Small Worlds and the Security of Ubiquitous Computing

Harald Vogt

Department of Computer Science
ETH Zurich, Switzerland
E-mail: vogt@inf.ethz.ch

Abstract

This paper considers the small-world phenomenon in two contexts: ubiquitous computing and security. We argue that small-world properties emerge in ubicomp systems due to their natural occurrence in human-centered networks and their meaning as a system design principle. The relationship of small worlds and security is examined from two viewpoints. First, results on the security of small-world networks are reviewed. Second, we look into how small-world properties are obstacles to, or can be helpful in maintaining security guarantees. We then give an example of how to improve the communication security in large-scale, self-organizing networks.

1. Introduction

The small-world phenomenon was first investigated by Milgram in the 1960s [7] in a study showing that it was possible to draw a connection between any two people in the U.S. with only few intermediaries. This phenomenon has become widely known as "six degrees of separation". Recent studies have examined it more thoroughly and it keeps appearing in many distinct areas where large numbers of interconnected nodes are involved, such as the brain and the Internet [14, 2]. It might be promising to investigate areas such as ubiquitous computing as well and try to identify small-world properties of such systems. This could lead to a better understanding of these systems and help to improve on properties related to performance, security, usability etc.

Intuitively, living in a "small world" means that interaction with any other entity is—in principle—always possible. It also means that you can get access to important information by asking the right people (who might not have the information that is asked for, but might know somebody who has). Even if you live on an isolated island, a single contact to the world outside would be sufficient to grant you access to the whole world. Technology supports this principle: it is easy to reach an acquaintance in a far country by telephone, or write an e-mail to somebody once met at a conference. This could be the first step towards something that might result in a big change in life, or an important decision being taken, or the acquisition of a piece of information that turns out to be valuable later. The view of human society as a "global village" results from the increasing interconnectedness of global processes. In practice, these theoretical possibilities are often curbed by social borders, ignorance, or the lack of (financial) resources. It is certainly not easily possible that a kid from Mongolia gets invited to an Argentinian barbecue just by asking a Spanish tourist he ran into in his home town, though it might happen.

Ubiquitous computing is the result of technological, social and economic developments. Its concepts are tightly connected to processes going on in human-centered societies. Computational entities, such as wearable computing devices, web services, electronic media, communication equipment, sensor networks, smart objects etc., are increasingly pervading all aspects of life. Every human being has his own (small-world) view on the world, with strong ties to closer, more important people and things, and weak ties to more distant, less influential entities. Similarly, in a ubicomp world each computational device or other entity has its own view of the world, comprised of interactions and relationships to other ubicomp entities scattered throughout the physical world. This view is sketched in Fig. 1.

When relating the small world model to ubicomp, it is important to identify the association of ubicomp entities to small-world roles. One such role are *hubs*, which are heavily connected entities. Hubs are often responsible for the existence of small-world properties, connecting parts of a network to each other. For example, a small child is connected to most of the "real" world through its parents, who serve as hubs, relaying important information to their child. Other entities serving as hubs may include physical objects such as a house or a car, institutions like a company or a sports club, or simply well-connected people.

A (subjective) sense of security can emerge in a small-world society from two possible perspectives. One of them regards entities in a small world as a group in which all members pursue common interests. Collaboration of this type has advantages, e.g., when dealing with rivaling interest groups. The second perspective is concerned with the distribution of individual risks amongst the members of a group. In both views, it is desirable that the group has the largest possible size. This increases the chances to outrival a competitor and minimizes the impact on the individual in case of disaster.

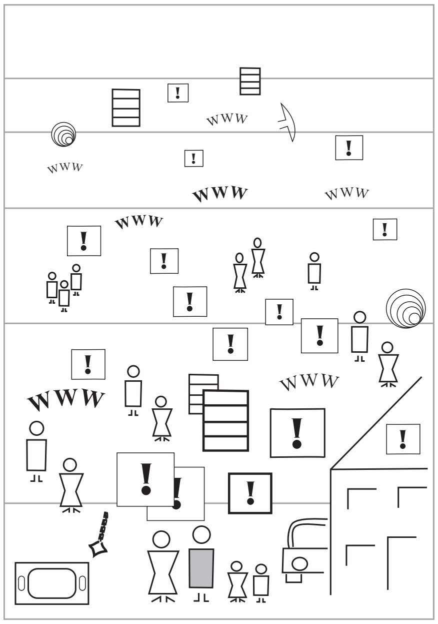

Figure 1. Ubiquitous computing from a single entity's point of view. Objects nearby are closely related to each other, while distant objects (towards the top) are more sparse. (Inspired by Saul Steinberg's cover of the 29 March 1976 issue of the *New Yorker* magazine)

A common theme in ubicomp are the tight restrictions on resources available on devices. This concerns mainly energy supplies, but also compuational and other capabilities. Thus, two principles guide the design of such systems to compensate for these limitations: collaboration and reliance on infrastructure. Self-organized collaboration seems to be a key requirement when building large-scale, networked systems. Access to infrastructural services, provided by more powerful yet less flexible devices, is often a prerequisite for implementing certain functionality.

These constraints have manifestations in security mechanisms. Consider the prominent *resurrecting duckling* security policy [11]: small, powerless devices are controlled by more powerful devices. The controlling device can be regarded as a "security hub" for the smaller device, connecting it to the network. In wireless sensor networks, base stations are often employed to provide for security services [8], serving a similar purpose. As we will see later, such hubs introduce vulnerabilities and therefore require special protection. It is challenging to provide security services in a totally self-organizing fashion.

In this paper, we try to make the point that ubicomp security can be approached through the small-world properties of ubicomp systems, both by avoiding vulnerabilities that are introduced through these properties and by exploiting the strengths that are offered by them.

2. Ubiquitous Small Worlds

2.1. Small-world properties in ubicomp

A "small-world" graph is characterised by the following properties [5, 14]:

- The graph is sparse, i.e. on average, a node has few connections.
- The distance between any two nodes is small related to network size, e.g. logarithmic in the number of nodes.
- The graph is clustered, i.e. nodes that have a common neighbour tend to be neighbours of each other as well.

An additional property is added to explain effects of some real-world networks more realistically [2]:

- The degree distribution scales according to a power law, allowing for a small but significant number of highly-connected nodes, called *hubs*.

There are various different models of graphs exhibiting small-world properties [5]. One family of such models starts with a regular (lattice) graph and adds shortcut links randomly for each vertex. These models differ in how the target node for a new link is selected, for example depending on the distance from the source node. Another family of graph models is assembled from scratch. Starting with a few connected nodes, new nodes are successively added and linked to existing nodes, preferring such nodes that already have a large degree. The resulting networks are called "scale-free". All of these models result in graphs with small diameters, which is due to the existence of shortcut links that connect remote clusters to each other.

We argue that many ubicomp systems are instances of small-world networks due to one of the following reasons:

1. Small-world properties have been discovered in the topologies of human relationships, such as collaboration networks or professional groups, and of artificial, self-organizing structures such as the World Wide Web. Since ubicomp comprises technologies

that strive for seamless, hidden, universal interaction between humans and between humans and machines, it would be hard to believe that these systems could ignore the fact that their operating environments are organized according to small-world principles, and still be useful.

2. Small-world techniques have been deliberately applied in improving the performance [15] and security [12] of self-organizing networks. Small-world properties are also occurring in natural, complex structures such as the brain [10]. Thus, it seems that small-world design principles can be useful in the design of technical systems. Ubicomp systems seem to be especially suited, as we will show next.

Small-world properties are already apparent in some classes of ubicomp system architectures. The first we consider are **smart object environments** [9]. Smart objects are (mainly) physical entitites that are equipped with or have access to a computational unit and sensoric input. They form communities through interaction with each other. Some of them have distinctive features such as user interfaces, storage capacity, actors, access to a backbone, gateway to other networks, or excess energy supplies. Smart objects are clustered according to *local* communication, leading to a high clustering coefficient. Since there may be a backbone in operation, shortcuts can exist, lowering the characteristic path length in the overall system. If very powerful devices are available, these can act as *hubs*. In **wearable computing** systems, similar relationships exist between personalized devices that are integrated in a body-area network and other entities such as home and workplace equipment, or the networks belonging to other people. An example application field is healthcare, where wearables contribute to therapy and communicate regularly to distant sites. Yet another class are **sensor networks** where large numbers of resource-restricted nodes communicate, collaboratively working on tasks. They can communicate to each other only locally, but shortcuts may exist, either physically [4] or in virtual overlay networks [12]. Clustering in such networks can be affected by high-powered nodes such as base stations.

Dynamic ubicomp networks are formed according to the requirements of their application. It could be valuable to find out whether existing models for network creation are valid for ubicomp scenarios, or new models have to be built. Such models could be useful for simulations, for example.

2.2. Small worlds and security: problems and possibilities

There are few studies on the security properties of small-world networks. In [1], the resiliency of networks against random failures and educated attacks is examined. Generally, the basic operation of a small-world network is ensured in the face of such threats. The failure of large numbers of nodes can be tolerated, the network stays connected, while performance may suffer slightly. While scale-free networks perform even better than “pure” small-world networks under failures, they are more vulnerable to a knowledgeable attacker who can identify and target the hubs that hold together the network.

The spreading of infectious diseases is facilitated in small-world networks [2]. This does not only apply to human societies (as could be observed in the recent outbreak of the SARS epidemic) but to structures such as the Internet as well (exploited for the distribution of computer viruses).

In contrast, living in a small world can give a subjective impression of security as has been noted in [6]. This probably relates to the distribution of individual risks over larger groups, a technique institutionalized for example by insurance companies or in military formations.

It turns out that the self-organized PGP certification graph (consisting of peer-certified instead of centrally certified public keys) exhibits small-world properties [3] such as a big clustering coefficient and a small diameter. We conjecture that the PGP web of trust wouldn’t be acceptable as a “trust enhancer” if it didn’t show these small-world properties.

We now review some fundamental security properties and relate them to small-world properties of complex networks. The goal is to prepare the grounds for exploiting (or avoiding them, for that matter) these properties to improve security features of systems.

Confidentiality. Just like the spreading of a disease, the distribution of information can be highly efficient in a small-world network. This suggests that once a piece of information has been revealed, it is virtually impossible to revoke it – if the information is useful to lots of people, it will circulate amongst many people within a short time. That means that a single byzantine node can reveal supposedly confidential data to lots of other nodes and permanently compromise the security of that data. On the other hand, the short paths in small-world networks suggest that if a message is to be relayed to a remote node, it has to travel only over a small number of hops. If all of the relaying nodes behave according to the security policy (maintaining the message’s confidentiality), the message will be kept confidential. If the number of byzantine nodes in the network is small, most paths will be secure. As simulations suggest, however, only very few failing nodes can be tolerated if no further precautions are taken [12].

Integrity. In a highly clustered network, nodes can mutually assure themselves of the integrity of data (i.e., the fact that these data conform to a reference model) since their data sources are likely to be correlated [16]. Additionally, information can be "inherently" verified in a small-world network. An important piece of information will eventually reach a node that is able to verify this information. The information will reach this node either by accident or by a purposeful contact from another node. The small-world properties suggest that a verifying node can be found by other nodes with relatively little effort. Thus, false information can be identified and warnings could be issued.

Availability. This property has already been considered above. Small-world networks can be resilient against attacks if hubs are avoided on which the connectedness of the network depends. The key to availability is the redundancy and diversity of paths between parts of the network.

3. An Application to Communication Security

Communication in large, homogeneous, self-organizing networks is restricted by the technical capabilities of nodes, which are typically equipped with short-range radio communication only. Messages therefore have to be relayed by other nodes in order to reach distant targets. Secure (authenticated and private) communication requires mutually shared keys or a public-key infrastructure. Since public-key operations are often too costly for small nodes and the capacity for storing keys is limited, we propose a different kind of communication scheme. This scheme is based on the small-world properties discussed previously.

Fig. 2 shows part of a network from the perspective of a single node A. The edges in this graph denote communication links. Several nodes throughout the network are trusted by A, i.e. they share a key. Here, we are not concerned with how to obtain such keys. We simply note that these could have been obtained through previous physical contact between these nodes, for example. These keys can be used, among other purposes, for authenticated message exchange. We will therefore call the graph induced by these trust relationships the *authentication graph*. Note that most of the nodes in its neighbourhood (only few hops distance) are trusted by A, plus some nodes randomly scattered throughout the network. This corresponds to a high clustering coefficient and random shortcut links.

A can directly exchange authenticated messages with trusted nodes, for example with node B. If A wants to use a service that is only offered by an untrusted node, say C, there are several possibilities. First, A could try to establish a new trust relationship with C. This is a costly procedure, possibly involving several message exchanges with a trusted third party. This will only pay off in case of a long-term relationship. For short term communication, A can try to contact a trusted node that is "closer" to C. One such node is B. B and C are immediate neighbours, so they probably share a common key and B could engage in a secure message exchange with C. B then relays the messages between A and C.

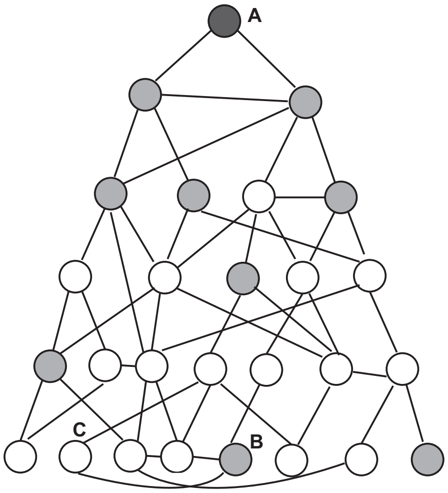

Figure 2. A network from a single node's point of view. Trusted nodes are shaded

Finding a "closer" node requires an underlying notion of distance. A natural choice for many systems is geographical distance, for example in location-aware sensor networks, due to its correlation with short-range radio.

It may not be possible to find a trusted node located in the neighbourhood of the target node, especially in large networks. Therefore, we need a means to span larger distances. This works as follows. The source selects its trusted peer that is closest to the target and sends the message on its way, together with a MAC that allows the trusted node to verify the integrity of the message. Each node on the path checks if it has an own trusted node that is even closer to the target. If successful, the message will be sent to that node next, together with an according MAC. This procedure is followed by every node on the path, resulting in an interleaved authentication path as in Fig. 3. Note that only shaded nodes are participating in the authentication and verification process of the message. The target (rightmost) node in this case receives, strictly speaking, an unauthenticated message and may not be willing to accept it. However, it is likely that the last shaded node is very close to the target and, as a neighbour, is trusted by the target.

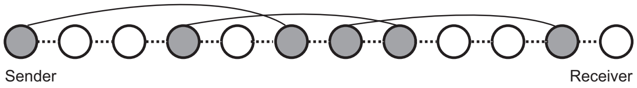

Figure 3. An interleaved authentication path

Simulations we have performed show that such paths exist in most cases, even if the number of trusted peers per node is rather small, and that these paths can be found using only local knowledge. Such paths can get longer than the shortest path between source and target nodes, but the overhead tends to be small. Detailed discussion of simulation results is out of the scope of this paper and will be presented elsewhere.

What are the security properties of such a communication scheme? First we have to fix a threat model. We assume that the only reasonable way of attacking the network is by seizure of nodes. We will not consider attacks from the outside or on different abstraction levels. The only action an attacker can attempt is compromising and taking control of already participating nodes. In this setting, a single malicious node is not able to alter the message on its path from the source to the target. Only collaborating nodes may be able to do so, but only under restricted conditions. In the path shown in Fig. 3, for example the 3rd and 6th node together are able to introduce false data: The 3rd node makes up the data which the 6th node accepts despite of the discrepancy to the source's MAC. The 4th node unwillingly helps to promote the authenticity of the falsified message.

In combination with local interleaved authentication [13], the demands on an attacker are further raised and the confidence that the message will not be altered on the last few steps before it reaches the target is increased. Regarding the example, this would mean that the 4th node is not so easily fooled into accepting the falsified message from the 3rd node and authenticating it further. Note that local interleaved authentication is always possible, since nodes in the close neighbourhood are trusted by default.

This Section has shown that by exploiting small-world properties of a system, or by deliberately introducing them, we can improve security guarantees within the system. Instead of providing end-to-end communication security means with all the attached overhead, we introduced a lightweight, collaborative authentication scheme that provides security guarantees approximating end-to-end guarantees.

4. Discussion

Small-world properties are inherent features of many societal, technical and natural networked systems. It is therefore almost unavoidable that ubicomp systems exhibit these properties, too. From a subjective point of view, a small world provides a sense of security, increasing the capability to withstand threats. On the other hand, some properties might make such a system more vulnerable. The challenge in system design is to avoid these dangers but instead exploit the small-world features for improving on the dependability of ubicomp systems.

We expect that the properties of small-world networks can provide guidance in the design of secure communicating systems, especially for large-scale, self-organizing systems occurring with the advent of ubiquitous computing. We have given an example how communication in such networks can be made more resilient against attacks exploiting small-world properties. We hope that these principles can be extended to other security features of ubicomp systems as well, such as privacy protection, access control, key management and others.

References

[1] R. Albert, H. Jeong, and A.-L. Barabási. Attack and Error Tolerance of Complex Networks. *Nature*, (406):378, 2000.

[2] A.-L. Barabási. *Linked*. Plume Books, 2003.

[3] S. Capkun, L. Buttyan, and J.-P. Hubaux. Small Worlds in Security Systems: an Analysis of the PGP Certificate Graph. In *New Security Paradigms Workshop*. ACM, 2002.

[4] K. A. Hawick and H. A. James. Small-World Effects in Wireless Agent Sensor Networks. Technical Report CSTN-001, Massey University, New Zealand, May 2004. `http://www.massey.ac.nz/~kahawick/cstn/001/cstn-001.html`.

[5] B. Hayes. Graph Theory in Practice: Part II. *American Scientist*, 88(2):104–109, 2000.

[6] J. S. Kleinfeld. Could it be a big world after all? *Society*, 2001. `http://www.uaf.edu/northern/big_world.html`.

[7] S. Milgram. The small world problem. *Psychology Today*, (2):60–67, 1967.

[8] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler. SPINS: Security Protocols for Sensor Networks. *Wireless Networks*, 8(8):521–534, 2002.

[9] F. Siegemund. A Context-Aware Communication Platform for Smart Objects. In *Pervasive Computing: Second International Conference, PERVASIVE 2004*, volume 3001 of *LNCS*. Springer-Verlag, 2004.

[10] O. Sporns. Network Analysis, Complexity, and Brain Function. *Complexity*, 8(1):56–60, 2003.

[11] F. Stajano and R. J. Anderson. The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks. In *Proc. of the 7th Int. Workshop on Security Protocols*, volume 1796 of *LNCS*, pages 172–194. Springer-Verlag, 1999.

[12] H. Vogt. Exploring Message Authentication in Sensor Networks. In *Proc. of European Workshop on Security of Ad Hoc and Sensor Networks (ESAS)*, LNCS. Springer-Verlag, 2004.

[13] H. Vogt. Integrity Preservation for Communication in Sensor Networks. Technical Report 434, ETH Zürich, Institute for Pervasive Computing, Feb. 2004.

[14] D. J. Watts. *Small Worlds*. Princeton University Press, 1999.

[15] H. Zhang, A. Goel, and R. Govindan. Using the Small-World Model to Improve Freenet Performance. In *INFOCOM*. IEEE, 2002.

[16] S. Zhu, S. Setia, S. Jajodia, and P. Ning. An Interleaved Hop-by-Hop Authentication Scheme for Filtering False Data Injection in Sensor Networks. In *IEEE Symposium on Security and Privacy*. IEEE, 2004.