AFOSR Final Report

Grant/Contract Title: "Low-Complexity Methods for Provably Good Information Transmission and Network Anomaly Detection via Packet Timings in Networks"

Grant/Contract Number: FA9550-08-1-0079

Principal Investigator: Todd P. Coleman

Dates: 01-MAR-2008 to 30-NOV-2010

My interest in understanding fundamental limits and practical algorithms/codes to instantiate such limits – within the context of communication based upon packet timings – was inspired by my understanding how neurons signal to one another with the timings of their spikes. Remarkably, using this modality to communicate over packet networks has not been explored in nearly as much depth as other communication modalities. Applications of this approach are abundant: including unequal error protection, covert communications, and computer network security.

Although an 1990s paper titled ``Bits through queues'' characterized fundamental limits of communication with packet timings over queuing channels, my curiosity led to a very different – and in some sense simpler - way to arrive at the fundamental limits of such complicated queuing channels (that are nonlinear and have memory). By considering the queuing system as a stochastic dynamical system with feedback coupled to a memoryless channel, I was able to provide a much simpler explanation of the fundamental limits of such a system [1]. Also, this new proof methodology can be extended to many other information-theoretic contexts, which I have also been actively pursuing [2,3]. Lastly, this new proof technique has led to the development of low-complexity, provably good, error-correcting codes that can be used over computer networks [4]. Extensions of this approach and its industrial potential have led to us winning UIUC's Grainger Award in Emerging Technologies in 2008.

[1] T. P. Coleman, "Asymptotic Equipartition Achievability Theorems for Single and Tandem-Server Queuing Timing Channels" submitted to IEEE Transactions on Information Theory, August 2009.

[2] T. P. Coleman, N. Kiyavash, and V. Subramanian, "The Rate-Distortion Function of a Poisson Process with a Queuing Distortion Measure" submitted to IEEE Transactions on Information Theory, May 2008, revised December 2009.

[3] S. K. Gorantla and T. P. Coleman, "The exponential server timing broadcast channel", in preparation for submission, IEEE Transactions on Information Theory, Spring 2010.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **30 MAR 2011**	2. REPORT TYPE	3. DATES COVERED **01-03-2008 to 30-11-2010**

4. TITLE AND SUBTITLE **Low-Complexity Methods For Provably Good Information Transmission And Network Anomaly Detection via Packet Timings In Networks**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Illinois at Urbana-Champaign,Coordinated Science Laboratory,1308 W. Main St,Urbana,IL,61801**	8. PERFORMING ORGANIZATION REPORT NUMBER **; AFRL-OSR-VA-TR-11-035**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-OSR-VA-TR-11-035**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

[4] T. P. Coleman and N. Kiyavash, "Practical Codes for Queuing Channels: An Algebraic, State-Space, Message-Passing Approach", IEEE Information Theory Workshop (ITW), May 5-9, 2008.

My curiosity about developing canonical design principles for brain-computer interfaces has led to an interpretation of the brain-computer interface as two cooperating agents (the brain and computer) with different pieces of information, who are attempting to cooperate with one another via noisy communication channels between them. The user's intent – in short – must be communicated to the interface via recorded brain signals – and given feedback from the interface's dynamics - so that it may help the user accomplish a task. This new interpretation, wed with recent results on implementable, optimal, feedback information theoretic coding schemes (titled "posterior matching schemes"), led us to design brain-machine interface paradigms from decentralized control and feedback information-theoretic perspective (see Section B.3). The symbiotic relationship between my neuroscience and information theory interests came full circle when the brain-machine interface paradigm engaged my curiosity to understand posterior matching feedback information-theoretic schemes from control viewpoints.

The essence of the optimal feedback-based posterior matching scheme – to communicate to the decoder information about the message that is "statistically independent of everything the decoder has seen so far" – led to my hypothesis that its structure is intimately related to stochastic control. Indeed, I was able to demonstrate that it indeed is acting like an optimal stochastic controller, over the space of posteriors, or decoder's beliefs about the message, acting to geometrically "steer" the decoder's belief towards certainty at the message point. It turns out that a simple reward function in this setting pertains to the "reduction in distance towards certainty", and this defines a canonical stochastic control problem with the PM scheme as an optimal solution. Moreover, a Lyapunov function exists, whose structure encodes how fast error probabilities tend to 0 [5].

I further became intrigued about dynamical systems interpretations of the posterior matching encoding scheme – which has an iterated function system structure with a positive Lyapunov exponent. This observation leads to a reverse iterated function type of decoder, with a negative Lyapunov exponent equal to the negative mutual information. These observations provide an alternative, dynamical systems-based proof that the posterior matching encoding scheme, wed with this dual decoder, achieves capacity. Further analysis of this scheme, and understanding dualities between channel coding with feedback and source coding with feedforward, has led to an optimal, low-complexity recursive scheme for source coding with feedforward causal side information [6].

By understanding the nature of optimal communication with feedback, and considering timing channels – which are channels that have state and internal feedback – I have become generally interested in high-level principles that underlie optimal structures of sequential, causal, information theoretic problems with feedback. This has led to a focus on a very general formulation of such problems and the interplay between information theory, control theory, and thermodynamics. This approach has been able to unify a collection of seemingly disparate results in the separate communities. With this, we have a general proof of a large variety of results composing information theory of timing channels, capacity of certain channels with internal state and feedback, rate-distortion theorems, fundamental limits of decentralized control, with practical implications for the design of brain-machine interfaces [7].

[5] T. P. Coleman, "A Stochastic Control Viewpoint on `Posterior Matching-Style' Communication Schemes", IEEE International Symposium on Information Theory (ISIT), Seoul, Korea, June 2009.

[6] H. Ebeid and T. P. Coleman, “The Posterior Matching Scheme for Source Coding with Feedforward: Lyapunov Exponents Analysis”, IEEE International Symposium on Information Theory (ISIT), Austin, TX, June 2010.

[7] S. K. Gorantla and T. P. Coleman, "Information-Theoretic Viewpoints on Causal Coding/Decoding Problems", submitted to IEEE Transactions on Information Theory, January 2011.

Many current viewpoints about how neural processes integrate to elicit complex brain function posit that populations of neurons in the human brain are connected to form functionally specialized assemblies. With the increasing ability to record multiple neural signals at different brain areas simultaneously, one core issue in neuroscience research isunderstanding the mechanistic phenomena and how to analyze these ensemble recordings and infer the structure of these mechanisms. One such approach to attempt to understand this mechanistic phenomena is from a causal network viewpoint.

The directed information – independently developed by information theorists Massey and Rissanen, is an information-theoretic quantity analogous to mutual information that encodes the fundamental limits of communication with feedback. From the viewpoint of a log likelihood ratio, it can be shown to be philosophically consistent with “Granger causality”, in that it measures directionality of causality (e.g., X causing Y) by assessing whether or not past values of X and Y help to predict the future of Y better than only past values of Y [8]. Remarkably, the directed information has been used very sparingly on simultaneously recorded data sets to answer questions of causality – despite the growing importance of these questions in a variety of fields, including economics, computer security, systems biology, and neuroscience. In [9] we demonstrate provably good directed information estimation algorithms. We have further attempted to develop parsimonious approaches approximations to infer causal networks with tree-like structures, and have developed a provably good methodology to infer causal directed dependence trees that can be efficiently implemented [10]. Because the brain is inherently dynamic and time-varying, we are tailoring these algorithms to understand time-varying causal influences that are efficiently solvable and provably good [10].

[8] C. Quinn, T. P. Coleman, N. Kiyavash, and N. G. Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings", Journal of Computational Neuroscience, June 2010.

[9] C. Quinn, T. P. Coleman, N. Kiyavash, "Causal Dependence Tree Approximations of Joint Distributions for Multiple Random Processes", IEEE Transactions on Information Theory, submitted January 2011.

[10] S. Kim, M. Aguilar, and T P. Coleman, "Analyzing Time-Varying Neural Causal Connectivity Using Sequential Probability Assignment", submitted to IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), October 2010.