Information Theory of Covert Timing Channels[1]

Aaron B. Wagner [a,2], and Venkat Anantharam [b]

[a] *Coordinated Science Laboratory, University of Illinois at Urbana -Champaign, Urbana, Illinois 61801, USA and School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA.*

[b] *Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 USA.*

Abstract. We review our recent work on the reliability function of the timing channel associated to the first in first out exponential-server queue. This result may be of use in understanding the limits to communication over covert timing channels arising in networks.

Keywords. covert communication, error exponent, point process, Poisson channel, queueing theory, timing channel, zero-rate reliability

1. Introduction

A communication network shared by users with different levels of security clearance will be called a *multi-level* communication network. A user with a high level of security clearance will be called High and a user with a low level of security clearance will be called Low. A *covert* communication channel in a multi-level communication network may be defined as one whose presence was not envisioned in setting up the security boundaries, and which therefore may permit unauthorized means of communication. A common security problem in multi-level communication networks is the existence of covert timing channels by which information can be leaked across security boundaries. High may be able to communicate classified information to Low without this security violation being detected.

The information capacity of covert timing channels in a network can in principle be arbitrarily large even if the available raw bit rates in the network are bounded: consider the theoretical limit when the covert transmitter and receiver are able to agree on a time reference with arbitrary precision and the transit times in the network are deterministic. Anantharam and Verdú [1] studied the information capacity of communicating via timing over a first in first out (FIFO) single-server queue with independent service times from

[1]This research was supported by DARPA Grants F30602-00-2-0538 and N66001-00-C-8062, by ONR Grant N00014-1-0637, by the NSF TRUST program, by NSF Grant ECS-0123512, and by a Graduate Research Fellowship from the NSF.

[2]E-mail: aaronw@uiuc.edu

customer to customer, and demonstrated that this capacity is smallest, equal to μ/e for queues with service rate μ, when the service time is exponentially distributed.

In this document we will restrict attention to the case of [1] with exponentially distributed service times. The covert timing channel in this case is called the exponential-server timing channel (ESTC). In exploiting this timing channel, High codes information into the times at which he/she sends packets into the queue, and Low runs a decoding algorithm based on the times at which he/she receives the packets. The content of the packets may be completely innocuous: they may consist entirely of unclassified material or material which is publicly available (e.g. jokes). While we focus on the ESTC, it is easy to believe that it may be possible in practice to covertly communicate information at very high rates across security boundaries in ostensibly secure multi-level networks, given the extremely high raw data rates possible in today's networking technology. Indeed, this should be possible even in the presence of cross traffic (other communication sessions using the same resources) and even though delays due to resource contention in the network result in Low only seeing a noisy version of the timing information conveyed by High. Fundamental investigations of such channels, and of techniques to detect their use and to reduce their capacity, appear to be worthwhile.

There is already an extensive literature on covert channels in computer and communication systems. It is not our purpose to give a survey of this area here. However, we will now selectively mention some excellent articles one might profitably read to gain perspective on this area, with apologies to authors whose work we have not cited. This will allow us to establish a context for our problem formulation and results.

A workable, if not entirely satisfactory, definition of a covert channel is that it is one that permits communication by exploiting entities not normally thought of as data objects. This paraphrases definitions proposed by Lampson [10] and Kemmerer [9]. The timing channel associated to a FIFO queue certainly appears to satisfy this definition. The existence and importance of covert channels, according to this working definition, has long been recognized. The communication strategy of the mole who visits the grocery store every evening with the understanding that he enters between 6 p.m. and 7 p.m. if he wishes to set up a drop for new information he has learned – while only visiting the store after 8 p.m. if he has no information to convey – is just one example from the popular spy literature.

The granddaddy of all papers on the information theoretic view of secrecy systems is that of Shannon [19]. While this paper does not explicity address covert channels, it is well worth reading for its broad discussion of the meaning of secret communication (now of course a bit outdated by new developments such as quantum cryptography). The suggestion in [19] to discuss the quality of a cipher in terms of the work required to decipher it is broadly in the same spirit as our (informal) interpretation of the reliability function in terms of the probability of detection of the attempt to communicate covertly, see Section 2. Exploring the latter connection more formally would be an interesting topic for further research.

The paper [10] is one of the earliest and most cited papers highlighting the importance of studying covert channels. Much of the literature on the subject also cites the Bell & LaPadula model [3] for access control policies in multi-level systems, which was aimed at ensuring the existence of security boundaries in such systems. Wray [24] points out, through an example, the slipperiness of a commonly perceived distinction between covert storage channels and covert timing channels. The paper of McLean [12] gives a good overview of the area as of the time of its writing, as does [9].

Interesting papers that discuss the use of bit fields in the headers of packets in a packet switched network for covert communication include those of Servetto and Vetterli [18] and Jones et al. [6]. If one believes that the header field is “not normally thought of as a data object" it is reasonable to call these channels covert. [1] The discussion in [6] of using this kind of covert channel in a constructive way to detect the presence of distributed denial of service attacks is particularly notable.

A fascinating paper of Simmons [20] reminisces on the history of strategic arms limitations negotiations between the United States and the former Soviet Union, and in particular on the discovery and role of certain covert channels that appeared to be present in some of the proposed verification protocols. Recently Li and Ephremides [11] have discussed a covert channel arising from the decisions made by the agents involved in the kinds of splitting algorithms commonly used in some multiple access protocols.

Finally, there has been a steady stream of work on covert timing channels, which is closest in spirit to the concerns addressed by our talks at the workshop. Information theoretic analyses of the capacity of timing channels, in the context of some bare-bones models, may be found in the papers of Moskowitz and Miller [15], Moskwitz and Kan [14] and Moskowitz et al. [13], among others. An idea to mitigate the capacity of such covert timing channels is discussed in [14], and in the papers of Kan and Moskowitz [7] and Kan et al. [8], among others. A noteworthy recent contribution to the problem of mitigating the capacity of covert timing channels is that of Giles and Hajek [5], who study this problem in the framework of a game between the covert transmitter-receiver pair and a jammer who might have delay or buffer size constraints. The jammer attempts to retime the packets from the transmitter. The mutual information between the input and output processes is taken as the objective of the game: the jammer wishes to reduce this, while the covert transmitter-receiver pair would like this to be high. The existence of a value for this game, in the sense of zero-sum game theory, see e.g. Owen [17], is proved in certain cases. Coding-decoding and jamming schemes that realize the value are also provided in some cases.

2. The relevance of the reliability function

When communicating covertly over the ESTC, the incentive for the sender to use a code with a short blocklength goes beyond minimizing the coding delay: a short blocklength decreases the likelihood of detection by the observer. A shorter blocklength generally comes at the expense of a higher error probability. Since the reliability function of a communication channel describes the tradeoff between the blocklength of a code and its error probability, see e.g. Gallager [4] for more details, it also captures the tradeoff between the probability of error and the probability of detection by the observer, assuming that the sender uses codes that are optimized for communication. Motivated by this (and other considerations) we carried out a study of the reliability exponent for communica-

[1] One might likewise wonder if it is reasonable to use the term “covert channel" for the transfer of information by steganographic techniques such as watermarking (see e.g. Moulin and O’Sullivan [16]). Is the image being watermarked a “data object" ? Is the paper carrying the watermark (in addition to the visible information) a “data object" ?

tion over the ESTC. This work has already been published in Wagner and Anantharam [22] and Wagner and Anantharam [23], and was the work we discussed at the workshop. We summarize the main results here. For proofs and other contextual details, see [22,23].

3. Summary of results

We assume that the queue is initially empty and that the service discipline is FIFO. We work with the definition of capacity region used by Sundaresan and Verdú [21], which is technically more convenient than the one in [1]; under both definitions the capacity of the channel is μ/e nats per unit time when the service rate is μ. Random-coding and sphere-packing bounds on the reliability function of the ESTC are proved by Arikan [2], and these coincide at rates between $(\mu/4)\log 2$ and μ/e (the capacity).[2] The lower bound for the reliability function, i.e. the random coding bound, equals $\mu/4$ at zero rate, while the upper bound, i.e. the sphere packing bound, equals μ at zero rate. Indeed, there is a gap between these bounds at rates below $(\mu/4)\log 2$. We proved that the *zero-rate reliability* of the ESTC, defined as the limit of the reliability function as the rate approaches zero, equals $\mu/2$. We also proved an improved upper bound on the reliability function of the ESTC at positive rates up to rate $(\mu/4)\log 2$. The proofs use some novel point-process techniques. Specifically, we need to define a distance metric over inputs to timing channels, which parallels Euclidean and Hamming distance for conventional channels, which we then use to bound the error probability of a pair of codewords, when used over the ESTC, in terms of the distance between them.

Formal statements and proofs of our results, as well as some parallel results for general service time distributions, are available in [23], to which we refer the reader. We have also determined the reliability function of the ESTC at rates above the capacity (this is defined as the exponent of the average probability of correct decoding); the precise statement of these results were reported in [22], and a journal paper on this topic, with complete proofs, is under preparation.

References

[1] V. Anantharam and S. Verdú, Bits through queues, *IEEE Transactions on Information Theory* **42:1** (1996), 4 –18.

[2] E. Arikan, On the reliability exponent of the exponential timing channel, *IEEE Transactions on Information Theory* **48:6** (2002), 1681 –1689.

[3] D.E. Bell and L.J. LaPadula, Secure Computer Systems: Unified Exposition and Multics Interpretation, *Technical Report* **MTR-2997**, MITRE Corporation, 1976.

[4] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, New York, NY, USA, 1968.

[5] J. Giles and B. Hajek, An information-theoretic and game-theoretic study of timing channels, *IEEE Transactions on Information Theory* **48:9** (2002), 2455 –2477.

[6] E. Jones, O. Le Moigne, and J.-M. Robert, IP Traceback Solutions based on Time to Live Covert Channel, *Proceedings of the 12th IEEE Conference on Networks*, (2004), 451 –457.

[2]Here $\log$ denotes the natural logarithm.

[7] M.H. Kang and I.S. Moskowitz, A pump for rapid, reliable and secure communication, *Proceedings of the 1993 ACM Conference on Computer and Communication Security*, (1993), 119 –129.

[8] M.H. Kang, I.S. Moskowitz, and D.C. Lee, A network pump, *IEEE Transactions on Software Engineering* **22:5** (1996), 329 –338.

[9] R.A. Kemmerer, A Practical Approach to Identifying Storage and Timing Channels: Twenty Years Later, *Proceedings of the 18th Annual Computer Security Applications Conference*, (2002), 109 –118.

[10] B.W. Lampson, A Note on the Confinement Problem, *Communications of the ACM* **16:1** (1973), 613 –615.

[11] S. Li and A. Ephremides, A Covert Channel in MAC Protocols Based on Splitting Algorithms, *Proceedings of the IEEE Wireless Communications and Networking Conference*, (2005), 1168 –1173.

[12] J. McLean, The Specification and Modeling of Computer Security, *Computer* **23:1** (1990), 9 –16.

[13] I.S. Moskowitz, S.J. Greenwald, and M.H. Kang, An analysis of the timed Z-channel, *Proceedings of the 1996 IEEE Computer Society Symposium on Security and Privacy*, IEEE Computer Society Press, Oakland, CA, USA, (1996), 2 –11.

[14] I.S. Moskowitz and M.H. Kang, Discussion of a statistical channel, *Proceedings of the IEEE-IMS Workshop on Information Theory and Statistics*, IEEE Press, New York, NY, USA, (1994), 95.

[15] I.S. Moskowitz and A.R. Miller, The Channel Capacity of a Certain Noisy Timing Channel, *IEEE Transactions on Information Theory* **38:4** (1992), 1339 –1344.

[16] P. Moulin J.A. O'Sullivan, Information-theoretic Analysis of Information Hiding, *IEEE Transactions on Information Theory* **49:3** (2003), 563 –593.

[17] G. Owen, *Game Theory*, Academic Press, San Diego, CA, USA, (1968).

[18] S.D. Servetto and M. Vetterli, Communication Using Phantoms: Covert Channels in the Internet, *Proceedings of the IEEE International Symposium on Information Theory*, Washington DC, USA, (2001), 229.

[19] C.E. Shannon, Communication Theory of Secrecy Systems, *Bell System Technical Journal* **28** (1949), 656 –715.

[20] G.J. Simmons, The History of Subliminal Channels, *IEEE Journal on Selected Areas in Communication* **16:4** (1998), 452 –462.

[21] R. Sundaresan and S. Verdú, Robust decoding for timing channels, *IEEE Transactions on Information Theory* **46:2** (2000), 405 –419.

[22] A.B. Wagner and V. Anantharam, Feedback, queueing, and reliability of the ideal Poisson channel above capacity, *Proceedings of the IEEE Symposium on Information Theory*, (2004), 447.

[23] A.B. Wagner and V. Anantharam, Zero-rate reliability of the exponential-server timing channel, *IEEE Transactions on Information Theory* **51:2** (2005), 447 –465.

[24] J.C. Wray, An Analysis of Covert Timing Channels, *Proceedings of the 1991 IEEE Computer Society Symposium on Security and Privacy*, IEEE Computer Society Press, Oakland, CA, USA, (1991), 2 –7.