded. Further aggressive optimizations are possible (e.g., propagating hashes of public keys instead of public keys themselves). We showed [13] that in a million-node system with• Bad sample probability in SybilGuard. When estimating the random route length, the probability of a bad sample,Thus, to allow for larger , SybilLimit needs to resolve all three.

REFERENCES

[1] A. Fiat, J. Saia, and M. Young, "Making Chord robust to byzantine attacks," in Proc. ESA, 2005, pp. 803–814.
[2] E. Damiani, D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Vi-
olante, "A reputation-based approach for choosing reliable resources in peer-to-peer networks," in Proc. ACM CCS, 2002, pp. 207–216.
[3] L. von Ahn, M. Blum, N. J. Hopper, and J. Langford, "CAPTCHA:
Using hard AI problems for security," in Proc. IACR Eurocrypt, 2003, pp. 294–311.
[4] A. Ramachandran and N. Feamster, "Understanding the network-level behavior of spammers," in Proc. ACM SIGCOMM, 2006, pp. 291–302.
[5] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman, "SybilGuard: Defending against sybil attacks via social networks," IEEE/ACM Trans. Netw., vol. 16, no. 3, pp. 576–589, Jun. 2008.
[6] M. Mitzenmacher and E. Upfal, Probability and Computing. Cam- bridge, U.K.: Cambridge Univ. Press, 2008.