The Dining Freemasons
(Security Protocols for Secret Societies)

Mike Bond and George Danezis

Computer Laboratory, University of Cambridge,
JJ Thompson Av., CB3 0FD, UK
{Mike.Bond, George.Danezis}@cl.cam.ac.uk

Abstract. We continue the popular theme of offline security by considering how computer security might be applied to the challenges presented in running a secret society. We discuss membership testing problems and solutions, set in the context of security authentication protocols, and present new building blocks which could be used to generate secret society protocols more robustly and generically, including the *lie channel* and the *compulsory arbitrary decision* model.

1 Introduction

Offline security has become a matter of study and interest in the academic computer security community, with aspects of physical lock and safe security being presented by Matt Blaze [1] and air travel security by Bruce Schneier [2]. They have argued quite convincingly, that that security outside the computer world, would benefit from the methodology, analysis and techniques that have been developed to protect computer systems, such as the careful threat modelling, security policies, understanding the strength of mechanism, and relying on small secrets rather than obscurity, as well as these worlds having a few lessons for computer security too.

In this paper we consider the *secret society* – an enterprise in which there has been much security-related innovation historically, but is largely overlooked. We present and categorise techniques that can be used by people, without the assistance of a computer, to authenticate their membership of a secret society and discretely exchange information past a warden. We link these real-word applications with the corresponding fields of authentication protocols and steganography models, but also embrace the constraints of the physical world – which sometimes leads to more elegant solutions.

Secret societies are a common subject matter of fiction novels[1], but of course many secret societies do exist in real life. Other closed-membership associations, from spy rings to the mafia, share the need for secrecy and covert authentication and communication, and would benefit from a more principled approach to security. To a first approximation, a secret society has three functions:

[1] For instance consider Dan Brown's recent bestseller "The Da Vinci Code"

– to recruit the worthy,
– to pass on a secret doctrine,
– and to reward its members.

Each area presents intruiging challenges, but cruicial to each aspect is membership testing – society members must be able to identify each other in order to pass on the doctrine, to confer rewards and to consider new applicants.

2 Membership Testing

How do you determine if someone is a member of your society? Societies with cell structures preclude full knowledge of membership; a member might have to search for another member when travelling in a new region. Alternatively there may be a full membership list, but a given member may not have full access, or the authentication might need to be performed in anonymous circumstances.

The simplest technique is *broadcast*. Each member of the society advertises their membership in a straightforward way to all who care to hear: overt societies use uniform or insignia to achieve this very purpose. Secret societies do the same, but attempt to to hide their broadcast.

2.1 Steganographic Broadcast

A *steganographic broadcast* is a signal visible to all but understood by few. In our model it requires no challenge whether issued deliberately by another member, or conincidentally by a stranger. In Roman times Christians doodled the sign of the fish in the sand to broadcast their membership; done casually with the sandal this is covert. For the signal to remain hidden even when repeated, it must have a low information content. The secret society must make a trade-off in their signal S between certainty of authentication and probability of discovery[2], or in other words choose between false positives and plausible deniability.

However, steganographic broadcasts can conceivably be replayed: the outsider carefully observes a suspected member, then repeats their set of actions exactly. Should the secret signal be discovered, all members of the society are quickly exposed and the society cannot observe that they are under attack.

To overcome the shortcomings of steganographic broadcast, we next consider what can be gained from adding interactivity, and in this case the verbal channel rather than the physical channel is better suited, and will be the focus of our discussion.

[2] Equally important in broadcast signal design is the standard deviation of the information content in the signal. A well designed signal cannot be poorly executed thus confirming membership with high certainty, but revealing the signal to all. The fish in the sand unfortunately has a high standard deviation.

2.2 Interactive authentication

Suppose we permit interaction between the prover and the verifier. Interactive proof is appealing as it reduces the workload on society members who need not keep up the effort of a constant broadcast. It is natural to apply this extension symmetrically, thus we arrive at a *steganographic simultaneous interactive proof*. The verifier uses a key phrase within conversation and the prover then must formulate the correct reponse. The trick of course is to design code phrases effectively, to achieve the usual balance between false positives and deniability. At one end of the scale, in WW2 British Sitcom *'Allo 'Allo* a member of the french resistance plans to authenticate himself to a cafe owner with the following exchange:

LeClerc: "Do you have a light?"
Artois: "I have no matches"
('Allo 'Allo, Pilot Episode)

When a fellow cafe customer lights LeClerc's cigarette before he can utter the first phrase, then leaves his matches at the bar, confusion ensues! An alternative is to use a signal with high information content, for instance when Bond authenticates a CIA agent:

Bond: "In Moscow, April is a spring month."
CIA Agent: "Where as here in St. Petersburg, we're freezing our asses off."
(James Bond, "Goldeneye")

Here the phrase is innocuous, as the high information content is in the exact wording. However, demanding exact wording on repeated authentications quickly damanges deniability, so here the phrase is a session authentication key rather than a long term one. A better approach for repeated authentication is to use multiple rounds of low-information response each slowly adding to the certainty of authentication for both parties.

Interestingly, bi-directional interactive authentication seems more natural to conceive here than the uni-directional counterpart. Uni-directional authentication differs from broadcast as it requires a challenge from the verifier. In fact, totally uni-directional interactive authentication is pointless – the challenges must yield no information as to whether the verifier is a member, thus there will be excessive false positives. However, partially balanced authentication may be a useful primitive: the goal of such an interaction could be for the verifier to deduce that the prover was a member with high probability, whilst the prover may only be able to authenticate the verifier with low probability. The prover enjoys knowledge that someone is *probably* testing their membership, but cannot achieve certainty.

Finally, mixing broadcast and authentication strategies will further reduce the workload on a society member. A well-crafted steganographic broadcast could reduce the number of candidates a member considers for performing full authentication, whilst not marking anyone definititively as a member of the society.

2.3 The Lie Channel

We have summarised the basic structure of steganographic mutual authentication protocols, but an important practical question remains open: how can one design a robust set of hidden phrases for gradual authentication which will work for repeatedly, and endure over a considerable period of time? Furthermore, how can the members commit this to memory?

We suggest exploiting the ability of the human brain to detect lies – to rapidly match a statement against a body of knowledge and determine whether or not it is contradictory. A concrete protocol for gradual mutual authentication can be built as follows. Assume the members of the society share a key, in the form of a 'holy' book B that is only known to the members of the society. Any such book can be used to provide authentication, yet to maintain deniability its subject matter should be appropriate to discuss over dinner (A cooking book, might be perfect, although a play by Shakespeare, or a crime fiction, could also be fine). Let B contains a set of true statements denoted $F_1, F_2, \ldots, F_n$. In the case of *Macbeth* some true statements could be:

> The characters are Macbeth, Lady-Macbeth, Duncan and Macduff. Macbeth is an evil noble^0. Lady-Macbeth is a greedy ambitious woman^1. Duncan is a king^2. Macduff is a loyal noble^3. Macbeth is weak because Macbeth married Lady-Macbeth and because Lady-Macbeth is greedy^4. Lady-Macbeth persuades Macbeth to want to be king.^{5} Macbeth murders Duncan using a knife because Macbeth wants to be king and because Macbeth is evil.^{6} Lady-Macbeth kills Lady-Macbeth.^{7} Macduff is angry because Macbeth murdered Duncan and because Macduff is loyal to Duncan.^{8} Macduff kills Macbeth.^{9} [4]

To initiate the authentication protocol Alice states a true or false fact C_0 from the set of facts $F_{0\ldots n}$. Bob has to reply with a true or false fact R_0, matching the challenge, and provide a second challenge C_1. Alice replies with a true or false fact matching the second challenge. As an example:

Alice: How is Duncan?
I hear he was the king of the casino last night! ($C_0 = F_2 = \text{True}$)
Bob: Another player has come to town, and he is the king now.
But he dominates his wife completely, won't let her play at all.
($R_0 = F_6 = \text{True},\ C_1 = \neg F_4 = \text{False}$)
Alice: Yes I know her, she's so generous though – she'd be useless
as a gambler! ($R_1 = \neg F_1 = \text{False}$)

Alice and Bob simply repeat this protocol until they are certain that the answers they got match, or do not match, the statements in B. Alice could include a new challenge in the third step, making the repeated protocol take on average two steps per round. Their certainty increases exponentially with each round. Note also that it is quite difficult to replay the conversation, since the

challenges that the parties are exchanging are fresh, and will on average require good knowledge of B to determine if they are true or false and answer correctly.

Note also that the protocol, correctly executed, protects the key B. An adversary observing the conversation does not know if a statement is true or false, and often will hear contradictory statements in different conversations. Therefore it is not trivial to reconstruct B fully. In practice marshalling the set of facts from B might seem cumbersome, but indeed many secret (and not-so-secret) societies to require their members to committ to memory large parts of their doctrine. It can often be part of a rite of initiation to recite some true facts, or even to participate in a ritualised (non-steganographic) authentication protocol similar to the above above. Learning a set of true statements along with a set of false statements also seems to be common practice according to the Fishman affidavit [3].

2.4 Deniable authentication

The basic authentication methods described above are straightforward and have certainly been used in practice. However, some situations demand an extra component from the authentication process – forward plausible deniability. Consider a defendant in court who might broadcast his membership of the secret society in the hope that jury or judge would hear. If the key has already leaked outside the society or some member chooses to leak it subsequent to his broadcast, then this broadcast could be used against him. Furthermore if it is the judge or jury members themselves who wish to reassure the defendant that they will support him, how can they do this and not risk incriminating themselves in the event of key compromise?

The solution we require is a *steganographic deniable authentication*: a judge can then authenticate himself to the defendant, but neither defendant nor outsider can ever prove that the authentication took place. If we assume the existence of a *deniable* covert channel – that is a channel that neither party can prove exists – the protocol becomes relatively straightforward.

$$\begin{aligned} A &\xrightarrow[\text{covert}]{} B: \quad N_A \\ B &\xrightarrow[\text{covert}]{} A: \quad N_B \\ A &\xrightarrow[\text{deniable}]{} B: N_A \oplus N_B \end{aligned} \tag{1}$$

Deniable channels generally have very limited bandwidth, and may not be covert, so if some static secret K is transmitted, a replay attack on the channel would be easy. Provision of a challenge by B prevents an attacker recording A's actions in minute detail then performing a replay attack. A's challenge ensures suspected members cannot be linked through giving the same response to sending of a fixed challenge if the attacker repeats B's actions in minute detail. A is authenticated to B as A proves the ability to recover B's nonce. To gain a concrete implementation of this protocol we next need to consider some real world covert and deniable channels.

3 Covert and Deniable Channels

If a channel between two parties cannot be observed by others or provably recorded by either party, then it is a *deniable channel*. Such channels are sometimes *covert channels* in that their existence is not known, but as deniable channels are harder to create than covert channels, a small number of well-known (ie. no longer covert) deniable channels may have to suffice for creation of the protocol.

3.1 The Compulsory Arbitrary Decision Model

The *Compulsory Arbitrary Decision (CAD) Model* is a generic template for creating covert and deniable channels. It carries only a single bit, and this represents a decision chosen, or within reasonable control of the sender, which is *compulsory* – it has to happen one way or the other, and *arbitrary* – the sender might reasonably be expected to choose either way.

This model contrasts building protocols on arbitrary inclusion models, where a particular phrase is said or not said, or where an action is performed or not performed. It is much more difficult to bound the information content of the inclusion of phrases within a conversation, so it is a poorer choice as a building block for secret society protocols.

A good example of a compulsory arbitrary decision (CAD) is when two people leave a room: one of them must walk through the door first. It is quite reasonable for one of the parties to be able to control this, and individuals are unlikely to have strong preferences either way. Interestingly, this channel can be duplex, but collisions will occur if both parties try to transmit the same bit value at once. This channel is covert, but in the presence of a video camera is not deniable.

Other examples of channels with similar characteristics include:

The Chinese Menu Channel. High-capacity CADs are also conceivable, for instance, the choice of item on a menu. When seated at a restaurant, ordering food is as good as compulsory, and if a dish is chosen by number at a chinese restaurant from a selection of maybe one hundred dishes, this could contain at least six bits of information. The channel may be covert, but as the request is verbal, and an itemised receipt is provided it is not deniable.

The Handshake Channel. The most imfamous deniable channel is the *handshake channel*, through which the freemasons allegedly signal their membership through adding pressure with their thumbs on or between particular knuckles of the recipient during a handshake. The appeal of using skin-to-skin contact is that modest pressure applied is easily detected, but difficult to observe through even close visual surveillance. This means that only the giver and the receiver can observe the channel, thus one could easily frame the other. Direct measuring of the handshake pressure would require equipment difficult to conceal on the hand. The handshake is thus an excellent deniable channel, though these days it is not particularly covert. Determining the practical information capacity per handshake is an open question, though it seems it is at least one to two bits.

In cryptology, a quantum cryptography channel exploiting polarisation of photons transmitted down a fibre-optic link represents the ultimate deniable channel. It is an open question whether or not a deniable channel can be created between two parties using purely verbal communication.

4 Other Challenges

Knowledge Set Bootstrapping. It is an interesting challenge to agree upon a shared key between two parties who are not already members of the same society, in the presence of a passive adversary who is observing the entire conversation. Whilst cryptographic solutions involving number theory such as Diffie-Hellman key exchange are fine for computers, they are not much good for humans.

The NSA may be able to intercept any US phonecall, and they may have formidable computing facilities, but if the computation lies in the human world, then their computational bounds are severely reduced.

Two parties can discuss their common knowledge, wheeling through books, moveies, music, religious texts, cryptographic standards – flagging each time when they hit a common area of knowledge. Key material can be efficiently collected from this area using the lie channel (see section 2.3) and then combined with existing key material, for instance using XOR. If their conversation crosses outside the bounds of knowledge of the adversary even for one category, then they have a short key, which can be later used to transmit a message maybe confirming a meeting location.

Counter-Surveillance A variation on this bootstrapping is to work in a common source of data which the attacker will find hard to record or memorise. e.g. the stream of traffic flowing past a window. Alternatively two parties under audio but not video surveillance could point at simple card with "Truth" and "Lie" printed on it to allow them to selectively mislead the eavesdropper.

Semantic Encoding There are already plenty of adequate proposals for hiding information within text, which are of use to those trying to transmit data across a monitored channel [5]. In interesting question is to consider what steganographic techniques could encode a very small amount of data in a body of text that will persist despite radical transforms which only preserve top-level semantics. For instance, suppose the doctrine of the secret society contains a parable. How can we encode in the detail of the story a byte of information in such a way that it will survive translation, summarisation, and incidental errors in the telling of the story, even exaggeration? A parable is an interesting choice as it is clearly an arbitrary story made to illustrate a point, yet because it can be read on many levels details which are not understood are not necessarily perceived as inconsequential by the storyteller.

5 Conclusions

Just as principles of computer security can teach offline security a lot, we have plenty to learn from the offline world too. We believe a study of secret societies

performed in greater depth would reveal some interesting new protocol challenges, and some real-world tools and constraints for which there may not yet be a cryptographic analogue. Furthermore, developing a full set of practical primitives for constructing secret society protocols would be of some use to todays existing secret societies in this age of increasing surveillance.

References

1. M. Blaze, "Toward a Broader View of Security Protocols", Twelfth International Workshop on Security Protocols, Cambridge UK, 26-28 April 2004
2. B. Schneier, Crypto-Gram Newsletter, August 15, 2004
3. The Fishman Affidavit http://www.xs4all.nl/ kspaink/fishman/index2.html.
4. W. Clocksin, University of Cambridge, Computer Science Tripos 2000, Paper 9, Question 9, http://www.cl.cam.ac.uk/tripos/y2000p9q9.pdf
5. Mikhail J. Atallah, Victor Raskin, Michael Crogan, Christian Hempelmann, Florian Kerschbaum, Dina Mohamed and Sanket Naik. "Natural Language Watermarking: Design, Analysis, and a Proof-of-Concept Implementation." Ira S. Moskowitz (Ed.): Information Hiding, 4th International Workshop, IHW 2001, Pittsburgh, PA, USA, April 25-27, 2001, Proceedings. Lecture Notes in Computer Science 2137 Springer 2001, ISBN 3-540-42733-3, pages 185–199