INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 2 (2016) PAGES 649-653
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016

DETECTING SYBIL ATTACK IN E-COMMERCE BY USING NST APPROACH

[1]B.Gowri, [2]S.Bhavani, [3]K.Ramadevi,
SKP Engineering 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.

Keywords – Sybil nodes, Sybil guard, peer to peer.

1. INTRODUCTION

Attack edge and enter the honest region. Notice that here Sybil- Limit reduces the number of such routes by using a that is much smaller than l. Furthermore, because we are concerned only with tails will . With , the adversary will have such slots total for all the sybil nodes.This reduction from to slots is the first key step in SybilLimit.However, doing random routes introduces two problems.The first is that it is impossible for a degree-d node to have more than d distinct random routes if we directly use SybilGuard's ap- proach. SybilLimit observes that one can use many independent instances of the random route protocol while still preserving the desired convergence/back-traceability property. This section highlights the key novel ideas in SybilLimit that eventually lead to the substantial end-to-end improvements over SybilGuard.

Theorem 3: Assume that the social network's honest region is fast-mixing and . For any given constants For the remaining small fraction of honest verifiers, Sybil-Limit provides a degraded guarantee that is not provable. Be- cause of space limitations, we will provide mostly intuitions in the following and leave formal/complete proofs to our technical report [43].

We adopt the phi- losophy that all guarantees of SybilLimit must be proven mathe- matically because experimental methods can cover only a subset of the adversary's strategies. Our proofs pay special attention to the correlation among various events, which turns out to be a $\bar{k}$ey challenge. We cannot assume independence for simplicity because, after all, SybilLimit exactly leverages external correla- tion among random routes. The following is the main theorem on SybilLimit's guarantee. requires that accepting S should not result in a large "load spike" and cause the load on any tail to exceed . Here, is the current average load across all 's tails, and is some universal constant that is not too small (we use h in our experiments). In comparison, SybilGuard does not have any attack.

Notice that here Sybil- Limit reduces the number of such routes by using a that is much smaller than l. Furthermore, because we are concerned only with tails will . With, the adversary will have such slots

.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744 VOL 2 ISSUE 2 (2016) PAGES 649-653
RECEIVED : 18/03/2016. PUBLISHED : 25/03/2016 March 25, 2016

total for all the sybil nodes.This reduction from to slots is the first key step in SybilLimit. However, doing random routes introduces two problems.The first is that it is impossible for a degree-d node to have more than d distinct random routes if we directly use SybilGuard's ap- proach. SybilLimit observes that one can use many independent instances of the random route protocol while still preserving the desired convergence/back-traceability property. The second problem is more serious. Instead, it use a novel and perhaps counterintuitive benchmarkingtechnique that mixes the real suspects with some random bench- mark suspects that are already known to be mostly honest. The technique guarantees that a node will never overestimate regardless of the adversary's behavior. If the adversary causes an underestimation for , somewhat counterintuitively, e technique can ensure that SybilLimit still achieves its end guarantees despite the underestimated. We will leave the detailed discussion to Section VII.

Condition: To help convey the intuition, we will assume in the following. In SybilLimit, each node random routes of length of all honest nodes will entirely determine whether 's tail is escaping and in the case of a non-escaping tail, which edge is the tail. Thus, the adversary has no influence over non-escaping tails.Since the distribution of the non-uniform tails is unknown, few probabilistic properties can be derived for them. Escaping tails are worse because their distribution is controlled by the ad- versary. We thus would like to first quantify the (small) frac- tion of non-uniform tails and escaping tails. Assuming that the honest region of the social network is fast-mixing, our technical report [43] proves that for most honest nodes, most of their tails are uniform tails.

simultaneously needed to ensure the following:

- Sybil nodes accepted by SybilGuard. The total number of sybil nodes accepted, is .
- Escaping probability in SybilGuard. The escaping probability of the verifier's

random route, is slots for the sybil nodes in SybilGuard.

In SybilLimit, the tail of each random route corresponds to a "slot" for registration. In any given s-instance, the adversary can fake distinct random routes of length that cross the (potentially close to zero) and , there is a set of onest verifiers and

universal constants and , such that using and in SybilLimit will

guar- antee that for any given verifier

in the set, with probability of at least.

As a reminder, the probability in the above lemma is defined

over the domain of all possible routing table states—obviously, if all routing tables are already determined, the tail will be some fixed edge.

It still possible for the tail of a non-escaping node to be escaping or non-uniform—it is just that such probability is

$o(1)$ for $o(n/\log n)$. We will not ignore

this fraction of tails, but knowing that they

fraction will facilitate our proof later. An honest node that is not non-escaping is called an escaping node. By Lemma 4, we have at most escaping nodes; such nodes are usually near the attack edges. Notice that given the topology of the honest region .a verifier in SybilLimit needs to do such routes, it remains quite likely that some of them are escaping. In fact, with and , the probability of at least one of the routes being escaping in SybilLimit is even larger than the probability of the single length-l random route being escaping in SybilGuard. Thus, so far we have only made the "all-or-nothing" effect in SybilGuard fractional. SybilLimit relies on its (new) balance condition to address this fraction of escaping routes. To obtain some intuition, let us imagine the verifier V's tails as bins that can accommodate up to a certain load. When V accepts a suspect S, out of all of V's tails that intersect with S's tails, S conceptually increments the load of the least loaded tail/bin. Because of the randomness in the system, one would conjecture that all of V's tails should have similar load. If this is indeed true, then we can enforce a quota on the load of each tail, which will in turn bound the number of sybil nodes accepted by V's escaping tails. Later, we will show that the balance condition bounds the number .

2) User and Node Dynamics: Most of our discussion so assumes that the social network is state and all nodes are online. All techniques in SybilGuard to efficiently deal with user/node dynamics, as well as techniques to properly overwrite stale reg- istration information for preventing certain attacks [13], apply to SybilLimit without modification. We do not elaborate on these due to space limitations. random route. This means that the adversary can register only public keys for all the sybil nodes combined. Inorder to accept a suspect S, must find an intersection between its random route and S's random route and then confirm.

The intersection condition requires that S's tails and 's 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. Under the same parameters, an average node using

SybilLimit would send around of data every few days, which is still quite acceptable. We refer the reader to [13] for further details. For much larger social networks (e.g., with bil- lions of nodes), the overhead may become substantial. Further reducing such overhead (e.g., via indirect multihop validation) is part of our future work.

3) 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 incur

extra storage or communication overhead by itself. First, a node does not need to store routing tables since it can keep a single random seed and then C. 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. Con- sider Fig. 2, where there is only a single attack edge.

4) Basic Security Properties: The secure random route pro- tocol provides some interesting basic security guarantees. We first formalize some

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

Theorem 1: Consider any fast-mixing graph with nodes. A random walk of length is sufficiently long such that, with probability of at least $(1/n)$, the last node/edge tra- versed is drawn from the node/edge stationary distribution of the graph.

In SybilGuard, a random walk starting from an honest node in the social network is called escaping if it ever crosses any attack edge.

Theorem 2: (From [13]) In any connected social network with nodes and attack edges, the probability of a length-l random walk starting from a uniformly random honest node being escaping is at most .

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

Honest nodes obey the protocol. The system also has one or more malicious human beings as malicious users, each with one or more identi- ties/nodes. To unify terminology, we call all identities created by malicious users as sybil identities/nodes. Sybil nodes are byzan- tine and may behave arbitrarily. All sybil nodes are colluding and are controlled by an adversary. A compromised honest node is completely ontrolled by the adversary and hence is consid- ered as a sybil node and not as an honest node.

There is an undirected social network among all the nodes, where each undirected edge corresponds to a human-established trust relation in the real world. The adversary may create ar- bitrary edges among sybil nodes in the social network. Each honest user knows his/her neighbors in the social network, while the adversary has full knowledge of the entire social network. The honest nodes have undirected edges among themselves in the social network. For expository purposes, we sometimes also consider the undirected edges as directed edges. Themixing, an assumption that had not been validated in the real world.been studied in sensor networks [35], [36], but the approaches and solutions usually rely on the unique properties of sensor net- works (e.g., key predistribution). Margolin et al. [37] proposed using cash rewards to motivate one sybil node to reveal other sybil nodes, which is complimentary to bounding the number of sybil nodes accepted in the SybilGuard uses a special kind of random walk, called random routes, in the social network. In a random walk, at each hop, the current node flips a coin on the fly to select a uniformly random edge to direct the walk (the walk is allowed to turn back). For random routes, each node uses a precomputed random permutation—“$x_1 x_2 \ldots x_d$,” 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 i will always exit via edge x_i. 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.

CONCLUSION

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 time [14] describes how fast we ap- proach the stationary distribution as the length of the walk in- creases. 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.

REFERENCES

[1] Ramachandran and N. Feamster, “Understanding the network-level behavior of spammers,” in Proc. ACM SIGCOMM, 2006, pp. 291–302.

[2] 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.

[3] M. Mitzenmacher and E. Upfal, Probability and Computing. Cam- bridge, U.K.: Cambridge Univ. Press, 2008.