Authenticated Names

Stanley Chow Christophe Gustave Dmitri Vinokurov

Alcatel-Lucent Canada
600 March Rd., K2K 2E6 Ottawa,
Canada

ABSTRACT

Currently, most means of communication include some form of identification of the sender/originator, but none of these identifications are securely authenticated (at least not conveniently or in wide use). Legitimate business entities can be misrepresented by their name, and this creates opportunities for various scams known as phishing. We propose a new end-to-end authentication scheme that can be used to authenticate companies over many means of communication including telephony, email, web, and Instant Messaging. The framework is flexible and gives concerned legitimate institutions the ability to delegate their authenticated names to employees outside the office as well as outsourcing companies.

Categories and Subject Descriptors

K.6.5 [**Management of computing and information systems**]: Security and Protection – *authentication.*

General Terms

Security, Human Factors, Standardization.

Keywords

Phishing, telephony, VoIP, email, web, instant messages, authentication, identity.

1. INTRODUCTION

Currently, most means of communication have some form of identification of the sender/originator, but none of these identifications are usefully authenticated (at least not conveniently or in wide use). Many problems are really different aspects of this same problem. For example, the Caller ID Name/Number feature in telephony is easy to forge [1]; this leads to a great chance of phone phishing - e.g., making a call with the spoofed displayed bank's number/name and asking for account information. Similarly, with no email sender authentication, we have more phishing attempts via spam email.

For many purposes, such as anti-phishing, it is not necessary to be able to authenticate all individuals. We propose a new authentication scheme for concerned entities which are primarily businesses and not individuals. This scheme can be used to authenticate over many means of Internet communication including telephony, email, web, and Instant Messages. The proposed "RealName" scheme is modeled after the Trademark system and specifically aims at only solving the identification problem. We also propose an extension to allow for legitimate delegation of authenticated names.

Our proposal solves only one facet of the problem – that of authenticating corporate entities. We do not provide a way to look up a name in real time. The usability aspect of this scheme is limited to solving the phishing problem and, eventually, to consumers' protection. This is useful against phishing spam (since they typically forge the sender name) but is not useful against other spam.

The rest of this paper is organized as follows: section 2 discusses the difficulty of identifying a company, section 3 contains our RealName proposal for a registry scheme, section 4 has our proposal for using RealName registries to authenticate Instant Messages, HTTP sessions and Caller ID Name in telephony, section 5 contains an extension to allow "delegation" of RealNames, section 6 discusses the trust model and policies, section 7 elaborates on security considerations inherent to the scheme, section 8 has an overview of related work, section 9 has an introduction to a prototype featuring RealName for IM, section 10 discusses the advantages of the RealName scheme and section 11 has concluding remarks and future work.

2. IDENTIFIYING A COMPANY

Many factors combine to make it hard to identify a company:

- Many companies have complex structures (for legal, financial, jurisdictional reasons) that are largely unknown outside the company. Entities like subsidiaries, joint ventures come and go – they are freely created, closed down, sold, transferred and renamed. It is essentially impossible for outsiders (sometimes even insiders) to track these changes.
- Most people don't really think of companies by the legal name(s); instead, people think of "The Brand" and associate the companies with the brand. Indeed, companies are only interested in brand advertising and not much in the presentation of corporate structure and the legal names of corporate entities. For instance, the only way for consumers to get to the web site of Matsushita Electric Industrial Co., Ltd. is the www.panasonic.net domain.
- There are many distinct (and overlapping) jurisdictions that control the uniqueness of names. Jurisdictions may be defined by a combination of subject matter and territorial area. For example, companies in Canada can be registered federally or provincially; but trademarks are registered only federally. So it

NSPW'07, September 18–21, 2007, North Conway, NH, USA.

is possible for there to be two companies to have the same name as long as they are registered in different provinces. On the other hand, even though trademarks are registered federally, there can be multiple registrations of the trademark as long as each of them is in a different business – for example, according to Wikipedia, "Mustang" can refer to a car, a plane, and a motorcycle amongst other things. Considering the whole world as a Global Village, the number of overlapping yet distinct jurisdictions is quite high.

This means the "identity" of a company is a rather nebulous concept and it is pointless to associate identity to the legal name of a company. (In some ways, the problem of identifying individuals is even worse; this is partly why we don't address the problem for individuals.)

The above problems are mostly externally imposed, but there is another problem that is self-inflicted: each company may choose to have multiple "addresses" for any single channel of communication (even if we ignore physical addresses). For example, a hypothetical large multi-national company branded as BigCo could be actually "BigCo Inc.", "BigCo Corp.", "The Big Company", "BigCo Enterprises Inc." or even all of them at the same time (for different entities in the corporate structure). There are even more possibilities in other channels:

- Many domains on the web:
  - BigCo.com, BigCo.biz, BigCo.org
  - BigCo.ca, BigCo.ru, BigCo.fr, …
  - BigCo-TV.com, BigCo-cars.com
  - TheBigCompany.com, …
- Many phone numbers:
  - Local numbers in each city
  - Toll-free numbers for service, sales, …
  - Numbers for special promotions
- Many addresses and domains for sending and receiving email
  - Incoming complaints, sales queries, web support, etc.
  - Outgoing newsletters, invoices, shipping notices, etc.
  - Bulk mail advertisings
  - Emails may be sent by third parties from other domains

In practice, new legitimate addresses (be they DNS domains, physical offices, phone numbers, etc.) are created at the drop of a hat and advertised to consumers, so consumers are conditioned to accept them (just as few consumers would question a new office address being used). This means a new domain like "BigCo-tv-monday-special.com" raises no alarms – unless one digs deep enough to find that it is registered to a "Big Crook's Cheap TV Inc." in some offshore haven.

2.1 DNS

The Domain Name System (DNS) is a flat name space (at least for the most used top-level domains like .com, .net and .org) that is not linked to company names or brands. This is a recipe for name conflicts when multiple legitimate registrants compete for the same domain name. This means only one of them will get the most desirable obvious domain name, while others will have to settle for less obvious domain names. For example, "pioneer.com" is owned by "Pioneer Hi-Bred International, Inc.", a DuPont Company – supplier of advanced plant genetics to farmers; so "Pioneer Electronic (USA) Inc." uses "pioneerelectronics.com" instead. People looking for a particular company have no way to know, without actually looking it up with the search engine, whether the obvious domain name is the right one. As an anti-phishing tool, search is not always practical. Search engines are not always immediately available given the resources or time constrains (PDA, bulk of emails to respond, etc.) Secondly, the interpretation of search results by individuals may vary and may be incorrect. It is also difficult to tell if a domain name that appears to belong to a company is legitimate or is a phishing attempt. Ultimately, the search engines are not in the business of authentication and may not take the steps necessary to provide authentication.

Another major problem is that many characters look alike. One classic trick is substituting 0 (digit zero) for O (letter), or 1 (digit) for l (letter). In recent days, there are much more sophisticated ruses using Cyrillic characters or Unicode characters. This allows the crooks to have fake domain names that are visually indistinguishable from the real names. There is a whole class of software that tries to use blacklists as well as heuristics to identify these fake domain names, but they suffer from the problems of network overhead as well as taking time to add rogue domains to the blacklist.

2.2 Caller ID

Empirically, Caller ID has not prevented phishing attacks over the phone. There are two flavors – Caller ID number and Caller ID name. The more commonly available Caller ID number is essentially useless as an anti-phishing tool. Ignoring that fact that Caller ID number is not secure and is easily spoofed, the consumer would have to know (and track changes to) the actual outgoing phone numbers used by each bank, etc. This is simply not possible.

As for Caller ID name, since the name is picked by the subscriber with no authentication, there is no point in looking at the security of how that name is looked up, transported and displayed. For example, in SIP, Display-Name can be picked arbitrarily for every call by the originator. Moreover, Request-URI and From: username@realm field values contribute to the SIP Digest Authentication hash input, whereas displayable info does not, i.e. it is not even a subject for authentication.

2.3 SSL/EV Certificate

The SSL certificate for HTTPS can be used to confirm that the web pages are coming from the owner of the URL. Unfortunately, as we see above, several factors conspire to make this not enough.

The SSL certificate only guarantees that the web server is the one that is associated with the domain name; this is a useful guarantee but is essentially redundant if DNS has not been subverted (SSL certificates also enable encrypted channels which is useful, but not the subject of this paper). As experience tells us, SSL certificates are not a big obstacle for phishing perpetrators.

The new EV certificates add "Extra Validation" to ensure that corporate entities are correctly identified; that is, to continue with our earlier example, the BigCo-tv-monday-special.com domain is certified because the owner (Big Crook's Cheap TV Inc.) has been absolutely confirmed. This is surely a useful fact, especially in after-the-fact investigation, but it is not particularly useful to a consumer. Ultimately, the problem is that EV strengthens the binding of the domain to the company and the identity of the company, but does

not help with the binding of the domain name to the Brand (EV certificates still rely on DNS ownership rules) As stated in the announcement [9], the improvement is about verifying "that the entity has exclusive right to use the domain specified in the EV Certificate". Even their proponents state that EV certificates cannot prevent phishing, they can only help track down the perpetrator (provided the corporate entity is not just a shell). An empirical study [2] measuring the impact of browser antiphishing techniques on attack success rates shows that overall EV does not provide a significant advantage in detecting phishing attacks.

2.4 Economic Incentives

Ross Anderson argues in [3] that in any system with multiple players, it is important that each player have the right incentive to bear the pain of providing protection and "liability for the failure of that protection". Many systems fail due to "perverse incentives" – not only do the players have no incentive, they are actively encouraged to ignore the system. We argue that DNS and SSL certificates, even if they worked as identification, have perverse incentives.

Consider the domain name registrar: it is in business to register domain names – that is the only way they get paid. Indeed, the system encourages each domain owner to preemptively register similar domain names. With the fees down to as low as US$5 a year, there is no way it can afford to do any authentication.

On the other hand, a corporate entity such as a bank cares very much about its reputation and that there are no imposters. In the traditional world, banks go after cheque forgers and any imposters, but in the current cyber world, there is little that they can do. Any legal action is likely to be multi-jurisdictional, i.e., expensive and slow. Actions under the DNS UDR are essentially gambles and can only knock out imposters one domain at a time. Any education of consumers is also expensive and slow, as well as annoying to their customers.

Consumers care that they don't become victim of fraud, but they also want convenient usage and so on. It is unrealistic to expect each and every consumer to be vigilant on every transaction, especially if vigilance means examining the details of each X.509 certificate and understanding the current legal landscape as well as the latest incarnations of computer fraud.

3. AUTHENTICATED NAMES FOR COMPANIES

We propose to go back to locally identifying names as opposed to globally identifying addresses. That is, we de-couple the identification function from DNS and explicitly do not accept certificates globally. We introduce a RealName registry (that is essentially a database of names plus a certificate authority). Clearly, as we have seen with DNS, this is difficult on a world-wide basis, as the registry has to respect jurisdictional boundaries since each jurisdiction is likely to allow names already registered in other jurisdictions. Even within a single jurisdiction, there are many cases where company names can be highly similar – in some cases, the names belong to related companies; in other cases, the name may have been licensed (a well known example involves "Apple Computers"); in yet other cases, the companies don't even know about each other; in still other cases, one registrant may be actively attempting to "pass off" itself as the other. It is almost impossible for anyone, even governments, to prevent duplicate or near duplicates.

Fortunately, there is a model that has mostly worked for many years – the trademark registries. Each jurisdiction has its own trademark registry, with different rules for resolving ownership of a trademark, and different rules for determining whether a proposed name infringes an existing trademark. Each registration is respected only within the jurisdiction of the trademark registry; consumers care only about their local jurisdictions and companies have to figure out which jurisdictions in which to register. (There are duplicate trademarks even within a single jurisdiction – if the trademarks are for different types of "goods" and "services". This is handled by the trademark always being used in context. We can rely on the same set of name conflict resolution rules.)

We propose a RealName registry system that operates like the trademark registries. In fact, we suggest the registry to be even more decentralized – each jurisdiction can operate its own registry; each professional association can operate its own, each trade association can operate its own. Registrants have to prove ownership of the name (to the satisfaction of the Registrar) and then get a X.509 certificate (with the registrar acting as a Certificate Authority). This RealName certificate, along with possession of the private key, will authenticate the registrant's claim to the name within the boundary of the registry. For example, if a company does business in Canada, then it would register its name in the Canadian registry. Note that the corporate entity is irrelevant – every subsidiary can use the same RealName.

Each end-user can pick and choose "local" registries he/she is willing to trust. Local in this context meaning in the sphere of interest for this end-user – typically the territorial jurisdictions and the professions that he/she deals with, etc. This trust is embodied by including the registry's root certificate in the "trusted" list. To carry on the example, most Canadians would import the root certificates for the registries they trust – which is most likely the Canadian registry and the provincial registries.

RealName registries sidesteps many problems:

- The many legal disputes that plague the DNS system
- The fake (but visually identical) domain names
- Ambiguous ownership rules (who owns the boycottXYZ.com site?)
- Existing system for handling disputes

Note that we have not solved the problems – we have merely made sure that each problem is now the responsibility of a motivated party. For example, a bank is very motivated to protect its name and to prevent fraud, so it is willing to spend considerable sums of money to register and protect its trademarks; we suggest they are willing to similarly register/protect its RealName and to make sure there are no infringing impostors. The consumer is motivated to generally authenticate businesses in her local area, so she imports the root certificates of the local jurisdictions. She can also (optionally) import root certificates of other places where she has dealings. The registries make money (and are perhaps regulated by law) from the trust of the public, and since each registry makes its own rules about the names, the registry must stand behind each registration of a RealName.

With the registries in place, we can then proceed to authenticate many channels of communications, including email, web pages, IM as well as authenticating parties of a phone/VoIP/SIP call.

One major advantage of the overall scheme is that no global cross-provider PKI support is required. Independent, locally-managed PKIs in each service provider's jurisdiction do all the work. Each RealName registry can start independently and be useful from day-1.

3.1 Registration Process

Registration is done by each jurisdiction having a public registry of names along with an X.509 certificate authority. Interested parties, such as banks, will register their names and monitor other names that could be confusing. Companies already do something like this for their trademarks, logos, domain names and so on.

Each registrant must provide a proof-of-identity to the registry to be added to the registry. To sidestep the many issues of duplicate names, cybersquatting, multiple jurisdictions, we propose to make each jurisdiction run its own registry. These registries are used mainly to bind the name to the certificates, and to make the list public, and are not used for lookup. Each Registrar is authoritative within its name space even when multiple Registrars overlap physically. When Registrars have similar/confusing names (e.g., NY City and NY State), this is the Registrars' responsibility to make sure the registry names are easy to distinguish.

At this stage, the organizations register their names in the registry. The registered names are the only ones that can later appear as "Authenticated by <registry>". Note that the organizations must register in each territory or jurisdiction of interest, just like they have to register their trademarks, domain names and so on. The organization must decide which registries are important – typically, this will be the set of geographical areas in which it does business and/or the list of professional or trade associations to which it belongs. We address internet shopping in section 7.

Ideally, each jurisdiction, probably a nation, a province or state, should have a single shared registry accessible to everyone. This means each entity needs to register only once in each jurisdictions of interest. On the other hand, these registries may be run by the local telephone carrier, the local government, a non-profit organization, or be a parallel function of the trademark office, DNS registry, or such; there could even be multiple registries for any particular jurisdiction. The interested entities need to register with each registry (just like they have to advertise on multiple phone directories). There are also advantages to having parallel hierarchies of registries, such as one geographical hierarchy with a level for global coverage, a level for nations, and one for each province or state, along with another "business" hierarchy for banks, stock brokers, etc.

The registrants must also actively monitor each registry to find potentially confusing names and to object to these confusing names. Some registries may just make the data publicly accessible, other registries may notify registrants of any new name and only enter the new name if there are no objections. Clearly, there are many ways that have been used in the registration of Trademarks, Domain Names, and so on; each registry can pick its own method. (If a registry picks a particularly poor method, it is likely that another registry will spring up to replace it.)

Each registry also acts as an X.509 Certificate Authority. Each registrant gets a certificate signed by the CA that is eventually used to authenticate the registered RealName. The certificate can include the trademark (along with areas of goods/services), logo, even HTML. Note that the meaning of each certificate is very limited – it only means the holder is entitled to that name/logo in that jurisdiction. The certificate does not care about who owns that name, where the company is registered, or whether there is any company of the same name in another country. The registrar is not responsible to sort out any disputes; if necessary, the parties can take it to court (and by definition, the case should be local to the jurisdiction, avoiding many of the worst problems). The registrar only needs to confirm authority to act for the company, and confirm ownership of the RealName by the company.

Since the Domain Name System already has country codes and most countries already have state/province codes, we can establish a convention of "abcxyz.RealName.bc.ca" for the Caller ID certificate of abcxyz in the province of British Columbia. Technically, this is appealing since we can reuse the whole DNS infrastructure; but administratively, this is problematic. One problem is that DNS has a whole name dispute resolution system that is not what we want. Another problem is that registries may be run by just about anyone, and there may well be multiple registries in a single province, so it is advisable to set up a parallel system.

4. AUTHENTICATING NAMES WITH REALNAME REGISTRIES

Authentication should be an end-to-end process – the claimant sends a RealName certificate and somehow proves ownership of the associated private key. For different channels, we propose different protocols. Note that there is no look up of the RealName and the registry/CA does not participate (except as the holder of revocation list).

4.1 Authentication for Instant Messaging

The informal nature of instant messaging makes this type of communications a target of choice for social engineering and identity theft. This is especially true in corporate environments where sensitive business data may be exchanged between colleagues via this means and end up in the wrong hands due to the lack of authenticated IM user session. Yet, IM users can define an arbitrary "screen name", most often a pseudonym, that makes impersonation a trivial task.

From our perspective, the problems specific to IM are:

- Most IM systems have a central server that "mediates" between clients. In this case, clients do not communicate directly with other clients.
- Two clients may not be online at the same time and once a central server is used, off-line messages are exchanged.
- Typically, client pairs have to mutually add each other into their own authorized contact lists prior to starting any dialog, as opposed to email where total strangers can communicate with each other.

Despite the reliance on a server, we propose a scheme that ensures end-to-end authentication of the IM user's screen name (recall that we only authenticate the company). The RealName certificates may be exchanged between IM clients either on a per IM dialog session,

or at the time of adding to the contact lists. While the first option has the benefit of allowing an IM client to use different authenticated screennames for different sessions, the latter option saves bandwidth.

We assume each user has already downloaded the root certificates of interest and obtained a RealName certificate. Whenever UA2 (User Agent 2) acknowledges the add request initiated by UA1, the screenname of UA1 is checked against the one embedded in the certificate and the result displayed in a user dialog window. Presumably the user will then decide accordingly whether to add UA1 to his normal contact list or a specific quarantine group in the contact list, etc. We can choose to confirm the private key and assume subsequent messages are authentic (this relies on the IM server to be well behaved).

To authenticate a message sent by UA1, the important parts (including the actual message, the sender screenname and the date/time) are cryptographically signed with private key and the signature is attached to the message.

The recipient user agent replicates and confirms the checksum; if the authentication fails, the message is flagged as unauthenticated and a notification is sent accordingly to the end user.

4.2 Authentication of Web Pages

For web page authentication, HTTP/HTML imposes a number of constraints:

- Most web pages are composed of many different pieces (each with a different URL). Frequently, the pieces of a single page come from different servers; e.g., many sites have advertising that is controlled by companies like Google.
- Many web servers handle multiple domains. For example, it is common for www.company.com, www.company.net, www.company.ca to all be served from a single server. In theory, the server setup should separate the multiple domains but frequently, these servers treat everything as www.company.com. One consequence is that many web sites end up using the wrong SSL certificate (since each SSL certificate specifies the domain).
- Most major web sites use load balancing and replication for reliability. This means multiple machines (of possibly different OS and Web server versions) may answer to the same URL. In extreme cases, different pieces of a single page, even just the pieces belonging to a single domain, could come from different servers in the load balance set.
- Many web pages are not static, but generated dynamically. These dynamic pages may be generated by very complicated systems. This means we don't want to force each page to be modified to have special tags, etc.
- Because SSL operates below HTTP and has no knowledge of the higher level protocol, SSL servers can only present one certificate for a particular IP/port combination. This means that with HTTPS, in most cases, it is not feasible to use name-based virtual hosting (TLS 1.1 does enable it).

One solution is to ignore SSL/TLS and operate at the HTTP level (this still allows HTTPS to use SSL/TLS for establishing encrypted channel). The web page will refer to a RealName certificate that is checked by the browser and then displayed in a separate unforgable area of the browser. The detailed steps for the browser are:

- At the start of a web page, fetch the HTML file using HTTP GET/PUT. We call this URL and the returned file "Top URL" and "Top HTML" respectively. (For the moment, we ignore the case where the return file is not HTML.)
- The Top HTML file includes a tag that has a serial number, a URL pointing to a RealName certificate along with a fingerprint of the certificate. (Note that this should be done only for the top level HTML. Allowing RealName certificates in lower level HTML opens holes for variations of cross-site scripting attacks.) The serial number is set so that the server knows which page is which.
- Fetch the RealName certificate as normal (using HTTP or HTTPS), confirm fingerprint of the certificate and validate it with pre-stored root certificate of registry.
- Confirm that the web server has the matching private key and that the HTML file actually came from the web server. Start by sending the serial number to the server; the server computes the checksum of the Top HTML file, encrypts it with the private key and sends it back. The client decrypts and compares the checksum against the locally computed checksum of the Top HTML file.
- If the authentication passed, display the name, logo, etc. retrieved from the certificate in a separate area of the browser window; that is, an area that the web-server cannot write to. We can possibly allow the certificate to include HTML to control the display of that area.

The advantage of this solution is that only top-level HTML files need to be changed. By assuring that the top-level HTML file is correct, we can be sure that we have correct URLs for all the subsidiary pieces. Thus, only the pages considered "important" may need to have a certificate, saving bandwidth, cycles, etc. Also, this does not require HTTP 1.1. On the other hand, the Top URL may return a Top HTML that contains a re-direct or time-out, in which case, the browser may need to authenticate the redirected page.

For dynamic pages, the server cannot compute the hash until the whole file has been computed and sent. We could specify that the encrypted hash be sent after the whole Top HTML file is formed by making the hash part of the </body> tag or the </html> tag, or even after the </html> tag.

Another solution is to make the certificate-handling be independent of the served URL and associated HTML pages. This solution requires no changes to any web pages (static or dynamic) so it is an attractive option. The certificate handling can be done as a normal request to a pre-defined special URL decoupled from the serving web site. Web servers hosting multiple domain names for a single corporate entity will particularly appreciate the easy configuration. We can also optimize this by only doing this check when new connections are opened (so visiting several pages with the same HTTP 1.1 connection will not recheck the RealName certificate).

We can combine both of these: try for the first solution; if the top-level HTML does not point to a RealName certificate, try the second solution (these two solutions can be tried in either order).

4.3 Authentication of Caller ID Name

4.3.1 Failure of Caller ID

In telephony, Caller ID has not been successful as a means of authentication. For example, the "Orange Box" [11] can often work against people with Call Waiting Caller ID. Large companies with lots of lines will have ISDN PRI or PBX that allow them to insert arbitrary number as the Caller ID; this is to allow the use of Direct Inward Dial numbers instead of the main company number. It is also possible to use VXML scripts [12] to spoof Caller ID.

Even if a phone service provider tries to control the displayed information with the existing means, there are problems. First, authentication of a company having multiple phone numbers by Caller ID Number is impractical. Numbers are added and subtracted, employees moved, branches opened and closed, etc. To simplify this, phone companies delegate the Caller ID work to the owner of the PRI/PBX. Second, telcos do not, typically, control the binding of a number to the associated name chosen at the time of registration. It is at best difficult to verify the name for an established landline telco; this is practically impossible for new VoIP providers. Third, there may be a jurisdiction issue. For example, there is nothing to prevent someone in Canada to register "Texas Capital Bakery" that truncates on 16 or less characters displays as the bank name. Alternatively, "Rational Business Consulting" can be registered in Texas and recorded in the local directory as RBC, which is one of the major Canadian banks. Thus, it is possible to carry out a phishing attack on a call with completely legitimate Caller ID Name and Number.

There have been proposals to authenticate the caller identity as opposed to the line identity, in particular in IP Telephony. As suggested in [13] and [10], caller authentication can be performed in the caller's domain. The authentication server would validate that the caller is authorized to assert the presented identity. A cryptographic signature (token) is created over the caller identification data and the recipient verifies it. Therefore, to authenticate the caller name, the recipient has to rely on the policy and diligence of all the remote service providers. It was admitted in [10] that even "a signature over the display-name does not prevent impersonation".

As the name implies, Caller ID can only identify the caller (to the callee). There are many scenarios where it is also desirable for the caller to authenticate the called party to prevent scam attempts, even though the caller controls the number that was dialed:

- The called number may have been wrongly associated to a 3rd party entity, for example via an email phishing scam.
- The called number may be a shared number (e.g., in a university dormitory, or on factory floor).
- The caller reaches an operator who may not be trustworthy.
- The caller reaches a home but a specific individual is the actual intended called party. This can be due to confidential information (financial or medical) handling regulations.
- Cell phones are often left unattended on a desk.
- The called number may be unknown; e.g., someone leaves a message "I am calling on behalf of John Smith. He is traveling and staying with a friend, call at this number."
- The called number may have been accidentally or maliciously forwarded to another number.

4.3.2 Authentication

In our proposal, the Private Key is independent of the phone line - it may even be stored in a portable device; thus both the calling and called party's authenticated identity is separated from the phone line identity.

For TDM lines, without assuming the end-user device has been upgraded, the authentication function is expected to be delegated to the upstream network equipment such as PBX, signaling gateway, or the phone company's central office. In general, for a SIP phone, it can be a first mile SIP proxy; for a landline phone, the authentication has to be done at the SS7/VoIP gateway which can perform the Caller ID authentication in addition to the protocol translation. In the latter case it is assumed that the Root Public Key of the recipient's local phone service provider is uploaded to the proxy and trusted. The proxy also must be able to attach the authentication result (flag) to the Caller ID information received from the caller and then transmit it to the called terminal using existing Caller ID technology. Authentication can be done at any point in the network as long as there is a reasonably secure (trusted) path from there back to the called party. Figure 1 illustrates these three options.

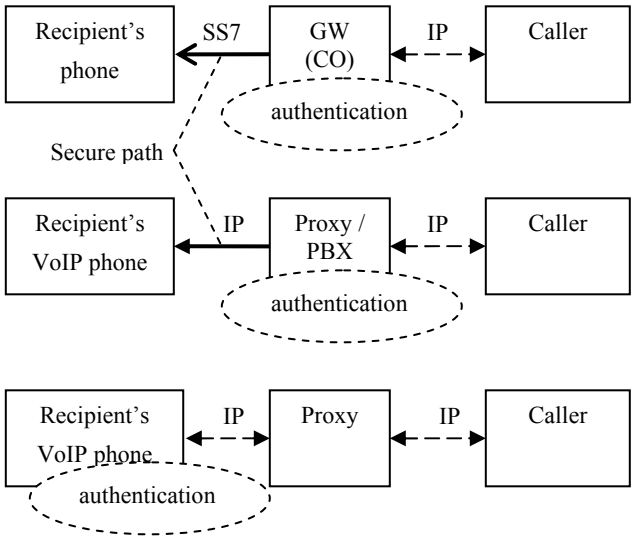

Figure 1. Options for the authenticator

4.3.3 On-demand authentication

The ultimate goal of name authentication is to help the callee to make an informed decision whether to trust the remote counterpart in this particular call. From this prospective, "delayed" authentication does not weaken the phishing protection; if the conversation just started, it is not yet too late to learn that the presented caller name was or was not properly verified.

For IP Telephony, delayed authentication would mean that the authentication protocol messages and data including certificate are sent over either the signaling port used for the call or an RTP port negotiated and opened after the call is set up. This delayed authentication has clear advantages from the performance prospective: call setup procedure and time are not affected;

authentication can be done "on-demand" (or "as needed") basis, i.e., only for answered calls and not for those terminated at the ringing stage or ended up in the voice mailbox. Moreover, only selected calls needing authentication (e.g., not those from the family members) will utilize the authentication data exchange, and this will lessen the network and processing loads on the switch. For two-party calls, during the call, whenever the user wants to authenticate the other party, he triggers a function on his device.

For TDM line end-devices, the request can be triggered by dialing a special short combination on the keypad and generated in the form of few DTMF tones sent towards the upstream equipment. This instructs the equipment to perform the authentication and to re-present to the user the Caller ID Name and the authentication status associated with the party on the call. For end-to-end IP telephony communication, the authentication messages exchange can be done in multiple ways: in the signaling path (piggybacking on the signaling protocol), in the voice (also called data or media) path, or in an independent IP channel. The request must contain enough information to associate it with the correct call.

5. REALNAME DELEGATION

There are times when we want to delegate the name usage, in that a company X may ask company Y to do things on X's behalf. Although described in terms of telephony, this also applies to email and so on.

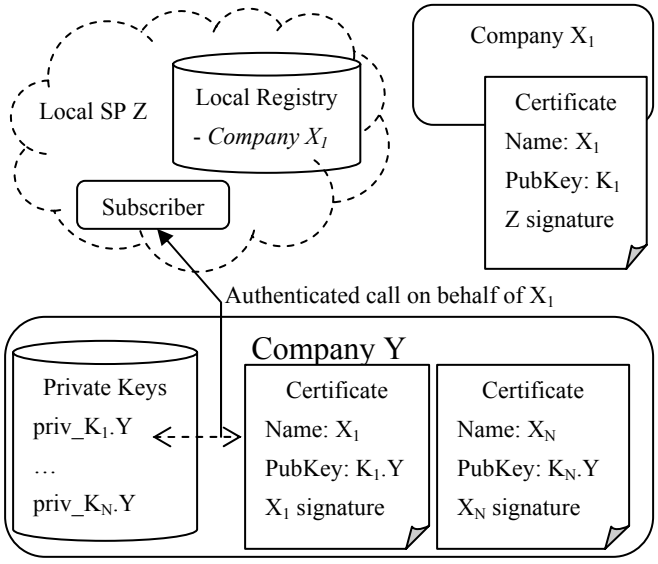

Figure 2. Name delegation mechanism

Imagine a client placing a troubleshooting case to the technical support center of company X and therefore, is expecting the call from company X. However, the support service may be outsourced to company Y (perhaps in another country). The authenticated Caller ID Name would be shown as "Y" in the client's display. The client would not trust company Y and may refuse to discuss his personal profile in company X. Moreover, the outsourcing location Y may provide the service to many companies like X. In this situation, there is a need for "authorized name forge", or what we call "delegation". This section describes how the authenticated Caller ID Name can be delegated to the third party. We propose that X should have the right to sign a certificate for another company Y, i.e. to create a certificate chain. This can be achieved simply by setting the keyCertSign bit of the keyUsage X.509 extension in the certificate given by SP (service provider) Z to X, as prescribed by [4]. Then company X can play the CA role and sign the certificate of trustee Y. Having Root CA certificate of SP Z, subscriber can verify the whole certificate chain. Therefore, the Caller ID Name presented by Y can be reliably verified by authentication mechanism supported by both Y and subscriber (or subscriber's gateway on his behalf).

Thus, the name of X may be legally used by Y as the display-name under control of company X. The general scheme is illustrated in Figure 2. The approach is scalable in that Y may be given the certificates from various entities {Xi} so that, for example, the operator of outsourced technical support centre would have a choice of the name to pick depending on the context of the prospective call or the recipient's profile.

6. TRUST MODEL AND POLICIES

In the handling and usage of certificates, many schemes have been proposed. In some sense, most of the schemes try to solve the "Global Problem" so that a certificate is acceptable to everyone and certificates from different registrars are interchangeable.

The RealName scheme is different in several ways:

1. Certificates attest to the Brand as opposed to the address (like a domain name) or the entity (like a company name).
2. Name space for Brand is not global but is inherently partitioned, with each registry owning its name space. There is no requirement for the global name resolution system (like Federated naming system).
3. Each registry is authoritative over its own name space.
4. Certificates are not globally accepted. The burden of clear presenting of the naming space is on the RealName owners and Registrars, and not on users.

The trust model is very simple – I trust a certificate from a registry if and only if I trust the authority of that registry over its name space. Conversely, if I don't deal with companies (with brands) in a particular name space, then I don't have to trust the registry; indeed, for all practical purpose, that registry (and name space) does not exist for me. There is no cross-certification, no extended trust, no dispute over ownership of names. This model is currently used for many things. For example, the ownership paper from the local car registry is a certificate from the authority attesting to your ownership of the car. This model is typically backed by legislation and is in use for cars, land/buildings, trademarks, etc.

As an anti-phishing defense, this simple trust model is sufficient for most consumers. For example, it is now difficult for an imposter to pretend to be a bank in the "local" registry (where local is relative to the consumer in the geography sense and/or the associations sense). For consumers that deal with non-local businesses, they need to be more careful and/or more knowledgeable (but still less than with existing schemes).

7. SECURITY ANALYSIS AND LIMITS

This section reviews the ways that could be attempted to circumvent the proposed scheme.

7.1 Impersonation in a local registry

Assume a Canadian user trusts and installs on his terminal the Root Certificates of two local RealName registries: "Canadian Hospitals" (Registry A) and "Canadian Law Offices" (Registry B). A legitimate organization (e.g., MyLocalHospital) is appropriately registered in Registry A, and a scammer registers the same or similar (like MLH) name in Registry B. Then the caller from MLH could pretend that the call is made from the hospital and ask for personal data, and the MLH name displayed on the user's phone will be authenticated by RealName.

This can happen fraudulently as well as legitimately when a trademark is used in multiple businesses. Since RealName is modeled after trademarks, we should expect RealName to inherit these problems; but at least we have limited the problem to the locality. We further minimize the impact of this problem by displaying the registry name as part of the authentication information. Thus the user knows which registry issued the certificate and is able to detect a fraud. Note that this demands much less of users than other schemes – there is no need to understand X.509, certification policy, and so on. The user only needs to know the purposes of the few trusted registries.

7.2 Legitimately conflicting names

Assume another Canadian user trusts and installs on his terminal the Root certificates of two RealName registries: "Canadian Hospitals" (Registry A) and "New York Hospitals" (Registry C). Registry A is local (one of those set by default) for the user, and the root certificate for Registry C is "imported". This creates an opportunity for confusion since both registries may certify different entities to the same name (legitimately or otherwise).

This ambiguity is easily resolved by looking at the name of the Registry, but again, we cannot expect all users to consistently do this correctly. We help by alerting users on these rare occasions when they should pay extra attention. We do this by introducing an extra level of authentication indication. In addition to the – *failed* (represented by red) and *authenticated by <Registry Name>* (represented by green)*;* we add the *authenticated by* ***Imported*** *<Registry Name>* (represented by yellow).

The intent is that for the vast majority of cases, people will be dealing with local entities. People with no imported root certificates will only see green or red. People who do have imported root certificates will be notified when an authentication does rely on that certificate – this alerts them to check that the displayed Registry-RealName combination is actually known and makes sense to them.

7.3 Call Waiting Caller ID spoof

Spoofing Call Waiting Caller ID during the phone call, known as Orange Box [11], works by emulating Caller ID signal in the established call. Essentially, Caller ID transmits the information in a safe channel – between the first and second ring *before* the phone is answered; phones that are vulnerable to Orange Box accept information *after* the phone is answered. Once the attacker overwrites the true authentication info, there is little can be done.

Note however, that "delayed" or "on-demand" RealName authentication protects against this attack, since the user is paying attention after the real information overwrites the fake one.

7.4 Multiple parties sessions

So far the presented approach was illustrated with one-to-one communication sessions. In many conference call systems, participants just dial to the conference and use a code to gain access to particular conference call sessions. It would be easy for an un-invited participant to join a conference call and listen. This can be a phone conference or an IM session.

We assume the Chair controlling the conference has the responsibility for security and that the goal is to make sure that all participants are invited ones. This is accomplished by the Chair authenticating each participant as one-to-one. Along with a count of the participants, this guarantees the desired security.

7.5 Internet shopping

We can break Internet Shopping down to several cases:

1. eShopping at a local company. In this case a local company would be certified by the local registry that is already trusted by the user.
2. eShopping at a global company that has local presence. For example, amazon.ca is the Canadian identity for amazon.com. This is effectively the same as case 1.
3. eShopping at an unknown company. This could be a global company that has no local presence, or more likely, a totally unknown company. We will concentrate on this case.

What does it mean to authenticate an unknown company? This question appears to have no meaningful answer. The existing authentication schemes also seem to have no good answer. For example, SSL/EV certificates will attest to the fact that it is the legitimate owner of the domain but says nothing about the owner. RealName is slightly better in that the user will know that the company has no local presence.

On the other hand, phishing attacks need the victim to have some prior relationship with the impostee. By impersonating a company unknown to the victim, the imposter gains no real advantage.

We propose that users have to establish trust out of band – by having dealt with the company before, by having recommendation, etc. Once the user has made the decision, the company's RealName certificate can be imported (as opposed to importing the root certificate of the registry). This means subsequent interaction with the company will be authenticated but not other companies from that locality.

8. RELATED WORK

Many different anti-phishing approaches have been tried. Ref. [5] presents a phishing detection technique targeting phishing sites with visually similar content web pages than the protected ones. Ref. [6] is an approach relying on a third party trusted device such as a cell phone embedding public key cryptography credentials to help in the authentication of a web site. Yee and Sitaker in [7] present a different approach for mitigating phishing threats, enabling users to assign "site labels" with visited web sites. In the same vein, [8] presents an approach that involves a web browser with local password storage protection when logging to web sites, and a heuristic engine determining if a visited web site is potentially fraudulent. Ref. [15] proposes to incorporate the company's logo into its X.509 certificate ("Secure Letterhead").

The common thread is that they are implementing new ways/channels to authenticate a website independent of the domain name. We take this as corroboration of our position.

8.1 Qualified Certificate

Qualified Certificates (QC) are defined in RFC 3039 [14] and the "qualified status" is specifically tied to "applicable governing law" in a way similar to our RealName proposal. The goals of the two schemes are disjoint. Other differences include:

- QC's aim to "identify a person … in public non-repudiation services". RealName attests only to the ownership of a name (with no information on the owner).
- QC's aim to "uniquely" identify a person. RealName explicitly intends that a single trademark may be registered in multiple CAs, possibly by different holders – due to the trademark's different owners in different jurisdictions.
- QC is intended to be (possibly) ubiquitous while RealName is intended only for companies that care. Hence higher cost and complexity are acceptable to the certificate holders.
- RealName is explicitly local in scope – users are expected to only accept RealName certificates from a few CAs; no checking on any extended policy is required. If the certificate is not from one of the trusted CAs, the certificate is rejected.

8.2 X.500

X.500 is a distributed directory system that may look like what we are proposing. In fact, there are some major differences:

- We solve only a single facet – that of authentication for institutions. We do not authenticate individuals.
- X.500 has a single root divided into countries; we allow multiple independent roots (each RealName registry is essentially a root) that may arbitrarily overlap.
- End users choose which registry to trust based on a combination of need, locality, and policy. It is explicitly expected that most users will trust a very few registries.
- Different expected usage – no directory look-up, end-users just validate the certificate presented by the claimant.
- No global PKI infrastructure, no cross certification.

9. PRELIMINARY PROTOTYPING

We have developed a first prototype featuring mutual RealName authentication capabilities. This application allows registration of new users to a registry authority of their choice delivering X.509 certificates with certified name and logo. Registry certificates and user certificates are deployed on the network equipment hosting the IM application featuring RealName functions. Figure 3 shows an outline of this prototype at the presentation layer.

Figure 3. IM User interface illustration

10. ADVANTAGES OF REALNAME

10.1 Authentication to the person or function

A company may decide to have sub-certificates only for roles that deal with the public but not other departments that are internal. For example, they can have a sub-certificate for customer service department but not for individuals. Each sub-certificate could include a description of the limits of authority.

10.2 Costs and incentives are aligned

In this proposal, each player has sensible incentives:

- DNS registrar only performs the lookup function; similar to the phone books that do not authenticate the number.
- SSL certificate vendor guarantees that the web server is owned by the legitimate owner of the domain name (which may or may not be the company intended by the consumer).
- The RealName registrar guarantees that only the legitimate owner can use a RealName (or the "brand"), so consumers can easily tell if they reached the right intended company.
- Entities that want to be authenticated spend money to register and check for imposters. They are motivated by preserving their business (reputation) by minimizing fraud.
- Consumers are interested in authenticating some entities, so they have to pick which registries are trusted.

10.3 Flexible deployment

One major advantage of the overall scheme is that no global cross-provider PKI support is required. Independent, locally-managed PKIs in each jurisdiction do the work. The framework is scalable and flexible by nature. A registry can be set up to address a local business need, and the solution provides third party delegation capabilities whenever an institution needs it.

A local service provider can introduce this type of service to its subscribers without being dependent on the other SPs' policies or willingness to support authenticated name service. Another advantage is the inherent ease of use of the authentication function.

In a lot of cases, end-users are the weak link in the information security chain. Phishing schemes rely on users. Thus, the presentation layer of the authentication scheme should be simple and intuitive, giving no opportunity for ambiguous identification. The solution fits well with these requirements, providing to users three basic types of information, namely the authentication status, the registry provider and the RealName.

10.4 Intuitive understanding of identity

With SSL certificates (including new EV certificates), the ownership of a domain name is determined by the DNS registrar and an SSL certificate merely certifies that the holder of the certificate owns the domain name. It is implicitly assumed that only the "legitimate" owner can own a domain name. This assumption leads to many of the disputes that are resolved under the Uniform Dispute Resolution (UDR). Many of these disputes come from conflicts where both parties have legitimate claim to the domain, or where different jurisdictions have different rules on typosquatting, etc. It may be quite surprising to the average consumer to learn that the certificate does not guarantee that the owner, while legitimate in DNS terms, is the "intended" owner. A RealName certificate, on the other hand, certifies that the holder owns the RealName in a particular jurisdiction. By respecting jurisdiction boundary, we sidestep many of the disputes. Any disputes will have a defined authoritative body with rules that are, by definition, local and well-known to everyone involved. This means the "intended" company will always be the only legitimate owner of that RealName in that jurisdiction. Consequently, it is much more difficult to (successfully) masquerade as a legitimate company.

11. CONCLUSION AND FUTURE WORK

We reviewed the concept of the identity of a company and why this is a nebulous concept. We also reviewed why some of the obvious solutions don't actually solve the identity problem. Our RealName scheme is modeled after the trademark systems and uses an identification model that is independent of DNS. This means there is always a "local" authority to arbitrate name ownership within a single jurisdiction.

There is no dependence on a global PKI. Deployment can start as independent islands and grow to global scale. The underlying authentication technology is flexible enough to allow for name delegation, and permits user anonymity when required.

Even though we believe that DNS is probably not the vehicle for this scheme, it is almost certain that each registry will have a domain name. We still have to make sure that consumers are not fooled by rogue registries. This problem does not arise in DNS since there are "well known" root servers. With RealName, a consumer wishing to import the root certificate of a registry at a remote place (e.g., a Canadian doing business with companies in Nigeria) may have difficulty knowing how to find the correct registry. This problem is likely compounded by each country taking a different approach and/or naming convention. It will probably be useful to establish some convention to mark whether a domain name is a RealName registry.

12. ACKNOWLEDGMENTS

This paper has benefited enormously from comments by the anonymous reviews. We also thank Peter Gutmann and Brad McFarlane for reviewing this article.

13. REFERENCES

[1] Sherr, M. et al. Signaling Vulnerabilities in Wiretapping Systems. *IEEE Security & Privacy, November/December 2005* (Vol. 3, No. 6) pp. 13-25.

[2] Collin Jackson, Daniel R. Simon, Desney S. Tan, Adam Barth. An Evaluation of Extended Validation and Picture-in-Picture Phishing Attacks. In *Proc. Usable Security, 2007*. http://www.usablesecurity.org/papers/jackson.pdf

[3] Andersen, R. Why information security is hard - an economic perspective. In *Proc. Of Computer Security Applications Conference, 2001*. ACSAC 2001.

[4] Housley, R., Polk, W., Ford, W., Solo, D. Internet *X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile*. IETF RFC 3280 (2002).

[5] Anthony Y. Fu, Liu Wenyin, Xiaotie Deng. Detecting Phishing Web Pages with Visual Similarity Assessment Based on Earth Mover's Distance (EMD). *IEEE Transactions on Dependable And Secure Computing, October/December 2006* (Vol. 3, NO. 4).

[6] Parno, B., Kuo, C., Perrig, A. Authentication and Fraud Detection: Phoolproof phishing prevention . In: *Proc. Financial Cryptography and Data Security, 2006*.

[7] Ka-Ping Yee., Kragen Sitaker: Passpet. Convenient Password Management and Phishing Protection. In *Proc. Symposium on Usable Privacy and Security, 2006*. http://cups.cs.cmu.edu/soups/2006/proceedings/p32_yee.pdf

[8] Min Wu, Robert C. Miller, Greg Little, Web Wallet. Preventing Phishing Attacks by Revealing User Intentions. In *Proc. Symposium on Usable Privacy and Security, 2006*. http://cups.cs.cmu.edu/soups/2006/proceedings/p102_wu.pdf

[9] CA/Browser Forum: Guidelines for the Issuance and Management of Extended Validation Certificates, Version 1.0, 2007.

[10] Peterson, J., Jennings, C.: *Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)*. IETF RFC 4474 (2006).

[11] Orange Boxing / Caller ID Hacking FAQ: http://artofhacking.com/files/ob-faq.htm

[12] Voxeo VoiceXML Development Guide http://docs.voxeo.com/voicexml/2.0/frame.jsp?page=t_7.htm

[13] Sawyer, et al.: *System and method for authentication of caller identification*. US Patent and Trademark Office, Patent 6,324,271. 2001.

[14] Santesson, S., Polk, W., Barzin, P., Nystrom M.: *Internet X.509 Public Key Infrastructure Qualified Certificates Profile*. IETF RFC 3039 (2001)

[15] Phillip Hallam-Baker, Outbound Authentication on the Users Terms. In *Proc. 2nd TIPPI Workshop, 2006*.