Some limitations of web of trust models

Stephen Wilson
Senior Manager, KMPG Certification Authority, Asia Pacific, Sydney, Australia

The "web of trust" is one approach to the problem of trusted exchange of public keys in a public key security system. In a web of trust, individuals accept the bulk of the responsibility for identifying and authenticating each other and subsequently swapping their keys. This trust model is supported by some commercial products and some industry standards. The main alternative is the Public Key Infrastructure (PKI) where key holders are identified and authenticated by third-party Certification Authorities (CAs). Rather than personally swapping keys, participants in a PKI obtain one another's public keys from one or more CAs in the form of digital certificates. These two trust models have, for some time, been vying for selection internationally in both policy and commercial forums. In Australia, the debate has been spurred on by recent deliberations over the possible form of a national peak authentication body, and by spirited discussion of the privacy impacts of a national hierarchy. There appears to be a view emerging that a web of trust might be easier to constitute than a hierarchy and that it may be inherently less intrusive. On closer inspection, however, these promises prove to be unfounded. This paper discusses certain limitations of any web of trust model, with particular reference to scalability, uniform standards of identification, auditability, and the protection of personal identification data.

Information Management & Computer Security
6/5 [1998] 218–220

[ISSN 0968-5227]

What is a web of trust?

In a public key security system, Alice needs the public key of all other users – let us call them correspondents – to whom she wishes to send confidential messages. And she needs to make her own public key available to anyone to whom she plans to send a signed message. Moreover, Alice needs to be able to trust the identity of purported holders of public keys, and likewise she needs to make sure that her own identity can be trusted by others. In a web of trust, correspondents undertake to establish trust in one another for themselves, basically on a pairwise basis. Once trust is established, respective public keys are exchanged and secure communications can follow.

In the diagrams that follow, correspondents who have established mutual trust are shown as linked together. For example, Alice and Bob trust each other, as well as certain other users (Figure 1). The web of trust is the set of all links that have been established in a given population of users. Links are only required when secure correspondence is planned between a pair of users.

Proponents of the web of trust model prefer retaining responsibility for identifying correspondents. They argue that this process mirrors the way trust is usually established in paper-based business. The system is basically self-reliant, and, for some people, appealingly flat in structure.

Scaling up a web of trust

The best known problem with the web of trust is that by itself it does not scale efficiently. That is, the work needed to maintain the web increases per user as the total user population grows.

For everyone to trust everyone else in a group of N correspondents, there are W links to be made to form a complete web of trust, where $W = ½N(N – 1)$. Thus the scale of the web of trust increases a little faster than the square of the population: if you double the population, there are just over four times as many links to be made. So for example, for ten users, a complete web has 45 links; for 100 users, there are 4,950 links; and for 1,000 users, there are nearly half a million links.

Scaling up through introductions

The scaling problem has of course long been recognised. The commonest solution is to allow for introductions. If Alice wishes to trust some stranger Steve, and she knows that Bob already trusts him, then she can have Bob "introduce" Steve to her. The dashed link below indicates that trust between Alice and Steve has been inherited from extant first-hand links. By allowing introductions, the number of first-hand links needed to complete a web of trust can be made much smaller than $½N(N – 1)$, depending on just who does the introductions. More on this point later.

Now note that even a first-order introduction depends on Alice trusting more than just Bob's identity. She must also trust that Bob knows Steve as well as she thinks she knows Bob (Figure 2). That is, she must trust Bob's processes for identifying people. This is a radical jump from needing to trust identity alone. And even if Alice can solve that problem, it becomes very unlikely that she can extend similar trust to Steve, to allow for a second-order introduction.

The problem of uniform identification of users in a web of trust cries out for standards. But the necessary corollaries of a standards authority and third-party administration are usually anathema to the model's proponents.

Managing introductions for maximum scaling

We will now consider how best to organise a system of introductions. In order to maximise scaling, and to reduce the possibility of uneven standards of identification, there is pressure to seek out a minimal number of introducers.

Let us say Alice delivers electronic services to a large population of users and she wishes to efficiently establish a web of trust. As best she can, Alice should seek introducers who each know as many other potential users as possible. The situation on the right below

Figure 1
Basic web of trust – Alice and Bob trust each other

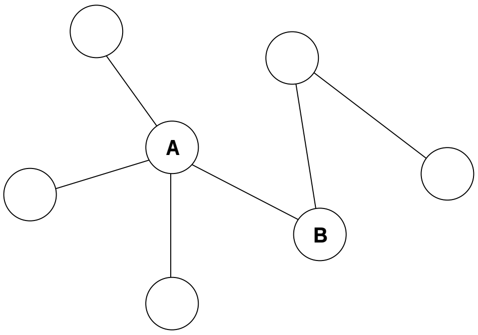
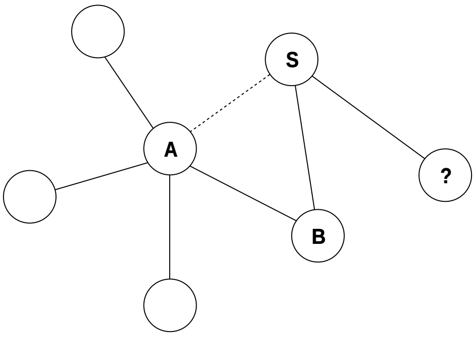


Figure 2
Introduction – Bob introduces Steve to Alice

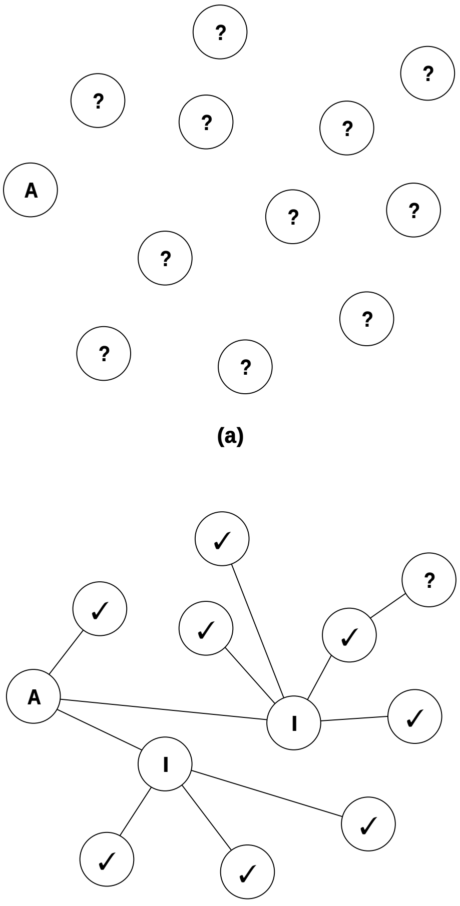


looks pretty good: Alice trusts two introducers and there is only one stranger needing a second-order introduction (See Figure 3).

But let us look closely at what is going on here. The introducers are becoming *de facto* representatives of sub-groups of users within the population. Indeed, since communities of interest abound in the real world, and since each typically has a recognised head, Alice may seek out those heads when constructing her web of trust. What is more, because each community has common interests (that is what makes them communities!) Alice will usually prefer to target her services at a group rather than individual level.

Now, if we rename the introducers registration authorities (RAs), in line with their representation of their respective communities, we can see (in Figure 4) Alice's pragmatic web of trust transform into the lower level of a PKI! Hierarchies after all are simply a natural result of the way we tend to organise ourselves. And in business it is common to have an abstracted identity based on one's membership of a certain community, be it a particular regional or focus group, a professional association, your employer or whatever.

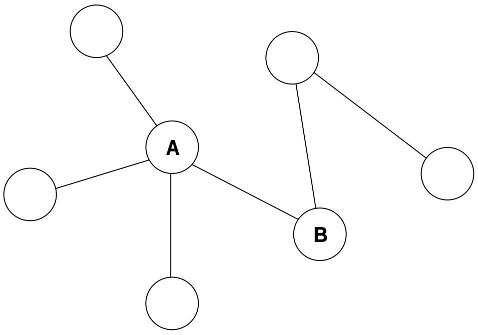
Figure 3
Alice optimises her introductions (before and after))



(a)

(b)

Figure 4
Introductions as registration authorities



Localisation of personal data

Finally let us look at one key aspect of privacy, the protection of personal identification data. If you are using an abstract business identity – say your employer and position – there is usually no need to make known your personal details, such as home address or employee number. Indeed, you would usually wish to keep these secret when conducting electronic business.

While a registration authority within a PKI might need to verify personal data before certifying one's identity and membership, it can keep the data secret in the process of issuing a certificate. The certificate need only make public your membership of the given community of interest. The RA can even allocate pseudonyms to mask the true names of end users.

Mutual recognition of communities within a PKI is typically ensured by higher-level certification authorities, sometimes known as Policy CAs (PCAs). Rather than identifying individuals, PCAs are generally restricted to identifying RAs or other CAs, and most importantly, vouchsafing their conduct. PCAs are therefore to be selected according to their ability and trustedness to audit organisations rather than their capacity to vet individuals. Audit and quality control process can be efficiently implemented at the PCA level, in the interests of maintaining uniform standards of identification for all users.

PKIs are frequently identified with flat, highly centralised management structures, wherein individuals and communities are subsumed into a uniform identification regime. But a multilevel PKI actually helps to preserve the structure and identity of lower-level communities. PCAs have no interest in the individuals who present themselves at lower levels of the hierarchy. All personal data can remain isolated from higher or central authorities. Responsibility for screening individuals should never be taken away from their respective communities of interest.

Conclusions

We have seen that natural pressures exist to create groups within groups, or hierarchies. In realising a practical web of trust, these pressures are manifest in the selection of minimal numbers of introducers, to reduce the number of first-hand links that need to be formed, and to ensure uniform identification of all users. Reliance on introductions requires users to trust others' processes as well as their simple identities, but ordinary users are not well equipped to audit others. Standardisation and third-party oversight are clearly called for. However, while these are part and parcel of most paper-based business, they are not usually welcomed in the web of trust model.

Multilevel public key infrastructures are perhaps not so much an alternative to the web of trust as an extension of it. By recognising trusted registration authorities, not only is the trust infrastructure made more manageable, but also privacy is actively enhanced, through abstraction of business identities and the localisation of personal identity data.