Towards Social Profile Based Overlays

David Isaac Wolinsky, Pierre St. Juste, P. Oscar Boykin, Renato Figueiredo
University of Florida

Abstract

Social networking has quickly become one of the most common activities of Internet users. As social networks evolve, they request more information from the users and thus requiring the users to place more trust into the social network. Peer-to-peer (P2P) overlays can return ownership of information and system control to the user as they can be constructed in a way to not require third party proxies.

In this paper, we present a novel concept known as the structured social overlay that applies social networks to structured P2P overlays to provide ownership, scalability, reliability, and security. Each user's profile is assigned a unique private, secure overlay, where members of that overlay have a friendship with the overlay owner. The profile data is stored using the profile overlay's distributed data stores. To ensure privacy, the profile overlay employs the public key infrastructure, where the role of certificate authority (CA) is handled by the overlay owner and each member of the overlay has a CA signed certificate. Each member of the social network, joins a common public overlay, which provides services to discover friends and bootstrap connections into existing private overlays through a distributed data store. We define interfaces that can be used to implement this system as well as explore some of the challenges related to it.

1 Introduction

Social networking has become pervasive in daily life, though as social networks grow so does the wealth of personal information that they store. Users become more dependent on social networks as users surrender tasks such as communication and identity to the social network. Once information has been released to a social network, known as a user's profile, the user is at the mercy of the social network. If the social network engages in activites disagreeable to the user, the user has the option to leave the social network surrendering their identity and features provided by the social network or to accept the disagreeable activities.

Structured P2P overlays provide a scalable, resilient, and self-managing platform for distributed applications. Structured overlays provide means by which users can easily create their own decentralized systems for the purpose of data sharing, interactive activities, and other networking enabled activities. In recent work **??**, we have implemented mechanisms that allow users to create and manage their own private overlays using a common public overlay to assist in discovery and NAT traversing of the private overlay. In this paper, we further this work by an indepth discussion on how to apply this technique to social networks.

Social networks consist of users, each of whom typically has a a profile, a set of friends, and a private message inbox. The profile contains the users personal information, status updates, and public conversation with the friends. Friends are individual which the user trust sufficiently to view the profile. The private message inbox allows users to send messages between each other without leaking any information to other friends. Using this model, we describe a common directory or public overlay which allows peers to provide services where peers can find friends and join overlays where there already exists an established relationship. Each user has their own profile overlay, where the members of the overlay are limited to the current friends of the profile owner. The profile overlay is secured by a public key infrastructure (PKI) with the profile owner being the certificate authority (CA). The profile information is stored information is stored in distributed datastores, allowing profile information to be accessed in scalable mechanisms regardless of the profile owner's online state.

The rest of this paper is organized as follows. Section 2 provides background and related work. Section 4 and Section 3 describe our contributions, namely, a private, profile overlay and a public, directory overlay. We conclude the paper in Section 5.

2 Background

In this section, we review structured P2P overlays and distributed and decentralized social network techniques.

2.1 Structured P2P Overlays

Structured P2P systems provide distributed lookup services with guaranteed search time with a lower bound of $O(\log N)$, in contrast to unstructured systems, which rely on global knowledge/broadcasts, or stochastic techniques such as random walks [1]. Some examples of structured systems can be found in [9, 10, 4, 5, 7]. In general, structured systems are able to make these guarantees by self-organizing a structured topology, such as a 2D ring or a hypercube.

In the overlay, each node is given a unique node ID drawn from a large address space. Each node id must be unique otherwise address collisions will occur, which can prevent nodes from participating in the overlay. Furthermore, having the node IDs well distributed assist in providing better scalability as many shortcut selection algorithms depend on having node IDs uniformly distributed across the entire address space. A simple mechanism to ensure this behavior is to have each node use a cryptographically strong random number generator to generate the node ID. Another mechanism involves the use of a trusted third party to generate node IDs and cryptographically sign them [2].

As with all P2P systems, in order for an incoming node to connect with the overlay, the node must know of at least one active participant. A list of nodes that are running on public addresses is typically distributed with the application, available through some out-of-band mechanism, or possibly using multicast to findpools [9].

In dealing with ring based overlays, a node must be connected to closest neighbors in the node ID address space; optimizations for fault tolerance suggest that it should be between 2 to $\log(N)$ on both sides. Having multiple peers on both sides assist in stabilizing the overlay structure when experiencing churn, particularly when peers leave without warning.

Overlay shortcuts enable efficient routing in ring-structured P2P systems. Different shortcut selection methods include: maintaining large tables without using connections and only verifying usability when routing messages [9, 5], maintaining a connection with a peer every set distance in the P2P address space [10], or using locations drawn from a harmonic distribution in the node address space [4].

Most structured P2P overlays support decentralized storage/lookup of information by mapping keys to specific node IDs in an overlay. At a minimum, the data is stored at the node ID either smaller or larger to the data's node ID and for fault tolerance the data can be stored at other nodes. This sort of mapping and data storage is called a distributed hash table (DHT).

In [3, 6, 8], the authors discuss the concept a single overlay supporting services by additional overlays that use the underlying overlay to assist in discovery. In [**?**], we provided a reference implementation of a multiple overlay system that supported the use of a public overlay's DHT to store currently active peers in the private overlays. Whereby users could create their own overlays without having to create their own bootstrap network. In addition, our system provides both relay and hole-punching NAT traversal techniques and supports point-to-point PKI based security.

2.2 Social Networks

3 The Directory Overlay

The directory overlay supports two features: 1) a directory for friend discovery and verification and 2) lists of peers currently active in each profile overlay.

3.1 Finding and Verifying Friends

3.2 Active Peers

4 The Profile Overlay

Each profile overlay resembles a private overlay as discussed in [**?**]. In this section, we focus on the tasks of controlling overlay membership, i.e., handling friendships; distribution and retrieval of profile information; and private messaging.

4.1 Handling Friendships

4.2 Storing and Retrieving Profile Data

4.3 Private Messaging

5 Conclusion

References

[1] M. Castro, M. Costa, and A. Rowstron. Debunking some myths about structured and unstructured overlays. In *NSDI'05: Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation*, 2005.

[2] M. Castro, P. Druschel, A. Ganesh, A. Rowstron, and D. S. Wallach. Security for structured peer-to-peer overlay networks. In *5th Symposium on Operating Systems Design and Implementaion (OSDI'02)*, December 2002.

[3] M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron. One ring to rule them all: Service discover and binding in structured peer-to-peer overlay networks. In *SIGOPS European Workshop*, Sept. 2002.

[4] G. S. Manku, M. Bawa, and P. Raghavan. Symphony: distributed hashing in a small world. In *USITS*, 2003.

[5] P. Maymounkov and D. Mazières. Kademlia: A peer-to-peer information system based on the XOR metric. In *IPTPS '02*, 2002.

[6] Randpeer development team. Randpeer. `http://www.randpeer.com`, May 2007.

[7] S. Ratnasamy, P. Francis, S. Shenker, and M. Handley. A scalable content-addressable network. In *In Proceedings of ACM SIGCOMM*, 2001.

[8] S. Ratnasamy, M. Handley, R. M. Karp, and S. Shenker. Application-level multicast using content-addressable networks. In *NGC '01: Proceedings of the Third International COST264 Workshop on Networked Group Communication*, 2001.

[9] A. Rowstron and P. Druschel. Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems. In *IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, November 2001.

[10] I. Stoica and et al. Chord: A scalable Peer-To-Peer lookup service for internet applications. In *SIGCOMM*, 2001.