Something Old Is New Again: Reimagining the Oldest Social Networking Platform

Ivan Voras, Marin Orlić, Mario Žagar

Something old is new again: reimagining the oldest social networking platform

Ivan Voras, Marin Orlić, Mario Žagar

Faculty of electrical engineering and computing, University of Zagreb, Croatia
{ivan.voras, marin.orlic, mario.zagar}@fer.hr

Abstract. The phenomenon of social networking has entered the lives of millions of people via the ubiquitous Web-based platforms such as Facebook and Google+ which are centralized platforms completely controlled by single entities. However, a complete parallel message-passing infrastructure already exists and has the benefit of 30 years experience and investments: the SMTP e-mail network, which together with the latest technologies can be utilized to provide a completely decentralized, convenient, private and even fault-tolerant social networking platform. This work in progress aims to design and implement a proof of concept of such idea.
Keywords: social network, protocol, e-mail, SMTP, IMAP, Internet, Web

1 Introduction

Internet-facilitated social networking platforms are a natural extension of the human desire for communication and keeping in touch with people we care about (including to some extent people we want do business with). In the year 2012 it is already considered "normal" for people to participate in and communicate by using large web-based social network services and both "Facebook" and "Twitter" are well on their way of becoming generic words. However, the popular social Web sites today are not communicating and exchanging information in a way even approaching seamless. They are also highly centralized and often highly invasive in their data privacy politics, but due to their huge popularity, even knowledgeable users who would otherwise not participate are drawn into using them by peer pressure. Every introduction of a new social network services (like Google+) contends for the users with every other existing service, and the competition is fierce.

This paper aims to suggest an idea for the implementation of a new social network platform on top of existing infrastructure, leveraging an existing large user base with minimal investments. The core idea is to adapt the service which already has a huge backing, in a way which is mostly seamless with its existing usage; this service is the venerable electronic mail (e-mail).

We discuss the capabilities of e-mail for content delivery in §2, the expected features of such a social network service, with a proposal for such a service implemented over existing e-mail infrastructure (specifically, SMTP and IMAP) in §3.

1.1 The role of e-mail and social networks in modern communication

The electronic mail service is one of the oldest services on the Internet – present in its codified form as Simple Mail Transfer Protocol (SMTP) since 1982 [1]. Though it was originally created to support several transport schemes (TCP, NCP, NITS, X.25) and with addressing mechanisms which do not completely rely on the Domain Name Service (DNS) for e-mail routing, the dominant modern implementations of the SMTP have practically removed all older transport and addressing schemes and support only the TCP/IP variant with simple addressing which is well known to all Internet users today.

The volume of e-mail messages exchanged over the Internet can only be estimated, but it is in any case simply impressive. Estimations for year 2010 [2] set the total number of e-mail messages exchanged daily at 294 billion, of which a staggering 89% (approximately 262 billion) are "spam mail". The yearly total of messages exchanged is somewhere in the range of 90 trillion. Not only are these numbers impressive in themselves but they convey how much e-mail has become a part of the Internet infrastructure on which people rely in their daily lives. They also indicate how robust this simple message passing service is (or must be) to allow such high volume of messages to be globally routed. We can contrast this to the recently prominent "social network" services. The published statistics of Facebook [3][11] speak of around 845 million monthly active users and 1 billion pieces of content (Web links, news stories, blog posts, notes, photo albums, etc.) shared daily, which is around 2% of daily world-wide non-spam e-mail traffic. Of course, while e-mail is strongly used for business communication, Facebook is still used almost exclusively for private social interaction, and Facebook does not have to deal with so many spam messages.

2 E-mail as a Distributed Content Distribution Service

Like many of the early described Internet protocols, SMTP in a large part shares a notion of equality between nodes connected to the network. Each node can receive all messages and route them to other nodes, but SMTP does not specify how the users of those hosts can access these messages. This need is addressed by separate protocols like POP3 [4] and IMAP4 [5], and recently, HTTP via "webmail" services. SMTP servers without additional internal rules (e.g. access lists, firewalls) can accept e-mail messages from any user and route it to any other user on the Internet, users being identified by their e-mail addresses. Today's SMTP implementations basically utilizes the DNS as a routing table, looking up mail exchanger entries (the "MX records") or host name entries ("A records") to get to the IP addresses of destination hosts.

In the modern SMTP sessions, servers do not respond to message reception to the sender until the message is safely written to a permanent storage. This "store-and-forward" scheme helps to achieve robustness. Two types of error messages are provided (for permanent and transient errors), which allow an implementation of a "wait-and-retry" protocol for transient errors, which goes a long way toward ensuring that errors like server overload and storage space exhaustion don't affect message delivery if resolved reasonably quickly. However, the SMTP does not itself actually guarantee message delivery, just that there should be reasonable effort on the side of the SMTP

server to do so. In this, it is similar to e.g. the IP. At best, in case of a permanent error during message delivery, the user indicated as the sender of the message in the message headers will receive a notification about this error.

2.1 Comparison to modern peer-to-peer sharing networks

Most of the similarity between the SMTP-based e-mail message delivery service and peer-to-peer content sharing networks is in how they are formed as *overlay networks* over existing infrastructure and in the way messages are routed between distant nodes. The e-mail service uses the ever-changing, almost *ad-hoc* information from the DNS for message routing. The e-mail service is distributed and peer-to-peer because of the protocol-level equivalence of all the SMTP hosts on the Internet. Barring internal rules and firewalls, each SMTP host can route a message to any other.

An important concept in overlays networks is the separation of addressing between the layer. In e-mail, the user addresses are from a completely separate addressing scheme from the one used to route TCP/IP network packets. In fact, e-mail addresses can reference users and hosts for which there is no 1:1 mapping to the TCP/IP network. One important consequence of this scheme is its independence in the case the transport technology changes (e.g. the switch from IPv4 to IPv6).

Modern e-mail messages contain rich content (images, multimedia, rich text formatting) by making use of the multi-part message format - "MIME" [6].

3 Merging e-mail and social networking services

There is a large volume of papers and best practices on the backend architectures that power large Web sites such as the social networking behemoths, which is very interesting from an engineering point of view but the users are generally attracted to only two features: the ease of use and the size of the community [7]. Among the basic functions offered by social networking services, we will concentrate on these:

- Inviting friends / building a social network
- Exchanging private messages (asynchronous / non-real-time messages)
- Exchanging personal information
- Publishing messages (status texts) and content objects

This basic set of functions seems to support the majority of the social interaction on Facebook and Google+, but many other forms of interaction can be implemented using this basic set. With the provision that "messages" in this case can be complex, carrying multimedia and other complex content, functions such as "events", "fan pages" and even social applications can be implemented using only message passing. We expect the following useful properties from the implementation of a social networking platform over e-mail:

- The easy, unambiguous mapping of persons (or at least their user accounts) to their network addresses by using e-mail addresses, which are well defined, ubiquitous and in inexhaustible supply.
- The ability for Internet Service Providers, other companies and all other entities on the Internet to easily augment their own infrastructure to support the new service and offer it to their users on as-needed basis, while still being independent entities with the ability to offer locally valuable additions to the service. In effect, the proposal will enable the creation of "mini-Facebooks" with well-defined interoperability, while simultaneously mostly solving the biggest obstacle: the huge infrastructure needed for such endevours.
- The robustness which comes from its decentralized nature and years of experience with e-mail infrastructure.
- The increased privacy which comes from moving some of the social interaction features to the user agent, from not having any single Internet entity capable of processing or storing global messages and from a standard which allows users to migrate between different service providers.
- The increased competition between service providers and between user agents in the quality of service, features and presentation, which arises from having a decentralized system.

Within the framework for this new service, the role of the SMTP will be in the transport of specially formatted (but completely valid and interoperable) e-mail messages between hosts on the Internet. The user agents (Mail User Agents – MUA), which may be implemented in any of the currently popular technologies (e.g. as desktop or as Web applications), would distinguish these messages from regular e-mail traffic, parse them and present them in with an augmented user interface which supports advanced features normally not available in e-mail. Practical implementations may or may not use the auxiliary protocols like POP3 and IMAP to retrieve the messages from users' mailboxes (i.e. just as they are currently implemented).

We envision to build the additional features into e-mail messages by making use of the standard MIME specification for multipart messages, formatting added data as XML documents, which has the additional benefit of allowing for simplified presentations of the content so that user agents without the support for the new features can display a basic, simplified e-mail message.

We imagine that the popular practical implementation of this new service will provide user interfaces similar to those found today in popular social networking Web sites and will completely hide the underlying e-mail based nature of the service. That said, the intention behind this idea is *not* to replace the current e-mail service and clients with something new, but to expect that the regular and "social" e-mail messages will coexist but be slightly separately treated by the user agents. One possible implementation of this coexistence we can propose may be in the form of separate user agents observing the same mailbox, with the augmented agent recognizing and, after fetching and parsing the message, hiding it or moving the message from the mailbox. An extension of this idea may be to do this on the server side, with user agents using an advanced protocol like IMAP to monitor [10] separate folders within

the mailbox (this ability is already present in most e-mail servers deployed today in the form of message filters or even spam filters).

An important difference between social networking services and e-mail services is in their privacy assumptions. As opposed to the "free-for-all" nature of the e-mail, social networking services deliberately limit the interaction to that between persons which are inside the network (i.e. "added as friends"). We address this issue by including simple, well-known and standardized public key cryptography technologies as an unavoidable part of the service. Currently, the best candidate for this type of technology is PGP (described in RFC 2440 [8][9]).

3.1 The basic features of social networking services implemented with message passing

We start by requiring each person using the service to have at least one cryptographic public-private key pair per e-mail address usable for the service. This key pair will in many cases be used to support the privacy and confidentiality within the social network. All content-carrying messages are required to contain at least one valid digital signature. In this scheme, the user's public key is a stronger identifier of the person's identity than the person's e-mail address. Joining two users in the commutative "friend" relation can basically be reduced to the two users exchanging their public keys. This is effectively the only point where it is possible for users to receive unsolicited e-mail messages (i.e. spam, junk mail), and an effective set of rules can be used to filter out unwanted message (e.g. by setting the limit on the number of characters in the message, requiring only plain-text introductory messages, etc.). Once a "friend request" is accepted, all further communication between the users will be either signed or signed and encrypted using public key cryptography technologies. Messages received by user agents without valid signatures must be discarded.

Private message exchange is conceptually the simplest case for both the sender and the client. Properly formatted messages are digitally signed by the sender and encrypted so only the receiver can read them.

Personal information is voluntarily entered by users of social networking services and all of it (except for basic information such as the e-mail address and the user's "name") is optional. Exchange of such information can be implemented by broadcasting a message to their "friends," which can be implemented incrementally to increase communication efficiency. Such arrangement also allows tight control of personal information sent to specific users.

The most widely used feature of social networking sites (at least as observed by these authors) is the message publishing feature by which users exchange messages and content within their social network (either the entire network or a subset of it - "circles" or "groups", depending on specific privacy settings). By taking advantage of the MIME message format to contain multiple embedded objects, the messages can contain arbitrary objects in arbitrary numbers, referenced and optionally described in the main message (e.g. a snapshot of a photo album with all the album's photo's thumbnails). Large multimedia objects (e.g. YouTube videos) can be linked to.

We suggest that the user agents' user interfaces follow similar design for functionalities and have overall look and feel similar to current implementations of social net-

work Web sites. In case the user agent is implemented as a Web application, it can also implement “apps” in a way analogous to today's popular products. Other features such as real-time instant messaging can be implemented with protocols such as Jabber (XMPP) [12].

3.2 Future work

As this paper proposes only an idea of a service for which this authors believe it will be useful to a global audience, future works must design and describe specific details usable for its implementation. An effort is under way to create such a specification in parallel with an experimental implementation. There are many fields left unaddressed in this paper, among which are the possibilities offer state management services to other applications, and the opportunity for data mining. These will get addressed in our future work.

The authors wish to encourage communication and discussion which would help develop this idea and solve its edge cases. We invite input from all interested parties.

References

1. J. Postel, “RFC 821: Simple Mail Tranfer Protocol”, IETF Network Working Group, 1982. Available online at: http://www.ietf.org/rfc/rfc821.txt (Accessed 2011-11-15)
2. Pingdom, Inc., “Internet 2010 in numbers”, business report by Pingdom, Inc. Available online at: http://royal.pingdom.com/2011/01/12/internet-2010-in-numbers/ (Accessed 2012-01-05)
3. Facebook, Inc.: “Facebook Statistics”, business report by Facebook, Inc. Available online at: http://www.facebook.com/press/info.php?statistics (Accessed 2010-11-15)
4. J. Myers et al, “RFC 1939: Post Office Protocol – Version 3”, IETF Network Working Group, 1996, Available online at: http://www.ietf.org/rfc/rfc1939.txt (Accessed 2010-11-05)
5. M. Crispin, “RFC 3501: Internet Message Access Protocol – Version 4rev1”, 2003, Available online at: http://www.ietf.org/rfc/rfc3501.txt (Accessed 2011-11-05)
6. N. Freed at al, “RFC 2045: Multipurpose Internet Mail Extensions”, 1996, Available online at: http://www.ietf.org/rfc/rfc2045.txt (Accessed 2011-11-06)
7. C. Lampe, N. Ellison, C. Steinfield, “A face(book) in the crowd: social Searching vs. social browsing”, In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work (CSCW '06)*, published by ACM, 2006.
8. J. Callas et al, “RFC 4880: OpenPGP Message Format”, 2007, Available online at: http://www.ietf.org/rfc/rfc4880.txt (Accessed 2011-11-15)
9. M. Elkins et al, “RFC 3156: MIME security with OpenPGP”, 2001, Available online at: http://www.ietf.org/rfc/rfc3156.txt (Accessed 2010-11-15)
10. B. Leiba, “RFC 2177: IMAP4 IDLE command”, 1997, Available online at: http://www.ietf.org/rfc/rfc2177.txt (Accessed 2011-11-15)
11. Facebook, “Fact Sheet”, Available online at: http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (Accessed 2012-01-05)
12. P. st Andre, “RFC 6120: Extensible Messaging and Presence Protocol (XMPP): Core”, 2011, Available online at: http://www.ietf.org/rfc/rfc6120.txt (Accessed 2012-02-15)