Thoughts on Identity

Bryan Turner
Sept, 2006

I was asked about my thoughts on the Identity debate. This is a (very) brief summary of my viewpoints on Identity in secure systems.

Terminology

Entity

An atomic, unique presence with the authority to make decisions for itself. ie: a person, or an autonomous control program.

Identity

The projection of an entity into some system. This may take the form of a public key, an email address, a physical key, drivers license, etc.

A Story to Set the Scene

Imagine you go to a bar, have a drink and notice a few people. That chick with the striped dress, that guy with the funny mustache. The next week you're back at the bar. Everyone else is different, but that chick with the striped dress is here again, and the guy with the funny mustache is in the corner. Week after week you return to the bar and easily identify the two characters.

One day you're at a friend's house having dinner and the discussion turns to people at the bar. You mention the chick with the striped dress, but no one recognizes her. You describe her physical features, her mannerisms, etc, but still there is no recognition.

Next you mention the guy with the funny mustache, surely they have seen him! You describe in exquisite detail the features of the man and his oddly shaped facial hair. Still, no one recognizes your description.

To you they are easily recognized, easily identified, yet somehow you cannot transfer this knowledge to others.

Every real system works this way; Drivers Licenses from out-of-state, passports from a country you've never heard of, a VIP membership to a club you don't recognize. What good is an identity or a credential that does not convince the verifier?

The naive solution is to create a "global identity". Obviously if everyone held a Drivers License from Tasmania, then everyone would know how to verify it. But is Tasmania the appropriate issuer for Drivers Licenses for the entire world? Why Tasmania anyway?

Do they have better drivers there?

So perhaps the global identity should be split, naturally, into smaller groups. But big ones, like countries. Yet, now we have the passport problem - I would not recognize, nor be able to verify a passport from Nauru.

Would you?

Yet somehow we do this every day - that chick with the dress, the guy with the mustache. Each Entity separately and individually establishes a Relationship, including an Identity, with each other Entity it interacts with. It does not matter that this Identity is not transferable, it is authentic and secure between any pair of Entities.

Through the established Relationships, all other higher-level protocols may be built; authentication, authorization, introduction, transitive trust, reputation, presence, social networking, etc.

Facts

1) Any mapping of an Entity into an Identity is partial; some information is lost.

a) It is impossible to tell if there is only one Entity controlling an Identity.
b) It is impossible to tell if an Entity has multiple Identities.
c) It is possible for two Identities to be equivalent in the same system.
d) It is impossible to assign a globally unique name to all Entities.

2) Identity exists separately from all other attributes;
ie: you can have Identity separate from anything else, and you can have attributes (2a)-(2e) without Identity. They are *independent*.

a) Presence
b) Location
c) Authority
d) Metadata
c) Social connectivity
d) Classification/Grouping
e) Etc..

3) Authentication is a separate, but necessary component in *secure* systems.

a) This differentiates a car key from a drivers license;
- a car key has no authentication
- a drivers license has (very poor) authentication; pictures, dates, address, etc.

b) If no security is needed, no authentication is needed.
c) Security is impossible without authentication.

4) No single Entity is trusted by all Entities.

a) A centralized system cannot work (requires a single trusted Entity).
b) A decentralized system is necessary.
c) Must not rely on a trusted Entity for security.
(may be used to improve efficiency though)

5) "Real Life" occurs offline. Think: Two iPods exchanging music at the beach.

a) An online Identity System cannot work.
b) An offline (disconnected) solution is necessary.

6) Users desire privacy and control over personal information gathering.

a) Tracing an Identity to an Entity should be impossible.
b) Tracing an Entity to an Identity should be impossible.
c) All transactions should be Forward Secure (implies Zero-Knowledge Protocols).
d) Required information must be supplied for each transaction,
and cannot be linked to an Entity or Identity.

Requirements for an Identity System

1) Each resource must be able to identify and authenticate each use of the resource.

a) It is not *necessary* to know which Entity is requesting the resource.
b) It is not *necessary* to know which Identity is requesting the resource.
c) Only knowledge of Authentication and Authorization are *necessary*;
I) Is the Entity using the resource Authentic within the System?
II) Is the Entity using the resource Authorized to do so?
III) (Advanced) Is the Entity acting on its own behalf?
ie: not under duress, being lied to, etc.

2) Each resource must be able to Account for each access.

a) Resource owner must be able to prove the resource was *accessed*.
b) Resource owner must be able to prove an *authentic* Entity accessed it.

c) Resource owner must be able to prove an *authorized* Entity accessed it.
d) Resource owner must be able to assign a billable Entity to each access.

Solutions

My research on this topic turns up three main solutions to Identity:

- Global Identity: MS Passport, National ID, etc..
- Public Key Infrastructure: RSA/PKI, Decentralized PKI, etc.
- Anonymous/Pseudonymous: Digital Cash, Freenet, Lyskanskya, etc.

We can generally rule out Global Identities as we know they can never work (Facts 1, 4, 5, 6). We can also rule out "pure" PKI as defined by RSA for similar reasons. This leaves us with two possible solutions:

- Decentralized PKI
- Anonymity/Pseudonymity

Decentralized PKI

Decentralized PKI is where every Entity chooses a public/private key on its own, and begins using it in the system. When it wants to access a resource, it identifies itself via its Public Key, and must establish an "account" with every resource provider. This is similar to the problem we have with today's websites (every site keeps its own database, has a separate userid/password combo, etc). Decentralized PKI eliminates the need for the userid/password combo, as it is the PK/SK combo instead.

Major issues with Decentralized PKI include complete failure of security when the Secret Key is revealed (which is guaranteed to happen eventually, even if that is 50-100 years after the fact). Loss of control over gathered information, essentially no privacy, and fractured support for higher-level attributes (each site must support each attribute separately). This leads to low-integration among participating sites. Decentralized PKI is also generally an online system.

Decentralized PKI runs against the grain of Facts 5 & 6. It is therefore unlikely to work in practice.

Anonymity/Pseudonymity

Anonymity/Pseudonymity allows each Entity to assume any number of Identities and operate behind these "masks" when requesting resources. Identities can be discarded at will. All protocols operate through Zero-Knowledge Proofs which are both privacy-preserving and Forward Secure.

As all information must be supplied at each transaction, every site receives the information it needs. This smoothes out the variations in supported features of different sites. Entities retain privacy and control over information gathered by resource owners (an Entity can switch “masks” and the resource owner cannot track who is behind the mask).

Entities receive authority via credentials, which are passed through the "mask" by an organization. Later, possession of the credential may be proven through any other "mask" of the same Entity. Organizations gain confidence in the Entity's authority without tying the Entity to a concrete Identity.

Organizations and Entities cannot collude to unmask any other Entity, so long as the honest Entity has not shared their master key.

Extended Discussion on Pseudonymity

The best way to understand Anonymity/Pseudonymity vs. Global Identity is to think about how agents in a system are named. If you don't have a name, you are anonymous. This is good in some cases, but without a name no one can reply back to you efficiently, they would have to "shout" to everyone in order to reach you. If the system is large, this becomes ineffectual.

If you take a name, but change it frequently, you get Pseudonymity. For any given conversation you have the same name, and agents can reply back to you directly. After the conversation is over you change your name. This leads to some interesting tactics; for instance you could use the same name for everyone and change it occasionally, or you could choose a different name for each agent you meet and never change it.

Using a different name for each agent you meet leaves you "anonymous" in the system (no one can collude to track you down) yet each agent you interact with knows you "by name". You can also choose on any particular transaction to generate a new name, even when talking to someone that has already met you under a different name - they will think you are a new agent.

Finally, if you select a name once and for all and never change it, you get Global Identity. So a Pseudonym system is really a slider-control on the level of precision other entities have of collaborating together to share information about you. It leaves you in control of the transaction and the dissemination of information.

The other major difference between Pseudonym systems and Global Identity systems is the type of security protocols used. Basically, if someone were to record your conversation, could they later use this to reveal who you are? A Global solution doesn't care; you are already known within the system, while a Pseudonym system attempts to conceal your Identity. Typically this means all the protocols are passed through a Zero-Knowledge Proof, which acts like a "filter" to remove forward-sensitive data from the transaction log. The data is only valid while the protocol is executing, revealing the log to someone at a later date is as effective as any arbitrary person typing in the log manually, anyone could generate the same log easily.

> Does Pseudonymity have a specific goal such as the ability "follow the
> money" for purposes of criminal investigation or is there something
> else involved?

Money is a whole different ballpark that is too long to detail here. Most Anonymity research targets Digital Cash or Filesharing, depending on which side of the fence the researcher sits on. Lysyanskaya neatly rides the fence, contributing some excellent work to both sides.

I do not support "revocable anonymity", as all the major techniques require a trusted entity to "benignly" hold your identity secret. The correct approach (as Lysyanskaya details) is to build revocation into the protocols directly. This is done by embedding globally identifying information into the protocol using a linear coding scheme. If you "cheat" the protocol, the mathematics behind it will allow your identity to be decoded (only by those you cheated). But if you are honest (within the protocol) it is impossible for any set of colluding parties to reveal your identity.