On the channel capacity of network flow watermarking

Amir Houmansadr* Todd Coleman* Negar Kiyavash† Nikita Borisov*
*Dept. of Electrical and Computer Engineering
†Dept. of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana–Champaign
{ahouman2,colemant,kiyavash,nikita}@illinois.edu

ABSTRACT

We consider the information theoretical capacity of the network flow watermarking problem. Finding capacity bounds of the watermarking problem can lead to compare different watermarking schemes and help in designing more efficient network flow watermarks.

1. INTRODUCTION

Recently, network flow watermarking has been used as an active notion of traffic analysis to provide better detection efficiency compared to traditional passive traffic analysis [7, 8] . Network flow watermarking has been suggested to be used for detection of stepping stone attacks [8, 4] as well as compromising anonymous networks [5, 6]. Different watermarking schemes have been designed, mainly by borrowing ideas from multimedia watermarking and applying them to the context of network flows.

The proposed watermarking schemes have been evaluated considering different features of the system, e.g., watermark invisibility, detection efficiency, and robustness to perturbations. However, the literature lacks information theoretical evaluations on the network flow watermarks. In this research, we aim in modeling the network flow watermarking problem for the context of information theory, and investigate the information theoretical bounds on the watermarks inserted into network flows. We consider the *private watermarking* problem, as would be discussed in the next section, and leave the *public watermarking* problem for the future research.

2. MODELING NETWORK FLOW WATERMARKING

Network flow watermarking schemes can be classified in two main categories: private watermarking and public watermarking. Private watermarking is the focus of this research and is described in the following.

Network flow watermarking is done by making perturbations on the timings of network lows. A watermark message, W, is generate and shared between watermarkers and watermark detectors. Figure 1 sketches the model of private flow watermarking. In private watermarking scheme, in contrary to the public watermarking, the timing information of the flows to be watermarked is also shared between watermarker and detector in addition to the watermark message. We model this as the side information, S, for the watermarking channel which is shared between watermarking parties. As shown in Figure 1, encoder generates the watermarked timing information, X, using the side information S and the watermark message W. This watermarked sequence passes through the channel and is used by detector to extract the watermark sequence using the shared side information S. The flow is declared to be watermarked if the extracted watermark $\tilde{W}$ matches the shared watermark W.

The described private watermarking problem resembles the *channels with side information* problem in the communications literature. The model is analysed in the following.

3. CHANNEL CAPACITY

Wolfowitz considers the capacity problem for a channel with side information between sender and receiver in [9], and Caire et al. considers a more general capacity problem [1]. In this research, we take a different proof for the capacity of a channel with shared side information (Figure 1).

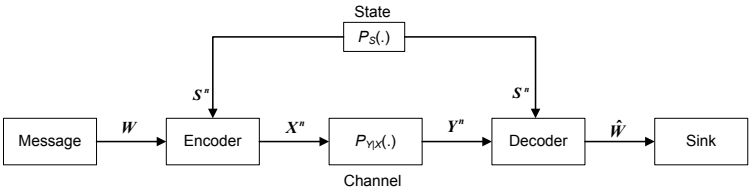


Figure 1: Channel with shared side information

3.1 Achievable region: memoryless channel

We consider the following random coding scheme to find an achievable region for the private watermarking model of Figure 1.

Codebook– Make a table with $M = 2^{nR}$ rows and $N = |\mathcal{S}|^n$ columns, where $|\mathcal{S}|$ is the size of side information S. To each cell of this table assign n bits randomly from the distribution $P_{X|S}$.

Encoding– For $W = w$ and $S^n = s^n$ send content of the cell at (w, s^n) in the table as the $X^n = x^n$.

Decoding– For any Y^n recieved, detector has knowledge about the corresponding $S^n = s^n$ and also the Codebook. So for any Y^n detector tries to find the cell in the codebook so that $(X^n(\hat{w}, S^n = s^n), Y^n) \in A_\epsilon^{(n)}(P_{XY|S})$. Detector returns $\hat{w}$ as the detected symbol. $A_\epsilon^{(n)}(.)$ is the *typical set* which is defined in [3].

The mentioned coding scheme results in the capacity of the private channel to be

$$R_{private} = \max_{P_{X|S}} I(X;Y|S) \tag{1}$$

where $I(;)$ is the mutual information [3].

3.2 Achievable region: non-memoryless chanel

The random coding scheme used in finding achievable region for the memoryless channel is based on having jointly typical sets and holding the *AEP* property [3]. To use the same random coding approach for the case of non-memoryless channel one should demonstrate how the AEP holds in order to find the achievable regions. Coleman et al. have recently used the same approach to prove the channel capacity for the exponential server timing channels [2].To do so, they consider the queue state of the non-memoryless channel, trying to model it as a non-memoryless channel. We use the same approach to prove the channel capacity for the non-memoryless case of the private watermarking problem. This results in the similar results for the capacity as in the case of the memoeyless channel.

4. REFERENCES

[1] G. Caire and S. Shamai. On the capacity of some channels with channel state information. *IEEE Transactions on Information Theory*, 45(6):2007–2019, 1999.

[2] T. P. Coleman. A simple memoryless proof of the capacity of the exponential server timing channel. In *IEEE Information Theory Workshop*, 2009.

[3] T. M. Cover and J. Thomas. *Elements of Information Theory*. New York: Wisley, 2nd edition, 2006.

[4] A. Houmansadr, N. Kiyavash, and N. Borisov. Rainbow: A robust and invisible non-blind watermark for network flows. In *ndss*, Feb. 2009.

[5] X. Wang, S. Chen, and S. Jajodia. Tracking anonymous peer-to-peer VoIP calls on the Internet. In C. Meadows, editor, *ACM Conference on Computer and Communications Security*, pages 81–91, New York, NY, USA, Nov. 2005. ACM.

[6] X. Wang, S. Chen, and S. Jajodia. Network flow watermarking attack on low-latency anonymous communication systems. In B. Pfitzmann and P. McDaniel, editors, *IEEE Symposium on Security and Privacy*, pages 116–130, May 2007.

[7] X. Wang, D. Reeves, and S. F. Wu. Inter-packet delay based correlation for tracing encrypted connections through stepping stones. In D. Gollmann, G. Karjoth, and M. Waidner, editors, *European Symposium on Research in Computer Security*, volume 2502 of *Lecture Notes in Computer Science*, pages 244–263. Springer, Oct. 2002.

[8] X. Wang and D. S. Reeves. Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays. In V. Atluri, editor, *ACM Conference on Computer and Communications Security*, pages 20–29, New York, NY, USA, 2003. ACM.

[9] J. Wolfowitz. *Coding Theorems of Information Theory*. Springer-Verlag, 3rd edition, 1978.

ON THE CHANNEL CAPACITY OF NETWORK FLOW WATERMARKING

Amir Houmansadr Siva Gorantla Todd Coleman Negar Kiyavash Nikita Borisov
University of Illinois at Urbana-Champaign

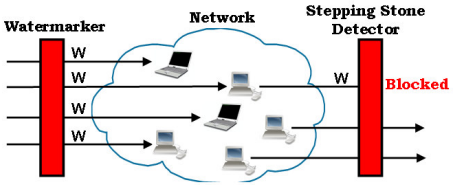
Network Flow Watermarking

Network flow watermarking is manipulating content independent patterns of network flows, e.g., packet timings, in order to perform traffic analysis.

Applications

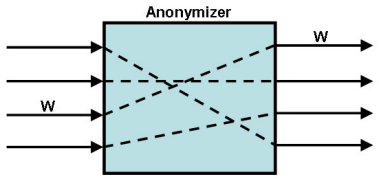
1- Stepping Stone Detection

Stepping stones are relays used by network intruders in order to conceal their identities.
• Flow watermarking can be used to detect relayed traffic.



2- Compromising Anonymity

Anonymous networks try to hide the relation between senders and receivers of network flows, e.g., TOR.
• Colluding attackers can use flow watermarking to break anonymity promises by linking senders/receivers.



Problem statement

Lack of information theoretical analysis of network flow watermarking in the literature

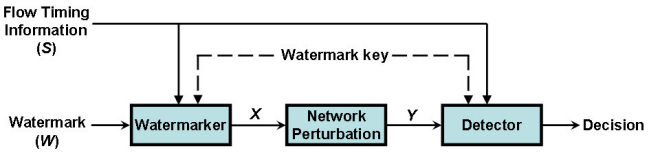
Challenge:

Non-memoryless behavior of timing channels

Flow watermarking types

1. Private watermarking: access to original flow at the detector
   • Focus of this research
2. Public watermarking: no access to original flow at the detector

Private watermarking model



• System modeling:
  • Flow timing information: side information *S*
  • Watermark sequence: message to be communicated
  • Computer network: a non-memoryless communication channel for timing information
  • Side information is shared between watermarker (encoder) and watermark detector (decoder)
  • Watermark key shared between sides

Capacity of memoryless channel

A. Asymptotic Equilibrium Property

• **Asymptotic Equipartition Property** (AEP) holds for a random process *X* if the empirical entropy is ε-close to its true entropy [2]:

$$-\frac{1}{n}\log p(X_1, X_2, \ldots, X_n) \xrightarrow{prob.} H(X)$$

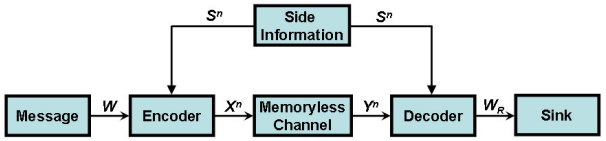
➢ Analog of the law of large numbers in information theory
• Theorem: AEP holds for i.i.d. processes
• The high probability region containing such sequences is called the **typical set.**

Joint AEP: holds for two random processes *X* and *Y* if the empirical marginal and joint entropies are ε-close to the true entropies [2].
• Theorem: Joint AEP holds for (X^n, Y^n) drawn i.i.d. according to

$$p(x^n, y^n) = \prod_i p(x_i, y_i)$$

B. Communication channel with shared side information

• This problem is equivalent to the private watermarking problem except for the channel
• Communication channel is memoryless



• The channel capacity is found to be [1]: $C=I(X;Y|S)$
• Proof methodology:
  • Generating a random codebook with rows and columns corresponding to messages and side info, respectively.
  • Encoder sends the appropriate cell content from the codebook
  • By receiving the altered message, Y^n, receiver looks for a cell from the codebook whose content is *jointly typical* with the received sequence. The index of this cell is returned as the message.

Non-memoryless approach

• We use the Exponential Server Timing channel (ESTC) to represent the flow watermarking channel.
➢ The system model is the same as 6.2 except for the channel which is non-memoryless → so, AEP does not hold generally.
❖ We intend to leverage the results of 6.2 by:
  ❖ using the observation of [3] that the ESTC channel is memoryless conditioned upon intermediate queue states
  ❖ show that AEP holds for some coding scheme

References

[1] J. Wolfowitz. Coding Theorems of Information Theory. Springer-Verlag, 3rd edition, 1978.
[2] T. M. Cover and J. Thomas. Elements of Information Theory. New York:Wisley, 2nd ed., 2006.
[3] T. P. Coleman, "A Simple Memoryless Proof of the Capacity of the Exponential Server Timing Channel," IEEE Information Theory Workshop, 2009.