Architectures for Person to Person Communications in Disaggregated Networks

Robert McAdoo, Patroklos Argyroudis, Linda Doyle, and Donal O'Mahony
Centre for Telecommunications Value-Chain Research
Department of Computer Science
University of Dublin, Trinity College, Ireland
{mcadoor, argp, ledoyle, omahony}@cs.tcd.ie

Abstract

Current telecommunications systems allocate identifiers to devices, and end-points. While this has worked well thus far, we believe that as more devices are used to access these systems, people will have a tendency to be mobile across these devices. In this article we review current systems and introduce a new model for telecommunications, which we term Person to Person Communications. Under this new model people are allocated identifiers, which they may associate with any number of devices. We outline why we believe this to be a more effective model than the current one.

1. Introduction

Future telecoms networks are likely to differ from today's networks in that they will facilitate communication on a person to person basis. These networks will allow people to be contacted once they have a point of attachment to the network. They will be organised as a network of networks, which are heterogeneous and not under the control of a single entity. We use the term disaggregated to describe these networks. Whereas today's networks are, for the most part, still aware of geographic location for identification and call routing, this will not be the case in future networks. Instead the focus will be on communication between people, who may be highly mobile. Presence will have an important part to play in these networks as people may not be available for receiving calls at all times. A new network architecture is needed to facilitate these requirements. Such an architecture, however, presents new problems to be addressed. The rest of the paper is structured as follows: Section 2 outlines the issues to be resolved, Section 3 takes a look at relevant telephony and network systems, and finally we will outline a high-level design in Section 4. As our design is currently in its infancy, the focus of this paper will be on the state of the art rather than the design of the system.

2. Issues

Our investigation of the problem domain revealed that the main issues to be addressed were as follows:

- Unique identification of users
- Allocation of identifiers
- Resolution of identifiers to routable tokens
- User privacy
- User control over session initiation
- Security of communications
- Interoperability with other networks

Our goal is to enable communications between people who are mobile across different points of attachment to a decentralised network. For this reason we propose utilising user identities, which are devoid of locational information. As with any network each participant in the network needs to be identified uniquely. Since these identifiers will be used by people to contact others they should be somewhat memorable, as people would be likely provide them to people who they wish to be contacted by. We must, thus, choose a method for identifying users in a unique, and memorable fashion.

These identifiers must be allocated by some authority. We wish for our network to be as decentralised as possible, and so we should not rely on a single authority to allocate all identifiers. We will investigate identifier allocation methods for this purpose in future work, and evaluate which would prove most effective for this purpose.

Once a user has been allocated an identifier, they will wish to be contactable. To achieve this,

0-7695-2650-0/06/$20.00 (c) IEEE

they will need to make their identifier available to others. When two people wish to contact each other today, one will typically obtain a unique identifier for the other using a directory, business card, or some other out-of-band means. They will then provide this to an application, which may resolve it to a routable identifier, if it isn't one already, and utilise that information to contact the other. Upon joining the network a user will associate their identifier with their location information. We must, thus, provide a means to resolve identifiers to routable tokens. In a decentralised architecture this may involve routing across different network types. We must thus have a routing agnostic architecture, which is capable of routing across different underlying network technologies.

Resolving this identifier to a routable token should not however provide any information, which may compromise a user's privacy. For instance, it should not be possible to establish a user's location by constantly performing lookups on their identifier to obtain location data. The user should also be protected from unsolicited calls, and other misuses of their identifier. For example, with the lack of cost involved in making a VoIP call, Spam over Internet Telephony (SPIT) has been recognised as a problem that must be overcome in the future. Thus providing the user with control over who can contact them would be desirable.

For this reason we envision users being able to create profiles, which define their availability to callers. They should be able to define parameters such as who they will accept calls from, at what times they will accept calls, and how these calls should be routed. A user's profile should be easy to update and must be managed in a way that guarantees enforceability.

The decentralised nature of our network creates a number of security issues. Since users of the network are to be allocated their identifiers in a decentralised fashion, we must guard against attacks on the allocation mechanism e.g. the Sybil attack [1] and ensure that only authorised users are allowed access to their identifier. Some infrastructure is, thus, required to provide authentication between users and the network. In addition to authenticating users we need to ensure that the data that is transmitted is protected from eavesdroppers. This is especially important as, in our network, data may be routed over links belonging to untrusted organisations. To this end communication should be protected by an encrypted channel. We must first, however, solve the problem of sharing keys across an insecure channel.

Our system should inter-operate with current telephony networks. Voice over IP (VoIP), for example, has seen sizeable take up in recent years. This has been facilitated in a large part by network operators offering the means to contact users in the Public Switched Telephone Network (PSTN), thereby creating a migration path for users, without losing any of the functionality that they currently enjoy, until VoIP achieves a critical mass of users. Similarly our system must provide a seamless migration path from current telephony systems in a manner that ensures that there is no loss in functionality.

3. State of the Art

In this section we will look at relevant current and emerging telecommunications and network systems. In each case we will take a look at the issues they address and how well they succeed in addressing those issues.

3.1 Public Switched Telephone Network

The identifier used in the Public Switched Telephone Network (PSTN) is a telephone number whose form is governed by ITU Recommendation E.164. The recommendation specifies that a telephone number has a maximum length of 15 digits, and usually takes the form ⟨country code⟩ + ⟨national destination code⟩ + ⟨subscriber number⟩. A National Destination Code together with a Subscriber Number comprise a Nationally Significant Number. Telephone numbers are allocated in a top-down manner. The ITU Telecommunication Standardization Sector (ITU-T) has responsibility for allocating country codes, while allocation of Nationally Significant Numbers is the responsibility of individual countries [2]. Usually there will be a regulatory body, which will allocate National Destination Codes, and ranges of Subscriber Numbers to Service Providers, who will then allocate subscriber numbers to their customers. While PSTN lines are fixed, there are services which allow for portability and mobility. Mobility is typically achieved using call forwarding, which is a service offered by telephone companies and allows a user to specify a telephone number to which their calls should be forwarded. This is a cumbersome way of achieving mobility as a user must enable the service themselves, and then change the details whenever they move between places. Number Portability, usually synonymous with Service Provider Portability, allows a subscriber to keep their old telephone number when they switch to a different Service Provider. There are a number of different techniques for achieving number portability, however

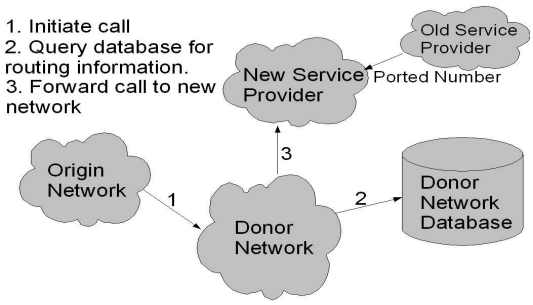

Figure 1: Onward Routing

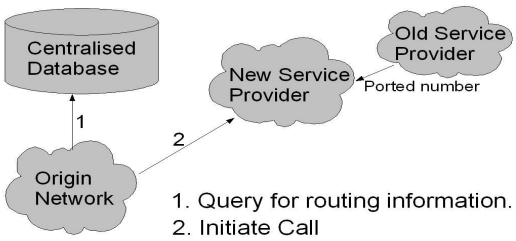

Figure 2: All Call Query

the premise is the same. A database stores a list of all ported numbers, and their routing numbers. This database may be hosted at either the donor network, i.e. the network that was originally assigned the number, or it may be centrally administered [3]. Two Number Portability algorithms are depicted in Fig. 1 and Fig. 2, namely Onward Routing and All Call Query respectively. In this scenario the Origin Network is the network from which the call originates, the new service network is the target network, where the subscriber is contactable, the old service network is the last network from which the number was ported, and the donor network is the network to which the number was originally allocated. Onward Routing makes use of an Internal Number Portability Database (INPDB), stored at the donor network, while All Call Query makes use of a Centralised Number Portability Database. Relying on the donor network is problematic, since if the Service Provider operating that network closes their operations the database will need to be moved, or may be lost altogether. However, relying on a centralised entity is also not strictly advantageous since access to that database is subject to their regulations. Global Number Portability, where a telephone number may be ported to a Service Provider in a different country remains an unsolved problem.

3.2 Universal Personal Telecommunications

Universal Personal Telecommunications (UPT) was introduced by the ITU in 2001, as a service where a subscriber is given a single number, at which they could be contactable regardless of their point of attachment to the telephone network [4]. The ITU allocated the 878 country code for this purpose. A subscriber would be allocated a UPT number from a service provider under this country code. They could then create a

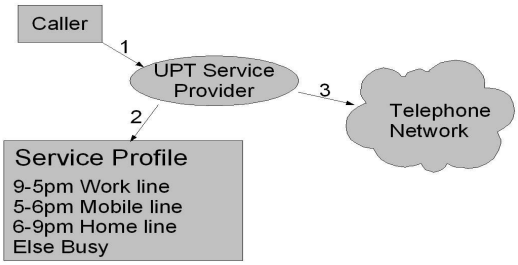

Figure 3: High-level view of UPT

Service Profile, which would specify to the Service Provider where their calls should be forwarded to, and at what times to allow calls through etc. A sample call setup is shown in Fig 3. When the caller phones the UPT number their Service Provider will route the call to the callee's UPT Service Provider, who will then consult the callee's Service Profile, and finally route the call to the correct location based on the rules defined in the profile. UPT solves the issue of identifying users independent of their location by effectively making the identifier a pointer to the location defined in the user's profile. UPT, however, has not attracted the interest of Service Providers, and so remains largely unused.

3.3 Global System for Mobile Communications

Global System for Mobile Communications (GSM) is a Mobile Communications standard, which is currently deployed in over 200 countries. Users are identified by a telephone number known as a Mobile Subscriber ISDN Number (MSISDN). GSM facilitates roaming between different geographic areas by maintaining the current location of a subscriber in a central database known as the Home Location Register (HLR). When a sub-

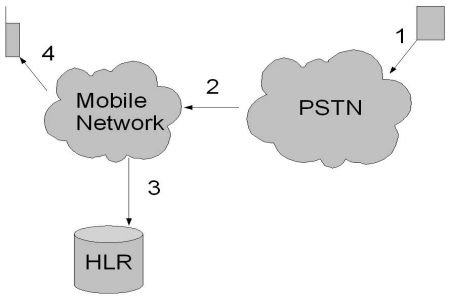

Figure 4: Locating subscribers in the GSM network

scriber is to be contacted the HLR is consulted to find their current location and the call may thus be routed. GSM subscribers are identified in the HLR by an International Mobile Subscriber Identity (IMSI) number. The IMSI number protects the subscriber's privacy from casual monitors of GSM traffic, since their telephone number is not transmitted [5].

Several components of the GSM security system have been shown to be vulnerable. The unpublished cryptographic algorithms used by GSM, the A5/1 stream cipher and the A3/8 Message Authentication Code (MAC), were successfully cryptanalysed by Biryukov et al [6]. The roaming security subsystem was also attacked based on the fact that data exchanged between base stations over microwave links are in practice transmitted in the clear, despite the provision for encryption specified in the GSM standard [7]. Since the home network does not want to reveal the secret key it shares with a subscriber to a visited network, it releases a number of precomputed authenticators, called "triples" in GSM terminology. The visited network uses these to authenticate a roaming subscriber without gaining access to the secret key the subscriber shares with her home network. As links between base stations are not encrypted an attacker can intercept the authenticators exchanged over the microwave transmissions between the home and visited networks. These can then be used to charge telephone calls to the real subscriber.

3.4 Skype

Skype is a Peer-to-Peer internet telephony service. Users of the network contact each other primarily using the Skype software client. The client behaves like an Instant Messaging application, with users receiving notification of their contacts presence information. Skype provides additional services that allow users of the PSTN to contact, and be contacted by Skype users. Users are identified by a Skype-issued username. A P2P network, which they call their Global Index (GI), is used to perform lookups on these usernames to retrieve location and presence information and facilitate call routing. Due to the proprietary nature of Skype little is known about how this lookup is performed, how data is routed, the structure of the Skype network, or its security model although there have been studies into these areas such as [8], [9] and [10].

3.5 Session Initiation Protocol

Session Initiation Protocol (SIP) is a general purpose signalling protocol developed by the IETF. SIP has established itself as the dominant protocol in the VoIP industry, with the majority of VoIP providers using it to initiate their calls, which would then be transported using a media transfer protocol, typically RTP. SIP peers communicate in a Peer-to-Peer manner, and are identified by a URI of the form *sip:username@example.com*. Calls to a SIP URI will mean have the signalling data transported in plaintext, while calls to a SIPS URI will ensure that all SIP data will be transported securely, using TLS [11]. However, because the media stream is separate to the signalling stream in a SIP system, this must also be encrypted. This is generally achieved using a secure media transport protocol such as SRTP [12]. As a text protocol SIP suffers from some performance issues, and the IP Multimedia Subsystem, which we will introduce later in the article, employs header compression to mitigate this.

3.6 ENUM

ENUM is a system for resolving E.164 numbers to URIs. It was designed to use DNS for name resolution, and a separate DNS domain, e164.arpa, was delegated for this purpose. Since DNS uses domain names as identifiers, an E.164 number must be converted to a domain name in order to allow it to be resolved. This works by taking the number, removing any non-digit characters, placing a period between each number, reversing the order of the digits, and appending the ".e164.arpa" string to the end of it. For example the number +35312345678 would be converted as follows:

1. Remove '+' - 35312345678
2. Insert periods - 3.5.3.1.2.3.4.5.6.7.8
3. Reverse order - 8.7.6.5.4.3.2.1.3.5.3
4. Append ".e164.arpa" - 8.7.6.5.4.3.2.1.3.5.3.e164.arpa

5. Result:
   $ORIGIN 8.7.6.5.4.3.2.1.3.5.3.e164.arpa
   IN NAPTR 100 10 "u" "sip+E2U"
   "!.*$!sip:me@example.com!"

This string may then be resolved according to the usual DNS algorithms. DNS was designed to be an open system, and as such there was a conscious decision not to limit access to DNS records in any way. DNS records are typically straight forward to update, and provide a maximum length of time for which they should be cached, the time to live (TTL). Systems such as Dynamic DNS take advantage of this to provide mobility for internet hosts by setting the TTL to a low value, and providing software to users which updates the address records for a specific domain name. Mobility for ENUM users could be provided in a similar manner.

Mockapetris identifies some shortcomings in DNS that ENUM must contend with if it is to be successfully deployed as a bridge between the PSTN and VoIP networks [13]. The main issues he highlights are scalability and latency issues with current DNS infrastructure. He explains how research undertaken by his company, Nominium, demonstrates that the current popular DNS servers would be unable to scale to support the addition of the large number of fixed and mobile telephones to the Internet. He explains that a current latency of a few seconds is not unusual for a response to a DNS request, but that such a delay is not acceptable as an additional delay to the setup time required for a telephone call.

3.7 IP Multimedia Subsystem (IMS)

The IP Multimedia Subsystem (IMS) is a network architecture for enabling the convergence of fixed and mobile telecommunications networks and the Internet. It was proposed by the 3rd Generation Partnership Project (3GPP), and the Internet Engineering Task Force. The aim of IMS is to facilitate access to Internet services for users of cellular networks, and as such IMS runs on IP and uses SIP as its signalling protocol. A user of IMS can access the network using a variety of devices e.g. PDAs, mobile phones, or computers. The designers of IMS recognised that there would likely be a large number of different devices that may be used to access the network, and that people could possibly access the network using more than one device. They also realised that a user may wish to be contactable under more than one identity as they may want to have separate identities for personal and business use. This was the motivation behind the design of Public User Identities, and Private User Identities. A Private User Identity is unique to a device, and plays a similar role to an IMSI in a GSM network. Private User Identities take the form of Network Access Identifiers (NAI), which are strings that are of the format: *username@operator.com*. A Public User Identity is analogous to an MSISDN in a GSM network. They may be either SIP or TEL URIs e.g. *tel:+3531234567*, so that compatibility with the PSTN may be maintained. A Private User Identity will be stored, along with one or more Public User Identities, on a smart card. A user may have multiple smart cards, for use in different devices, thus allowing them to be mobile across devices. In addition, the 3GPP have defined a presence architecture for IMS, which allows users authorised by a subscriber to obtain their presence information. This presence architecture would allow for the deployment of applications such as Instant Messaging, and enable them to combine with other services such as voice. All applications would then benefit from the mobility provided by the network architecture itself.

Due to the centralised nature of IMS, however, the range and number of applications that will be deployed on the architecture is likely to be dependent on the network operators. The operators will most likely exert full control over these applications. This is in contrast to the Internet where networks are generally much more open.

3.8 Google Talk

With their foray into the world of VoIP and Instant Messaging, Google have taken the opposite approach to Skype, and have used open protocols namely the Extensible Messaging and Presence Protocol (XMPP) [14], and Jingle [15], an extension to XMPP that adds support for multimedia streams. XMPP is the protocol used by Jabber, a decentralised Instant Messaging implementation. Jabber identifiers take the form of email addresses i.e. *username@server.com*. A user will register an account with a server, which keeps track of the users location. It will then assume responsibility for delivering that user's outgoing and incoming messages. A message will be delivered to the server indicated in the user's identifier, which will in turn deliver it to the individual user. Jabber is decentralised and allows anyone to setup their own server. In addition to delivering messages to other Jabber users, a server may also run a number of transports. Transports act as gateways between different networks enabling a Jabber user to send messages to people on these networks. At the time of writing Google Talk was still in beta, and as such not all the intended

features have been implemented. For this reason encryption for calls is currently not supported. Google have also not provided a means for users to communicate with people on the PSTN.

4. General Model

In this section we will present our own proposal for addressing the issues identified in the first section of this paper. We will outline the functionality of our system on a high-level. However our system is still under active development and therefore presented in brief.

We will start by discussing the topology of the network. Our network will consist of a set of enterprise networks, which are interconnected by the Internet. Enterprise networks represent administrative domains, and each will have its own centralised control. However, no central entity will exist, that has control over all of the networks. Nodes in different administrative domains will communicate through gateways.

Users may roam between these administrative domains, and for this reason should be identified in a location independent way. Upon joining the network the user will register their location, or a token that makes it possible to route data to their location, and their identifier with the network. Other users may then contact them by performing a search for this identifier, which will return the token that will allow a session to be initiated.

We mentioned in Section 2 that we envision support for the creation of user profiles. In their profile a user will be able to define details such as which users are authorised to contact them, what times other are allowed to contact them at, and what to do when they're unavailable to receive calls. Since this profile affects whether or not a user is capable of contacting the other, it should affect the result of their search. No result, or an unavailable message should be returned when contact is expressly forbidden by a profile. This will prevent circumvention of the profile by malicious users.

Assuming that the profile doesn't expressly forbid a call initiation, and the user is available to receive calls,a call may be initiated. This call will be initiated by SIP or some other session control protocol. As both hosts can route packets to each other the call will commence.

There will be a number of gateways in the network, which will provide call termination and transit to, and from the PSTN to our network. Calls destined for the PSTN can be terminated by any gateway with a PSTN connection. To make calls more cost effective for the user however, gateways should advertise number ranges that they are willing to terminate for, and at what cost. This would enable a client to choose the best match for call termination. Calls from the PSTN would require a gateway, and a number for the user. This number could be their conventional telephone number ported to a company that provides gateway access to the network, or could be a UPT number with a provider that can terminate to the network.

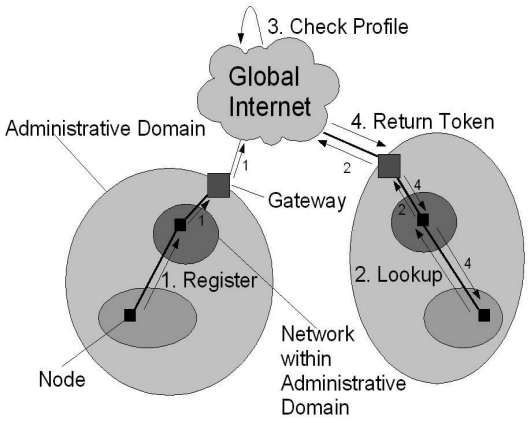

Figure 5: A step by step illustration of communications between two nodes.

Currently we are at the design stage, and we still need to address many issues such as privacy, allocation of identifiers etc. We will be using the NTRG stack as the basis for our development since it provides facilities for decentralised name resolution, routing, and addressing [16].

5. Conclusion

Under the current telecommunications model, a person will contact another at a given identifier. Identifiers, however, are allocated to devices and as such a user may have a number of different identifiers, and may be contactable at all or some of these at any given moment. We propose a new model, where an identifier is allocated to a person and may be associated with any number of devices. A user will then define how this identifier should route their calls according to a profile. We believe our model to be a more flexible user-friendly version, which allows for portability, mobility, and gives more control to the user. We have outlined the high-level issues to be resolved, given an overview of current systems and some of the issues they fail to address, and provided a high-level outline of the functionality of the sys-

tem.

6. Acknowledgements

This material is based, in part, upon works supported by Science Foundation Ireland under grant number 03/CE3/I405.

7. References

[1] J. Douceur, "The Sybil Attack", In the Proceedings of the IPTPS02 Workshop, Cambridge MA (USA), March 2002.

[2] ITU, *Telephone Network and ISDN - Operation, Numbering, Routing, and Mobile Service*, vol. II, CCITT, November 1988, Fascicle II.2.

[3] M. Foster, T. McGarry, and J. Yu, "Number Portability in the Global Switched Telephone Network (GSTN): An Overview", RFC 3482, 2001.

[4] The International Telecommunication Union ITU, "First Mobile Phones, Now Mobile Numbers", ITU Press Release, 2001.

[5] Jorg Eberspacher, Hans-Jorg Vogel, and Christian Bettstetter, *GSM - Switching, Services and Protocols*, John Wiley and Sons Ltd., Chichester, England., 2001.

[6] Alex Biryukov, Adi Shamir, and David Wagner, "Real Time Cryptanalysis of A5/1 on a PC", *Lecture Notes in Computer Science, Vol. 1978*, 2001.

[7] Ross Anderson, "GSM Hack - Operator Flunks the Challenge", The RISKS Digest, Vol. 19, No. 48, 1997.

[8] Salman A. Baset and Henning Schulzrinne, "An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol", Columbia University Computer Science Department Technical Report CUCS-039-04, 2004.

[9] Saikat Guha and Neil Daswani, "An Experimental Study of the Skype Peer-to-Peer VoIP System", *IPTPS '06*, February 2006.

[10] Simson Garfinkel, "VoIP and Skype Security", http://www.simson.net/ref/2005/OSI_Skype6.pdf, 2005.

[11] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, 2002.

[12] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, 2004.

[13] Paul V. Mockapetris, "Telephony's Next Act", IEEE Spectrum, Apr 2006.

[14] P. Saint-Andre, "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 3920, 2004.

[15] Scott Ludwig and Peter Saint-Andre, "JEP-0167: Jingle Audio Media Description Format", February 2006.

[16] Donal O'Mahony and Linda Doyle, *Mobile Computing: Implementing Pervasive Information and Communication Technologies*, chapter An Adaptable Node Architecture for Future Wireless Networks, Kluwer Publishing, 2001.