PRIDE: Peer-to-Peer Reputation Infrastructure for Decentralized Environments

Prashant Dewan & Partha Dasgupta
Department of Computer Science and Engineering
Ira A. Fulton School of Engineering
Arizona State University
{dewan,partha}@asu.edu

ABSTRACT

Peer-to-peer (P2P) networks use the fundamental assumption that the nodes in the network will cooperate and will not cheat. In the absence of any common goals shared by the nodes of a peer-to-peer network, external motivation to cooperate and be trustworthy is mandated. Digital Reputations can be used to inject trust among the nodes of a network. This paper presents PRIDE, a reputation system for decentralized peer-to-peer networks. PRIDE uses self-certification a scheme for identification of peers using digital certificates similar to SDSI certificates, an elicitation-storage protocol for exchange of recommendations and IP Based Safeguard (IBS) to mitigate a peer's vulnerability to 'liar farms.'

Categories and Subject Descriptors: C2.3Distributed SystemsPeer-to-Peer Networks, Security, Identity

General Terms: Peer-to-Peer Networks, Security, Identity, Reputations

Keywords: Reputation Systems, Security, Peer-to-Peer

1. INTRODUCTION

In Gnutella[1] 70% of the users do not share any files and nearly 50% of the queries are answered by the top 1% of the nodes [1]. Free riding in Gnutella leads to degradation of the system performance and the absence of common goals among peers makes the network vulnerable to malicious behavior. *Digital Reputations* can be used to inject the necessary motivation in the Gnutella servents. Digital reputation systems like EBay and P2Prep[4] assume that peers will value their reputations because of the benefit that can be derived from a good reputation, and will continue performing transactions in the same 'fashion', as they have performed in the past. PRIDE complements Gnutella with minimal modification of the Gnutella protocol. PRIDE does not need any central server to identify the peers. Peers generate their own identities using self certification and locally store the recommendations received by them. A peer (requester) uses the Gnutella query for locating content providers. Subsequently it selects the 'best' peer (provider) from the list (obtained in the query phase), based on the reputations of the providers. The requester downloads content from the provider and grants a recommendation to the provider. Additionally the requester signs and stores the transaction number of the provider in the network. The provider stores the recommendations locally and shows them to the next requester as a proof of its reputation. As a result the requesters do not have to perform a network search for the provider's reputation information. The salient features of PRIDE are self-certification, IP Based Safeguard and the elicitation-storage protocol.

[1]one of the largest p2p networks

WWW2004, May 17–22, 2004, New York, New York, USA.
ACM 1-58113-912-8/04/0005.

2. THREAT MODEL

One of the main goals of a reputation system for a p2p network is to reduce the chances of a peer getting cheated in a transaction. For example, if a requester downloads music from a provider, it should be able to ascertain the probability of the fact that the music files are complete. Similarly the provider would like to ascertain that the recommendation it will receive from the requester, will at least be in conformance with the quality of the service provided. Digital reputations can mitigate both of the above threats. A reputation system brings in its own set of issues and threats. It is important to have a identifier allocation strategy for peers in order to restrict any peer from generating a liar farm (multiple identities) [5] to raise its own reputation. In addition, the reputation information has to be tamper proof and easily retrievable while upholding the accountability of the reputation issuer. The other challenges for any reputation system are "ballot stuffing," "bad mouthing," negative and positive discrimination [3].

3. SELF-CERTIFICATION

Each peer runs its own certificate authority (CA) which signs the identity certificate(s) of the peer. All the certificates used in self-certification are digitally signed statements, similar to SDSI certificates [6]. An identity certificate acts as a 'proxy' to the peer and binds the public key of a peer to the other information of the peer. Additionally, the IP-ADDRESS field, a mandatory field for the identity certificate, specifies the IP address range from which this identity can be used. If the identity is used from an IP address out of the range, the other party in the transaction suspects foul play and aborts the interaction . From here on, the word 'identity' is used to refer to identity certificate. Identities are needed to link together (at least some of the) transactions performed by the peer. Two transactions performed using the same identity can be traced back to the

identity. The identity may or may not be traceable to the peer. Only when the peer calculating the reputation of an identity knows the transactions of the identity, the reputation of the identity can be calculated. IP addresses cannot be used as identities because IP addresses are shared among peers in two different time spans. In addition, the peers generally do not manage their own IP addresses. A peer can perform two roles in the network, a requester: a peer that requests for service or a provider: a peer that provides the service. At the end of each transaction the requester issues a *recommendation* (SDSI certificate) to the provider. A positive recommendation increases the reputation of the identity of the provider by one while a negative recommendation diminishes it by one. The requester also submits the corresponding identity certificate to the provider. An example of a transaction initiated when the requester downloads some files from the provider.

4. LIAR FARMS & IP BASED SAFEGUARD

Using self-certification any peer can generate a large number of identities, feign a large set of peers and maliciously increase the reputation of one or more of its identities by giving false recommendations (similar to ballot stuffing). Such a farm of identities is called a 'liar farm.' A liar farm can be countered if all the identities of a peer can be mapped back to the peer. If self certification is used, a peer's identities cannot be mapped back to it without its consent. IP Based Safeguard uses security zones that are the subsets of the IP space. We assume that only the information provider receives the recommendation, and the information requester provides the recommendation. The requester selects a provider with which it wants to perform a transaction and verifies the identity of the old requesters (recommendation issuers) of the provider. Each peer maintains a local database of verified identities. The requester first checks its local database for a valid identity, matching the identity associated with the recommendation. If it does not find any such valid identity, it verifies the identity of the old requesters of the provider, by performing a cryptographic challenge-response at the IP address in the identity certificate of the previous requester(s). If the challenge-response fails for a particular recommendation, that recommendation is not included in the reputation for the provider identity. Dellarocas recommends [2] peer annonymity to prevent bad mouthing and discrimmination. This is an ongoing area of research.

Once the requester has a list of (apparently) valid recommendations, it sorts them by the order of the IP addresses (in the corresponding identity certificates), determines a security distance, d and divides the linear IP space into slices (security zones) of length d. It averages all the recommendations received by the provider from identities whose IP is in the same security zone. Finally, it adds the averages of each security zone to calculate the reputation of the provider identity. IBS is based on the fundamental assumption that it will be difficult for any malicious peer to generate identities that have totally non-contiguous IP addresses. By increasing the security distance d, a requester can reduce the probability of the provider having an identity farm. This reduced probability is at the expense of reduction in accuracy of the reputation of the provider. By decreasing d, the probability of a peer being victimized to a liar farm rises, although the accuracy of the reputation of the provider is higher. The experiments show that the mean variation of ranks of peers in a network that uses IBS is 13.2 ± 0.2 with a 95% confidence level. IBS and liar farms have been extensively discussed in [5].

5. ELICITATION-STORAGE PROTOCOL

A requester obtains a list of providers who have the required content. The requester also receives the reputation of each of the providers in the list. The requester uses IBS to normalize the reputation of the peers and selects the 'best' peer based on the reputation of the possible providers and initiates the ES protocol. On the requester's initiation, the provider that is selected by the requester on the basis of its reputation, generates a new transaction id (TID) by using the last transaction id as a seed for a one-way function. The requester verifies if the same TID has been used for any other transaction by the provider. Once the TID is verified, the requester checks (at least some of) the past recommendations[2] received by the provider and, once satisfied, performs the transaction. Following IBS the requester recalculates the reputation of the peer by averaging the recommendations received in each security zone. If the recalculated value of peer's reputation is above the requester's threshold it performs the transaction and if not it contacts the second peer in the list and so on. Once the transaction (file download) is complete, the requester gives a signed recommendation to the provider, which is stored by the provider. In addition, the requester signs the TID and stores it in the P2P network. The details of the protocol can be found in [4]

6. CONCLUSION

PRIDE increases the satisfaction level of the peers from the system by directing the requesters to the providers who have a history of better performance. The peers in PRIDE weed out the rogues by giving them bad recommendations, thereby lowering the reputations of bad peers. It provides the providers an incentive to provide accurate and timely information in order to obtain good recommendations from requesters. The current domain of transaction is confined to downloading files. In the future the domain of a transaction can be expanded to include higher stake transactions in E-Markets like E-Bay. More work needs to be done for integrating globally trusted Certificate Authorities and corresponding hierarchies in the name spaces into local name spaces used by PRIDE

7. REFERENCES

[1] E. Adar and B. Huberman. Free Riding on Gnutella, 10/02/00.

[2] C. Dellarocas. Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior. In *ACM Conference on Electronic Commerce*, pages 150–157, 2000.

[3] C. Dellarocas Building Trust On-line : The Design of Reliable Reuptation Reporting System, *eBusiness@MIT working paper 101. MIT Sloan School of Management*, Cambridge, Mass., 2001.

[4] P. Dewan. Injecting Trust In Peer-to-Peer Systems. Ph.D. Dissertation Proposal, Arizona State University, June 16,2003

[5] P. Dewan. Countering identity farms in reputation systems for P2P network. Technical report, Arizona State University, Tempe, January 2004.

[6] R. L. Rivest and B. Lampson. SDSI A Simple Distributed Security Infrastructure. In *Crypto*, Santa Barbara ,USA, 1996.

[2]Recommendation is the reputation information pertaining to one interaction between two distinct peers