Listen Too Closely and You May Be Confused

Eric Cronin, Micah Sherr, and Matt Blaze

Department of Computer and Information Science
University of Pennsylvania
{ecronin,msherr,blaze}@cis.upenn.edu

1 Introduction

Among the most basic simplifying assumptions of modern communications security is the notion that most communication channels should, by their very nature, be considered vulnerable to interception. It has long been considered almost reckless to suggest depending on any supposed intrinsic security properties of the network itself[1], and especially foolish in complex, decentralized, heterogeneously-controlled networks such as the modern Internet. Orthodox doctrine is that any security must be either end-to-end (as with cryptography), or not considered to exist at all.

While this heuristic well serves cautious confidential communicators, it is unsatisfying from the point of view of the *eavesdropper.* Paradoxically, while end-to-end security may be a prerequisite to robust confidentiality in most networks, it does not follow that a *lack* of end-to-end security always makes it possible to eavesdrop.

In this position paper, we investigate whether the very properties that make it unwise to depend on the network for security can be turned on their head to effectively frustrate eavesdropping. We observe that while the Internet protocol stack and architecture make no confidentiality or authenticity guarantees regarding the traffic that passes across it, neither do they make any guarantees to those who wish to intercept this traffic (whether authorized to do so or not). While interception is often feasible on a benign network if performed with some care, we propose that it may be possible to artificially exacerbate the eavesdropper's problem, to the point of introducing ambiguities that make it unclear how to reconstruct the actual messages passed between targeted parties even when the cleartext is accurately captured.

In particular, at least six properties of the Internet protocol stack and architecture might make it difficult for an eavesdropper to reconstruct a data stream: decentralized control and heterogeneous implementations; "best effort," as opposed to reliable, message delivery that allows data to be re-ordered, duplicated or dropped in transit; shared state and context between communicating parties (e.g., in TCP, and particularly end-to-end error correction); dynamic (and often asymmetric) routing that can change during a flow's lifetime; lack of sender and receiver authentication; and ambiguities in protocols, implementations, and configurations.

[1] Quantum cryptography represents a respectable counterexample, of course.

B. Christianson et al. (Eds.): Security Protocols 2005, LNCS 4631, pp. 245–249, 2007.

These properties mean that a great deal of state information is involved in the correct interpretation of any given packet, and this state is spread across many places, including each of the communicating parties and the network itself. Without complete knowledge of this state, the mere presence of a packet somewhere on the network does not automatically imply that it will be accepted by the recipient given in its header, that it came from the supposed sender, or that it has not been (or will not be) altered, duplicated, or deleted somewhere along its path.

Any intercept system must take into account these properties (and all the corresponding state) in order to ensure not only that it is sufficiently *sensitive* (that it receives all data exchanged between the targets), but also that it is sufficiently *selective* (that it rejects spurious data that is not actually part of the targets' exchange). There are quite a few degrees of freedom that affect how intercepts must be performed. The figure of merit most often considered in judging intercept systems is sensitivity; adequate selectivity, on the other hand, is generally thought to be easily achieved by cursory examination of, e.g., packet headers. In fact, selectivity may be a far more difficult problem than most intercept systems recognize, especially in the presence of deliberate countermeasures.

Depending on the network configuration, many ambiguities can be easily induced, either by one of the communicating parties or by a third party altogether. In fact, as we look at in detail in a forthcoming companion work [8], across much of the protocol stack, from the physical layer to the applications, it is surprisingly simple to introduce data that appears entirely valid but that might not be received and processed by the purported recipient. The Internet appears almost to have been designed to maximize uncertainty from the point of view of those eavesdropping on it.

We believe these properties can be formalized, and, furthermore, that they can often be exploited quite effectively to render it difficult for many interception configurations to obtain reliable transcripts of networked sessions, even when no end-to-end security is employed. Surprisingly, this does not appear to require any bilateral countermeasures to be employed by the communicants, and in some cases, can be performed entirely by a third party.

In particular, we observe that a single party, which we call a *confuser*, can introduce traffic directed at an eavesdropper but that is never actually received (or if received, is rejected) by the ostensible recipient. Depending on the eavesdropper's configuration, and, especially, its position on the network, this traffic can be made indistinguishable from legitimate traffic. In the presence of sufficient confusion, an eavesdropper may be able to be made arbitrarily uncertain as to whether a given intercepted message was real or spurious. We call this indistinguishable-but-spurious traffic *confusion*.

1.1 Related Literature

Investigating the problems associated with electronic eavesdropping is not a new area of research. However, most work in this area has focused on narrow subdomains of the problem space, especially Network Intrusion Detection Systems

(NIDS) [13,14,10,18]. Other related work is found in the field of privacy enhancing technologies, which obscure endpoint identity and evade traffic analysis (but where the content is generally encrypted end-to-end) [15,9]. Both commercial and open source developers have produced a number of tools for eavesdropping, and we look to these as example practical adversaries for evaluation [1,2,11,19,12]. Finally, it is instructive to examine how some of the most important users of eavesdropping, law enforcement and the courts, view the potential for problems with intercepted communications [3,6]. But again, there has been surprisingly little investigation of the relationship between eavesdropping and active third parties, and little work on the "fidelity" of intercepted traffic.

Finally, Rivest's work on Winnowing and Chaffing [16], which uses chaff "noise" to create privacy without encryption is fundamentally related to confusion, albeit in a different context. There, a cooperating sender and receiver achieve end-to-end confidentiality without the use of encryption by sending noise data that is rejected by the receiver by failing a cryptographic checksum that cannot be computed by an eavesdropper. In this proposal, on the other hand, the sender and receiver need not themselves participate in the scheme, and the network topology and protocols themselves frustrate the eavesdropper's ability to distinguish real data from noise.

2 Confusion and Interception Fidelity

Any communications interception system must deliver a sufficiently accurate reproduction of the communications between the targeted parties to satisfy its requirements. That is, the quality of an interception must be of sufficient "fidelity" to be useful, much as the quality of an audio recording must be of high enough fidelity to satisfy its listener. How high the fidelity must be depends on the application; in a non-targeted diagnostic system, such as might be used in a network operations center, low fidelity may be acceptable, while in a law enforcement intercept conducted under a court order intended for producing legal evidence, higher fidelity (and with known properties) may be required.

In a digital network, several factors affect interception fidelity, including whether the relevant data are captured ("sensitivity") and whether irrelevant noise is correctly discarded ("selectivity").

The fidelity of eavesdropping systems, especially those operating in the presence of active attacks, has not been extensively investigated[2]. Instead, much of the prior work relating to electronic eavesdropping has focused on the problem of sensitivity; that is, how to adequately capture the data that is transported across the communication medium. In particular, much attention has been made to preventing *evasion* attacks [14,17] in which an attacker attempts to bypass an electronic wiretap by crafting abnormal traffic that escapes interception.

Our focus is on analyzing and developing attacks (and countermeasures to attacks) against selectivity, rather than simply evasion. An interception system

[2] We are not, however, the first to question the reliability of eavesdropping systems. See [4,5].

is susceptible to *confusion* if it is possible to direct apparently valid traffic to the eavesdropper that is not detected (or, if detected, is rejected) by the targeted receiver. In fact, across much of the protocol stack, from the physical layer to the applications, it is surprisingly simple to introduce data that appears entirely valid but that might, or might not, be received and processed by the intended recipient.

3 Preliminary Steps

Our exploration into confusion to date has just scratched the surface, but already we have seen interesting and confirming results [8]. We have developed a basic model of confusion, and used this to analyze how "confusable" different protocols are. To explore the more practical side of confusion, we have also constructed several confusion generators for Internet communications. Surprisingly, even the least confusable of protocols contains enough ambiguity to thwart current eavesdropping tools.

4 Conclusion

As we have seen, accurate capture of all cleartext is not by itself sufficient to ensure accurate message reconstruction by an eavesdropper, and rejection of spurious traffic is not a trivial problem. To build an accurate and reliable message stream, an eavesdropper must simulate correctly the entire delivery process, about which there can be significant uncertainty in real networks.

Our analysis and preliminary experimental work suggest that unilateral and third party countermeasures can greatly increase this uncertainty, and relatively simple confusion injection techniques can effectively thwart many Internet interception systems.

There has been remarkably little research investigating the fidelity that can be achieved in eavesdropping on digital networks, and existing work does not consider second- and third-party active countermeasures that thwart the eavesdropper without the use of end-to-end techniques.

From a communications security standpoint, confusion is a potentially interesting "third" technique, alongside cryptography and steganography for increasing confidentiality. Current network architectures, exemplified by the Internet and many sensor network systems, appear to have structures that make eavesdropping vulnerable to our approaches.

Confusion may also have legal and public policy implications. Network-based interceptions are said to be increasingly important as evidence in (criminal and civil) legal cases and as an intelligence tool for law enforcement and national security. But how reliable are network wiretaps? Can they be thwarted by active techniques such as confusion? Are there systematic techniques for analyzing an interception system to measure is reliability and fidelity? How should the courts treat interception evidence?

Current US law on the treatment of electronic evidence is remarkably inconsistent; depending on context, computer-based data is either accepted almost uncritically or rejected out-of-hand as unreliable. For example, the Department of Justice prosecutor's manual on computer evidence [7] cites at least four contradictory, yet controlling, cases on this issue. Well-understood answers to the questions above could help inform the increasingly important legal and policy debate on wiretap evidence and how it should be collected and treated.

References

1. NetIntercept. http://www.sandstorm.net/products/netintercept/
2. NetWitness. http://www.forensicsexplorers.com/
3. Electronic Crime Scene Investigation: A Guide for First Responders (July 2002), http://www.ojp.usdoj.gov/nij/pubs-sum/187736.htm
4. Bellovin, S.M.: Wiretapping the net. The Bridge 20(2), 21–26 (2002)
5. Blaze, M., Bellovin, S.M.: Inside RISKS: Tapping, tapping on my network door. Communications of the ACM 43(10) (December 2000)
6. Casey, E.: Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet (2004)
7. Computer Crime and Intellectual Property Section. Criminal Division. United States Department of Justice. Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations (July 2002), http://www.cybercrime.gov/s&smanual2002.htm
8. Cronin, E., Sherr, M., Blaze, M.: On the reliability of Internet eavesdropping. (submitted for publication, February 2005)
9. Dingledine, R., Mathewson, N., Syverson, P.: Tor: The Second-Generation Onion Router. In: Proc. of the 13th Usenix Security Symposium, pp. 303–320 (August 2004)
10. Handley, M., Kreibich, C., Paxson, V.: Network intrusion detection: Evasion, traffic normalization, and end-to-end protocol semantics. In: Proc. of the 10th Usenix Security Symposium (August 2001)
11. Jacobson, V., Leres, C., McCanne, S.: tcpdump. http://www.tcpdump.org/
12. Lightfoot, C.: Driftnet. http://www.ex-parrot.com/~chris/driftnet/
13. Paxson, V.: Bro: a system for detecting network intruders in real-time. Computer Networks (Amsterdam, Netherlands: 1999) 31(23–24), 2435–2463 (1999)
14. Ptacek, T., Newsham, T.: Insertion, evasion, and denial of service: Eluding network intrusion detection. Technical report, Secure Networks, Inc. (1998)
15. Reiter, M.K., Rubin, A.D.: Crowds: Anonymity for web transactions (1998)
16. Rivest, R.: Chaffing and winnowing: Confidentiality without encryption (March 1998), http://theory.lcs.mit.edu/~rivest/chaffing.txt
17. SANS. Intrusion detection FAQ: How does fragroute evade NIDS detection? (2002), http://www.sans.org/resources/idfaq/fragroute.php
18. Shankar, U., Paxson, V.: Active mapping: Resisting NIDS evasion without altering traffic. In: Proc. of the 2003 IEEE Symposium on Security and Privacy, IEEE Computer Society Press, Los Alamitos (2003)
19. The Ethereal Project. Ethereal: A network protocol analyzer. http://www.ethereal.com/