

# Statistical Mechanics-Inspired Modeling of Heterogeneous Packet Transmission in Communication Networks

Soumik Sarkar, *Member, IEEE*, Kushal Mukherjee, *Member, IEEE*, Asok Ray, *Fellow, IEEE*, Abhishek Srivastav, and Thomas A. Wettergren, *Senior Member, IEEE*

**Abstract**—This paper presents the qualitative nature of communication network operations as abstraction of typical thermodynamic parameters (e.g., order parameter, temperature, and pressure). Specifically, statistical mechanics-inspired models of critical phenomena (e.g., phase transitions and size scaling) for heterogeneous packet transmission are developed in terms of multiple intensive parameters, namely, the external packet load on the network system and the packet transmission probabilities of heterogeneous packet types. Network phase diagrams are constructed based on these traffic parameters, and decision and control strategies are formulated for heterogeneous packet transmission in the network system. In this context, decision functions and control objectives are derived in closed forms, and the pertinent results of test and validation on a simulated network system are presented.

**Index Terms**—Communication network, heterogeneous packet transmission, phase transition, statistical mechanics.

## I. INTRODUCTION

CONCEPTS of statistical mechanics have been extensively used to model thermodynamic characteristics of physical phenomena in terms of their relationships between micro- and macrobehaviors [1], [2]. From this perspective, tools of statistical mechanics are appropriate for modeling the behavior of multiagent systems (e.g., communication networks) because their global behavior emerges from the local dynamics of the participating agents. This phenomenon has inspired many researchers to investigate complex networks from the mathematical perspectives of thermodynamics and statistical mechanics [3]. Specifically, Hui and Karasan [4] have addressed the thermodynamic formalism of communication networks. More

recently, Albert and Barabasi [5] have reported a comprehensive review of the recent literature in this field.

The behavior and topological organization of communication networks [6] and sensor networks [7] have similar characteristics as many natural and human-engineered systems, such as those found in the disciplines of sociology [8], biology [9], and finance [10]. Phase transition is a characteristic phenomenon of complex systems, consisting of interacting and interdependent dynamics, where a nonsmooth change in the output behavior may take place with a relatively small variation of the system parameter(s). In this context, tools of statistical mechanics are useful for characterization of critical phenomena corresponding to the dependence of the global behavior (e.g., connectivity, average rate of change of queue length, and average packet drop rate) of communication networks on their local parameters (e.g., communication radius, packet load, and transmission probability) [11]–[14].

A key task in the analysis of phase transitions is the identification of the system behavior in the vicinity of a critical point, where a global parameter quantifies the presence of order in the underlying system. Usually, this *order parameter* [1], [2] is zero in the disordered phase and may have increased to a significant nonzero value in the ordered phase. In essence, a phase transition is realized as a discontinuity in the zeroth or a higher derivative of the order parameter for a change from zero to a nonzero value if an intensive parameter (e.g., temperature  $T$ ) of the system is perturbed from its critical value. In general, the nature of phase transition is classified into two types: 1) first-order phase transition, where the order parameter changes discontinuously with the intensive parameter at the critical point and 2) continuous (also called second or higher order) phase transition, where the order parameter varies continuously with an intensive parameter during the phase transition but its first or a higher derivative with respect to the intensive parameter at the critical value is discontinuous [15]. For example, ferromagnetism is a well-known case of continuous phase transition, where the order parameter is magnetization  $M(T)$  with a nonzero value leading to spontaneous magnetization at temperatures below the critical temperature, called the Curie point [1], [2].

The congestion phenomenon in communication networks is an example of a phase transition. Typically, the concept of bottleneck buffers (routers) is used to detect and control (i.e., mitigate or prevent) congestion in communication networks.

Manuscript received June 3, 2011; revised October 21, 2011; accepted January 24, 2012. Date of publication March 9, 2012; date of current version July 13, 2012. This work was supported in part by the Office of Naval Research under Grant N00014-09-1-0688 and in part by the Army Research Laboratory and the Army Research Office under Grant W911NF-07-1-0376. This paper was recommended by Associate Editor S. X. Yang.

S. Sarkar and A. Srivastav were with the Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA 16802 USA, and are currently with the United Technologies Research Center, East Hartford, CT 06108 USA (e-mail: sarkars@utrc.utc.com; srivasa1@utrc.utc.com).

K. Mukherjee and A. Ray are with the Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA 16802 USA (e-mail: kum162@psu.edu; axr2@psu.edu).

T. A. Wettergren is with the Naval Undersea Warfare Center, Newport, RI 02841 USA (e-mail: t.a.wettergren@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSMCB.2012.2186611

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>AUG 2012</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2012 to 00-00-2012</b>	
4. TITLE AND SUBTITLE <b>Statistical Mechanics-Inspired Modeling of Heterogeneous Packet Transmission in Communication Networks</b>				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) <b>Pennsylvania State University, Department of Mechanical Engineering, University Park, PA, 16802</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT <b>This paper presents the qualitative nature of communication network operations as abstraction of typical thermodynamic parameters (e.g., order parameter, temperature, and pressure). Specifically, statistical mechanics-inspired models of critical phenomena (e.g., phase transitions and size scaling) for heterogeneous packet transmission are developed in terms of multiple intensive parameters, namely, the external packet load on the network system and the packet transmission probabilities of heterogeneous packet types. Network phase diagrams are constructed based on these traffic parameters, and decision and control strategies are formulated for heterogeneous packet transmission in the network system. In this context, decision functions and control objectives are derived in closed forms, and the pertinent results of test and validation on a simulated network system are presented.</b>					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

The network structure could be reduced to a multisource and single destination (or vice versa) through the usage of bottleneck buffers, and thus, mean-field models have been developed for such cases [16]. However, analysis of certain problems (e.g., distributed decision making, perimeter surveillance by static sensor networks, and statistical estimation of parameter distribution over the entire space) is apparently very difficult and often intractable. Moreover, in many applications, on-line adaptation of the packet routing sequence is required to maintain acceptable performance in the presence of uncertain exogenous perturbations and channel fluctuations. For example, sensor networks are often employed in environments with limited communication capability. In such applications, the information travels between sensor nodes through a combination of relay nodes and other sensor nodes. For communication-constrained sensor networks, the information that is transmitted is usually event driven (e.g., detection of intruders) and is thus intermittent in nature. For this reason, it is important to have a controllable routing procedure that takes advantage of the available network structure to avoid bottlenecks and other congestion issues.

From the aforementioned perspective, this paper presents a statistical mechanics-inspired concept, namely, equilibrium thermodynamics formalism, that is more suitable in event-driven situations compared to the traditional time-driven modeling approaches. Phenomena of phase transitions and size scaling are characterized in large-scale communication networks. The underlying theories have been tested and validated on a simulation test bed that consists of a 2-D square-grid network model. The objective here is to obtain an unambiguous understanding of the critical behavior on a relatively simple network with tractable dynamic properties. The simulation experiments would help establishing a knowledge base for the synthesis of decision and control laws in different classes (e.g., wired and wireless) of real-life communication networks. Network phase diagrams are constructed based on the network parameters to formulate control strategies for heterogeneous packet transmission. In this regard, closed-form solutions for a simulated network system are presented.

This paper is organized in five sections, including the present one. Section II describes the model of the network along with the architecture of the simulation test bed. Then, the qualitative nature of phase transition in the underlying system is characterized, and the effects of network size are analyzed. In these analyses, phase transitions are investigated based on a single intensive parameter, namely, the external packet load of the communication network system. Section III investigates the effects of two intensive parameters on the network performance. Phase diagrams are constructed for these cases by defining network analogs of thermodynamic quantities (e.g., order parameter, temperature, and pressure). Based on the equilibrium phase diagrams, Section IV presents a novel thermodynamic approach for controlling heterogeneous packet transmission. It also presents the basic philosophy of the control approach along with a representative numerical experiment on the network architecture of the current study. The paper is summarized and concluded in Section V along with recommendations for future research.

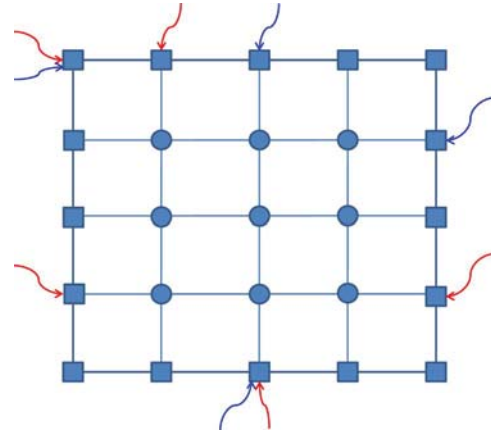


Fig. 1. Network structure in the simulation model.

## II. NETWORK SIMULATION TEST BED

The network test bed simulates a 2-D square grid [12] as shown in Fig. 1, where the nodes (routers) are placed at the grid points. For a square-grid network with  $N \times N$  nodes, there are  $(4N - 4)$  boundary nodes (shown as squares in Fig. 1) and  $(N^2 - 4N + 4)$  internal nodes (shown as circles in Fig. 1). Only boundary nodes are assumed to be the sources and/or the sinks for packet generation and destruction; internal nodes can only transmit the packets. Each node receives packets in a queue (with a finite buffer length) from its neighboring nodes, and packets are terminated after reaching their respective destinations. In each time unit, packets are created in the boundary nodes with a Poisson arrival rate  $\lambda$ . For simplicity, all packets are considered to be of unit length. The destination of a packet is chosen randomly from the set of boundary nodes, including its source node. Each node (boundary or internal) transmits one packet from the head of its queue to a deterministically chosen neighboring node at each time unit. The node chosen to forward a data packet is selected so that the packet travels via the shortest path to its destination. When there are more than one candidate nodes for the shortest path, the node with a smaller queue length is chosen to reduce the probability of early congestion in the network. A node drops the oldest packet from its fully occupied queue to accommodate a new packet.

### A. Network Congestion Modeled as a Phase Transition

Congestion is a phenomenon of significant importance in communication networks. Before a congestion occurs (i.e., the network is capable of handling the average number of arriving packets), the average packet drop rate remains negligibly small. However, as the packet influx rate crosses a critical threshold, the drop rate increases to a nonnegligible value. Thus, at a steady state, the nonzero average packet drop rate can serve as an index of the degree of congestion. The phenomenon of congestion can be viewed as a continuous phase transition from a steady state to an unsteady state of network communications, which is characterized based on the concepts of equilibrium thermodynamics in this paper. Previous papers [11], [17] reported similar characterization based on the average rate of change of queue length; however, they required the assumption

of infinite queue length. This assumption is removed in this paper to address the congestion problem in a more realistic scenario.

As discussed in Section I, a global order parameter of the system under consideration needs to be identified for investigating the occurrence of phase transitions in communication networks. The packet drop rate is a feasible candidate to serve as the order parameter in this analysis. Furthermore, global intensive parameters that trigger the network phase transition need to be identified, and in this problem, packet influx rate  $\lambda$  is one such parameter. Since only  $4(N-1)$  boundary nodes generate the packets in this simulated network and all  $N^2$  nodes receive these packets, a *surface correction* is needed to accurately define the intensive parameter, namely, effective load per node  $\lambda_{eff}$ , which is defined as follows to avoid the introduction of the surface effects:

$$\lambda_{eff} \triangleq \left( \frac{4(N-1)}{N^2} \right) \lambda. \quad (1)$$

A few scaling adjustments are made to define the order parameter. Let  $d(t)$  be the total number of packets dropped in the network at time  $t$ . Thus, the instantaneous packet drop rate (per node) at time  $t$  is  $d(t)/N^2$  and its time-averaged value is denoted as  $D \triangleq \overline{d(t)}/N^2$ . The order parameter  $M$  is defined based on the normalization of  $D$  with respect to  $\lambda_{eff}$

$$M = \frac{D}{\lambda_{eff}}. \quad (2)$$

Note that the order parameter  $M$  is the fraction of incoming packets that are dropped from the network nodes per unit time. Therefore, in the absence of congestion,  $M$  could be assumed to be negligibly small, i.e., there is no packet drop. As the network becomes congested, the worst case scenario is the dropping of all incoming packets, i.e.,  $M$  approaches 1.

Following the aforementioned procedure,  $M$  is computed for given values of  $\lambda_{eff}$ . To eliminate the effects of transients, the expected value of  $D$  is calculated using only steady-state time series data. The plot of  $M$  versus  $\lambda_{eff}$  for the  $10 \times 10$  simulated network is shown in Fig. 2. It is seen that there is a critical value  $\lambda_{eff}^c \approx 0.11$  of effective load per node such that, for  $\lambda_{eff} < \lambda_{eff}^c$ , the order parameter  $M$  is almost negligible; in contrast, for  $\lambda_{eff} > \lambda_{eff}^c$ ,  $M$  takes on nonzero values. This change of network behavior across the critical value  $\lambda_{eff}^c$  is identified as a continuous *phase transition*, where the network moves from an *uncongested phase* (U) of negligibly small  $M$  to a *congested phase* (C) of finite positive  $M$ .

### B. Construction of Size Scaling Laws

It is very important to understand the laws of size scaling for human-engineered multiagent systems (e.g., communication network systems) due to their inherent finite and smaller sizes, as compared to their natural counterpart. In general, for finite-sized systems, the critical value of the intensive parameter  $\lambda_{eff}^c(N)$  is a function of size  $N$ , and as  $N$  goes to infinity,  $\lambda_{eff}^c(N)$  converges to the  $\lambda_{eff}^c(\infty)$  of the corresponding

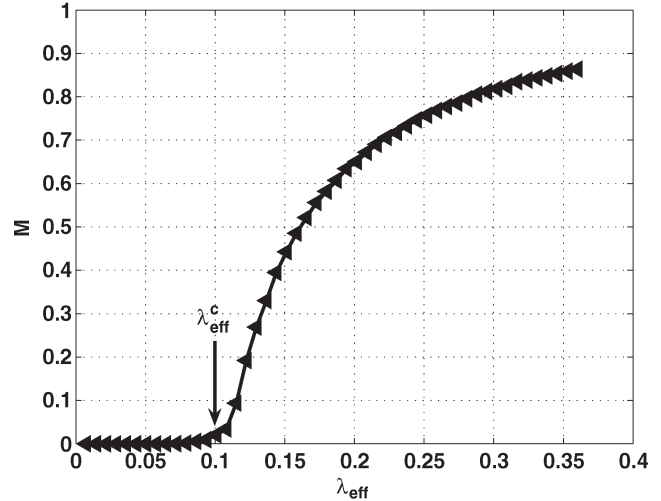


Fig. 2. Continuous phase transition in a square-grid communication network. The normalized packet drop rate ( $M$ ) is plotted against the effective load per node ( $\lambda_{eff}$ ).

infinite-sized system in the thermodynamic limit. In the statistical mechanics literature, the space correlation length  $\xi(\Delta)$  behaves as a function of  $|\Delta|$  in a power law. Here,  $\Delta$  can be considered as  $\lambda_{eff}^c(N) - \lambda_{eff}^c(\infty)$ . Also, the following size dependence form of a test function  $\xi(\Delta)$  is assumed [18], which conforms to the characteristics of  $\xi$ , and:

$$\xi(\lambda_{eff}^c(N) - \lambda_{eff}^c(\infty)) = pN \quad (3)$$

where  $p$  is the proportionality constant. Thus, the following law is postulated based on the structure of the 2-D Ising model (also known as the Onsager model) [1], [2]:

$$\lambda_{eff}^c(N) = \lambda_{eff}^c(\infty) + qN^{-\frac{1}{\nu}} \quad (4)$$

where  $q$  and  $\nu$  are constant model parameters that are called the coefficient and index, respectively. However, in the present problem,  $\lambda_{eff} \sim (1/N)$ . Thus, it is impossible to provide a finite effective load per node to an infinite-sized network, which implies that there cannot be any finite load phase transition in an infinite-sized network, i.e.,  $\lambda_{eff}^c(\infty) = 0$ . Hence, for this case

$$\lambda_{eff}^c(N) = qN^{-\frac{1}{\nu}} \quad (5)$$

where the model parameters (i.e., coefficient  $q$  and exponent  $\nu$ ) in (5) are identified by fitting with the simulation data generated on the network test bed. Fig. 3 shows the closeness of the fitted data with the model, where the identified parameters are as follows:  $q = 10^{0.034} \approx 1.08$  and  $\nu = (1/1.1) \approx 0.9$ .

Fig. 4 shows the possibility of generating a size-independent global model for phase transition in square-grid communication networks. Such a model can be constructed by using a reduced (i.e., normalized) effective load  $\lambda_{red} = \lambda_{eff}/\lambda_{eff}^c$  instead of  $\lambda_{eff}$ . This approach draws inspiration from classical thermodynamics, where compressibility curves for different pure substances are unified by using reduced pressure or reduced temperature instead of directly using pressure or temperature as the independent variables.



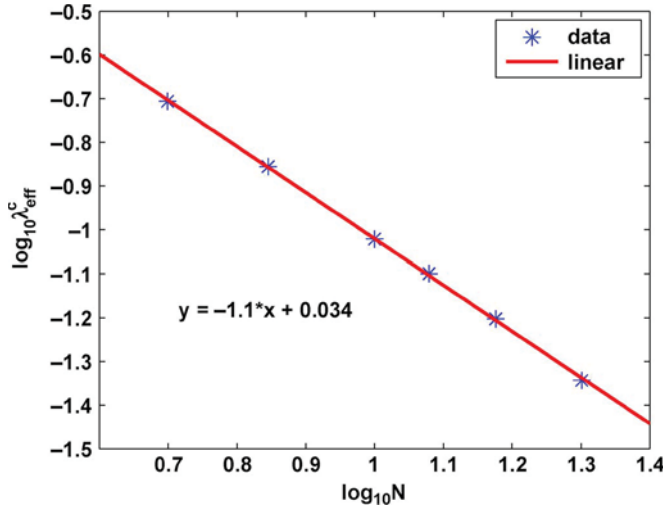


Fig. 3. Size dependence of critical network load. The critical value of effective load per node ( $\lambda_{eff}^c$ ) is plotted against the size of the network ( $N$ ) in the log-log scale.

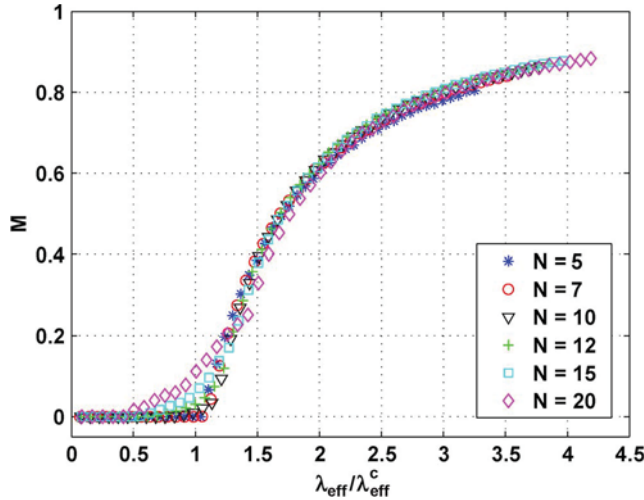


Fig. 4. Size-independent global model for network phase transition. The normalized packet drop rate ( $M$ ) is plotted against the effective load per node ( $\lambda_{eff}$ ), normalized with respect to its critical value ( $\lambda_{eff}^c$ ).

*Remark 2.1 (Critical Slowing Down):* Note that there is a slight mismatch among the phase transition curves in Fig. 4, which can be attributed to the fact of *critical slowing down* [18]. The normalized time-correlation function  $\phi_x(t_1 - t_2)$  of an observable  $x(t)$  is defined [19] as

$$\phi_x(t_1 - t_2) = \frac{[\langle x(t_1)x(t_2) \rangle - \langle x \rangle^2]}{[\langle x^2 \rangle - \langle x \rangle^2]} \approx e^{-\frac{|t_1 - t_2|}{\tau}} \quad (6)$$

where  $\tau$  is the slowest relaxation time or temporal correlation length of the observable  $x(t)$ . It is known that, near a phase transition point, the relaxation time  $\tau$  of the slowest mode of a system diverges, i.e.,  $\tau \rightarrow \infty$ . This phenomenon is known as the *critical slowing down*, and as a consequence, it takes a long time to make two consecutive independent observations near a phase transition point. Therefore, it is difficult to simulate a large system in the vicinity of the phase transition point.

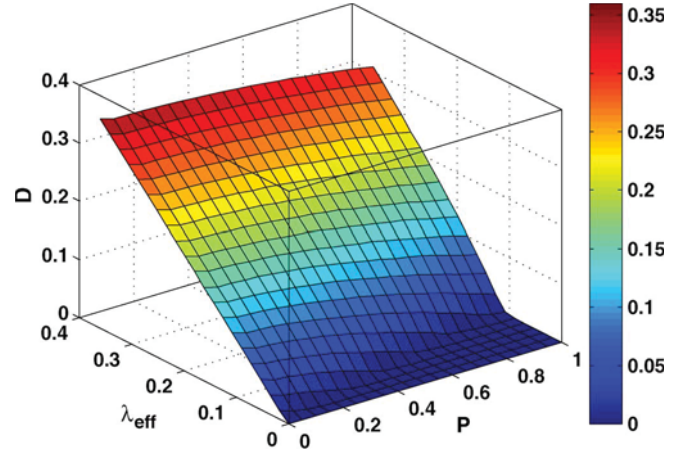


Fig. 5. Average (over time) packet drop rate ( $D$ ) as a function of packet arrival rate ( $\lambda_{eff}$ ) and transmission probability ( $P$ ).

### III. HETEROGENEOUS PACKET TRANSMISSION: INTRODUCTION OF A SECOND INTENSIVE PARAMETER

Previous analyses studied the effects of variations of a single parameter (e.g., packet arrival rate  $\lambda$ ) on the average packet drop rate in the network. This section introduces a second network parameter, namely, the packet transmission probability  $P$ , for networks transmitting heterogeneous packets. Given that each node of the network has multiple independent queues (buffers) for different packet types, packet transmission probability is defined as the probability of transmitting a particular packet in a time unit.

In the previous case, with a single packet type,  $P$  was set to unity. It is obvious that, when  $P$  is decreased from unity, the network is expected to move from the steady phase to an unsteady phase for lower values of the effective load  $\lambda_{eff}$ . Similar phenomena take place in two-phase thermodynamic (e.g., solid-liquid and gas-liquid) systems, where the critical temperature of solid-liquid (or gas-liquid) transition can be altered by changing the superincumbent pressure; lower pressures usually lead to lower values of melting or boiling point. In this context, the packet transmission probability  $P$  of the nodes is called the *network pressure*, while  $\lambda_{eff}$  is called the *network temperature*. Therefore, phase transition in the network is a function of a combination of network temperature  $\lambda_{eff}$  and network pressure  $P$ . Fig. 5 presents a 3-D plot of the average packet drop rate as a function of  $\lambda_{eff}$  and  $P$ , and it is observed that  $\lambda_{eff}^c$  decreases with a decrease in  $P$ .

#### A. Problem Formulation

The problem of heterogeneous packet transmission in networks is formalized in this section. Let there be  $K$  independent queues for  $K$  types of packets for each node of the network. For example, in the context of perimeter surveillance sensor networks, heterogeneity may accrue from differences in the sensor modality (e.g., audio, video, and magnetic sensor packets). In terms of packet arrival, let the effective Poisson arrival rate of packet type  $i$  be simply (subscript *eff* is omitted) denoted by  $\lambda_i$  ( $0 \leq \lambda_i \leq 1$ ) for  $i = 1, 2, \dots, K$ . However, only one channel is used for packet transmission, i.e., at a given

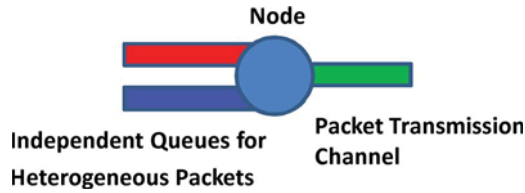


Fig. 6. Illustration for heterogeneous packet transmission mechanism.

time instant, a node can at most transmit one packet. A typical scenario is illustrated in Fig. 6. The probability that a node in the network transmits a packet of type  $i$  is denoted by  $P_i$ , where  $0 \leq P_i \leq 1$ . A single transmission channel leads to the constraint  $\sum_{i=1}^K P_i \leq 1$ . In this setting, the network is analogous to multicomponent materials in a thermodynamic sense. Thus, phase diagrams can be constructed for the equilibrium states for such networks which eventually lead to a probabilistic (static) control strategy of heterogeneous packet transmission. The next section presents the idea of phase diagrams for networks transmitting heterogeneous packets with a representative example of two packet types.

### B. Construction of Phase Diagrams

In the current setting of independent queues for different packet types, the uncongested phase  $U$  (respectively, congested phase  $C$ ) for a particular packet type is characterized by zero (respectively, nonzero) packet drop rate. Thus, with respect to a particular packet type, a network is in an uncongested or a congested phase. This leads to the existence of mixed phases, where the network remains in an uncongested phase for some packet types while being in a congested phase for the rest of the packet types. For a network carrying  $K$  packet types, there are  $2^K$  possible phases, which include the fully uncongested, fully congested, and mixed phases.

The purpose of a network phase diagram is to determine its phase for given distributions of  $\lambda_i$  and  $P_i$ . A representative example with two packet types is presented here. As discussed earlier, there can be the following four different network phases for this example.

- 1) *Completely uncongested phase ( $U1 + U2$ )*: Both packet types are in uncongested phase, which signifies negligibly small values of average packet drop rates  $D_1$  and  $D_2$  for both queues.
- 2) *Congested packet type 1 and uncongested packet type 2 ( $C1 + U2$ )*: Packet type 1 has a relatively large (i.e., nonzero) drop rate  $D_1$  while packet type 2 still maintains a negligibly small  $D_2$ .
- 3) *Uncongested packet type 1 and congested packet type II ( $U1 + C2$ )*: Packet type 1 maintains a negligibly small drop rate  $D_1$  while packet type 2 has a relatively large (positive) drop rate  $D_2$ .
- 4) *Completely congested phase ( $C1 + C2$ )*: Both packet types are in the congested phase, i.e., average packet drop rates  $D_1$  and  $D_2$  are large for both queues.

*Remark 3.1:* When compared to a multicomponent material in the thermodynamic sense, the network with two packet types has a remarkable similarity with two-component mixtures, e.g.,

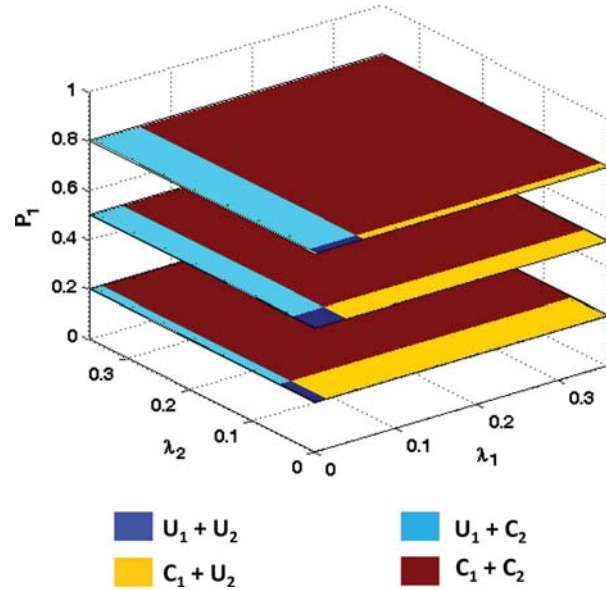


Fig. 7. Network phase diagram with two packet types. The transmission probability of packet type  $i$  is denoted as  $P_i$  under the constraint  $P_1 + P_2 = 1$ . The arrival rate for packets of type  $i$  is denoted as  $\lambda_i$ .

a mixture of olivine (i.e., an isolated tetrahedra) and pyroxene (i.e., single chain tetrahedra). The corresponding phase diagram is known as the *binary eutectic phase diagram* [20] that explains the chemical process of generating the two immiscible crystals from a completely miscible liquid based on temperature and pressure variations. This process also has four possible phases, which are the following: completely liquid phase (analogous to  $C1 + C2$ ), mixture of liquid olivine with pyroxene crystals (analogous to  $C1 + U2$ ), mixture of olivine crystals with liquid pyroxene (analogous to  $U1 + C2$ ), and, finally, the completely crystallized phase (analogous to  $U1 + U2$ ).

Fig. 7 shows a 3-D phase diagram for the aforementioned example, which is constructed via Monte Carlo simulation. Four different phases are color coded. Both  $\lambda_1$  and  $\lambda_2$  are varied independently from 0 to 1. Under the constraint  $P_1 + P_2 = 1$ , planes for three different values of  $P_1$ , namely, 0.2, 0.5, and 0.8, show changes in the sizes of the phases across the range of  $P_1$ . The regularity in shapes of different phase zones is attributed to the independence of the queues in different packet types. This type of phase diagrams leads to the concept of controlling heterogeneous packet transmission from the perspectives of statistical mechanics. The basic idea is to move as close as possible to the phase boundary of the  $U1 + U2$  phase from outside by choosing an appropriate  $P_1$ . On the other hand, if the network is already inside the  $U1 + U2$  phase, then the strategy should be to move as far away as possible from its phase boundary, also by choosing an appropriate  $P_1$ . The key objective is to achieve robustness against the uncertainties in  $\lambda_1$  and  $\lambda_2$  or any other internal disturbances.

## IV. CONTROL OF PACKET TRANSMISSION

The thermodynamic approach, presented earlier to analyze and understand the congestion phenomenon in communication networks, could lead to a novel means to control heterogeneous packet transmission by tuning the packet transmission

probability at a node level. This section presents the basic philosophy of such a control strategy along with representative numerical experiments on the network test bed. From a control perspective, the congestion states of a network are categorized into the following two types.

- 1) Phase Type I: At least a queue for one packet type is in the congested phase.
- 2) Phase Type II: Queues for all packet types are not in the congested phase.

The control strategies are fundamentally different for these two phase types. For Phase Type I, the drop rate of at least one type of packet is nonzero, where the objective is to choose a probability distribution of packet transmission to minimize the worst case drop rate among packet types with nonzero drop rates that can be scaled by respective packet arrival rates and also by a user-defined packet priority distribution. In contrast, in Phase Type II, drop rates for all packet types are already zero. Thus, any transmission probability distribution that keeps all the drop rates at zero is an optimal distribution according to the objective of Phase Type I. However, a unique solution can be achieved by the use of a different objective that would provide robustness against the uncertainties and/or fluctuations in packet arrival rates. The idea here is to choose a transmission probability distribution that pushes the network deeper into the steady phase (i.e., away from the phase boundary) so that, even with fluctuations or uncertainties in the packet arrival rate, the probability of being in the unsteady phase is minimized. In other words, the objective is to maximize the least uncertainty tolerance in packet arrival rate. As before, the uncertainty tolerance can be scaled by the respective packet arrival rates and the user-defined packet priority distribution. In the sequel, the control objectives are described formally along with optimal solutions for general and special cases.

### A. Control Objectives and Optimal Solutions

A user-defined packet priority distribution  $\{\alpha_i\}$ , with  $\alpha_i \in [0, 1]$  and  $\sum_{i=1}^K \alpha_i = 1$ , is considered for defining the control objectives in both phase types. A higher value of  $\alpha_i$  signifies higher priority for packet type  $i$  while lowering the drop rate in Phase Type I or increasing the uncertainty tolerance in Phase Type II.

1) *Control Objective for Phase Type I:* In Phase Type I, the following model for the average drop rate  $D_i$  of packet type  $i$  is assumed:

$$D_i = f_1(P_i, \lambda_i) \quad \forall i \quad (7)$$

where  $P_i$  and  $\lambda_i$  are defined as before. To find an optimal transmission probability distribution  $P_i$  for a given packet arrival distribution  $\lambda_i$  and a packet priority distribution  $\alpha_i$  that minimize the (weighted) worst case packet drop rate, the cost functional is chosen as

$$C(P_1, P_2, \dots, P_K) = \max_{i=1, \dots, K} \frac{\alpha_i D_i}{\lambda_i} \quad (8)$$

$$= \max_i \frac{\alpha_i f_1(P_i, \lambda_i)}{\lambda_i}. \quad (9)$$

The optimal packet transmission probability for each packet type  $i$  is denoted as  $P_i^*$  and is given by the constrained optimization solution as

$$(P_1^*, P_2^*, \dots, P_K^*) = \arg \min_{\sum_i P_i \leq 1} C(P_1, P_2, \dots, P_K) \quad (10)$$

$$= \arg \min_{\sum_i P_i \leq 1} \max_i \frac{\alpha_i f_1(P_i, \lambda_i)}{\lambda_i}. \quad (11)$$

Although the constraint on the transmission probability distribution is  $\sum_i P_i \leq 1$ , the inequality is replaced by an equality sign for optimality. The rationale for involving  $\lambda_i$  in the objective function is to normalize the drop rates for each packet type as it was done to define the order parameter  $M$  in Section II. In this context, the solution of the aforementioned optimization problem is fairly straightforward. For the optimal packet transmission probability distribution, a solution of the form

$$\frac{\alpha_i D_i}{\lambda_i} = \frac{\alpha_j D_j}{\lambda_j} \quad \forall i, j \quad (12)$$

is globally optimal if it exists. This is because, if one chooses a packet transmission probability distribution to reduce  $\alpha_i D_i / \lambda_i$  even further for some value of  $i$ , then  $\alpha_j D_j / \lambda_j$  would increase for at least one  $j$ . Thus, the cost functional would increase. The proof of optimality is presented in Appendix A which also introduces a sequential and recursive approach to min-max optimization when the condition described in (12) does not exist.

2) *Control Objective for Phase Type II:* In Phase Type II,  $D_i = 0$ , i.e.,  $f_1(P_i, \lambda_i) = 0 \forall i$ , which leads to the critical curve that yields the critical transmission probability  $P^c$  for a given packet arrival rate  $\lambda$ . Let the critical curve be described as

$$P^c = f_2(\lambda). \quad (13)$$

For a given packet arrival distribution  $\lambda_i$ , if the network is in Phase Type II, then

$$\sum_i P_i^c = \sum_i f_2(\lambda_i) \leq 1. \quad (14)$$

This implies that the packet transmission probability distribution  $P_i^c$ , where  $\sum_i P_i^c \leq 1$ , is sufficient for the network to be in Phase Type II. However, the queues for different packet types can still use the remaining packet transmission capability of the network expressed as  $(1 - \sum_i P_i^c)$ . The goal here is to distribute this remaining packet transmission capability among the queues in an optimal sense. To formalize this problem, a virtual packet arrival distribution  $\lambda_i^U$  is defined, where  $\lambda_i^U$  ( $\lambda_i \leq \lambda_i^U \leq 1$ ) denotes a distribution that includes uncertainty/fluctuations in packet arrival rates, for which the network just remains in Phase Type II. The corresponding packet transmission probability  $P_i$  is given as  $f_2(\lambda_i^U)$ , which leads to

$$\lambda_i^U = \lambda_i + \tilde{\lambda}_i \quad (15)$$

$$\Rightarrow \tilde{\lambda}_i = f_2^{-1}(P_i) - \lambda_i \quad (16)$$



where  $\tilde{\lambda}_i (\geq 0)$  uncertainty tolerance provided for packet type  $i$ , for which the network still remains in the steady phase (Phase Type II). The objective of the optimization problem for Phase Type II is to choose a transmission probability distribution  $P_i$  that maximizes the minimum uncertainty tolerance for packet arrival statistics. Thus, the objective functional is the worst case (weighted) tolerance  $\tilde{\lambda}_i$  that is given as

$$O(P_1, P_2, \dots, P_K) = \min_{i=1, \dots, K} \frac{\tilde{\lambda}_i}{\alpha_i} \quad (17)$$

$$= \min_i \frac{f_2^{-1}(P_i) - \lambda_i}{\alpha_i}. \quad (18)$$

The constrained optimization problem that has to be solved to obtain the optimal packet transmission probability is given as

$$(P_1^*, P_2^*, \dots, P_K^*) = \arg \max_{\substack{P_i \leq 1 \\ P_i \geq f_2(\lambda)}} O(P_1, P_2, \dots, P_K) \quad (19)$$

$$= \arg \max_{\substack{P_i \leq 1 \\ P_i \geq f_2(\lambda)}} \min_i \frac{f_2^{-1}(P_i) - \lambda_i}{\alpha_i}. \quad (20)$$

The solution for the aforementioned optimization problem is

$$\frac{\tilde{\lambda}_i}{\alpha_i} = \frac{\tilde{\lambda}_j}{\alpha_j} \quad \forall i, j. \quad (21)$$

The details are presented in Appendix B. Note the use of the packet priority distribution  $\alpha_i$  in the optimization problem, where the queue for packet type  $i$  gains more uncertainty tolerance for a higher value of  $\alpha_i$ .

### B. Approximation of Functional Forms

The formulation of a general control strategy is presented above. For a given network structure and routing strategy, one needs to estimate the functions  $f_1$  and  $f_2$ , for Phase Type I and Phase Type 2, respectively, in order to implement the control strategy (see Section IV-A). An example procedure is shown in this section with respect to the network considered in the current study. Fig. 5 generates an idea for constructing the functional forms for both  $f_1$  and  $f_2$ . For Phase Type I (i.e., average packet drop rate per node  $D > 0$ ), the function  $f_1$  can be approximated with a linear function, i.e., with a plane. As there is no bias, the plane equation for  $D$  as a function of  $P$  and  $\lambda$  is

$$D = aP + b\lambda. \quad (22)$$

The linear least square fit of the (numerical) experimental data is shown in Fig. 8, and the error is found to be sufficiently low. Note that the plane (22) only fits the data in Phase Type I and is not valid for Phase Type II data, where  $D$  is always zero. The error in the fit is observed to increase near the critical curve (see the circled zone in the figure). This can be attributed to the problem *critical slowing down* near the critical points as discussed earlier. Note that  $D$  decreases with an increase in  $P$  and  $D$  increases with an increase in  $\lambda$ .

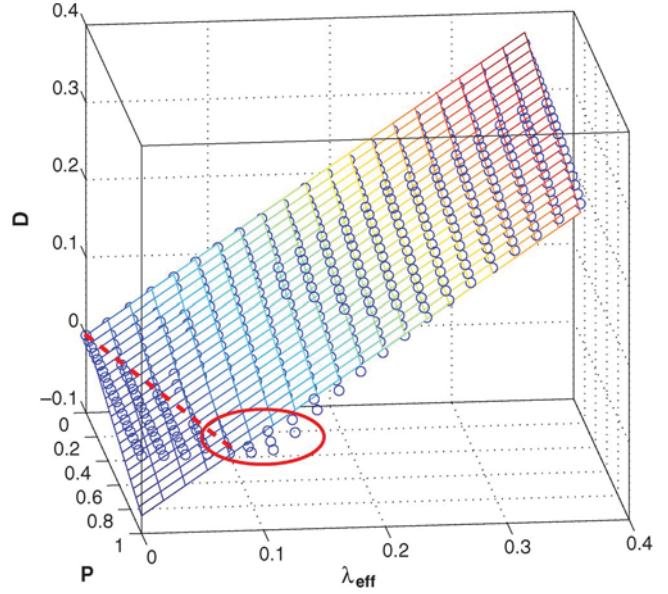


Fig. 8. Approximation of average (over time) packet drop rate ( $D$ ) as a function of packet arrival rate ( $\lambda_{eff}$ ) and transmission probability ( $P$ ) in square-grid networks. (Note: The region for  $D < 0$  is inaccessible because the packet drop rate is nonnegative.)

The critical curve equation (for Phase Type II) is found by setting  $D = 0$ . In other words, the critical curve is the straight line at which the least square fit plane intersects the  $D = 0$  plane. Thus, the following equation for the critical curve is obtained:

$$P^c = c\lambda \quad (23)$$

where  $c$  is a positive constant as  $P^c$  must increase with an increase in  $\lambda$ . Furthermore,  $c$  is expressed as follows:  $c \triangleq -b/a$ , where  $b$  is positive and  $a$  is strictly negative (i.e.,  $a = -|a| < 0$ ). Fig. 8 shows the critical line (dotted) for the current network model.

The aforementioned approximate functional forms allow the derivation of the closed-form solutions (for both phase types) of the optimal packet transmission probability distribution by solving algebraic equations.

1) *Solution for Phase Type I*: In Phase Type I, the sufficiency criterion for the optimal distribution is given by the solution of the following system of algebraic equations:

$$\frac{\alpha_i a}{\lambda_i} P_i + \alpha_i b = \frac{\alpha_j a}{\lambda_j} P_j + \alpha_j b \quad \forall i, j \in \{1, \dots, K\}. \quad (24)$$

For a network with  $K$  packet types, the above system provides  $(K - 1)$  independent equations, and the  $K$ th equation is provided by the constraint  $\sum_i P_i = 1$ .

Defining  $Q \triangleq (\alpha_i a / \lambda_i) P_i + \alpha_i b \forall i \in \{1, \dots, K\}$ , (23) and (24) are combined to obtain

$$Q = \frac{\alpha_i |a|}{\lambda_i} (P_i^c - P_i) \quad \forall i \in \{1, \dots, K\}. \quad (25)$$

It follows that  $P_i$  can be evaluated in terms of  $Q$  as

$$P_i = P_i^c - \frac{Q \lambda_i}{\alpha_i |a|}. \quad (26)$$



The parameter  $Q$  is evaluated by summing the expressions for all  $i$  and using the constraint  $\sum_i P_i = 1$

$$Q = \frac{\sum_{i=1}^K P_i^c - 1}{\sum_{i=1}^K \left( \frac{\lambda_i}{\alpha_i |a|} \right)}. \quad (27)$$

However, for certain choices of  $\alpha_i$ , the solution of  $P_i$  may be negative for a given distribution of  $\lambda_i$ , which does not satisfy the constraint  $P_i \geq 0 \forall i$ . This implies that a solution for the set of equations in (24) does not exist in the feasible region. Therefore, as described in Appendix A, the negative solutions should be constrained to zeros, and the optimization should be repeated for the rest of the packet types in order to satisfy the constraint  $\sum_i P_i = 1$ . The algorithm for such a sequential optimization procedure is given hereinafter.

#### Algorithm for Sequential Optimization

$$Q^{(1)} = (\sum_{i=1}^K P_i^c - 1) / \sum_{i=1}^K (\lambda_i / \alpha_i |a|);$$

$$P_j^{(1)} = P_j^c - (Q^{(1)} \lambda_j / \alpha_j |a|); \text{ (Evaluate } P_j^{(1)} \forall j)$$

$$i = 1;$$

**while**  $P_j^{(i)} < 0$  for some  $j$  **do**

**for all**  $j : P_j^{(i)} < 0$  **do**

$$P_j^{(i)} \leftarrow 0;$$

**end for**

$i \leftarrow i + 1$

$$Q^{(i)} = (\sum_{k=\{j:P_j^{(i-1)} \neq 0\}} P_k^c - 1) /$$

$$(\sum_{k=\{j:P_j^{(i-1)} \neq 0\}} (\lambda_k / \alpha_k |a|));$$

**for all**  $j : P_j^{(i-1)} \neq 0$  **do**

$$P_j^{(i)} = P_j^c - (Q^{(i)} \lambda_j / \alpha_j |a|);$$

**end for**

**end while**

Note that, as a consequence of the aforementioned algorithm, the transmission probability distribution may have  $P_i = 0$  for some values of  $i$  and the rest of the elements will have  $P_j > 0$  (where  $j \in \{1, 2, \dots, K\} \setminus \{i\}$ ). The scaled drop rates can be written as

$$\frac{\alpha_i D_i}{\lambda_i} = \alpha_i b - \frac{\alpha_i |a|}{\lambda_i} P_i. \quad (28)$$

Therefore,  $\forall j, k \in \{1, 2, \dots, K\} \setminus \{i\}$  (i.e., for all packet types with nonzero transmission probability), the solution obtained from the algorithm will satisfy

$$\frac{\alpha_j D_j}{\lambda_j} = \frac{\alpha_k D_k}{\lambda_k} \quad \forall j, k. \quad (29)$$

Thus, due to the constraints on the transmission probability distribution, the optimal condition described in (12) will only be satisfied on a subset (containing packet types with nonzero transmission probability) of the set of all packet types. The rest of the packet types will have zero transmission probability essentially due to their low  $(\alpha/\lambda)$  ratio. As an example, for a

network with two packet types, the optimal packet transmission probability for type 1 is given by  $P_1$

$$P_1 = \begin{cases} 0 & \text{if } \frac{\alpha_1}{\alpha_2} < \max \left[ 1 - \frac{1}{c\lambda_2}, 0 \right] \\ \frac{\alpha_2 + c\lambda_2(\alpha_1 - \alpha_2)}{\lambda_2 \left[ \frac{\alpha_1}{\lambda_1} + \frac{\alpha_2}{\lambda_2} \right]} & \text{if } \max \left[ 1 - \frac{1}{c\lambda_2}, 0 \right] \leq \frac{\alpha_1}{\alpha_2} \leq \frac{1}{\max \left[ 1 - \frac{1}{c\lambda_1}, 0 \right]} \\ 1 & \text{if } \frac{\alpha_1}{\alpha_2} > \frac{1}{\max \left[ 1 - \frac{1}{c\lambda_1}, 0 \right]} \end{cases}$$

and  $P_2 = 1 - P_1$ . Note that, in Phase Type I,  $\sum_i P_i^c = c \sum_i \lambda_i > 1$  which ensures that  $\max \left[ 1 - \frac{1}{c\lambda_2}, 0 \right] \leq \frac{1}{\max \left[ 1 - \frac{1}{c\lambda_1}, 0 \right]}$ .

2) *Solution for Phase Type II:* For Phase Type II, as discussed earlier, the remaining transmission capability (beyond the critical curve)  $1 - \sum_i P_i^c$  is distributed among the queues of different packet types. To achieve the closed-form solution, we begin with the solution for Phase Type II [given in (21)]. Using the model in (23), the optimal solution can be written as

$$\frac{P_i - \lambda_i}{\alpha_i} = R \quad \forall i \quad (30)$$

where  $R$  is a constant. Appendix B establishes the existence and optimality of the solution. Therefore

$$P_i = c(R\alpha_i + \lambda_i). \quad (31)$$

The value of constant  $R$  is evaluated using the constraint  $\sum_k P_k = c \sum_k (R\alpha_k + \lambda_k) = 1$  as

$$R = \frac{1}{c} - \sum_k \lambda_k. \quad (32)$$

Thus, the optimal packet transmission probability distribution for Phase Type II is given by

$$P_i = c\lambda_i + \alpha_i \left[ 1 - c \sum_k \lambda_k \right]. \quad (33)$$

For example, the exact solution for the network with two packet types is

$$P_1 = \alpha_1 + c[\lambda_1 - \alpha_1(\lambda_1 + \lambda_2)] \quad (34)$$

and  $P_2 = 1 - P_1$ .

#### C. Representative Experimentation

A representative experimentation on the simulated network test bed is presented in this section to validate the effectiveness of the control strategy developed earlier. The simulation test bed is the same network system analyzed in this paper. This network structure has the potential of serving as a fundamental building block for real-life (and hence more complex) networks. For example, the network structure used here can be considered as a part of a larger sensor network, where edges of the current (smaller) network interact with the rest of the network. In [21], an important observation is made that distinguishes sensor network routing from other classical IP-based routing schemes. In sensor networks, generated data have significant redundancy as multiple sensors in the vicinity of an event may

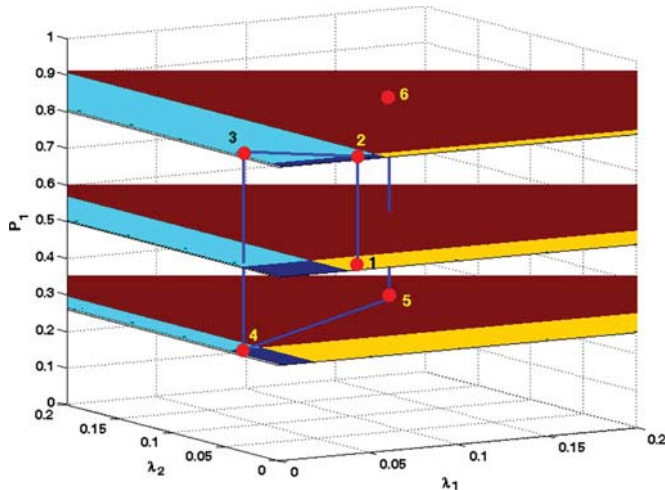


Fig. 9. System trajectory in the phase diagram for the representative experimentation. The transmission probability of packet type  $i$  is denoted as  $P_i$  under the constraint  $P_1 + P_2 = 1$ . The arrival rate for packets of type  $i$  is denoted as  $\lambda_i$ .

generate the same data. Routing protocols should use such redundancy to improve energy and bandwidth utilization. The redundancy can be used by implementing an activity schedule where a significant fraction of nodes are kept inactive at a given time. Furthermore, often events to be detected by large sensor networks are local, rare, and random. Such information needs to be propagated throughout the network for activity scheduling as well as for achieving robustness to node failures. This paper addresses the issue of information propagation by multisource and multidestination types of communication.

Two different types of packets, namely, packet type 1 and packet type 2 (e.g., video and audio sensors), are considered for the validation of the underlying concept. Following the notations used in this paper, the packet arrival statistics is described by  $\lambda_1$  and  $\lambda_2$ , packet priority distribution is denoted by  $\alpha_1$  (taken as 0.6) and  $\alpha_2$  (taken as 0.4), and drop rates are denoted by  $D_1$  and  $D_2$ . The goal here is to estimate  $\lambda_1$  and  $\lambda_2$  over a time window, then to detect the phase type with respect to the estimated arrival rates, and to implement the optimal packet transmission probability distribution  $P_1$  and  $P_2$ . Let the time scale of network dynamics be denoted as fast time scale (denoted by *Time*) and the time scale of arrival statistics estimation be denoted as slow time scale. In this experiment, a slow time-scale epoch contains a window of 2000 fast time-scale epochs.

Fig. 9 shows the steady-state position of the network in the equilibrium phase diagram at different slow time-scale epochs. The phase diagram is the same as the one shown in Section III. Also, the time-series response of  $D_1$  and  $D_2$  is shown in Fig. 10 along with the corresponding control set point  $P_1$ .

For model identification purpose, parameters are identified to be  $a = -0.0771$ ,  $b = 1.0404$ , and  $c = -b/a = 13.4942$ . Initially ( $Time = 1$ ), the network is at state 1 (as shown in Fig. 9) with  $\lambda_1 = 0.05$ ,  $\lambda_2 = 0.01$ , and  $P_1 = P_2 = 0.5$ . Over slow time-scale epoch 1,  $\lambda_1$  and  $\lambda_2$  are estimated, and at the beginning of slow time-scale epoch 2 ( $Time = 2000$ ), the network moves to the optimal location with  $P_1 \approx 0.8$  (state 2 in

Fig. 9). As a consequence, the very small ripples noticed in the  $D_1$  time series in slow time-scale epoch 1 disappear. (Note that this is a Phase Type II control.) Although lower  $P_1$  would have been permissible, it reached  $P_1 = 0.8$  to optimally move away from the phase boundary. At the beginning of slow time-scale epoch 3 ( $Time = 4000$ ), the packet arrival statistics is changed to  $\lambda_1 = 0.01$  and  $\lambda_2 = 0.05$  (state 3 in Fig. 9), and it can be observed that, after some initial transience response,  $D_2$  settles to a nonzero value during slow time-scale epoch 3 whereas  $D_1$  remains at zero. Upon estimation of  $\lambda_1$  and  $\lambda_2$  over slow time-scale epoch 3, the controller takes the network to the optimal location with  $P_1 \approx 0.26$  (state 4 in Fig. 9) at the beginning of slow time-scale epoch 4 ( $Time = 6000$ ). As a consequence,  $D_2$  response dies down, keeping the  $D_1 = 0$  response. The network is still in Phase Type II. Finally, at the beginning of slow time-scale epoch 5 ( $Time = 8000$ ), the packet arrival statistics is changed to  $\lambda_1 = 0.15$  and  $\lambda_2 = 0.15$  (state 5 in Fig. 9); both  $D_1$  and  $D_2$  responses settle in a nonzero value over the slow time-scale epoch 5. However, the steady-state value of  $D_1$  is higher than that of  $D_2$  as the packet transmission probability distribution is still in favor of packet type 2 (with  $P_1 = 0.26$  and  $P_2 = 0.74$ ). From the control point of view, the network is in Phase Type I, and the controller drives the network to the optimal location with  $P_1 \approx 0.8$  (state 6 in Fig. 9) at the beginning of slow time-scale epoch 6 ( $Time = 10000$ ). Higher value of  $P_1$  compared to that of  $P_2$  is due to the higher priority of packet type 1 over packet type 2 ( $\alpha_1 > \alpha_2$ ). Consequently,  $D_1$  settles to a lower value compared to  $D_2$  over the slow time-scale epoch 6.

## V. SUMMARY, CONCLUSION, AND FUTURE WORK

This paper presents statistical mechanics-inspired analysis of critical phenomena in square-grid wired communication networks and validates the pertinent results on a simulation test bed. A comprehensive finite-size scaling analysis has been performed for a specific network structure, where network analogs of intensive thermodynamic variables (e.g., order parameter, temperature, and pressure) are introduced to unambiguously explain the underlying concepts. Phase diagrams are constructed for networks transmitting heterogeneous packet types for the control of the network system. Control strategies are synthesized to achieve robustness in the fully uncongested phase and to mitigate the worst case packet drop rate in congested phases. Closed-form solutions for optimal control strategies have been formulated for the square-grid communication network structure. The analysis framework, reported in this paper, can be potentially extended to ad hoc wireless sensor networks for boundary surveillance problems, where spatially distributed autonomous agents with (possibly multimodal) sensing capabilities collaboratively monitor a given environment. In general, different sensors sense the environment in a distributed manner and need to communicate with other types of sensors for information fusion to make a comprehensive decision. Therefore, complex multihop routing of (possibly multipriority) information packets in a sensor network bears significant relevance. In addition, such a simple network structure may serve as subsets of large complex networks, where a thorough understanding of small subsets is necessary for the control of the entire network.

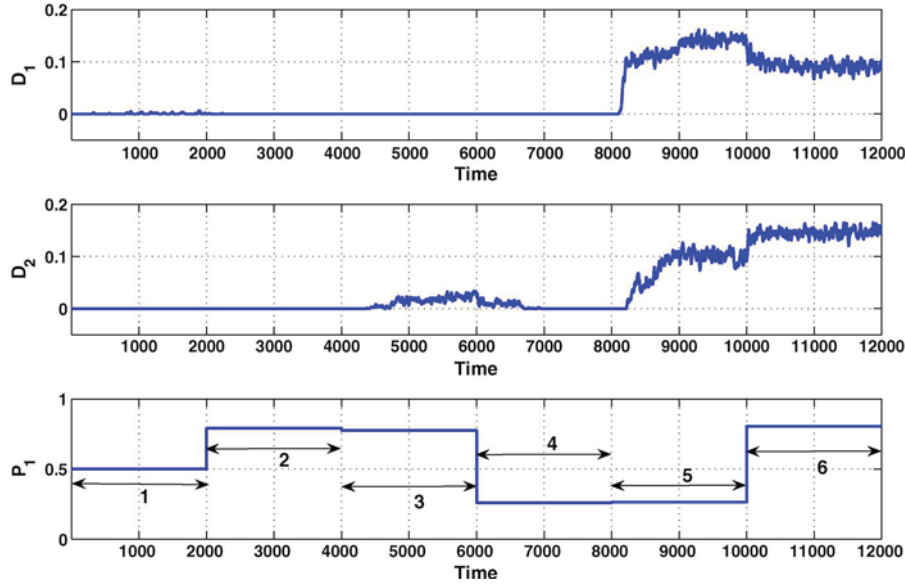


Fig. 10. Time series observation of packet drop rates for packet type 1 ( $D_1$ ) and packet type 2 ( $D_2$ ) and transmission probability for packet type 1 (control input)  $P_1$ .

In this paper, the packet arrival statistics is estimated in a centralized fashion, and a global packet transmission probability distribution is broadcasted for optimal control. In this context, the (statistical mechanics-inspired) *ensemble approach* may allow the control algorithm to be executed over a network in a communication-constrained environment via sparse statistical sampling for the estimation of network performance parameters. It will be more useful if each node takes its own decision based on the local (its own or of its neighborhood) estimation of packet arrival statistics and yet the global objective is satisfied. Due to the *ensemble approach* of the current study, such a distributed control problem is not intractable as well and will be an important topic of future research. This approach will also enable the routing strategy to handle different packet arrival statistics in different locations of the network.

This initial study develops a priority-based queuing mechanism and does not include the aspect of scheduling mechanism for temporal management of packet transmission. Therefore, from an application perspective, the dynamic aspects of the network behavior need to be addressed in future, particularly to handle distributed implementation of the algorithm and varying packet arrival statistics.

Apart from the issue of distributed realization, the following topics are recommended for future research from the perspectives of stability, performance analysis, and decision and control:

- 1) validation of the theoretical results in more complex and realistic network scenarios with various features (e.g., layered architectures and industry-standard protocols);
- 2) investigation of the effects of network topology (e.g. rectangular instead of square grid), packet arrival statistics, and distributions of source and destination nodes on performance analysis and decision and control;
- 3) dynamic analysis for convergence and stability of the network system as an augmentation of the equilibrium behavior analysis, reported in this paper;

- 4) handling of distributed implementation of the algorithms and varying packet arrival statistics.

#### APPENDIX A MIN-MAX OPTIMIZATION

Let  $g_i, i = 1, 2, \dots, K$ , be continuous and strictly monotonically decreasing functions defined as

$$g_i : [0, x_i^c] \rightarrow [g_i^{\min}, g_i^{\max}]$$

where  $0 < x_i^c \leq 1$ ,  $g_i(0) = g_i^{\max}$ , and  $g_i(x_i^c) = g_i^{\min}$ .

The optimization task is to minimize the cost functional

$$C(x_1, x_2, \dots, x_K) = \max_i g_i(x_i)$$

under the equality constraint  $\sum_i x_i = 1$  in the domain  $[x_1 x_2 \dots x_K] \in [0, x_1^c] \times [0, x_2^c] \times \dots \times [0, x_K^c]$ . It is noted that the feasible region is nonempty under the condition  $\sum_i x_i^c \geq 1$ . In this context, the following two cases are considered.

- Case 1) Let there exist a solution  $[x_1^* x_2^* \dots x_K^*]$  (in the feasible region and satisfying the equality constraint) such that  $g_1(x_1^*) = g_2(x_2^*) = \dots = g_K(x_K^*) = g^*$ , which implies that the cost  $C(x_1, x_2, \dots, x_K) = g^*$ .

Now let a nontrivial perturbation to another point in the feasible region be  $[\tilde{x}_1 \tilde{x}_2 \dots \tilde{x}_K]$ . Due to the constraints  $\sum_i x_i^* = \sum_i \tilde{x}_i = 1$ , there must exist at least one  $m$  such that  $x_m^* > \tilde{x}_m$ . Consequently,  $g_m(x_m^*) < g_m(\tilde{x}_m)$  implying that  $C(x_1^*, x_2^*, \dots, x_K^*) < C(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_K)$ . This establishes that  $[x_1^* x_2^* \dots x_K^*]$  is the globally optimal solution.

- Case 2) If a solution of the form  $[x_1^* x_2^* \dots x_K^*]$  such that  $g_1(x_1^*) = g_2(x_2^*) = \dots = g_K(x_K^*) = g^*$  does not exist, then it implies that the solution lies on

the edge of the feasible region. A constraint is activated by choosing an index  $m$  with the smallest  $g_m^{max}$  and constraining  $x_m = 0$ . The solution of the remaining problem is similar to the original problem except that it has one fewer function. This solution is locally optimal, and the process is iterated recursively to arrive at the final solution. In other words, the scheme is to sequentially activate constraints  $x_i = 0$  until a solution is found. Consequently, the maximum number of iterations that may be required for convergence is equal to  $K$ .

#### APPENDIX B MAX-MIN OPTIMIZATION

Let  $h_i, i = 1, 2, \dots, K$ , be continuous and strictly monotonically increasing homeomorphic functions defined as

$$h_i : [x_i^c, 1] \rightarrow [0, h_i^{max}]$$

where  $0 \leq x_i^c < 1$ ,  $h_i(x_i^c) = 0$ , and  $h_i(1) = h_i^{max}$ .

The optimization task is to maximize the cost functional

$$O(x_1, x_2, \dots, x_K) = \min_i h_i(x_i)$$

under the equality constraint  $\sum_i x_i = 1$  in the domain  $[x_1 \ x_2 \ \dots \ x_K] \in [x_1^c, 1] \times [x_2^c, 1] \times \dots \times [x_K^c, 1]$ . It is noted that the feasible region is nonempty under the condition  $\sum_i x_i^c < 1$ .

Let there exist a solution  $[x_1^* \ x_2^* \ \dots \ x_K^*]$  (in the feasible region and satisfying equality constraint) such that  $h_1(x_1^*) = h_2(x_2^*) = \dots = h_K(x_K^*) = h^*$ , which implies that  $O(x_1, x_2, \dots, x_K) = h^*$ . Let a nontrivial perturbation to another point in the feasible region be  $[\tilde{x}_1 \ \tilde{x}_2 \ \dots \ \tilde{x}_K]$ . Due to the constraints  $\sum_i x_i^* = \sum_i \tilde{x}_i = 1$ , there must exist at least one  $m$  such that  $x_m^* > \tilde{x}_m$ . Consequently,  $h_m(x_m^*) > h_m(\tilde{x}_m)$  implying that  $O(x_1^*, x_2^*, \dots, x_K^*) > O(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_K)$ . This establishes that  $[x_1^* \ x_2^* \ \dots \ x_K^*]$  is the globally optimal solution. The subsequent part shows the existence of such a solution.

Let  $y \in \bigcap_{i=1}^K [0, h_i^{max}] = [y_{min}, y_{max}]$ , where  $y_{min} = 0$  and  $y_{max} = \min_i h_i^{max}$ . Let a function  $J$  be defined on  $[0, \min_i h_i^{max}]$  as

$$J(y) = \sum_{i=1}^K h_i^{-1}(y) \quad y \in [y_{min}, y_{max}].$$

In the above equation, the inverse exists due to the continuity and strict monotonicity property of  $h_i$ . Under the assumption of homeomorphism,  $h_i^{-1}, i = 1, 2, \dots, K$ , are also strictly monotonically increasing functions.

The evaluation of  $J$  at  $y_{min}$  and  $y_{max}$  yields

$$J(y_{min}) = \sum_{i=1}^K h_i^{-1}(0) = \sum_i x_i^c < 1$$

$$J(y_{max}) = \sum_{i=1}^K h_i^{-1} \left( \min_i h_i^{max} \right) \geq h_m^{-1} (h_m^{max}) = 1$$

where  $m = \arg \min_i h_i^{max}$ . Therefore, the inequalities  $J(y_{min}) < 1 \leq J(y_{max})$  imply that there exists an  $h^* \in [y_{min}, y_{max}]$  such that  $J(h^*) = 1$ . The optimal solution is obtained as  $x_i^* = h_i^{-1}(h^*)$ , which satisfies the necessary condition  $h_1(x_1^*) = h_2(x_2^*) = \dots = h_K(x_K^*) = h^*$  under the constraint  $\sum_i x_i^* = 1$ .

#### ACKNOWLEDGMENT

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

#### REFERENCES

- [1] R. Pathria, *Statistical Mechanics*, 2nd ed. Oxford, UK: Butterworth-Heinemann, 1996.
- [2] K. Huang, *Statistical Mechanics*, 2nd ed. New York: Wiley, 1987.
- [3] V. Beneš, *Mathematical Theory of Connecting Networks and Telephone Traffic*. New York: Academic, 1965.
- [4] J. Y. Hui and E. Karasan, "A thermodynamic theory of broadband networks with application to dynamic routing," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 991–1003, Aug. 1995.
- [5] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.
- [6] B. Krishnamachari, S. Wicker, and R. Bejar, "Phase transition phenomena in wireless ad-hoc networks," in *Proc. IEEE Global Telecommun. Conf.*, 2001, pp. 2921–2925.
- [7] A. Srivastav, A. Ray, and S. Phoha, "Adaptive sensor activity scheduling in distributed sensor networks: A statistical mechanics approach," *Int. J. Distrib. Sens. Netw.*, vol. 5, no. 3, pp. 242–261, Jul. 2009.
- [8] S. Durlauf, "How can statistical mechanics contribute to social science?" *Proc. Natl. Acad. Sci. USA*, vol. 96, no. 19, pp. 10582–10584, Sep. 1999.
- [9] M. Millonas, "Swarms, phase transitions, and collective intelligence (paper 1); and a nonequilibrium statistical field theory of swarms and other spatially extended complex systems (paper 2)," Working Papers 93-06-039, Santa Fe Institute, Jun. 1993.
- [10] M. Schulz, *Statistical Physics and Economics: Concepts, Tools and Applications Springer Tracts in Modern Physics*, vol. 184. Berlin, Germany: Springer-Verlag, 2003.
- [11] S. Sarkar, K. Mukherjee, A. Srivastav, and A. Ray, "Understanding phase transition in communication networks to enable robust and resilient control," in *Amer. Control Conf.*, St. Louis, MO, 2009, pp. 1549–1554.
- [12] T. Ohira and R. Sawatari, "Phase transition in computer network traffic model," *Phys. Rev. E*, vol. 58, no. 1, pp. 193–195, Jul. 1998.
- [13] A. Lawniczak and X. Tang, "Network traffic behaviour near phase transition point," *Eur. Phys. J. B, Condens. Matter Complex Syst.*, vol. 50, no. 1/2, pp. 231–236, Mar. 2006.
- [14] R. Guimera, A. Arenas, A. Diaz-Guilera, and F. Giralt, "Dynamical properties of model communication networks," *Phys. Rev. E*, vol. 66, no. 2, p. 026704, Aug. 2002.
- [15] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*. Reading, MA: Perseus Books, 1992.
- [16] F. Baccelli, D. R. McDonald, and J. Reynier, "A mean-field model for multiple TCP connections through a buffer implementing red," *Perform. Eval.*, vol. 49, no. 1–4, pp. 77–97, Sep. 2002.
- [17] S. Sarkar, K. Mukherjee, A. Srivastav, and A. Ray, "Critical phenomena and finite-size scaling in communication networks," in *Proc. Amer. Control Conf.*, Baltimore, MD, 2010, pp. 271–276.
- [18] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics*, 3rd ed. Singapore: World Scientific, 2005.
- [19] D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [20] L. Fichter, "Binary eutectic phase diagram," Dept. Geol. Environ. Sci., James Madison Univ., Harrisonburg, Virginia, Tech. Rep., 2000.
- [21] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Netw.*, vol. 3, no. 3, pp. 325–349, 2005.





**Soumik Sarkar** (S'10–M'11) received the B.S. degree in mechanical engineering from Jadavpur University, Kolkata, India, in 2006 and the concurrent Masters' degrees in mathematics and mechanical engineering and the Ph.D. degree in mechanical engineering from The Pennsylvania State University, University Park, in 2009 and 2011, respectively.

He is currently a Senior Research Scientist with the United Technologies Research Center, East Hartford, CT. His research interests include cyber-physical systems, distributed control, network analysis, statistical signal processing and pattern recognition, sensor fusion, fault detection, and prognostics and health management in aerospace, building, and energy systems.

Dr. Sarkar is a member of the American Society of Mechanical Engineers.



**Kushal Mukherjee** (S'08–M'11) received the B.S. degree in mechanical engineering from the Indian Institute of Technology, Roorkee, India, in 2006 and the concurrent Masters' degrees in electrical engineering and mechanical engineering and the Ph.D. degree in mechanical engineering from The Pennsylvania State University, University Park, in 2009 and 2011, respectively.

He is currently a Postdoctoral Research Scholar with The Pennsylvania State University. His research interests include symbolic dynamics, sensor networks, and multiagent systems.

Dr. Mukherjee is a member of the American Society of Mechanical Engineers.



**Asok Ray** (SM'83–F'02) received the graduate degrees in each discipline of electrical engineering, mathematics, and computer science and the Ph.D. degree in mechanical engineering from Northeastern University, Boston, MA.

He joined The Pennsylvania State University, University Park, in July 1985 and is currently a Distinguished Professor of Mechanical Engineering and a Graduate Faculty of Electrical Engineering. He has authored or coauthored over 500 research publications, including about 200 and 50 scholarly articles in refereed journals. Further details of Dr. Ray's credentials are available in the web address <http://www.mne.psu.edu/Ray/>.

Dr. Ray is a fellow of American Society of Mechanical Engineers and the World Innovative Foundation.



**Abhishek Srivastav** received the B.S. degree in mechanical engineering from the Indian Institute of Technology, Kanpur, India, in 2003 and the concurrent Masters' degrees in mathematics and mechanical engineering and the Ph.D. degree in mechanical engineering from The Pennsylvania State University, University Park, in 2006 and 2009, respectively.

He is currently a Senior Research Scientist with the United Technologies Research Center, East Hartford, CT. His research interests include data mining, machine learning, prognostics and health management, and sensor networks.

Dr. Srivastav is a member of the American Society of Mechanical Engineers.



**Thomas A. Wettergren** (M'98–SM'06) received the B.S. degree in electrical engineering and the Ph.D. degree in applied mathematics from Rensselaer Polytechnic Institute, Troy, NY.

He joined the Naval Undersea Warfare Center, Newport, RI, in 1995, where he has served as a Research Scientist in the Torpedo Systems, Sonar Systems, and Undersea Combat Systems Departments. He currently serves as the U.S. Navy Senior Technologist for Operational and Information Science, as well as a Senior Research Scientist at the Newport laboratory. His personal research interests are in the mathematical modeling, analysis, optimization, and control of undersea sensing systems.

Dr. Wettergren is a member of the Society for Industrial and Applied Mathematics.