Network Information Theory

Faculty
Professor Lizhong Zheng

Visiting Scientists and Research Affiliates
[1]Dr. Emmanuel Abbe[1], Professor Chenggao Han[2], Professor Jun Tang[3]

Graduate Students
Baris Nakiboglu, Chung Chan, Shao-lun Huang, Mina Karzand, Hye-Won Chung

Administrative Staff
Michael Lewy

Introduction

The high lever goal of the research activities in our groups is to extend the conventional information theory to develop theoretical tools that can be applied to a broader range of problems, including dynamic communications, networks, real-time communications, cooperative communications, etc. and even beyond the realm of communications, to understand the information exchange in algorithms, social networks, or other physical or biological systems.

The conventional information theory, while had great successes in understanding static point-to-point communication problems, has some key assumptions and the fundamental mindsets following these assumptions, which is now becoming its obstacles to be applied to new fields.

One of these assumptions is the notion of reliable communication. The transmitted information, in the form of information bits, are recovered at the receiver perfectly, with the help of long codes, This greatly simplified the concept of communication, allowing the designs of physical layer signaling to be separated from the source, the network, and the application. The entire theory and practice of digital communication we see today is built on such a vision. In our mind, "bits" appears to be the only natural measurement of information. Unconsciously, we take for granted that the information we care about is in the form of bits, and can be stored or transmitted with no errors.

A related issue in information theory is the use of long codes. In classical treatment, long block codes are used to allow law of large numbers to eliminate the uncertainty of communication, and thus ensure reliable data recovery. Practically, this requires the channel statistic to remain constant for the duration of the coding block, and the delay requirement of the application is long enough for block transmissions. These assumptions are no longer valid in dynamic wireless networks, which is precisely the reason that the classical information theory faces difficulties in these problems. More importantly, the use of long codes and law of large numbers oversimplified the statistical nature of communication channels. This leads to the common approach of using "bits" as the universal measure of information, and using rate and mutual information as the only way to describe communication channels. In the cases that information are not conveyed as coded and reliable bits, our theory becomes powerless.

The main reason that the existing communication theory is based on coded perfect information recovery is that the problem otherwise is much more complicated. Without reliable decoding, it becomes meaningless to measure the observed information contents by number of "bits". Instead, one can only describe the knowledge obtained by looking at how much the posterior distribution is changed. For a nR-bit data message, if we cannot decode it with full confidence,

[1]EFPL/ Ecole Polytechnique Federale de Lausanne, Switzerland
[2]University of Electro-Communications (UEC), Japan
[3]Tsinghua University, China

then the only way we can characterize how much we know about the message is to give the distribution, conditioned on our observations, over the 2^nR possible values. The subject of interests is changed from nR binary digits, to a distribution, described by 2^nR real values. We often refer to the space of these high dimensional distributions as the belief space and its elements the belief vectors[?]. The key technical step, to generalize from coded reliable communication, is to track the movement of the belief vector from one time instance to another, and use that to describe information exchange in the process. Clearly, better mathematical tools are required to study this much more complicated problem.

The tool we need for this is a geometric structure on belief space. With the geometric approach, we can visualize any general information processing as a transform or a sequence of transforms, on the belief space, moving the belief vector towards a desired destination. This approach is analogous to a slow motion of information processing, allowing us to quantify what information is passed or dissipated throughout through the process, and hence understand the process better or come up with better designs.

This general research theme can be applied to many different topics, ranging from network communications to quantum detection problems. The several projects our group is working on cover some aspects of this. Our main contribution lies in the frontier of extending information theory from its conventional applications, to understand more general and more natural information exchanges.

Projects

1. Information Theory with Directions

Sponsor: NSF Grant No. CCF-0830100

Information theoretical quantities, such as entropy and mutual information, are all based on the key concept of Kullback-Leibler divergence, which in a sense measures the distance between a pair of probability distributions. Analysis based on such quantities have only limited success on problems that are static in nature, where relatively few distributions are involved in the problem. In dynamic networked applications, however, the change in the environment, the lack of coordinations, and the interactions among users in the forms of interference or cooperation, often complicate the problem by requiring the understanding the relation between many more probability distributions. Viewing distributions as high dimensional subjects, the analysis based only on a notion of distance is fundamentally insufficient. In this project, we study dynamic network information theory problems using a set of new geometric techniques. By developing notions of inner products, projections, and coordinate systems in the space of distributions, we add a sense of direction in information theoretic analysis.

The geometric approach we employed to study is in a sense a "calculus" on the space of probability distributions, by using properly defined inner products and coordinate systems. Different from the familiar Euclidean geometry, we will focus on an embedding that naturally gives rise to information theoretic quantities such as the K-L divergence. Our goal is to explore the use of such approach to offer new insights to high dimensional communication problems, often found in multi-terminal or dynamic scenarios. In this proposal, we will give a brief introduction to this approach, and show some preliminary results, where the geometric method is applied to some well known, in some cases open, problems in information theory. We demonstrated that some conventional solutions can be better understood and often extended, once we understand the underlying geometric structure of the problem. One key advantage of this approach is that by revealing the geometric structure, the analysis becomes more systematic, or sometimes even mechanical. It takes away the requirement of creative proving techniques such as cleverly chosen auxiliary random variables or genie, but rather rely on conceptually straightforward procedures to get definitive answers.

The general approach we propose to take is closely related to the study of information geometry. In a nutshell, information geometry is the method of using differential geometry to study probabilistic models, where the embedding is carefully chosen such that the K-L divergences is naturally related to the length of curves on a manifold. Connections can also be made between information theoretic quantities, such as mutual information and divergence, and local properties of a probabilistic model, such as Fisher information. This naturally connects information theory to estimation theory, and offers more insights of information processing in general. Consequently, there has been some active research using information geometry in statistical inference, robust hypothesis testing, as well as cognitive science.

The direct applications of information geometry on communication problems are however rather limited. The main reason is that most communication problems involve coding over long sequences. The precise characterization of the distribution of such long sequences, the geometric structure of the underlying manifolds, and the problem solutions are thus often too complicated. A major contribution of Shannon's information theory is to find the right approximations to such complicated problems, and reduce them to the level that the solutions are simple enough to be implemented in practice. This step of simplification and approximation is missing from the current research of using information geometry to study communication systems.

Our study is based on a key simplification of information geometry, where we focus on the cases that the distributions of interests lie in vicinity from each others. Mathematically, this means instead of studying the global behavior on a manifold, we first focus on the local behavior, where a manifold, no matter how complicated to start with, can always be approximated by its tangent plane. Consequently, the K-L divergence between these distributions can be approximated by a weighted Euclidean norm on a linear vector space, and many complicated divergence optimization problems can be reduced to linear algebra, allowing analytical solutions and geometric intuition. While this seems over simplifying, it turns out that the key issue we discussed above, a notion of direction in the space of probability distributions, is well captured by our approach. To be specific, we have already obtained exciting preliminary results and plan to continue working on the following main areas.

1. **Muti-user Information Theory:**

   In multi-user information theory, it is often di_cult to reduce the coding problem, which involves many symbols, into a single letter problem. Such single-letterization, if possible, reveals a nice memoryless feature of the solution. The known approaches to prove the optimality of single letter solutions requires careful choices of auxiliary random variables, which are often non-intuitive and limited to specifc problems. By reducing some of these problems with local approximations, we found that the key to solutions lies in a specific coordinate system in the space of distributions, which is related to the singular vectors of a particular matrix, which we call divergence transition matrix (DTM). With the help of this coordinate system, the multi-letter divergence optimization problems can often be reduced to standard linear algebra problems, allowing analytical solutions. The notion of divergence transition is in fact very general, and can be applied to many other network information theory problems. This geometric approach is particularly useful in generalizing the recent results on Gaussian broadcasting channels and interference channels. As many of these results are based on entropy power inequality (EPI), which is based on special properties of Gaussian distributions, we are in need of some systematic way to quantify the non-Gaussianness in an arbitrary distribution, in order to generalize these results. To do that, we use a special coordinate system in the space of distributions, defined in a neighborhood of a Gaussian distribution. It turns out that one natural coordinate system one could use is defined by the famous Hermite polynomials, whose elegant properties allows our approach to be very effective on some hard problems, such as the interference channel.

2. Compound Channel and Universal Receiver:

Even for single user channels, the geometric analysis can be valuable. For example, in dynamic wireless channels, decoding with limited channel side information (CSI) becomes important. The well-known "mismatched" decoding problem considers a receiver that is misinformed and decodes using an "incorrect" decoding metric. It turns out that the capacity loss by using the wrong metric can be easily visualized using the geometric analysis, in terms of a projection. We generalized this result to study compound channel, with a constraint that the decoder is "linear, i.e., uses an additive decoding metric. We showed that the compound channel capacity can be achieved by linear decoders, provided that the set of possible channels satis_es a geometric condition of being "one-sided" [5]. Based on this success, we propose to explore the designs of an "plug-and-play" receiver, which, at the cost of a small fraction of the capacity, decode the transmitted messages with a linear complexity, without requiring any knowledge of the channel. Moreover, throughout our preliminary study, we have developed a process, which we call "lifting", to generalize our results from local variational analysis to the general cases. A comprehensive understanding of lifting is of interests in order to generalize several of our results.

3. Error Exponent and Structured Codes:

Reliability function and error exponent characterize the tradeffo_ between the data rate and the error probability of a channel code. In conventional coding problems over static channels, error exponent is often regarded as a secondary issue comparing to the data rate, since in those systems, one can often employ a long code or rely on other error control mechanisms such as ARQ to improve the error performance. Starting from wireless systems, the notion of dynamic environment becomes a key challenge in the new systems, and the role of error exponents starts to become important. Diversity order, for example, as a widely used concept for wireless systems, is indeed one special form of error exponent. The commonly used method to compute error exponents, however, is sometimes hard to follow. Although the procedure based on Chernoff bounds allow the results to be evaluated in a fairly simple way, the intuition behind the steps is often obscure to even be best trained researchers. Furthermore, generalizing this method to multi-user coding or structured coding problems appears to be quite difficult. Again, our geometric analysis is a natural tool for this problem, since the key issue of the error exponent problems is the variation of probability distributions. Different from multi- user problems where the variations of the distributions of the source or the channel are due to the changes in side information or interference signals, here, the variation comes from the rare events governed by large deviation principles, which moves the empirical distribution of the noise away from its ensemble. Our geometric approach helps to visualize such variations. We have used our approach to re-prove most of the known results on error exponents, producing much more intuitive derivations. We plan to extend this line of work to study unequal error protection codes, where multiple data messages are joint encoded and transmitted over a single channel, but with different levels of error protections. Such heterogeneous information processing techniques is an important conceptual step towards understanding network information theory.

2. Capacity, Cooperation and Optimization in Mobile Adhoc Networks

Sponsor: DARPA ITMANET Program Contract No. 18870740

The DARPA program on Information Theory for Mobile Ad Hoc Networks (ITMANET) is one of the biggest effort from DoD in the recent years to advance the understanding of information flows in dynamic networks. Our effort is a part of the FLoWs (Fundamental Limits of Wireless Systems) team, which consists PI's from MIT, Stanford, UIUC, and Caltech. Under the support of this project, we focus particularly on the design paradigm of "layerless" networking. That is, to view dynamic communication as information transfer over the network viewed as a general probabilistic map, and deviate from the traditional layered design philosophy, to create new transmission schemes to improve the network performance.

Under the support of this program, we have conducted research in two major areas.

First, we worked on funding the receiver designs under channel uncertainty, which is a main challenge for dynamic wireless networks as frequent updates of channel state information is often difficulty. Such problems are often formulated as compound channel in the literature, and the general solution is to use maximum mutual information (MMI) receivers. That is, to compute the empirical mutual information between each codeword and the received word and pick the highest one. The practical difficulty of implementing MMI receiver is that to compute the "score" of each codeword, it requires exponentially large storage and computations. We focus on a particular type of receiver, namely "linear universal receivers". The advantage of our approach is that the linear receiver are not only easier to implement, but also reveals a nice geometric structure over the space of probability distributions, where each linear receiver above can be viewed as a hyper-plane in the space of empirical joint distributions between the input and the output. With such in-sights, we can better understand how to separate the correct and incorrect codewords based on their joint empirical distribution with the received words, which allows us to understand the behavior of generalized linear receivers such as the MMI receiver.

The main result of our work is the construction of generalized linear decoders that achieve compound channel capacity on most compound sets. This construction requires solving some rather complicated optimization problems involving the Kullback-Leibler (KL) divergence (like almost every other information theoretical problem). To obtain insights to this problem, we introduced a special tool: local geometric analysis. In a nutshell, we focus on the special cases where the two distributions in the KL divergence are close to each other, which can be thought in this context as approximating the given compound channels by very noisy channels. In this local setting, information theoretical quantities can be naturally understood as quantities in an inner product space, where conditional distributions and decoding metrics correspond to vectors; divergence and mutual information correspond to squared norms; ! and the data rate with mismatched linear decoders can be understood with projections. The relation between these quantities can thus be understood intuitively. While the results from such local approximations only apply to the special very noisy cases, we show that some of these results can be \lifted" to the naturally corresponding statements about general cases. Using this approach, we derive the following main results of the paper.

- First we derive a new condition on S to be one-sided", under which a linear decoder, which decodes using the log likelihood of the worst channel over the compound set, achieves capacity. This condition is more general than the previously known one, which requires S to be convex;

- We show that if the compound set S can be written as a finite union of one sided sets, then a generalized linear decoder using the log a posteriori distribution of the worst channels of each one-sided subset achieves the compound capacity; in contrast, GLRT using these worst channels is not a universal decoder.

Besides the specific results on the compound channels, we also like to emphasize the use of the local geometric analysis. As most of multi-terminal information theory problems involve optimizations of K-L divergences, often between distributions with high dimensionality, we believe the localization method used in this paper can be a generic tool to simplify these problems. Focusing on certain special cases, this method is obviously useful in providing counter examples to disprove conjectures. However, we also hope to convince the readers that the insights provided by the geometric analysis can be also valuable in solving the general problem.

For example, our de_nition of one-sided sets and the use of log a posteriori distributions as decoding metrics can be seen as \naturally" suggested by the local analysis.

Our second focus area under the ITMANET project is the study of feedback channels. In feedback channels, we assume a perfect real-time causal receiver feedback is available at the transmitter, helping it to construct the next symbol. Thus, at each time, instead of mapping the data message into the transmitted symbol at the time, the transmitter maps the message to the symbol based on the current history of the received signals.

The interesting part of this problem is that the coding problem is time varying, based on the accumulation of the receiver knowledge. On the other hand, the coding considered here is only for a single symbol, instead of over a long block of symbols like the conventional channels.

Our approach to study this problem is to develop a new way to describe communication, by monitoring the

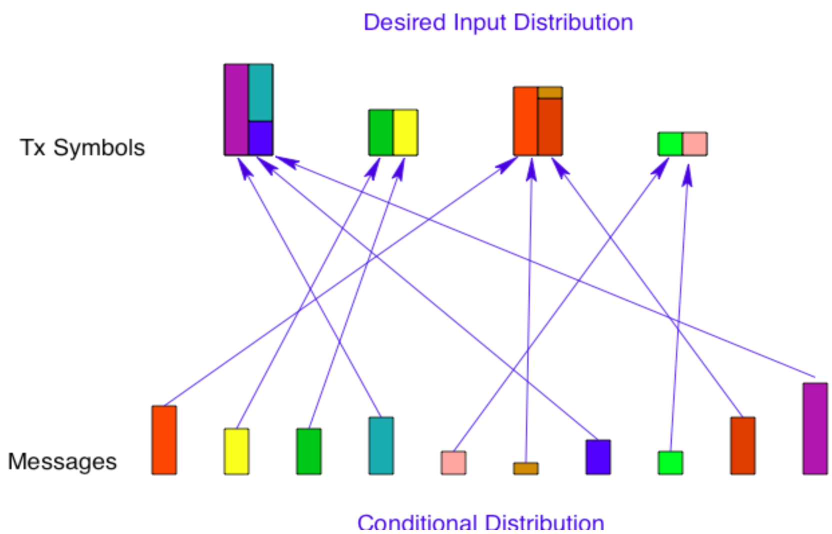

Figure 1: Optimal input for the feedback channel as matching the posterior distribution to the optimal input distribution of the channel

changes of the posterior distribution of the messages, conditioned on the receiver knowledge. This geometric approach is based on defining new metrics on the space of probability distributions, and optimizes the progress of the belief vector at each time. The specific metric we chose is directly related to an exponential bound of the probability of error at the end of the block. The optimal signaling for this problem turns out to be surprisingly simple. As illustrated in the following figure, one can think of the coding procedure as first exponentially tilt the posterior distribution over the messages, and then shape it according to the optimal input distribution to the channel.

3. Dynamic Wireless Networks Based on Open Physical Layer

Sponsor: AFOSR Young Investigator Program, Contract No. FA9550-08-1-0143

The conventional wisdom, that any physical media should be encapsulated into bit pipes shipping bits with perfect reliability, is now becoming the main conceptual obstacle of developing new generations of wireless networks, especially for highly dynamic military applications. We propose to study the designs of an "open" physical medium, and its implications to the higher layer

network architectures. An open physical medium allows several data messages to be encoded and transmitted over the same channel, and allows the higher layer applications to determine explicitly the reliability levels of the individual messages. A new notion of "capacity" of a channel is described by the optimal tradeoff between the rates and reliability of messages it can accommodate. We demonstrate that with this new interface to the physical media, a network can operate much more efficiently than the conventional ones based on bit rates, as more efficient coordination and cooperation among network nodes are enabled by the more
exible controls over the physical media.

Utilizing the "open" physical media, we envision a heterogeneous network, where data with different meanings and reliability are allowed to share the same media. We propose to study several aspects of such networks: 1) the use of layered source codes to guarantee the delivery of critical messages in dynamic/hostile environment; 2) embedding feedbacks, ARQ, and network control messages in data transmission to improve effciency, and to unity network controls and data; 3) contents driven networking that prioritize the scheduling and routing of data according to the precision and urgency of the contents.

The key concept in this work is that the reliability of a physical channel is NOT characterized by a single number such as the error exponent. As the total capability of error protection being limited in dynamic networks, it becomes valuable to tradeoff the resources and provide better protections to the more important messages. We study a layered coding structure for general physical channels, where several messages, including data from different sources as well as network control signals, are encoded and transmitted together. Superior to the current approach of allocating time, frequency, and signal power to different messages, layered codes allow several sub-messages are encoded together, thus the total resource in the channel can be optimally shifted from one layer to another seamlessly. Such new designs of coding structures makes the physical medium truly "open" to the higher layer network applications, allowing fine-tuned globally optimal controls of the transmissions of different kinds of data.

The concept of an "open" physical medium has profound impacts on the designs of the higher layer network architecture. The question how should the network operate with the new controls over the physical media involves nearly every aspect of networking. We propose to study the following topics as a first step

1. Unified data and network controls: as dynamic networks require frequent re-configuration, the overhead costs of controlling the data streams are dramatically increased in wireless networks. Layered coding structure allows the control signals to be embedded in data transmissions, satisfying the latency and reliability requirements of control messages while maximizing the data throughput in the same time. This way, the control and data messages are treated in a unified way, allowing simplified and efficient designs of networks. In particular, ARQ and feedbacks of the channel state information is frequently used in wireless networks to improve reliability and facilitate opportunistic communications. Efficiently embedding such signals in data streams can have impacts on practical designs in a short term.

2. Layered source channel code: coupling layered source code with layered channel code offers valuable features, especially in dynamic or hostile environment, where it is often desirable to provide extra protections to guarantee the transmission of crucial messages, while allowing further refinement of details when the channel is good. The applications of structured coding techniques are thus not limited to the physical layer, but also reach to source coding in audio, video, and sensing applications. The method of information geometry can also be used for a systematic extension to the problems of compress sensing and signal inference.

3 Contents driven heterogenous networking: the proposed research serves as a first step to pursue the ultimate vision of contents driven networking: data are divided into pieces, transmitted over the network separately; in blocks or in streams; with different reliability; fused or redivided at different locations of the network; prioritized in scheduling and routing according to the precision and urgency of the contents. Such new network architectures based on the open physical media allows the network resources to be allocated globally, giving wireless networks a new level of flexibility and efficiency to best serve the applications it supports.

Some of our key results in this area include

- Designs of Layered Coding with Optimal Reliability:

The key technique to facilitate the concept of open physical layer is to allow multiple layers of data to be encoded together. Different from the classical multiplexing problem, the theoretical challenge here is to provide these layers with different levels of reliability, with an optimal tradeoff between the sizes and the reliability of different layers. With system control signals, high priority data mixed with normal payload, allowing higher layer to seamlessly control the protection on different types of data is the central promise of our work. Unequal error protection (UEP) is a natural generalization to the conventional content-agnostic information processing, where the information bits are not treated equally, but with different priority levels. The simplest method of unequal error protection is to allocate different channels for different types of data. For example, wireless systems allocate a separate \control channel", often with short codes and low spectral efficiency, to transmit control signals with high reliability. The well know Gray code, assigning similar bit strings to close by constellation points, can be viewed as UEP: even if there is some error in identifying the transmitted symbol, there is a good chance that some of the bits are correctly received. Most of the existing approaches focus on designing good codes for particular channel models. The optimality of these designs was established in only limited cases. In this paper, we aim at providing a general theory of the fundamental performance limits in UEP.

In the existing formulations of unequal error protection codes, the information bits are divided into sub-groups, and the decoding errors in different sub-groups of bits are viewed as different types of errors. For example, one might want to provide a better protection to one subset of bits by ensuring that errors in these bits are less probable than the other bits. We refer to such formulation as \bit-wise UEP". In a more general setup, one might want to have better protections to a subset of messages, by ensuring a lower probability of error when one of these special messages is transmitted. For example, one might consider embedding a special message in a normal k-bit code, where the extra message has a special meaning, say, a _re alarm, and needs to have a smaller error probability when it is transmitted. Note that the error event conditioned on the special message is not associated to error in any particular bit. Instead, it corresponds to a particular bit-sequence (corresponding to the special message) being decoded as some other bit-sequence. We refer to such problems as \message-wise UEP".

Depending on the applications, it is sometimes meaningful to treat the missed-detection and false- alarm events differently. Here we borrow from the hypothesis testing terminology. We say that missed- detection of a particular message occurred when this message was transmitted but the receiver missed it by decoding to some other message. Similarly, a false alarm of a message is the event when the receiver decoded to this message although it was not transmitted.

Our main contribution in this area is a set of results, identifying the performance limits and optimal coding strategies, for several new formulations of UEP. In general, a UEP problem can be stated as a combination of a multiple requirements of protection, such as simultaneously protecting some bits in one way, other bits in some other way, some messages in some way, and some messages in another. Therefore, the general UEP problem could be too elaborate to describe. We focus on a few special cases, most with immediate practical applications, and try to illustrate the main insights and mathematical techniques involved in these problems.

- **Unified Transmission of Data and Control Signals in Networks:**

The overhead required to control the operation of a wireless network is a key design challenge. In addition to the common control signals such as congestion/routing controls, connection/handoff signals, and ARQ, frequent and timely exchanging channel and network state information is required for opportunistic communication; coordi- nations among wireless nodes for cooperative transmissions may indeed involve local data exchanging. The volume of control signals in dynamic wireless networks can be much larger than the conventional concept, as the fading environment and the geographic positions of the wireless nodes may changes rapidly. Furthermore, all these different kind of "control" signals have different priorities based on the different consequences of success/failure, latency, or quality of their transmissions. Clearly, giving each class a separated bandwidth makes the network design much more complicated, less effcient, and difficult to adapt to environment or task changes.

The concept of open physical media offers a possible remedy of this problem. Here, the different levels of quality of service (QoS) of individual messages are implemented at the stage of channel coding, and is much more efficient, and easy to change online. From the network point of view, there is no fundamental difference between most of the control signals and the data traffic: there are all just data, labeled with possibly different latency and reliability requirements, to be delivered over the same physical medium according to these requirements. Thus data and control are unified by the flexibility of QoS provision of layered codes. The question remains as what levels of QoS for the controls would be preferred by the network? This leads to a new breed of network optimization problems, based on the limitation of the physical channel, in terms of the optimal tradeoff of QoS it can provide.

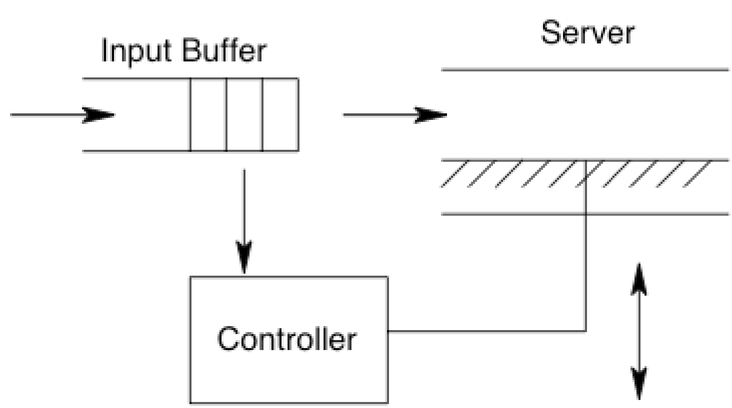

Figure 2: A Network Control Example with Imperfect "Lazy" Controller

As shown in Figure 2, we consider a conventional queuing problem with a variable rate server. It is well known that as long as the server rate is larger than the data arrival rate, the queuing system is stable. Furthermore, if a controller observes the buffer size and controls the server rate, which serves higher rate when the buffer is long, the over flow probability can be reduced while keeping the average server rate fixed. An ideal controller needs to perfectly observe the buffer size, and frequently adjust the valve around the balancing point. This requires a higher rate of control signals. Eventually, the cost of control is taken from the overall resource and reduces the server rate. We consider a "lazy" controller, which reduces the rate of control signals by increasing the server rate when the buffer is really long, and vice versa. Moreover, instead of assuming the perfect observation of the buffer size, and perfect transmission of the control signal, we assume both of these control signals to have limited precision and reliability. Clearly, with a fixed overall resource, the new approach improves the performance, especially when the data packets are small and frequent control is required. Such an algorithm can hardly be implemented in a conventional system, with a static separation between the data and the control channels. With the help of the open physical media, the global optimal solution becomes reasonable.

- **Contents Driven Heterogeneous Networks:**
The key point of network coding is to divide the data into small pieces, transmitted over different links, merged with other data, and relayed by various nodes in the network, until enough information reaches the destination and get decoded. In our vision, data should also flow through the network in a similar fashion. However, we take into consideration that data bits are not all perfectly reliable, and thus differentiate data into heterogenous classes. We particularly emphasize the capacity of dynamic processing, and incorporate into the overall network resource allocation problem the costs of control signals needed to operate the dynamic network. We study the inference problem to understand how to combine multiple observations of the data source, where each observation may have different levels of precision and reliability. This problem is relevant to the relay operation of a node that receives several data pieces about the same source from different network paths, where each piece may go through different lossy processing along its path. The goal isto efficiently summarize these correlated data, and forward these data messages with an appropriate labeling of confidence level. This problem is also closely related to the inference problems faced insensor networks.

References
[1] Krishna P. Jagnnathan, Eytan Modiano, and Lizhong Zheng On the Trade-off Between Control Rate and Congestion in Single Server Systems Infocom 09.

[2] Shashibhushan Borade, Baris Nakiboglu, and Lizhong Zheng, Unequal Error Protection: Some fundamental limits and optimal strategies, to appear, IEEE Transactions on Information Theory, 2010

[3] Emmanuel Abbe and Lizhong Zheng, Hermite Polynomials and Gaussian Noise, International Symposium on Information Theory (ISIT), 2009

[4] Baris Nakiboglu and Lizhong Zheng, Upper Bounds to Error Probability with Feedback, International Symposium on Information Theory (ISIT), 2009

[5] Emmanuel Abbe and Lizhong Zheng, Linear Universal Decoding for Compound Channels: a Local to Global Geometric Approach IEEE Transactions on Information Theory