Spontaneous Networking: An Application-oriented Approach to Ad Hoc Networking

Laura Marie Feeney
Bengt Ahlgren
Assar Westerlund

Computer and Network Architectures Laboratory
Swedish Institute of Computer Science
Box 1263, SE-164 29 Kista, Sweden
`http://www.sics.se/cna`

Abstract: An ad hoc network must operate independently of a pre-established or centralized network management infrastructure, while still providing administrative services needed to support applications. Address allocation, name resolution, service location, authentication, and access control policies represent just some of the functionality that must be supported – without pre-configuration n or centralized services.

In order to solve these problems, it is necessary to leverage some aspect of the environment in which the network operate. We introduce the notion of a *spontaneous network*, created when a group of people come together for some collaborative activity. In this case, we can use the human interactions associated with the activity in order to establish a basic service and security infrastructure.

We structure our discussion around a practical real-world scenario illustrating the use of such a network, identifying the key challenges involved and some of the techniques that can be used to address them.

Keywords: ad hoc networks, zero configuration, spontaneous networks, collaborative software

1 Introduction

Broadly defined, a mobile ad hoc network is a group of wireless nodes that cooperatively form a network that operates without the support of any fixed infrastructure. Such networks have been proposed for a wide variety of goals: data collection in sensor arrays; providing a communication infrastructure in hostile or disaster-stricken areas; forming the basis of a pervasive computing infrastructure, where everything is always connected. Ad hoc networking presents two distinct challenges.

First, the network must operate independently of an access point infrastructure, even though the connectivity among nodes changes rapidly and unpredictably. Connectivity over a larger area is obtained using various ad hoc routing techniques: Each node communicates directly with nodes within wireless range and indirectly with all other destinations using a dynamically-determined multi-hop path via other nodes in the network. Routing protocols for use in this demanding environment have been the subject of a great deal of research [2, 6].

Second, the network must operate independently of a pre-established or centralized network management infrastructure, while still providing administrative services needed to support applications.

In an infrastructure network, address allocation and name assignment are handled by a central authority on the global scale and properly configured automated services within an administrative domain. Alternative techniques must be used in an ad hoc environment.

Basic networks services such as name resolution, file system management and mail and web services are centrally administered. Many general purpose applications are also server based and must be initialized and and configured for each organization's environment by specifying resources and enforcing its security policies. These tasks are performed by human administrators and carefully configured automated network management tools.

In the ad hoc environment, these services cannot be centralized, because the network population and topology are not known in advance. Nor can they be pre-configured, since much of the configuration will not be determined until the network is instantiated and may need frequent updating due to the unpredictable nature of the network. Nodes may need to cooperate to provide functionality like group communication, replicated object storage, caching and remote object invocation.

In the fixed network, security depends on encryption keys, certificates, and authorization information stored in trusted repositories, both local and global. Other security mechanisms depend on firewalls, which block traffic from outside an organization's administrative domain.

In the ad hoc in the ad hoc environment, it is necessary to support key distribution for encrypting traffic in the wireless media, authentication of users and devices in the absence of a global identity database. Once it is possible to identify network participants, users must adapt authorization and access control policies in response to the changing network population. Firewall-based security is not applicable in this context because there is no administrative domain in which to distinguish inside and outside.

Addressing this second challenge is essential to realizing the potential of ad hoc networking. Consider an ad hoc network that acts as a substitute for a wireless communication infrastructure in a militarily hostile or disaster-stricken area. In this case, service location services such as "contact medical support" are particularly important. Additional complexity comes from the fact that network participants may originate from different administrative domains (e.g. relief organizations) and in the field, not only network connectivity, but also service availability, security policy and a host of parameters will change frequently and unpredictably. Clearly, the ability to route traffic from one node to another is not sufficient to obtain useful results. Despite its importance, the problem of administration and configuration in an ad hoc network has received little attention.

Solving this problem requires addressing many hard issues, not least of which is the diversity of application scenarios and requirements. The exact administrative functionality that is needed depends on the nature of the network and the applications that it will support. For example, pre-configured hardware may suffice for the simplest and most specialized networks.

It is not possible to generate something – especially sophisticated configuration infor-

Figure 1: **Spontaneous Networking:** Because there is no dependence on outside resources, people can collaborate anywhere, anytime – simply and securely.

mation – from nothing. It is necessary to leverage some aspect of the network, whether this is some property of the devices, the applications that are intended to run in it or the behavior of the participating users, in order to develop a mechanism for establishing its administrative infrastructure. Taking this application oriented approach, we have chosen as our starting point the configuration requirements posed by the network that is created when a group of people use a collaborative application. In this situation, the human interactions that take place in the context of such an application can be leveraged to create the necessary configuration infrastructure. We call such a network a *spontaneous network*.

The paper is organized as follows: Section 2 motivates the notion of spontaneous networking. Section 3 discusses five fundamental challenges in ad hoc and spontaneous networking. Section 4 presents an extended example of a collaborative application operating in such an environment. In section 5, we discuss key enabling technologies. Conclusions are presented in section 6.

2 Spontaneous networks

We have chosen to focus on a small, but practically relevant, subset of the ad hoc networking space. We are developing techniques for ad hoc networks that are created when a group of people come together and use wireless computing devices for some computer-based collaborative activity. We call these networks *spontaneous networks*. Users do not need to identify all the (human and computational) participants or to administratively configure their computers in advance, while they are still connected to their home networks. Nor are they constrained by the availability or affordability of a pervasive wireless infrastructure or the need for reliable access to their home networks or other infrastructure via the global Internet. Applications such as interactive presentations, collaborative editing, shared whiteboards, and games are ideal candidates for operation in such an environment.

A spontaneous network reflects intentional interactions among the users, who have chosen to collaborate for some purpose. It is this intentionality that can be leveraged in order to create an ordered method for initializing the network configuration.

By their nature, spontaneous networks have a limited extent in both space and time. They are comprised of powerful host machines, such as laptop computers or emerging high end PDA's and mobile phones. The nodes are connected using a variety of wireless technologies, such as IEEE 802.11, IR, or Bluetooth, meaning that bridging between heterogenous interfaces will be an important part of multihop routing. Although the population and topology of a spontaneous network are unpredictable and dynamic, it is expected that significant changes will be relatively infrequent compared to the lifetime of the network.

Because wireless connectivity is based on physical proximity, it reflects the way that humans interact. People who are near each other can communicate, can exchange things with each other, and can ask people to relay information to others. In particular, authentication and trust can be based on "first-person" interaction, rather than relying on remote, centralized administrative services. In addition, because collaborative applications are intended to reflect the users' interactions, making their basic structure and human-computer interface suitable for use in an ad hoc environment.

Our goal is to empower users to create computer networks with the same ease, flexibility and spontaneity as the human interactions they are intended to facilitate. The ability for users from different organizations to meet anywhere, open their laptops, and begin working together – simply and securely – has wide-reaching practical value. To achieve this, we attempt to leverage human interactions to create the needed shared context.

3 Five Challenges

In this section, we summarize five key challenges posed by the spontaneous networking environment. Each reflects some way in which a spontaneous network differs from other well-known infrastructure or mobile computing environments.

Network boundaries are poorly defined. In a conventional wired or wireless network, "coming up on a network" simply means enabling the network interface and connecting to (and possibly authenticating with) the networking infrastructure. This well defined step is expected to put the node into (reasonably reliable) contact with all necessary services. In a wireless ad hoc network, there is no natural equivalent: the network may experience arbitrary partitions and merges.

The network is not planned. Conventional networks are not built on an ad hoc basis: their design is part of the art and science of system administration. Logical and administrative boundaries suggest where services should be hosted and replicated.

In infrastructure networks, firewalls are often used to create security boundaries. The ad hoc networking environment highlights the inadequacy of such a model. A node must create a network infrastructure in cooperation with untrusted partners: it cannot define a (non-trivial) boundary inside which it accepts and outside which it filters. This suggests that both ACL-based security and simple and intuitive methods for manipulating ACL's will be of great importance.

Hosts are not pre-configured. It should be possible to create an ad hoc network anytime, anywhere and with any participants. This severely limits the amount of administrative or configuration information that can be pre-configured on a node.

Information such as the NAI's [1] or user principals of the participants, the host names and addresses of the nodes and available services and the nodes where they are hosted cannot be pre-determined. Therefore nodes and services cannot be pre-configured with this information.

This has especially important implications for security, which usually relies on availability of a trusted key management infrastructure.

There are no central servers. Servers are problematic because nodes that become partitioned from the server must agree to either promote a backup server or reinitialize the service. If two partitions merge, the servers in each partition may need to synchronize their state. The exact nature of this synchronization is service dependent.

This is not the same as operating while disconnected from a home infrastructure. Within an ad hoc network, there is no well defined primary version of an object, to which all modifications must eventually be committed. In fact, users can be disconnected from their home infrastructure and operating in a spontaneous networking environment at the same time.

- Consider a user who has files cached on her laptop (e.g. Coda [7]) before disconnecting from her home infrastructure to attend a conference. She participates in a number of collaborative editing sessions operating in a spontaneous networking environment. When the various authors return home, they independently reintegrate the documents into their home infrastructure.
- Consider a user who wishes to send mail to a colleague (who is also part of the spontaneous network). He may attempt to send it via the spontaneous network before queueing it to send via the infrastructure network.

Users are not experts. Operation must be intuitive to non-technical users. Users are notoriously bad at configuration, especially in an environment that poses complex security issues. Therefore, it is important to minimize the inexpert user's exposure

[1] Network Access Identifier, a network-global identifier for a user [1]

to the administrative infrastructure and make necessary activities as intuitive as possible.

Many of these challenges are also seen in the area of pervasive or ubiquitous computing [11]. A ubiquitous computing environment is a pervasive computing infrastructure in which people interact seamlessly with the many devices surrounding them. Although such an environment has many dynamic elements, work in this area often assumes a fixed infrastructure for purposes such as authentication, configuration or even attempting to anticipate the user's needs or actions based on previous history.

4 Example

In this section, we describe a scenario that illustrates the practical value of spontaneous networking as we have defined it here. It will also highlight the challenges that need to be solved in order to implement spontaneous networking.

At home

> Ms. Spont of Acme, Inc. is working from home. She makes some final adjustments to her presentation and makes sure that her laptop has the latest copies of all project documents. She then leaves for an off-site project meeting with Partner, Inc.

Ms. Spont has a small wireless LAN in her home, allowing her to work from anywhere in the home and possibly also to control household appliances such as her VCR. The wireless LAN is connected to the Internet with a fixed, high-speed connection.

When she accesses her corporate account, she is connected to a VPN maintained by Acme, Inc. The relevant security associations have been previously established between Ms. Spont, the service provider, and Acme Inc. Network services such as interface configuration also depend on the administrative infrastructure of these organizations, although Ms. Spont need not be aware of these issues. This is not spontaneous networking.

Coming to the meeting

> Ms. Spont arrives at Partner, Inc. for the project meeting. She is early and the rest of the project members have not yet arrived. While she waits, she uses her laptop to read this morning's e-mail.

Ms. Spont is not permitted to access the wireless LAN at Partner, Inc. because she is not an employee. Ms. Spont's GPRS mobile phone provides the Internet access necessary for accessing her mailbox at Acme Inc. This is also not spontaneous networking. Ms. Spont, or more specifically her GPRS phone hardware, is an established customer of her GPRS service provider. There is infrastructure support for interface configuration, to allow her authenticate herself to her employer's network, and, of course, for her account to be billed.

Setting up a secure association

> After a while, Mr. Host of Partner, Inc. arrives together with the rest of the project team, who come from a number of organizations. They all move to the meeting room.
>
> In the meeting room, Mr. Host uses his PDA to create a secret session key for the duration of the meeting. When Mr. Host uses his PDA's IR link to exchange electronic business cards with all participants, he also simultaneously distributes the session key. Using this key, the participants create a (moderately) secure infrastructure-less wireless network, or "spontaneous VPN". In the same process, addresses are assigned and useful names, such as "Mr. Host's PDA", are made available to the participants.

The session key is essential in the creation of a private network. The key is sent in clear text, once per receiving unit, using short range infrared communication. Because IR does not travel through walls, no-one outside the meeting room can overhear the secret session key. Once distributed, this session key can be used to secure RF communication,

While this mechanism is clearly not suitable for highly sensitive data, it conforms with most people's intuition: a closed room is a reasonably private setting. It also correlates a human interaction – exchanging business cards – with access to the data exchanged during the meeting.

Note especially that no infrastructure or pre-configuration of any kind is used in the establishment of this "spontaneous VPN". The participating nodes collectively perform address allocation, and establish a name database.

The project members' equipment does not need further authentication from any outside infrastructure. The human identification of the persons together with their equipment and the session key is sufficient to establish the secure association. It does not really matter what computer is present at the meeting. Ms. Spont can use any computer or PDA that she brought as long as she gives it the key she had obtained earlier.

Sharing documents

> The project team uses a collaborative application which allows the participants to both look at and edit specific shared files.

Collaborative applications have been the subject of research and commercial development for many years. In the case of a spontaneous network, the network must detect and respond to changes in connectivity and in service availability. It also necessary to provide facilities such that applications can be informed of and respond to such events.

Moreover, if the hosts have been configured to share documents in a "safe" – that is, firewalled – environment, they may be inappropriately configured for the spontaneous network.

Network partition and merge

> During the meeting, the project team is divided into two groups. One of the groups moves to another meeting room. The two groups lose wireless connectivity with each other, but continue to work on the shared documents.

The people who move may suspend their laptops while moving to the other room. When the laptops are resumed, they have no information about the state of the spontaneous network, or even whether they are still part of the same spontaneous network or have been resumed in an entirely different context.

The collaborative application must be able to handle multiple replicas in the partitioned network in order to recover the state of the group prior to partition.

> One more person, Ms. Late, joins the group in the original meeting room where Mr. Host is. He exchanges electronic business cards with her, transferring the secret session key, which includes her in the spontaneous VPN for the meeting.

Note that Ms. Late joins the group while the network is partitioned. Duplicate address detection can thus only reliably be done within the local partition. There could be more members joining in the other partition which the nodes in the local partition do not know about.

> When the team gets back to the meeting room, the two partitions of the spontaneous VPN merge. The collaborative application reconciles the document replicas that the groups have independently edited.

Ms. Late's equipment can now for the first time communicate with the whole network. Duplicate addresses and naming conflicts are issues. It is necessary to synchronize both the state of various network services and the objects maintained by the application. Naturally, the collaborative application may need manual intervention to be able to reconcile the documents.

5 Enabling Technologies

It is not necessary to invent completely new technology for realizing spontaneous networks, there are existent or emerging technologies that can be used. This also has the advantage of making the interconnection between infrastructure and spontaneous networks easier. We will review these technologies and try to explain how they fit in the larger picture, noting how they address some of the challenges discussed in section 3, as well as their current limitations for use in ad hoc networks.

Automatic/Dynamic configuration

Since the spontaneous network cannot rely on any central servers or on any pre-configuration and has to handle dynamic network topologies, common ways of handling address assignment such as DHCP [1] will not work. The assignment of addresses needs to be automatic. There are two existing technologies being developed within the Internet Engineering Task Force (IETF) that can be leveraged to implement this: IPv6 [1] stateless address configuration and Zero Configuration Networking [1].

IPv6 provides a mechanism for obtaining a unique address with or without the help of a router (stateless autoconfiguration). This address is derived from the low-level MAC address encoded in most network interface cards or is a randomly generated address, a technique which the 128-bit IPv6 address space makes feasible. In the IPv6 model, it is also natural to have a changing set of addresses associated with an interface, which is useful when a node is part of several different spontaneous (and non spontaneous) networks. For these and other reasons, IPv6 is widely expected to play a key role in the development of many aspects of the global wireless network.

The Zeroconf Working Group is also studying the problem of enabling networks that do not require any configuration or administration. The group's charter includes interface configuration, naming, service location (see below) and multicast address allocation. Unlike ad hoc networks, zero-configuration networks are defined as including at most one router, i.e. a star topology. Moreover, many zeroconf environments are envisioned as somewhat static networks, in which configuration results in a well-known, relatively steady state.

Because ad hoc But environments are not administered, applications and nodes cannot assume that they know what services are available and where they are hosted. This service location information has to be obtained at run-time. Both SLPv2 [1] and Jini [9] define decentralized service location mechanisms without relying on servers to store information about what services are available where.

However, service *location* protocols are not sufficient to enable network services in a spontaneous networking environment. Service management in decentralized environment also requires mechanisms to decide which nodes *should* host services, to determine how a service's initial configuration and databases are initialized, and to ensure a service's availability and database consistency in the face of partitions and merges.

Security

Security as implemented in a conventional network relies on central servers to certify parties that want to communicate to each other. This cannot be the case in a spontaneous network.

Authentication in the spontaneous network can leverage the fact that the network is created by people, who inherently implement complex trust models while interacting with each other. This means that a security association can be initiated using “first person” authentication, as seen in the example above. We anticipate that this principle can be extended to support larger and more complex situations.

Another example of a security model that initializes security associations via close physical contact is the “resurrecting duckling” [8] system, in which simple devices are imprinted with an association to another device.

Peer-to-peer operation

There is a substantial body of work on mobile computing environments that support disconnected or weakly connected access to a file system [7] or other system services [5]. Most work in this area has emphasized support for individual users who are disconnected from or have only intermittent or low bandwidth connectivity to their home infrastructure. Techniques are generally based on optimistic concurrency, using tentative updates that are committed when they have been confirmed at a “primary” server.

A more general case of disconnected operation uses a peer-to-peer model, as in the support for weakly replicated objects found in [10]. Multiple servers can maintain replicas of the same object, and writes can be made to any available replica. As with disconnected operation, changes are tentative until they have been committed at the primary server. This additional complexity leads to a requirement for strong ordering semantics between secondary servers, such as gossiping [10] or group communication [4]. These techniques may not be feasible in an ad hoc environment in which the node and server population cannot be determined in advance and from which primary or secondary servers can disappear from the network forever.

JetFile [3] is a distributed file system that is largely based on peer-to-peer communication. Nodes share data with each other without going through any central server. The function that is not delegated to the participating nodes is keeping track of the latest version of each file. Since everybody should have the same notion of version number, those are handed out by a server.

6 Conclusion

We have defined a spontaneous network as a small scale ad hoc network intended to support a collaborative application. In such a network, the human behaviors associated with collaboration can be leveraged in order to create the administrative infrastructure needed to secure the network and configure services. We have explored some of the unique challenges that need to be faced in building such environments and discussed some relevant technologies that can be used to further our ongoing work in building toolkits and prototype applications.

Acknowledgements

We would like to thank our colleagues in the CONNECTED project, particularly Mats Björkman of Uppsala University, for much helpful discussion. Conversations with Franklin Reynolds of Nokia Research Center, Boston have also positivly influenced this work.

References

[1] Internet Engineering Task Force. Working Group Charters, RFC’s, and Internet-Drafts describing IPv6, Service Location Protocol, MANET routing protocols and Zero Configuration networks are available at `http://www.ietf.org`.

[2] Laura Marie Feeney. "A Taxonomy for Routing Protocols in Mobile Ad Hoc Networks." SICS Technical Report T99:07. Swedish Institute of Computer Science, Kista, Sweden.

[3] Björn Grönvall, Assar Westerlund, and Stephen Pink. The Design of a Multicast-based Distributed File System. In Proceedings of the Third Symposium on Operating System Design and Implementation, New Orleans, USA, Feb. 1999.

[4] Dag Johanson, Robbert van Renesse and Fred B. Schneider. Operating System Support for Mobile Agents with Afterword. In Mobility: Processes, Computers and Agents. Dejan Milojicic, Frederick Douglis and Richard Wheeler, Eds. Addison Wesley, 1999. pp 535-556. (Originally appeared at 5th IEEE Workshop on Hot Topics in Operating Systems.)

[5] Anthony D. Joseph, Alan F. deLespinasse, Joshua A. Tauber, David K. Gifford, and M. Frans Kaashoek. Rover: A toolkit for mobile information access. In Proceedings of the 15th ACM Symposium on Operating Systems Principles, pages 156-171, Copper Mountain, Co., December 1995.

[6] Elizabeth Royer and C-K. Toh. A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks. *IEEE Personal Communications*, April 1999.

[7] M. Satyanarayanan, *Coda: A Highly Available File System for a Distributed Workstation Environment*, In Proceedings of the Second IEEE Workshop on Workstation Operating Systems, Sep. 1989.

[8] Frank Stajano. The Resurrecting Duckling: Security Issues for Ad Hoc Wireless Networks. Proceedings of the 7th International Workshop on Security Protocols. Lecture Notes in Computer Science, Springer-Verlag, 1999.

[9] Jini `http://www.sun.com/jini`.

[10] Douglas B. Terry, Marvin M Theimer, Karin Petersen, Alan J. Demers, Mike J. Spreitzer and Carl H. Hauser. Managing Update Conflicts in Bayou, a Weakly Connected Replicated Storage System. In Proceedings of the 15th ACM Symposium on Operating System Principles, pp. 173-183, December 1995.

[11] Mark Weiser. Some Computer Science Issues in Ubiquitous Computing. Communications of the ACM, 36(7):74-84, July, 1993.

Bengt Ahlgren is the manager of the Computer and Network Architectures Laboratory at the Swedish Institute of Computer Science in Kista, Sweden. He joined SICS in 1989 after finishing his M.Sc. in Computer Science at Uppsala University. In 1998, he completed his PhD at Uppsala University in the Dept. of Computer Systems. Bengt's interests include protocol implementation and its relation to operating systems and computer architecture, IP routing, multicast and network supprt for multimedia applications. (Email: bengta@sics.se)

Laura Marie Feeney has been a researcher at the Swedish Institute of Computer Science since 1999. She is participating in a number of projects in the area of mobile ad hoc computing, emphasizing energy consumption, routing, QoS and minimal-configuration systems. Her research interests include many topics in systems and networking. (Email: lmfeeney@sics.se)

Assar Westerlund joined the Swedish Institute of Computer Science in 1996 and is also a computer science student at the Royal Institute of Technology, Sweden. His research interests include mobile networks, file systems, security and cryptography, and scalable systems. (Email: assar@sics.se)