Trust problems in pervasive monitoring

Melinda Shore, Karen O'Donoghue

We open by noting that this workshop is titled "A W3C/IAB workshop on Strengthening the Internet Against Pervasive Monitoring." The word "passive" is not used. Similarly, while the IETF mailing list is called "perpass," for "pervasive, passive monitoring," the word "passive" also does not appear on the mailing list listinfo page[1]. This is fortunate, since while there may be extensive passive monitoring of internet traffic being done by public agencies[2], there has been a cascade of revelations that the US NSA has been taking positive, active steps to eavesdrop on internet traffic, including working with equipment manufacturers to weaken their product design, inserting physical taps in cables outside major data centers, subverting cryptographic keys, and installing radio transceivers inside USB cables.

Active monitoring changes the trust structure of the internet and of the mechanisms we use to implement trust (or rather, to use as the basis for trust decisions).

Trust mechanisms in the IETF have tended to rely on

- Shared secrets
- hierarchies of vouchsafing by trusted third parties
- networks of vouchsafing by trusted third parties
- trust on first contact, or opportunistic encryption (sometimes referred to in the IETF as "better than nothing security")

Which is to say that except for the last bullet point, our trust mechanisms have tended to rely on pre-existing relationships and identity. It's also the case that trust is transacted: credentials are presented, assertions are made, and each party makes a decision whether or not trust is warranted in this particular case. This trust is explicit. At the same time, when we make a decision to establish an unauthenticated connection to a remote server, we are also making a trust decision, as well, based on unprovable assertions (that the server is who it says it is, that it really is the server at a given address, that a given domain name resolves correctly to a given address, etc.). In these cases the trust is implicit.

One characteristic of the pervasive monitoring problem is that we have parties with whom there is no trust relationship, explicit or implicit, inserting themselves into a multiparty communication. They may be doing so by masquerading as one party, they may be eavesdropping on and decrypting traffic, and so on. This is happening outside the trust frameworks we've established and perhaps more significantly outside the assumptions that we make about who's actually participating in our communications.

[1] "perpass -- The perpass list is for IETF discussion of pervasive monitoring," https://www.ietf.org/mailman/listinfo/perpass.

[2] monitoring by private, non-government entities under agreed-upon terms of use is out-of-scope for the time being

Although opportunistic encryption is often mentioned as a remedy to some of the problems introduced by what we've recently learned about internet surveillance, it is clearly extremely fraught from a trust context. What opportunistic encryption offers is some assurance that the party with whom you're continuing to communicate is the one with which you initially established contact[3], and by providing a mechanism to encrypt traffic even in cases where there are loose (or no) assurances about who the corresponding party is, that should, in theory, provide cryptographic protection against casual snooping.

Another mechanism that's been discussed is to provide stronger end-to-end security, for example in the case of messaging privacy[4]. One of the trust problems being addressed by the Crocker paper is that it may be the case that intermediate nodes (for example, mail servers, xmpp servers, etc.) may be collaborating or compromised. You may have established a "trust" relationship with a server using standardized trust mechanisms but that doesn't mean that the server is actually trustworthy. However, end-to-end mechanisms increase the potential for exposing metadata that can lead to inferences about activities and intent, network analysis, and so on. They propose a Tor-like series of relays.

Work has been done recognizing that an authenticated and "trusted" server may not, in fact, be entirely trustworthy and that current mechanisms for bootstrapping encryption in HTTP leaves the server with disproportionate power in determining whether encryption is actually used[5], and proposes a mechanism for clients to use ALPN to identify the use of encryption. In this case, as in the Crocker paper, we have acknowledgment that a "trusted" server may, in fact, be a bad actor.

Unfortunately, however, nearly every internet transaction is leaking both data and metadata, from the 5-tuple in the IP header to DNS lookups to X.509 certificate (and other credential) validation to transaction duration and content length. So far it appears that many of the proposed mediations against surveillance introduce new tradeoffs, some made explicit and some not, and that in the future it may be necessary to reconsider how we evaluate trust decisions.

[3] Note that while ssh is often mentioned as an example of trust-on-first-contact, additional assurances may be provided in the form of server key fingerprints. There's still a bootstrapping problem but it increases the amount of effort required from a miscreant.

[4] Crocker, D. and P. Resnick, "STRINT Workshop Position Paper: Levels of Opportunistic Privacy Protection for Messaging-Oriented Architectures, January 2014, https://ietf.org/doc/draft-crocker-strint-workshop-messaging/

[5] Nottingham, M. "Opportunistic Encryption for HTTP URIs," October 2013, http://tools.ietf.org/html/draft-nottingham-http2-encryption-01