A Case for P2P Infrastructure for Social Networks - Opportunities & Challenges

Sonja Buchegger
Deutsche Telekom Laboratories
Ernst-Reuter-Platz 7,
D-10587 Berlin, Germany
sonja@ieee.org

Anwitaman Datta
Centre for Advanced Information Systems
School of Computer Engineering
NTU Singapore
anwitaman@ntu.edu.sg

Abstract

Online Social Networks like Facebook, MySpace, Xing, etc. have become extremely popular. Yet they have some limitations that we want to overcome for a next generation of social networks: privacy concerns and requirements of Internet connectivity, both of which are due to web-based applications on a central site whose owner has access to all data. To overcome these limitations, we envision a paradigm shift from client-server to a peer-to-peer infrastructure coupled with encryption so that users keep control of their data and can use the social network also locally, without Internet access. This shift gives rise to many research questions intersecting networking, security, distributed systems and social network analysis, leading to a better understanding of how technology can support social interactions. This paper is an attempt to identify the core functionalities necessary to build social networking applications and services, and the research challenges in realizing them in a decentralized setting. In the tradition of research-path defining papers in the peer-to-peer community [5, 14], we highlight some challenges and opportunities for peer-to-peer in the era of social networks. We also present our own approach at realizing peer-to-peer social networks.

1 Introduction

The landscape of Internet usage has changed dramatically in recent years, both in the way the computers connected to the network interact as well as the way the end-users using these computers interact - with the Internet, and with each other. On the networking layer, infused by the (somewhat infamous) success of P2P file-swapping software, the last decade has witnessed an increased emphasis of using resources available at the edge to perform tasks which would otherwise have heavily burdened any centralized infrastructure. Thus to say, there is an increased proliferation of peer-to-peer mechanisms to either replace, or more often supplement, the client-server paradigm.

On the application layer, with the advent of Web 2.0 and social networks, we witness end users participating not only as passive consumers of content provided by the web-sites (client/server), but also as a contributor creating content collaboratively with fellow users. Thus, at a logical level, many of these Web 2.0 applications are inherently peer-to-peer in nature. Nevertheless, somewhat ironically, all current Web 2.0 applications rely on an underlying infrastructure based on the traditional client-server model.

When the user interactions are peer-to-peer in nature, and while there is such a proliferation of unrelated P2P systems and applications, it is natural to ask if and how to realize a peer-to-peer underlying networking infrastructure for Web 2.0 applications. Perhaps in this irony lies the opportunity for P2P to redeem itself. The almost only well known popular P2P applications (besides Skype) are file-sharing and video streaming. Like many other technologies, file-sharing is susceptible to illegal use. The technology of P2P file-sharing, and even more generally the whole P2P paradigm has thus often been demonized. The question has often been, what is a legitimate P2P application? We believe that P2P infrastructure for Web 2.0 applications, particularly social networks, is one such crucial application where end users can benefit from using a P2P infrastructure. The match could not have been any better or more natural than when both the underlying network resources and infrastructure, as well as the content is provided and consumed by end-users.

Of course at this juncture it is legitimate to ask, why use a peer-to-peer infrastructure for supporting social networks, when the good old client-server architecture works fine. One can give the traditional arguments that P2P scales well, since a growing user base naturally brings in more infrastructural resources. This definitely can be a good incentive for people with good ideas but little money to support and expand overnight if their popularity increases. Also, if popularity declines over time, there is less exposure. However,

978-1-4244-3375-9/09/$25.00 ©2009 IEEE

given the success of numerous upstart companies, which have managed to scale well to not just millions but even hundreds of millions of users, the traditional scalability argument alone does not justify the hassles of a P2P infrastructure.

Even as social networking sites claim to grow their user base at great speeds, the paranoid among us have long been wary, and as people gradually get to understand the implications better [12, 9] and get more pragmatic, privacy will become a major concern. Particularly privacy and protection from massive data-mining and "big-brotherly" treatment of the users by the social networking service providers. This is expected eventually to lead to a significant population of users, who while they would like to enjoy the benefits and fun of social networks, may also want to restrict access to their personal data not only from fellow users who happen to be strangers, but also from any "big brother". This disaffected population is expected to be the early adopters of social networks which rely on a peer-to-peer infrastructure and encryption. In the long run, we believe that if comparable quality of service can be achieved, a significantly large, if not all users will indeed be inclined to use it. As an anecdote, one may consider email users who use encryption (like PGP). It may be a small fraction of all the people who use emails, but nevertheless, their need is genuine, and the population of such users in non-negligible. Besides privacy and other related security concerns, the P2P approach also provides content creators to execute greater control over their content, as well as avoid censorship either by the website owner, or censorship of the hosting website by a third party.

While some sites follow up with corrective measures because of users' outcry, e.g. [3], and one may also argue about legislative solutions to protect users' privacy, there is no guarantee that in the future the users' data will not be misused. The primary objective here is thus to aim for a system which makes it technologically harder (ideally, impossible) to violate the users' privacy and large scale data mining, even while the users continue to enjoy the advantages of social networking.

Essentially, a P2P approach seems promising to be the right technology to achieve both privacy and freedom of speech. For this reason, user-provided content and participatory media creation suit themselves better to a peer-to-peer rather than a client-server model. Another incentive for users to embrace such a model is to evade any constraints put by the service provider in the present or future (e.g. for the amount of storage space, or subscription fees, or service shut down).

Last, but certainly not least, realizing an application layer Internet on top of diverse networking infrastructure, including the Internet, but also mobile - cellular as well as ad-hoc, and supporting Web 2.0 applications on top of such an application layer Internet, can also help making them ubiquitous, as shown in Figure 1.

By supporting the direct exchange of information between devices, be it between users that meet or between adjacent nodes of a city mesh network, a peer-to-peer infrastructure can take advantage of real social networks and geographic proximity. In contrast to a centralized web server, local connectivity already facilitates social networking without Internet access.

We investigate here the (un)suitability of the current P2P technologies to support social networking applications, and try to identify the important outstanding issues. In order to do so, we also need to understand and classify better the salient properties of current social networking applications.

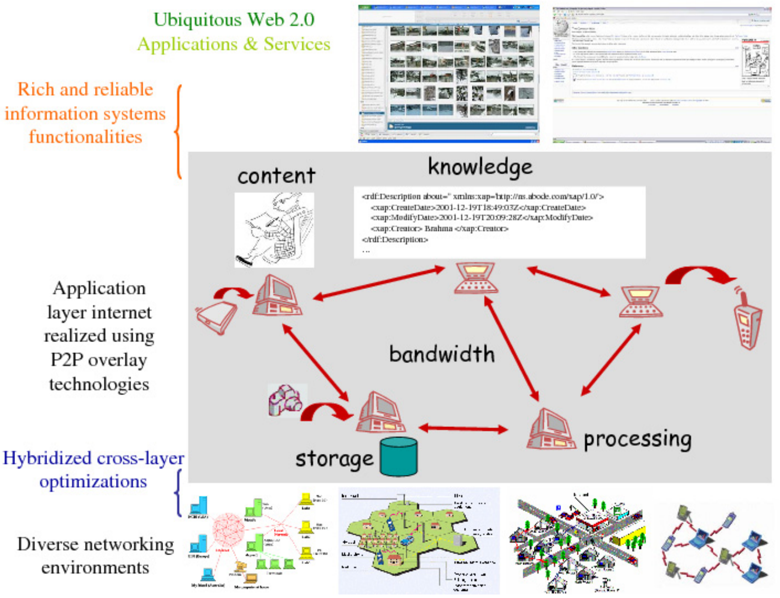
Over the last seven or eight years that P2P technologies have come into prominence, the technology as well as the research community has made several strides. At some crucial junctures along the line, researchers have tried to summarize the current results, their shortcomings and the next challenges [5, 14]. While these erstwhile challenges have since been mostly addressed, the current paper is a similar attempt to chart a path for the next steps that the research community as a whole may benefit from addressing. In the 2006 edition of IEEE P2P, the panel discussion was on the grand challenges for the community, and we think that this paper highlights some of these. The list is surely not exclusive nor completely novel, but we hope to have produced something exhaustive enough to support basic functionalities in contemporary social networking applications. In the process we are identifying some of the next open issues to deal with to extend the reach of the peer-to-peer paradigm.

2 P2P Social Networks Opportunities

What does a p2p infrastructure offer that a client-server one does not? How does this trade off the features provided by a client-server architecture? In this section, we address these questions.

An immediate advantage of a peer-to-peer infrastructure is rather straightforward: it is not centralized, not owned by a single entity. The central storage of user information and ownership by a company, along with commercial exploitation of this information e.g. for ad revenue, raises privacy concerns that could be better addressed by a peer-to-peer approach, with encryption and appropriate key management.

Centralized web-based social networks do not match the inherent peer-to-peer nature of both social networks themselves and of participatory media creation. By mapping a peer-to-peer application to a peer-to-peer infrastructure, direct connections can be exploited such that locality can be taken into account. This enables peers to be mobile and independent from Internet connectivity. Social networking applications can be run on small devices such as PDAs or


Ubiquitous Web 2.0
Applications & Services
Rich and reliable information systems functionalities
content
knowledge
Application layer internet realized using P2P overlay technologies
bandwidth
processing
storage
Hybridized cross-layer optimizations
Diverse networking environments

```
<rdf:Description about='' xmlns:xap='http://ns.adobe.com/xap/1.0/'>
   <xap:CreateDate>2001-12-19T18:49:03Z</xap:CreateDate>
   <xap:ModifyDate>2001-12-19T20:09:28Z</xap:ModifyDate>
   <xap:Creator> Brahma </xap:Creator>
</rdf:Description>
...
```

Figure 1. P2P Overlay Information Systems based Ubiquitous Web 2.0

phones as well as on home servers or access routers. Peers can carry information for each other in a delay-tolerant fashion and use local access points for local information.

Given the reality of privacy breaches by centralized online social network providers, as exemplified by Facebook's beacon application [12] among others [9], there is a motivation for giving the control over data back to the users and not have one entity access to all personal data of the participants in the social network. With a peer-to-peer approach, decentralization is a given, and combined with appropriate encryption users can determine whom they allow access to their data.

While access control by encryption for user data privacy would be possible in a centralized system, it does not go as far as a peer-to-peer approach in ensuring user control. First, whoever would be willing to provide the centralized service and infrastructure would also be able to cease to provide the service or change its terms. Second, due to the lack of data mining and advertising possibilities, there would be fewer incentives and means to provide a good service and all the servers necessary for a centralized solution. Third, a centralized service requires more trust by the users than a distributed system that limits the risk of privacy breaches by not providing a central repository of user data, so that only a small fraction of protected data may be exposed at any time, should the encryption be broken.

In addition to addressing the privacy aspects in general, there is an opportunity to support a non-commercialized self-organized service. Web-based centralized online social networks today bring together the social sphere of family and friends with the commercial sphere. This combination enables targeted advertising thanks to profile information and data mining and thus based on a person's revealed preferences and extends it to a more precise targeting by taking into account social information. We envision a peer-to-peer social network that separates these spheres and enables users to maintain their social network without commercial prompting by advertisement.

User control of data, as provided by a peer-to-peer and secured social network, has consequences beyond privacy and freedom from advertisement. One such consequence is that users can also exercise control over the content they create in terms of intellectual property. User control in this sense means control over who can access their content and what they are allowed to do with that content, e.g. access control can be combined by licensing models of the user's choice (e.g. creative commons licenses) allowing for flexible content rights, as opposed to the current practice of copyright for the online social network providers. Likewise, users also can enjoy freedom of speech, without fearing censorship or other obstacles (like a subscription fee), which a central service provider can impose at its whim.

Another aspect of control is how the social network can be accessed. Moving down the layers from application to network to physical access, there is another instance of peer-to-peer paradigm suitability that has been overlooked: decentralized access via various means (such as direct exchange, as in opportunistic networks) for ubiquitous social networks as opposed to those limited to the web.

2.1 Privacy Trade-Offs

Online social networks have several goals. Besides keeping track of the whereabouts and activities of members of one's social network, they provide spaces for identity creation and expression, a face to the public (with different levels of openness according to the specific social network). They also allow to reconnect people to one's social network, e.g. search for old classmates that are not in the active real-life social network. A centralized website can bring exposure more easily.

There are several attitudes to privacy, ranging from a concern to only allow access to a well-selected circle of friends, to unawareness or indifference versus privacy issues, to a rejection of privacy in favor of the opportunity to present one's chosen identity or persona to a large audience. Given these aspects of identity, public, and accessibility to find and be found, a web service seems to be a more natural fit than a peer-to-peer system. Yet it only addresses the exposure toward other users and offers no protection against the service provider itself.

The commercial benefit for a provider of social network services is more present and more easily available in a centralized system, where the application provider has control of the data. Data mining is both more efficient and more effective in a central repository where the structure is determined by the same entity that provides the application and mines the user data. These commercial reasons can serve as a disincentive for using a peer-based approach when a company wants to provide a social networking service. For community-driven efforts, this business model dependency is not there - it might even be preferable to not have such data-mining potential. In addition, a peer-to-peer infrastructure offers scalability and gradual growth by demand-based adding of resources when peers join.

Social networking sites are not restricted to interaction with friends one already has, but allow users to find others with similar interests, location or organization. By such means as collaborative filtering, affinity groups can be found. For instance, comparing explicit social libraries or implicit usage patterns of book-buying behavior give indications for shared interests. As in general with data mining, a central collection is more efficient. There is, however, a trade-off between search capabilities and privacy, so more efficiency also means a more effective potential breach of privacy.

3 Differences to Other P2P Applications

Peer-to-peer social networks so far mainly fall into two categories: media distribution and collaborative work. The social aspect of social networking, which is a focus on the online social networks, i.e. interacting through applications for fun, leaving publicly visible messages for each other, updating the user status, advertising events or other interests to a group, etc. are not emphasized. There is a chat facility integrated in most of them, but no recorded messages others can see.

Peer-to-peer as a paradigm still seems to be mainly associated with file sharing, even though in variations thereof that go beyond swapping music, such as collaborative work on distributed file editing. This notion of peer-to-peer is rather limited and limiting for social networks, which is reflected in the approaches tried so far.

We want to make the case of extending the peer-to-peer notion beyond file-sharing and take a fresh look at how we can use a peer-to-peer infrastructure to support social networking.

While there are properties common to peer-to-peer approaches that we can take advantage of for peer-to-peer social networks, in this section we focus on what makes the requirements for social networks different from other peer-to-peer services in order to point out where new solutions are needed.

Peer-to-peer storage has been done successfully for file-sharing. These results can, however not be directly applied to social networks: In P2P file-sharing, any copy of a music, video, media file, potentially present in high numbers, will do. These copies are usually not updated, although new versions or different content get added to the system. In social networks, information, such as the current status of a person, is updated often and it makes a difference which version gets downloaded, the value of outdated information is much less than that of timely information that enables users to react to content changes.

In most file-sharing systems, files are not encrypted. When they are [6, 4], mostly for reasons of plausible deniability, the keys to the files can be obtained. For peer-to-peer social networks privacy is even more important as it concerns personal information, so storing content unencrypted at other peers whom the user does not want to access personal information is not an option. Files should only be readable by peers that are specifically allowed to access them.

This access control has been also required for peer-to-peer collaborative work support (CSCW), albeit for smaller groups than the typical user base of online social networks. Work colleagues are added or removed from a working group at a lower rate than expected churn for social networks. For peer-to-peer social networks we need a way of dealing with dynamic relations, that is churn both in terms of online/offline behavior and of adding/removing friends and corresponding access rights.

This dynamic behavior coupled with a different distribution of interest add requirements to availability. In file-sharing, a file is potentially of interest to a large population of users, a Zipf-law distribution of file popularity and thus download availability has been observed. For social networks, this distribution is expected to be different and often limited to a number of friends that is not directly correlated with the network size. The information about a person is of interest to their social network, not typically the general public, there are different economies of scale and scope at play.

Who is interested and authorized to access in social networks also differs from peer-to-peer backup/storage systems, where there is typically one owner of data, so granting access rights and key management is much simplified. Peer-to-peer storage for file backup has been addressed by numerous systems such as Farsite [2] for individuals but these do not address the issues arising from social networks, nor utilize the opportunities of ingrained mutual trust.

For social networks, we need a large-scale peer-to-peer network with fine-grained access control for reading and writing, with changing files (versions), small number of interested peers compared to overall population, and enable a list of features including file-sharing, chat, news-feeds, public and private asynchronous messaging, search, notifications. This means that there are components we can take from other peer-to-peer services, but need to modify and extend them as described above.

4 Challenges for P2P Social Networks

In this section, we formulate goals, requirements and challenges for peer-to-peer social networks. While many of them are technical in nature, some are trade-offs that depend on preference, such as whether privacy should be prioritized or search. Such trade-offs are closely connected to technical questions and are thus discussed together next.

4.1 Application Level Goals

Realizing social networks in a P2P infrastructure paves the way for features that go beyond replicating what is offered in online social networks and that take advantage of the inherent distributed nature to enable user control and ubiquitous access, as follows.

- Social networks without centralized third-party repository,
- with data accessible - either for reading or manipulation - only to authorized members of the particular social network (not crawlers or companies, etc.),
- available even when individual members are offline,
- accessible via a variety of interfaces, online or using direct exchange between devices
- on an open platform with possibility to add applications.

4.2 Underlying Functionality Requirements

How can we have the benefits of existing online social networks, but preserve data ownership by the user and be accessible from anywhere? To that end, how do we store data, keep it up-to-date, and control access to the data?

4.2.1 Social Network Features in a P2P Infrastructure

A first step is to take the existing functionality of online social networks and map them to a P2P infrastructure. This requires finding ways for distributed storage of data, updates propagation and versioning, a topology and protocol that enables search and addressing, i.e. a mechanism to find friends in the topology, robustness against churn, openness for third-party applications, and means for content revocation (by encryption and/or time).

Storage. Where should content be stored? Only at friends? Encrypted and at random nodes? In a DHT? As in file-sharing, there will be several answers to this question. The requirement for redundancy to provide availability of data depends to a large extent on the duration and distribution of time peers are online. These activity patterns are also influenced by the geographic distribution of the peers and shifted by time zones. The distribution of interested and authorized peers and the desired probability of availability are to be traded off with storage requirements, especially if the system should allow for storing of media files and not only links to websites where such media files can be found. To quantify these factors and trade-offs, we are currently conducting experiments on PlanetLab with a simple prototype of social network peers. As a starting point for evaluation, we store data at peers that requested the data (friends).

Updates. How can we deal with updates, e.g. status updates of friends? In peer collaboration systems like Groove, updates, e.g. of a workplace, are sent to a small group of peers via a peer-to-peer synchronization mechanism. In P2P social networks, with distributed storage and replication - and a potential need for scalability, the requirements change. P2P publish/subscribe mechanisms are a possibility, but their security in terms of access control will have to be developed further. Unlike a traditional peer-to-peer environment, where many peers are involved, each of the subnetworks will be much smaller (though larger than typical collaborative groups), making it relatively simpler to realize quorum systems and deal with updates. We will evaluate the use of dynamic quorums [11] within each group of friends.

Topology. Should nodes be connected according to their social connection? This would cluster friends in the overlay network, which would facilitate updates. As a downside, given the possibility of a relatively small set of friends, this would limit the availability and robustness of data access. How can we build a peer-to-peer topology suitable for social networks? In pure file-sharing networks, the topology does not depend on whether the peers know each other and nodes exchange content with any other nodes in the network. At the other end of the spectrum, existing examples of peer-to-peer social networks (in the widest sense) are mostly platforms for collaboration or media sharing and they tend to consist of collaborative groups that are relatively closed circles, e.g. using a “ring of trust” or darknets [7]. In contrast, online social networking services have overlapping circles. A person is in many circles, proportional to the number of friends (one hop in the social network) and group memberships. Potentially multiple overlapping darknets is thus a potential primitive to build a peer-to-peer social network. For our first proof-of-concept implementation we chose a two-tiered approach, with peers connecting to each other for information exchange but also connecting to a DHT as a lookup-service.

Search, Addressing. Related to the topic of updates above, how can users find their friends from the real social network in the P2P virtualization thereof, and conversely, how can they discover new friends by virtue of common interests. Over multiple sessions, peers may change their physical address. In a typical file sharing network, this is not an issue. One just needs to find some peer with the content it is looking for. However friends and trust links of a social network are essential, and so it is crucial to both be able to find back friends even if they have changed their physical address, and also authenticate the identity. Traditionally, peer identity is tied with an IP address [13] which clearly is not sufficient. However, handling peer identity in a self-contained manner in a P2P system is also feasible [1] (also partially in Skype [15]). It may also be difficult to maintain a complete ring like in traditional structured overlays as an index structure, if the network is based on only social links. Recent advances in realizing distributed indexing with a ringless overlay [8] potentially holds the key to this issue. These mechanisms composed together potentially can help maintain social network links under churn. Another search issue is - how can users find out about information available concerning their interests? In social networks, tagging or folksonomies is the basic mechanism to annotate content. Recently, there have also been advances made in enabling decentralized tagging [10], which paves another step towards realizing social networks on top of a P2P infrastructure. Note that there is a trade-off between privacy protection and search capabilities.

Openness to New Applications. One of the most alluring features of current online social networks is that they are open to third-party applications, which enables a constant change of what a social networking service provides to the users. There is a core functionality for maintaining social ties, such as profile information, connection to friends, status updates, internal messaging, posting on each other’s sites, events notification. In addition, third-party applications provide more and unpredictable ways of contacting users, finding out about other users’ interests, forming groups and group identities, etc. This openness to extensions potentially provides great benefits for the users. The price for these benefits is the risk that comes with opening the service to untrusted third parties, extending the privacy problem from the single service provider to all application providers. In a peer-to-peer environment, if some users choose to enable a third-party application, their choice should not affect other users or even users connected directly to them. How to draw this boundary is an open and challenging question.

In this paper we have focused on how to replicate the core functionality of current online social networks and add user control of data and access.

4.2.2 For User Control

Keeping control over their data with the user implies the need for security support, so the classical requirements for security (confidentiality, access control, integrity, authentication, non-repudiation) apply, albeit modified for the context of peer-to-peer social networks.

Security. The main questions for user control are in the domain of access control, e.g. how can we ensure that only authorized friends can access content. For distributed storage with other peers that the user not necessarily wants to access data, the content has to be encrypted, as done for example for file backup [2] or anonymous peer-to-peer file-sharing [6, 4]. To manage access to encrypted data, key distribution and maintenance have to be handled such that the social network group can access data but be flexible enough to handle churn in terms of going offline and coming back, additions and removal to the user’s social network. Group membership research has dealt with questions of key management and renewal and how to give access to new members of a group by issuing new keys in rounds [16]. Likewise darknets [7] also share a key within the group. Such existing mechanisms are however grossly inadequate to meet the finer granularity of access control needs for social network features.

Even in the most simple scenario, where all members are allowed full access, if one wishes to realize control on membership itself, then sharing a secret key is not enough. Any member who already has the shared key can pass it on to new members. Therefore, keys and identities need to be combined for access control, but without access to a file system, mechanisms like access control lists are not feasible. In many online networks (for example Yahoo! Groups[1]), a smaller subset of members own and moderate the membership of a group. Thus even a minor variation of the basic

[1] http://groups.yahoo.com/

groups like darknets, to realize a group where all members still have equal access to content, but only a subset of members control the membership itself, is non-trivial in a decentralized setting.

There is on top of that the need for a finer granularity of access control, determining who can read, write or modify and delete each shared object, and how to enforce such access control in a decentralized setting, while still guaranteeing non-repudiation as well as preventing impersonation and replay. Achieving such finer granularity of access control in a decentralized manner is, we believe one of the hardest security challenges, and the biggest hurdle in realizing a P2P infrastructure for social networking applications.

Other security issues like prevention of DDoS and Sybil attacks, enforcing cooperation and preventing free-riding or content pollution, and establishing trust are also of course long-standing issues in the community, but since they have been in the spotlight for years now, we do not highlight these here. That of course does not mean that these are trivial, or even practically solved. However, in the social network context, some of these issues may actually get simpler to deal with [17].

For peer identities, we would like to take advantage of opportunistic networks and peer authentication by in-person contact, when friends meet in real life and exchange keys over their phones. For bootstrapping authentication, a central authority (trusted third party) seems hard to avoid. Again, authentication from darknets will be a starting point to work from.

Robustness. Against misbehavior: In a centralized system, one can turn to the provider in case of user misbehavior, there is usually a process defined for dealing with such complaints. In a peer-to-peer system, there is no authority that can ban users for misbehavior or remove content. Robustness against free-riding: Without the monetary incentive offered by advertising, other incentives have to come in to make users shoulder the responsibilities for keeping up the infrastructure, providing storage and ensuring availability by staying online. Robustness and Trust: Once access to content is granted, it is difficult to revoke that right. When a user allows a friend to see a message, the friend can store the message and keep access to it even after a change of key. Trust has to be at least equal to assigned access rights, due to this difficulty.

4.2.3 For Mobility, Ubiquitous Access

To take advantage of the peer-to-peer nature of social networks, a mapping of physical social network to virtual and vice versa enables extensions to offering access via web browsers by phone applications and direct exchange of data in physical proximity.

Limited Peers. A major impedance to widespread adoption of a system like we envision will be users' reluctance to install yet another software. Consequently, it will be essential to allow for two classes of users, a core network of users who run the P2P software as well as a web service front-end, and the other, who are essentially clients accessing this service. This of course throws open Pandora's box with lots of questions, including technological feasibility as well as game theoretic issues like incentives and fairness in such a two tier system. Another immediate benefit however of allowing such two-tier system is that users can then participate in the social network with resource constrained (e.g. mobile) devices, which they may use as an auxiliary, even when they contribute resources to the core P2P network with their primary device.

Locality. Using direct exchange between devices, real-life social networks can be used to support the peer-to-peer social networking application. In addition to such opportunistic networks between users, a distributed architecture also enables us to take advantage of geographic proximity and its correlation with local interests. For example, most access routers for home Internet access now come with USB slots where storage can be added or they already have unused storage on the device itself. These routers are typically always on and thus would provide some stability for availability of data of local interest. This local interest can arise from the locality of events but also from the locality of typical real-life social networks of friends and neighbors. How to best harness this locality remains to be seen.

5 Related Work

Although there have been some attempts at peer-to-peer social networks, the term has been used so loosely that these attempts have turned out to have a much narrower focus than the support of the equivalent of online social network services that we have in mind. The core aims of previous attempts at P2P social networks can be classified into file-sharing and collaboration. In the former, a social community is built to exchange files (mostly music, often created by community members) in a trusted circle of friends. An example for this category is Soulseek, which works similar to Napster, with a central directory and peer connections, is mostly used for music, peers as producers and consumers. Similarly, SpinXpress provides file-sharing and communication, a discussion forum and wiki, and search for creative commons licensed media. Examples for collaboration and communication are Groove, which provides shared workspaces and peer-to-peer synchronization, and Kerika, which uses JXTA and has a "ring of trust".

In the past, there have been some more attempts at using peer-to-peer infrastructure for social networks in the widest sense, that have been stopped. For instance, iMeem ceased to provide a peer-to-peer complement to its web access. Wisebook, which used to be integrated with Facebook for peer-to-peer file exchange, is now defunct. The most recent attempts that was ceased is Allpears, a darknet file-sharing system that planned to provide social network features and worked by browser extension and BitTorrent. The reasons for the discontinued attempts vary from legal (file-sharing) to commercial (lack of funding for decentralized system that reduces commercial use of user data) to pragmatic (users prefered web access to installing clients).

Research papers that address at least parts of our requirements are discussed in the relevant context in the previous sections. To recap, the following peer-to-peer applications share a subset of properties with peer-to-peer social networks as we envision them, but differ in fundamental ways that limit their applicability: P2P storage [2], P2P e-mail, P2P publish/subscribe, anonymous P2P file-sharing [6, 4], darknets [7], group communications.

6 Conclusions

In this paper, we make a case for using a peer-to-peer infrastructure for social networks to address problems stemming from a centralized service-provider owned approach, such as privacy and access limitations. We listed a variety of research challenges and opportunities that result from a shift to peer-to-peer, e.g. security issues to enable user control of data, storage, topology, search, and update management. A peer-to-peer infrastructure for social networks offers advantages and opportunities for several parties concerned with the phenomenon of online social networks. Users can keep the benefits from traditional online social networks while enjoying control over their privacy and intellectual property, and complement the web-based access by direct exchange. Communications providers can offer additional services that support the exchange of user-generated content without dependency on third-party application providers. For research, there are many intriguing questions on how to move peer-to-peer from its traditional application of file-sharing to social networks. For a paradigm such as peer-to-peer, most notorious for file-sharing under murky legal circumstances, supporting online social networks offers a fresh opportunity to show its strengths. Mapping the peer-to-peer nature of social networks onto an equally peer-to-peer overlay infrastructure and a diversity of underlying networks, including peer-to-peer opportunistic networks, provides a way of using the same paradigm on several layers, opening the door for interdisciplinary research ranging from social network analysis via peer-to-peer networks to opportunistic or delay-tolerant networks.

For details about our own effort on realizing peer-to-peer social networks see http://www.peerson.net.

Acknowledgements

We thank Balachander Krishnamurthy for his helpful comments on an earlier version of this paper, Doris Reim for her work on topology and a prototype, Youssef Afify for his work on access control, and Le Hung Vu and Niraj Murarka for discussions on the topic.

References

[1] K. Aberer, A. Datta, and M. Hauswirth. Efficient, self-contained handling of identity in peer-to-peer systems. *IEEE Transactions on Knowledge and Data Engineering*, 16(7), 2004.

[2] A. Adya, W. Bolosky, M. Castro, R. Chaiken, G. Cermak, J. Douceur, J. Howell, J. Lorch, M. Theimer, and R. Wattenhofer. Farsite: Federated, available, and reliable storage for an incompletely trusted environment, 2002.

[3] M. Aspan. Quitting Facebook Gets Easier, Feb. 2008. http://www.nytimes.com/2008/02/13/technology/13face.html.

[4] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong. Freenet: A distributed anonymous information storage and retrieval system. In *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pages 46–66, July 2000.

[5] N. Daswani, H. Garcia-Molina, and B. Yang. Open problems in data-sharing peer-to-peer systems. In *ICDT '03: Proceedings of the 9th International Conference on Database Theory*, pages 1–15, London, UK, 2002. Springer-Verlag.

[6] R. Dingledine, M. J. Freedman, and D. Molnar. The free haven project: Distributed anonymous storage service. In H. Federrath, editor, *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*. Springer-Verlag, LNCS 2009, July 2000.

[7] J. Frankel. Waste P2P Darknet, 2003. http://waste.sourceforge.net/.

[8] S. Girdzijauskas, W. Galuba, V. Darlagiannis, A. Datta, and K. Aberer. Fuzzynet: Zero-maintenance Ringless Overlay. Technical report, 2008.

[9] J. Golbeck. Quechup: Another Social Network Enemy!, Sept. 2007. Oreillynet.com.

[10] O. Gorlitz, S. Sizov, and S. Staab. Pints: Peer-to-peer infrastructure for tagging systems. *IPTPS*, 2008.

[11] N. Lynch and A. Shvartsman. Rambo: A reconfigurable atomic memory service for dynamic networks, 2002.

[12] J. C. Perez. Facebook's Beacon More Intrusive Than Previously Thought, Nov 2007. http://www.pcworld.com/article/id,140182-c,onlineprivacy/article.html.

[13] J. Pouwelse, P. Garbacki, J. Wang, A. Bakker, J. Yang, A. Iosup, D. Epema, M. Reinders, M. van Steen, and H. Sips. Tribler: A social-based peer-to-peer system. *Concurrency and Computation: Practice and Experience*, 20, February 2008.

[14] S. Ratnasamy, I. Stoica, and S. Shenker. Routing algorithms for dhts: Some open questions. In *IPTPS*, 2002.

[15] Skype.com. Skype P2P telephony explained, 2004. http://www.skype.com/intl/en/download/explained.html.

[16] M. Steiner, G. Tsudik, and M. Waidner. Key agreement in dynamic peer groups. *IEEE Transactions on Parallel and Distributed Systems*, 11(8), 2000.

[17] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. Sybilguard: defending against sybil attacks via social networks. In *SIGCOMM*, 2006.