End-User Concerns about Pervasive Internet Monitoring: Principles and Practice

W3C/IAB Workshop on Strengthening the Internet Against Pervasive Monitoring (STRINT)

28th February – 1st March 2014, London, UK

Tara Whalen, Stuart Cheshire and David Singer, Apple, Inc.
Opinions expressed here are solely the opinions of the authors

Abstract

This position paper will discuss pervasive monitoring on the Internet from the perspective of end users: what are overarching concerns around pervasive monitoring, and what are some steps that could be taken to address those concerns? We begin by exploring a preliminary set of *characteristics* of systemic surveillance, which can be used to pinpoint dominant concerns of end users that should be addressed through technical means. We then illustrate one specific significant problem facing end users, namely that of *certificate errors*, which can be exploited to facilitate pervasive surveillance. We suggest that users should not be required to determine whether a certificate error is valid, but instead to block access to websites that generate such errors. We believe this approach would be more effective in protecting end users in an environment of persistent network threats.

Principle: Some Characteristics of Pervasive Monitoring

Given that this workshop is squarely addressing the domain of pervasive monitoring, it may be helpful to look at what the defining characteristics of this concept might be. There is a wealth of research about monitoring in the *surveillance studies* field, which can provide us with a starting point for discussion. Legal scholar Susan Freiwald has proposed a useful framework for considering whether certain surveillance methods ought to require judicial supervision (e.g., a warrant) for their use, given the potentially significant impact of those methods [1]. This framework identifies four factors that should be taken into account when determining whether heightened judicial oversight is required, which should be the case when the method is *hidden, intrusive, continuous*, and/or *indiscriminate*.

The Internet monitoring context that we are discussing here is not identical to that of Freiwald's, in that we are not looking at judicial oversight considerations. Instead, we can broaden Freiwald's analysis, and consider: what aspects of pervasive monitoring demand that *additional countermeasures* be widely deployed, in order to counteract the risk of widespread exploitation? We examine these four characteristics in turn, and what their implications might be for digital countermeasures.

1. *Hidden* surveillance, by Freiwald's analysis, means that it cannot readily be detected by its intended target(s). Consider the situation when this surveillance is being carried out by law enforcement: if the target can determine that they are being watched, then they are provided with the opportunity to challenge the legal legitimacy of that action. In the absence of such indicators, a person would only know if they were informed after the fact; this would argue that strong oversight is necessary to curb possible abuse. In general, a person might also decide to take specific measures to thwart monitoring if they had an indication that such monitoring was happening (or likely to happen).

Within an Internet context, we would need to consider how we might make pervasive monitoring more evident. An example of such an existing transparency tool is the Electronic Frontier Foundation's SSL Observatory [2], which was developed to, among other things, reveal odd (and thus suspicious) behavior in the certificate infrastructure. Such tools provide a means for detecting the subversion of secure communications.

The following two characteristics are similar enough, within the context of Internet monitoring, to be considered jointly:

2. *Intrusive* surveillance is monitoring that yields up a great deal of information about a person. As Freiwald notes, "Because e-mails typically contain more personal data than analogous phone calls or even videos, acquisition of stored e-mail intrudes more on personal privacy than does a wiretap or video surveillance. A simple e-mail message has textual header information that discloses the time it was composed, its subject line plus any attachments, and the electronic addresses of the sender, the recipient, and any who receive courtesy copies of it." Similarly, the Supreme Court of Canada recently required that *specific*, prior judicial authorization was required to search a computer when executing a search warrant for the premises in which that computer is located, stating that "[it] is difficult to imagine a more intrusive invasion of privacy than the search of a personal or home computer." [3]

3. *Continuous* surveillance is of long duration, which means, according to Freiwald, that it is "more likely to be both incriminating and intrusive." This is quite similar to the previous element, although it can be seen as a *series* of intrusions—one can consider it as an escalation in scope. Any intrusive surveillance will be magnified if it persists over time. For network communications, it may not be overly difficult to leave a monitoring system in place for long periods.

These two characteristics map neatly to the workshop theme of *pervasive* monitoring: that which collects large amounts of data over an extended period. When so much personal data is at stake, one needs to ensure a proportional amount of protection for it. There will likely be a number of countermeasures suggested to address this overarching problem, both during this workshop and beyond. As a simple example, there have been calls for default encryption to be more widely deployed through the Internet. Similarly, the provision of perfect forward secrecy (e.g., as provided for through Off-the-Record Messaging [4]) could assist in combating continuous surveillance, by preventing an adversary from using only a single key to decrypt data collected across multiple sessions.

4. *Indiscriminate* surveillance is that which captures irrelevant persons or actions within its scope. The less targeted the monitoring is, the higher the risk is that it will be inappropriate. Matt Blaze recently wrote an article for The Guardian in which he discussed how the focused tactics of intelligence agencies were preferable to the approach of indiscriminate, massive data collection (of often irrelevant information) [5]. He stated that "... targeted operations…have the significant advantage that they leave the rest of us – and the systems we rely on – alone."

Finding countermeasures for indiscriminate monitoring is a challenge; suggested technical solutions generally try to "raise the bar" for untargeted collection—as discussed above, this generally takes the form of widespread (default) encryption and anonymous communications methods such as Tor that serve to confound trivial traffic analysis over wide swathes of digital communications.

Note that these four characteristics are not intended to be an exhaustive set of all aspects of pervasive monitoring, but they may serve as a useful initial step for further exploration and discussion. We will use them as a backdrop for the following section, in which we discuss one specific real-world example of an end-user problem.

Practice: Certificate Errors

Encrypted communications that uses certificates to ensure end-to-end behavior is compromised, and hence enables *hidden*, *continuous* and *indiscriminate* surveillance when the certificates cannot be verified. Worse, it does it when the user is under the impression that they are, in fact, protected (they are using SSL).

Allowing users to ignore X.509 certificate errors allows server administrators not to fix certificate errors promptly, leaving users unprotected and vulnerable. Instead of enabling *users*, this option really enables *sites* not to fix their errors.

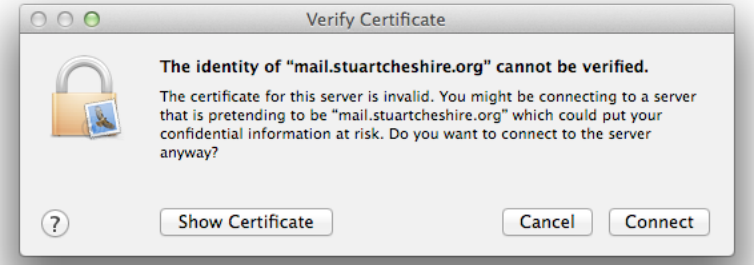
Background

Transport Layer Security and X.509 certificates provide the basis for most of today's secure web browsing, as well as secure access to mail servers and CalDAV calendar servers.

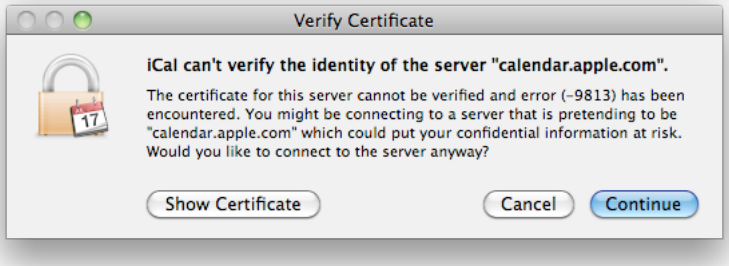
Web browsers, mail clients, and calendar programs offer the user the option of ignoring certificate errors, for reasons of expediency. The reasoning is that the user may not want a misconfigured or expired certificate to prevent them from reading their email, or checking their calendar. The logic is that we should trust the user to be smart enough to know when a certificate error is benign, and when it is cause for alarm. The problem is not simply one of naïve users who are not equipped to make this determination, though that is certainly an important issue. The real problem is that even the most educated and sophisticated user has no choice.

Examples

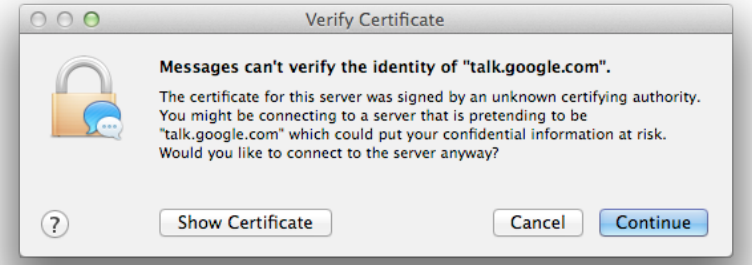
You're at the airport, about to board a flight, and your boss has asked you to send an email before you leave. Your mail program tells you there's a certificate problem. Do you want to connect and send your email, or do you want to explain to your boss why you didn't do what he asked?



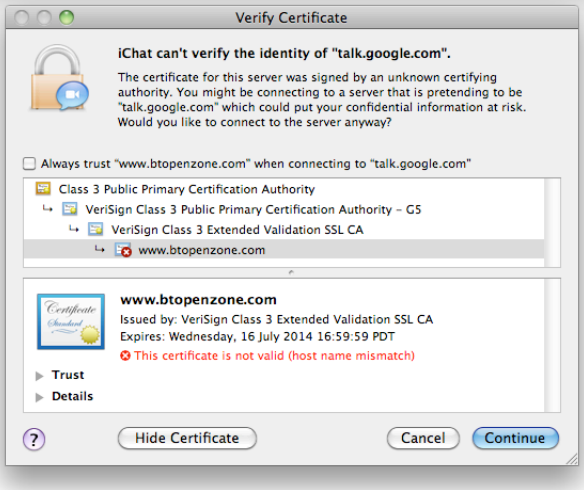
You need to schedule a meeting. Do you want to schedule it or not?



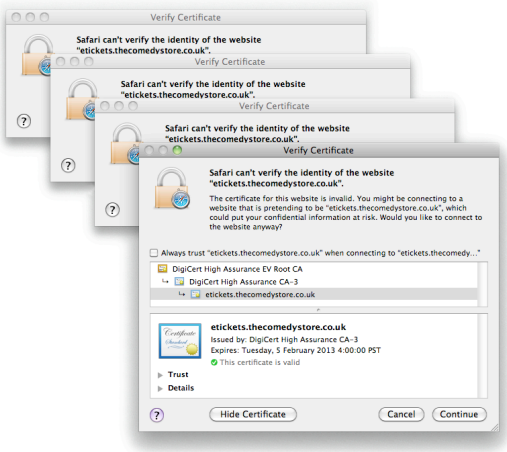
You need to ask a co-worker a question about the meeting, but your Jabber client says there's a certificate problem. Do you want to ask your question, or not?

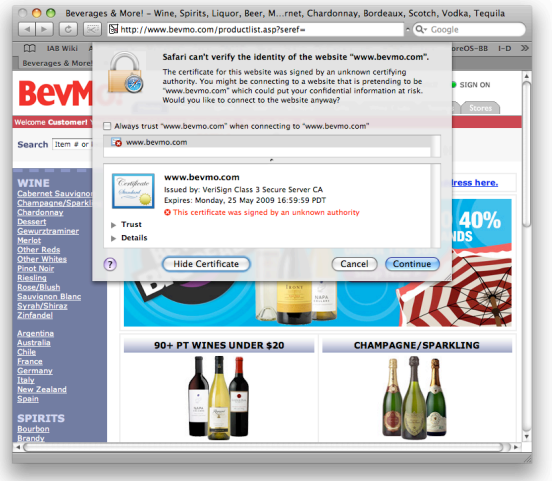


Actually, maybe you shouldn't hit the default "Continue" button quite so quickly in this case. You're at a Wi-Fi hotspot. Your TCP connection really is being intercepted, and you probably don't want to send your Google password to some unknown Wi-Fi hotspot operator.

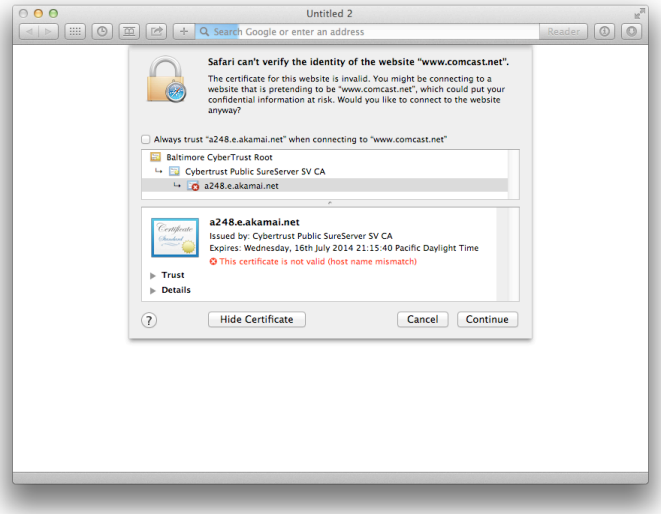


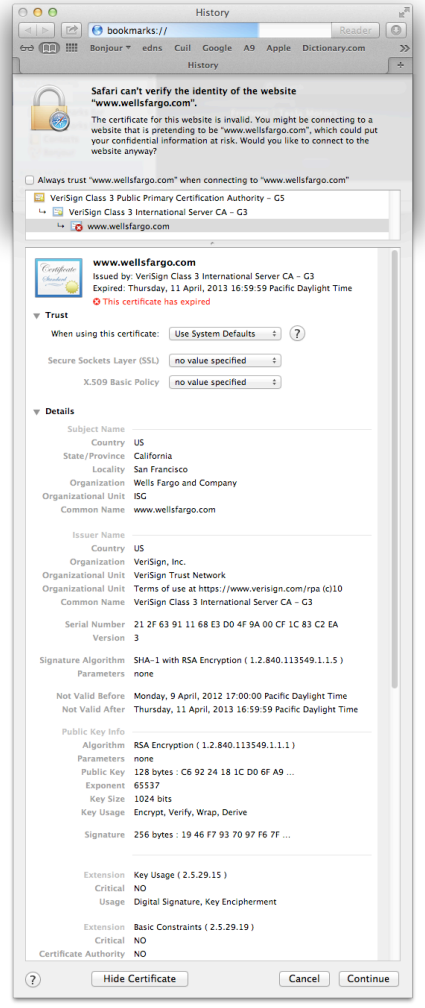
Other cases are not as easy. You're buying tickets for friends to see a comedy show. Do you want to buy the tickets or not? You've been asked to buy beer for a Superbowl party. Do you want to buy the beer or not?

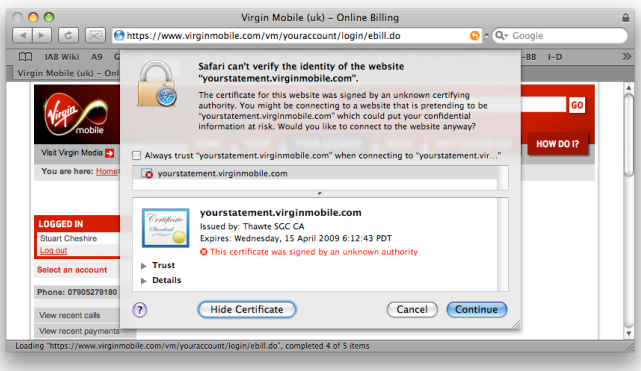




You don't want your Internet connection shut off. Do you want to log on to pay your bill or not? You don't want your mobile phone shut off. Do you want to log on to the web site to pay your phone bill or not? Actually, you have many bills to pay, but your bank web site itself has a certificate error. Do you want to pay your bills, or have services shut off and have to pay reconnection fees?



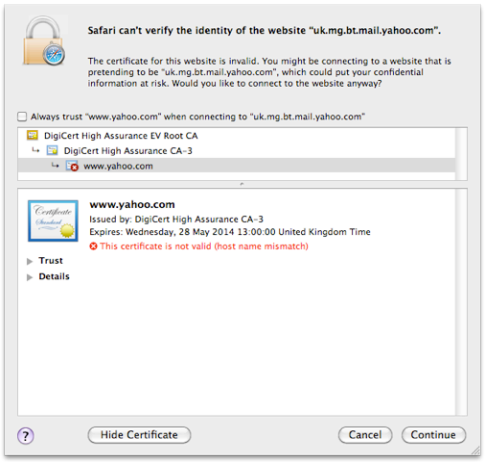




Users face these problems daily, and have little choice but to ignore the certificate errors and get on with their lives: a wealth of studies shows that users often dismiss warning pop-ups in order to get on with completing their actual work [6]. This is a quite rational choice, given the demands and priorities that users have, but it also carries a security risk. Sadly, the administrators running these web sites often don't make fixing certificate errors a high priority. In many cases, this is because they expect people to simply disregard the error and proceed regardless.

One of the authors [Cheshire] personally experienced an example in which a site that his employer asked him to visit had a certificate error. He called the support line of the site in question to report it, since the site could potentially compromise an individual's confidentiality. They actually had this situation covered in their telephone support scripts. He was told that this error message was perfectly normal, and was in fact how he could be sure that he was connecting to the right web site, because it always happened. What he was told to do was what everyone does: Click the blue "Continue" button. After exhausting that avenue, he called his company representative, who also explained that this was normal. He finally found someone who actually understood the issue he was raising, and made some calls to the site, and by the next day the certificate error had been fixed. That was a successful result, but it took an IAB member half a day on the telephone to get a single web site's bad certificate fixed. This is not a scalable solution.

A similar example is the web mail service British Telecom provides for its customers. At the time the service was actually provided by Yahoo under contract to British Telecom, and the web site certificate had a hostname mismatch. Every British Telecom customer checking their email had to ignore the certificate error, potentially exposing their email credentials to malicious attackers.



The fact that ignoring the certificate error was an option the customers had meant it wasn't urgent for BT and Yahoo to fix the problem themselves.

When users are forced to ignore certificate errors, they may as well have no security at all. Since the user has no assurance that they're not falling victim to a man-in-the-middle attack, the fact that the connection is encrypted means nothing, when it might well be an encrypted connection to a malicious attacker.

Conclusion

The time has come for us to add greater protection for users by removing the option that allows them to ignore certificate errors. If the administrators fail to put the hostname into the DNS server, then users will be unable to connect to the service, and their software doesn't give them the option to "Ignore the DNS error." If the server is down, then users will be unable to connect to the service, and their software doesn't give them the option to "Ignore the server-not-responding error." Certificate errors should be treated as similar "hard" failures.

This is not a change that can reasonably be made by one web browser (or other client application, like Mail, or Calendar) independently, since that change would create, in the users' minds, the impression that that particular client was defective, because it could not connect to services that competing clients could successfully connect to. We don't need to have telephone support scripts that tell customers to use a different web browser that does allow them to ignore certificate errors.

For this change to be effective, we need industry-wide support for the notion that security is important, and that ignoring security failures is no longer acceptable if we wish to provide truly effective safeguards for end users.

How do we get there?

References

[1] Freiwald, Susan. "First Principles of Communications Privacy," 2007 Stanford Technology Law Review 3 (2007) <https://ssrn.com/abstract=1132421>

[2] EFF SSL Observatory. <https://www.eff.org/observatory/>

[3] R. v. Vu, 2013 SCC 60. Available online at <https://canlii.ca/en/ca/scc-l/doc/2012/2012canlii31579/2012canlii31579.html>

[4] Off-the-Record Messaging. <https://otr.cypherpunks.ca/>

[5] Blaze, Matt. "NSA revelations: the 'middle ground' everyone should be talking about." The Guardian, 6 January 2014. <https://www.theguardian.com/commentisfree/2014/jan/06/nsa-tailored-access-operations-privacy>

[6] Sunshine, Joshua, et al. "Crying Wolf: An Empirical Study of SSL Warning Effectiveness." 2009 *USENIX Security Symposium*. <https://www.usenix.org/legacy/event/sec09/tech/full_papers/sec09_browser.pdf>