A Taxonomy for Congestion Control Algorithms in Packet Switching Networks

The authors propose a new taxonomy for congestion control algorithms in packet switching networks based on control theory. This taxonomy provides a coherent framework for the comparative study of existing algorithms and offers clues toward the development of new congestion control strategies.

Cui-Qing Yang and Alapati V. S. Reddy

Congestion control in packet-switching networks became a high priority in network design and research due to ever-growing network bandwidth and intensive network applications. Dozens of various congestion control strategies have been proposed, and more are forthcoming. This article proposes a new taxonomy of congestion control algorithms in packet-switching computer networks. Based on control-theoretic concepts, we view a congestion control scheme as a control policy to achieve prescribed goals (e.g., round-trip delay, or throughput) in a distributed network environment. Accordingly, a set of criteria for control systems can be used to classify characteristics of various congestion control algorithms. A taxonomy that follows such a theory not only provides a coherent framework for comparative studies of existing approaches, but also helps in future research and development of new strategies for congestion control.

Congestion in a packet-switching network is a state in which performance degrades due to the saturation of network resources such as communication links, processor cycles, and memory buffers. Adverse effects resulting from such congestion include the long delay of message delivery, waste of system resources, and possible *network collapse*, when all communication in the entire network ceases. Network congestion, like traffic jams in big cities, are becoming real threats to the growth of network interconnections and communication applications.

Studies on congestion control, which outline measures for controlling network traffics in order to prevent, avoid, or recover from network congestion, have long been considered significant for the future development of network communications. A large number of various congestion control schemes have been proposed, and a few mechanisms have been implemented in real networks, such as the control methods in IBM's System Networking architecture (SNA) [12], Digital's Networking Architecture (DNA) [7], and the Internet [19, 36]. However, despite years of research efforts, the problem of network congestion control remains a critical issue and a high priority, especially given the prospective of the continually growing speed and size of future networks.

The existing approaches for network congestion control cover a broad range of techniques, including window (buffer) flow control [24], source quench [37], slow start [19], schedule-based control [32], binary feedback [38], rate-based control [4], etc. [51].

It is often difficult to characterize and compare various features among different congestion control schemes. Current literature in the field classifies most congestion control approaches into two categories: approaches for congestion avoidance, and approaches for congestion recovery. Such a simple classification only provides a very general picture of common properties between separating groups of approaches. A detailed taxonomy is required in order to help researchers and engineers understand the similarities and differences among various schemes, and to decide which techniques are best suited for particular designs.

In this article, we propose, a new taxonomy for congestion control algorithms in packet-switching computer networks. We view a network as a large, distributed control system, in which a congestion control scheme is a (distributed) control policy executed at each node (hosts or switches) of the network in order to maintain a certain level of stable conditions. Although such a distributed network control system is too complex to be solvable based on traditional control theories, well-established control-theoretic concepts are qualified candidates for the classification of various congestion control policies. This article shows how a set of criteria for control systems is defined as a taxonomy of congestion control algorithms for

CUI-QING YANG is an associate professor in the computer sciences department at the University of North Texas, Denton.

ALAPATI V. S. REDDY works in the Cellular Digital Packet Data Group at BNR, Richardson.

 0890-8044/95/$04.00 1995 © IEEE

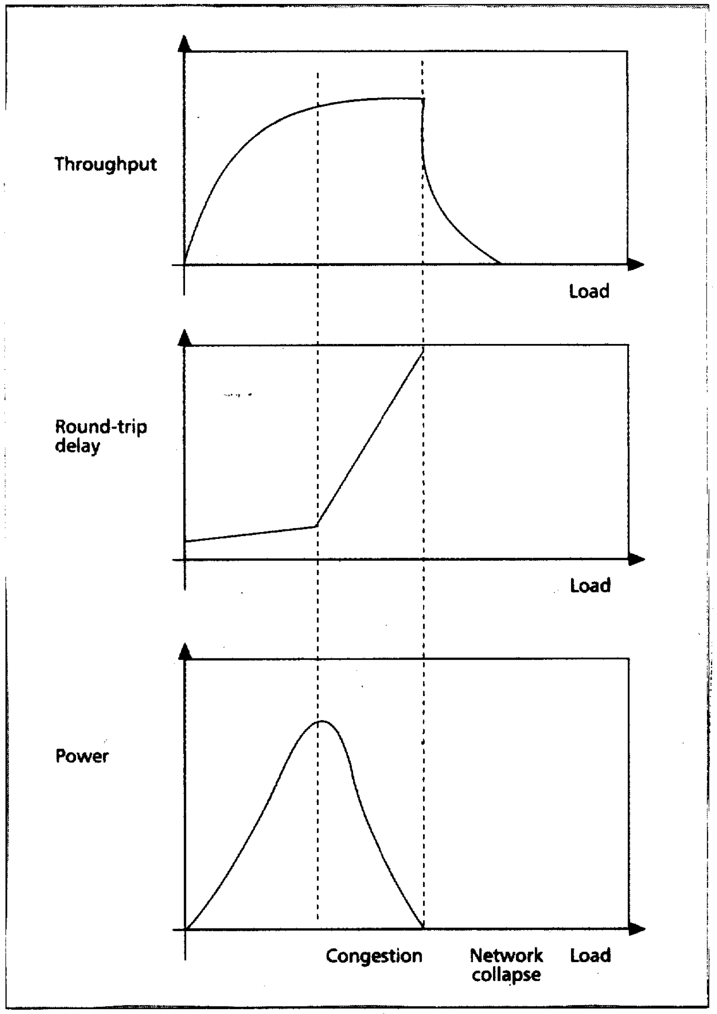
packet switching networks, and how this taxonomy is applied in characterizing individual features of existing congestion control algorithms. We believe that such a taxonomy not only provides a coherent framework for comparative study of existing approaches, but also can help future research in developing new strategies for congestion control.

After a brief discussion of basic concepts of network congestion and some related issues in the second section, the third section presents our new taxonomy of congestion control schemes for computer networks based on the control-theoretic concepts. A classification of most existing approaches in congestion control is then conducted using the new taxonomy in the fourth section. With the growing concern of congestion control in frame-relay and ATM networks, the fifth section classifies some congestion control strategies of frame-relay and ATM networks based on the framework of our new taxonomy. This demonstrates how a taxonomy can help in overviewing, characterizing, and comparing various features of different congestion control algorithms. Conclusions are made in the last section of the article.

Congestion and its Control in Packet-Switching Networks

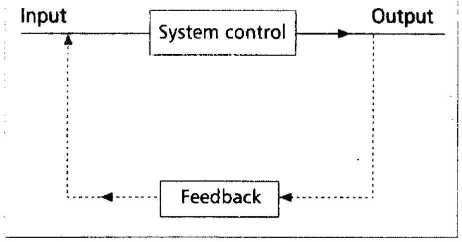
Network congestion has been well recognized as a resource sharing problem. In a packet-switched network, resources are shared among all the hosts attached to it, including switch processors, communication channels, and buffer spaces. These three driving forces of data transmission in network communication can also be potential bottlenecks that cause congestion in the network. On the one hand, networks need to serve all user requests for data transmission, which are often unpredictable and bursty with regard to transmission starting time, rate, and size. On the other hand, any physical resource in the network has a finite capacity, and must be managed for sharing among different transmissions. Consequently, network congestion will result if the resources in the network cannot meet all of the users' current demands.

A more formal and quantitative definition for network congestion is based on the performance behavior of a network. Figure 1a shows the throughput-load relationship in a packet-switching network without effective means of flow control. We see that, as the load is small, network throughput generally keeps up with the increase of the load until the offered load reaches to the *knee* point, where the increase of the throughput becomes much slower than the increase of the load. If the load keeps increasing up to the capacity of the network, the queues on switching nodes will build up, potentially resulting in packets being dropped, and throughput will eventually arrive at its maximum and then decrease sharply to a low value (possibly zero). It is at this point that the network is said to be congested [24]. Figures 1b and 1c illustrate the relationships between the round-trip delay, and the resource power with respect to the offered load. The delay (or response time) curve follows a similar pattern as the throughput curve. At first, the response time rises slowly with the load due to the fast increment of the throughput. Then after the knee point is reached, the delay curve jumps significantly while the throughput stays flat. Finally, the delay grows indefinitely when the network becomes congested. The resource power is defined as the ratio of the throughput to the response time. The resource power gets to its maximum value at the knee point, where the average queue size is close to one, including the packet in service [28].

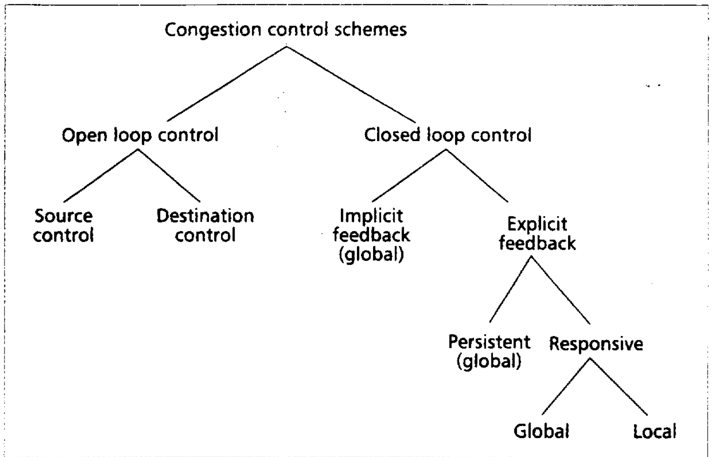


■ Figure 1. *Network performance vs. offered traffic load.*

In order to maintain a network always in a healthy working condition, certain measures or mechanisms have to be provided to prevent the network from operating in the congested region for any significant period of time. Such mechanisms are generally referred to as the congestion control of networks. The congestion control in packet-switching networks may involve different components in a network, including the host machines of sources and destinations, as well as switching nodes. Many congestion control algorithms have been proposed and developed, and may be divided into two categories: congestion avoidance and congestion recovery. The strategy of congestion avoidance is preventive in nature; it is aimed to keep the operation of a network at or near the point of maximum power, so that congestion will never occur. Whereas, the goal of congestion recovery is to restore the operation of a network to its normal state after congestion



■ Figure 2. *Diagram of a general control system.*



■ Figure 3. *Taxomony for congestion control algorithms.*

has occurred. Without a congestion recovery scheme, a network may crash entirely whenever congestion occurs. Therefore, even if a network adopts a strategy of congestion avoidance, congestion recovery schemes would still be required to retain throughput in the case of abrupt changes in a network that may cause congestion.

The problem of congestion control has long been considered an important topic in R & D of computer networks. With the recent development of network technology and the growth of network-intensive applications, the issue of congestion control becomes even more urgent. A great number of congestion control algorithms and strategies have been reported in literature. The earlier efforts of traffic control in communication networks were based on buffer management algorithms for flow control at the link level. The slide-window scheme of flow control was implemented on many networks (APARNET, TYMNET, and DECNET) for controlling the transmission rate of node-to-node or end-to-end [13]. However, the buffer based mechanism of flow control is not effective in preventing congestion to occur when the communication traffic becomes abnormally high at some hot-spots of a network. (There is no method in a flow control to stop the retransmission from a source that might cause a congestion.)

In recent years, a large number of different congestion control algorithms have been proposed and developed, ranging from Random Drop [27], Source Quench [10], Isarithmic scheme [5], Slow Start and Search [19, 46], Virtual Clock [52], Binary Feedback [38], to rate-based congestion control [4], and so on. All these algorithms vary in terms of their operating conditions, functional principles, and performance behaviors. Although a number of survey papers on a variety of congestion control algorithms have appeared in the literature [28, 50, 51], there is still no a systematic way for classification and comparison of so many diverse congestion control algorithms. A framework of taxonomy on congestion control algorithms in packet-switching networks will help people understand the major features of existing algorithms and similarities and differences among various control schemes, and formulate new control algorithms that can be better fit the characteristics of future network traffic.

A New Taxonomy for Congestion Control Algorithms

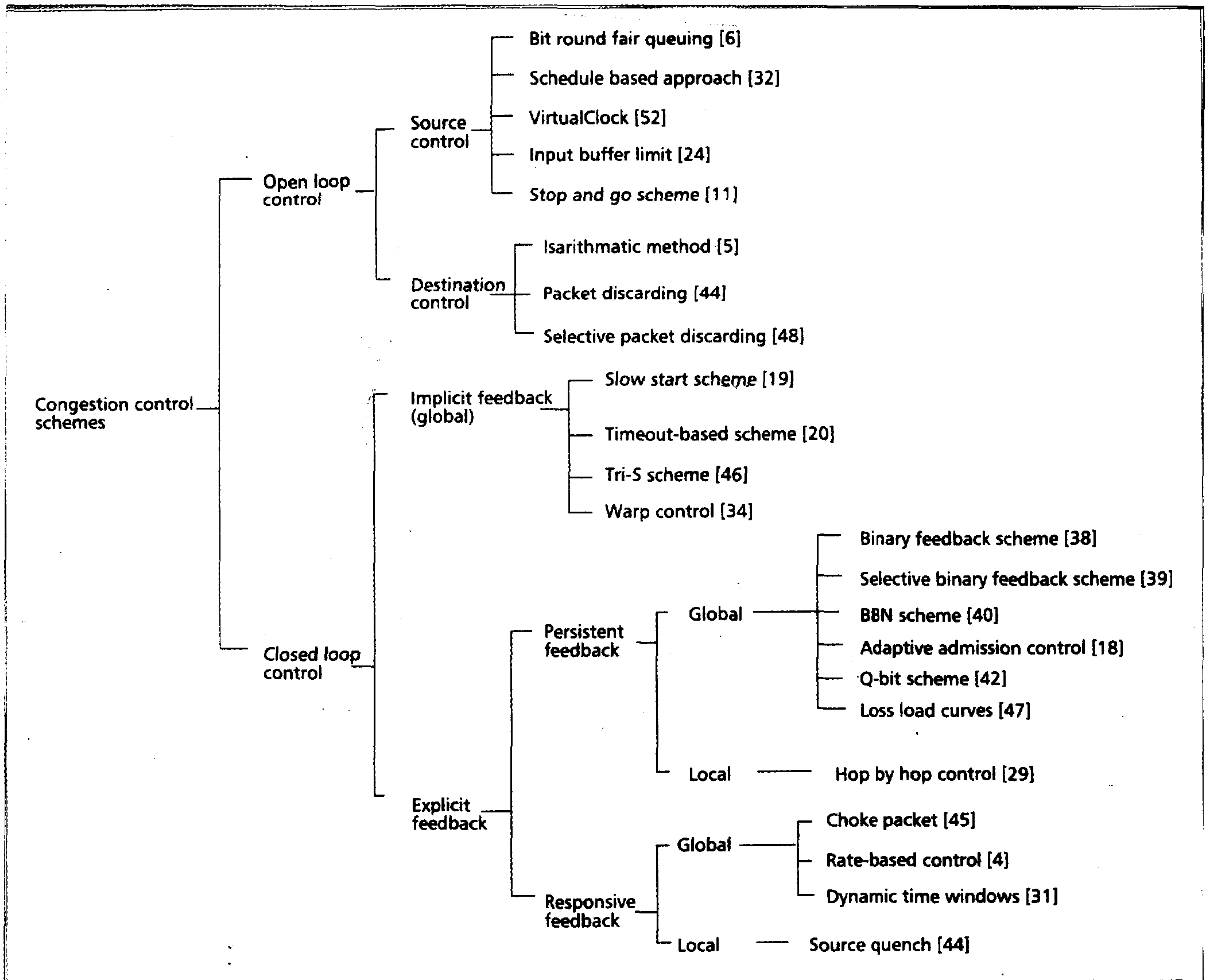
In this section we present a new taxonomy for congestion control algorithms in computer networks that is based on the control theory. For that purpose we will first draw analogies between a control system and a network system.

Computer Networks as Distributed Control Systems

A control system is defined as a collection of objects bonded by some form of interdependence. The objects comprising the system will not remain in a state of equilibrium relative to each other and the surrounding world. Under the influence of external stimuli, the state of the system will be changing with time in a manner which is entirely dependent on the characteristics of the stimuli and the bonds of interaction. It is possible to change the states of a system in a prescribed manner by properly choosing the inputs within some reasonable limits, i.e., one may exert influence on the system states by means of intelligent manipulation of the inputs. In general, the goal of a control system is to achieve dynamic characteristics of a system and to maintain desired system responses regarding various input stimuli. More often, since the desired system response is known, a signal proportional to the error between the desired and actual response is generated and sent back to the input. Utilization of this signal to control the process results in a closed-loop sequence of operations, which is what constitutes a feedback control system. The ability to dynamically adjust and maintain a steady-state performance via transient states is the distinctive advantage of feedback control systems. Figure 2 illustrates the configuration of a general control system.

Control systems have been with us for as long as life itself. Proper functioning of the biological systems clearly requires controls of a more or less complicated nature. A simple example of a manually controlled system is maneuvering an automobile. The vehicle operator, in a closed-loop fashion, continuously exerts control over various outputs of the system, such as velocities and orientations of the car, in a traffic lane. A guided missile that aims by means of a sensor-control device is another example of an automatic control system.

A computer network is an interconnected collection of autonomous computers. We model the computer network system as multiple users' generating jobs in a closed queuing network of servers



■ **Figure 4.** *Classification of existing congestion control algorithms.*

representing network routers. Therefore, we can view the congestion control in a computer network as a control system for the purpose of maintaining the overall traffic within certain normal levels. The entire network can be seen as a big system with inputs and outputs to/from each host machine. The state of the network system, for the purpose of congestion control, can be considered to be comprised of queue lengths at individual servers, such as routers and end-nodes. The feedback signals can be obtained from the differences between the current state of the system and some predefined limits, e.g., thresholds of queuing lengths. The decision function components at individual hosts can adjust sizes of the slide-windows or the rates of the input traffic to achieve a better performance for the entire network.

As with the case of any control system, in a network system the instantaneous state of the network varies dynamically. The purpose of the network control is to produce optimal throughput and overall delay for the communication traffic in the system, which is the ultimate goal in the design of any network congestion control algorithm. In general, the optimization criteria for any congestion control algorithm is to maximize the function power, which is defined as the ratio of the throughput to delay. The utilization of this function in the control of a network results in a closed-loop sequence of operations that justifies the analogy between a network and a control system.

Although we have drawn the analogy between a network and a control system, major differences between a network system and a traditional, centralized control system stem from the complexity and sheer size of the possible control space in a distributed network system. A store-and-forward computer network system consists of a group of geographically distributed, autonomous resources: a large number of communication links (probably composed of different physical media with different delay and throughput characteristics), switching nodes, and host machines connected by the links. The geographically distributed environment introduces unknown varying communication delays, as well as communication unreliability. Due to such communication delay and unreliability, controlling a distributed network system is much more complex than controlling a traditional centralized system. There is not yet (as far as we know) an established control theory that masters the behavior of computer network systems. However, the analogy

Control algorithms	Methods of control	Time of action
Bit-round fair queuing [6]	Switches maintain separated queues for each source. Data is transmitted packet by packet via bit by bit round-robin in order to achieve fairness.	Next packet to send is chosen from the queue with the smallest value of number of rounds at the completion of sending the previous packet.
Schedule based control [32]	Division of channel bandwidth into equal time frames where each frame has a fixed number of slots assigned to individual users.	A packet can be sent using a user's own slot or using a slot whose designated user has no packet to send at the moment.
VirtualClock scheme [52]	A virtual clock is assigned to each dataflow that ticks every time a data packet arrives. The tick step is equal to the mean inter-packet gap of the flow.	The difference between a flow's virtual clock and the real time indicates how far the flow deviates from the specified rate. If beyond some threshold, the rate should be reduced.
Input buffer limit [24]	Differentiate between input and transit traffic at each node and impose a limit on the fraction of buffers for input traffic in a node's buffer pool.	In case of congestion (low numbers of buffer available), input traffic is shut off and transit traffic can take all buffers in the buffer pool.
Stop and go scheme [11]	Using time frame as packet admission policy for each connection at the source node and a multiplexing service discipline enforced at the switching nodes.	Total length of admissible packets in one time frame is limited at the input of each connection, and a packet arriving in one frame at a switch is never transmitted over an output link during the same time frame.

■ Table 1. *Open loop with source control.*

Control algorithms	Methods of control	Time of action
Isarithmetic policy [5]	A finite number of "packet carriers" called permits are assigned to the nodes of network; therefore, the total amount of traffic in the entire network is limited.	A permit must be captured whenever a user's data is ready to enter in the network. Otherwise, the data is discarded.
Packet discarding scheme [44]	The number of buffers at each switch node in the network is fixed. No buffer reservation in advance is allowed.	Packets are simply discarded when no empty buffers are available.
Selective packet discarding policy [48]	The same as the packet discarding scheme.	When no buffers are available, packets are discarded selectively, e.g., first drop packets that entered the network recently.

■ Table 2. *Open loop with destination control.*

between network systems and a control system sheds light on our understanding the nature of controls in a computer network, especially the classification of various algorithms of congestion control.

Taxonomy Based on Control Theory

Based on the above concept of control theory we propose a new taxonomy for the classification of various congestion control algorithms. This proposed taxonomy focuses on the decision-making process of individual congestion control algorithms. The characteristics of how each algorithm extracts information for their control decisions are used as the basis for the classification. The main categories in the taxonomy are:

- Open loop congestion control algorithms.
- Closed loop congestion control algorithms.

Several subcategories exit under each category. Fig. 3 shows the classification tree of the taxonomy.

This section gives an overview of each category classified in the taxonomy. The detailed discussion of individual congestion control algorithms under each subcategory is presented in the next section.

Open Loop Congestion Control Algorithms — Open loop congestion control algorithms are the algorithms in which the control decisions of algorithms do not depend on any sort of feedback information from the congested spots in the network. These algorithms do not monitor the state of the network dynamically. The congestion control algorithm serves as a controller or control actuator purely based on its own knowledge of local node, such as the bandwidth capacity of the local links, and the available buffers in the system. These schemes have a continuous activation feature and have an admission handling mechanism, which has the advantage of stabilizing the traffic arrival process. Generally, these open loop schemes are not robust enough and therefore cannot guard the network against all traffic patterns. They can be further classified as control algorithms which exhibit control at the source or destination machines. The open loop congestion control algorithms at the source tend to control the rate of flow at the sources of traffic, whereas the destination control algorithms intend to control the network traffic either at the destination or some intermediate nodes along the path to a destination.

Closed Loop Congestion Control Algorithms — Closed loop control algorithms make their control decisions based on some sort of feedback information to the sources. This feedback can be either global or

Control algorithms	Methods of control	Time of action	Feedback information
Slow start scheme [19]	Window-based control: window size is slowly increased at the start of some traffic when no symptom of congestion appears; a multiplicative decrease in window size is followed when traffic becomes congestive.	Window size is increased by one (up to a certain limit) every time an ack is received if the round-trip delay is within some threshold. Otherwise, the window size is decreased multiplicatively (say, by half).	Round-trip delay of acknowledgments in the protocol serves as feedback.
Timeout-based scheme [20]	Sources follow a set of self-restraining rules to limit the packets entering into the network. Window-based control with policies specifying a minimum, a maximum, an initialization, an increase, and a decrease size of the sending window.	The window size is increased approximately by one, every round-trip interval up to the maximum. On a time-out, the windowsize resets to the minimum.	Timeout for the acknowledgment at the source serves as the feedback.
Tri-S states scheme [46]	Window-based control attempting to quickly establish an optimal and fair operating point rather than approaching it slowly in the slow start scheme.	Based on the metric of normalized throughput gradient (NTG), the allocation of network resources is adjusted when there are significant traffic changes, e.g., at the beginning or the end of a connection.	Round-trip delay of acknowledgments in the protocol serves as feedback.
Warp control [34]	This abstract model of congestion control is similar to the timeout-based scheme [20]. It uses rate -based admission control at the source of traffic. A time-stamp-based measure, called Warp, is defined to monitor network utilization, and is used to adjust to the new rate of traffic.	Synchronized clocks are assumed at each node in the network. Time stamps of data packets in opposite directions of a pair of source and destination nodes are used to calculate the warp that indicates utilization of the network.	Time delays of data packets are used as implicit feedback information between the source and destination.

■ **Table 3.** *Classic loop control with implicit feedback.*

local: global means the feedback information goes all the way from destination to source, whereas local means the feedback information comes only from immediate neighbors. With the provision of feedback, these algorithms are able to monitor the network performance dynamically. The feedback involved in these algorithms may be implicit or explicit. In the explicit feedback scheme, feedbacks have to be sent explicitly as separate messages (certainly, some of these messages can be *piggybacked*). If there is no necessity of sending the feedback explicitly, the scheme is said to be an implicit feedback scheme. Some examples of such implicit feedbacks are time delays of acknowledgments, and arrival rates of packets from the same machines as the destinations involved in the control scheme. Under an explicit feedback scheme, we further divide into persistent feedback and responsive feedback. Feedback that is available at all times it is called persistent feedback, and feedback that is only trigged under certain conditions (e.g., the traffic exceeds some threshold) is called responsive feedback.

Many explicit closed loop algorithms can be classified as algorithms that are anticipatory or reactive to congestion. They are anticipatory in the sense that they are congestion avoidance schemes, which tend to drive the network toward the optimal operating point but without falling into the danger of congestion. In the anticipatory stage, such algorithms tend to control traffic admission by an admission control policy at the entry. Whereas a reactive strategy is a congestion recovery scheme that responds to conditions of network congestion.

Classification and Comparisons of Congestion Control Algorithms of Packet Switching Networks

In this section, we apply the new taxonomy for congestion control algorithms discussed above to a group of existing congestion control algorithms, which can be found in the literature. The classification of these algorithms demonstrate, on one hand, how the new taxonomy provides a framework for classifying most of the existing algorithms, and on the other hand, how each individual algorithms can fit into various categories of the taxonomy. Figure 4 shows the table of various congestion control algorithms under each category.

Open Loop with Source Control

As mentioned earlier, algorithms in this scheme exert the control on traffic at the source end and use mainly the local knowledge of the network. These algorithms have admission policies that stabilize the traffic arrival process. Algorithms included in this category are: the bit-round fair queuing method [6], the schedule-based control [32], the VirtualClock scheme [52], the input buffer limit model [24], and the stop-and-go policy [11]. The features of each scheme are summarized in Table 1.

Control algorithms	Methods of control	Time of action	Feedback information
Binary feedback scheme [38]	A congestion indication bit is included in each packet which is set to 1 if a switching node detects it is in a congestive state (e.g., queue length greater than one). The destination sends back the congestion bit to the source that adjusts its window size accordingly.	If at least 50 percent of the congestion bits sent back from the destination are set, the current window size is reduced to 87.5 percent of its value.	The state of congestion indication bits in each packet.
Selective binary feedback [39]	A count on the number of packets sent within an average queue interval is kept for each user at a switching node. Only those users who are sending more than their fair share are asked to reduce the load.	The congestion indication bits are set only for users whose demand is higher than their fair share. Therefore, the source window size will be reduced accordingly.	The state of congestion indication bits in each packet.
BBN scheme [40]	A flow ration is maintained for each flow at the source and is calculated based on the feedback information of resource utilization (called resource ration) sent by each individual switching node.	Resource ration is computed at each switching node according to the utilization of each resource measured during a time period. Flow ration is enforced by using a set of "throttlers" that specify the rates at which each flow may submit packets bound for a given destination.	The resource ration computed at each switching node is added to the routing update packets that is propagated through the entire network via flooding.

■ **Figure 4a.** *Closed loop control with persistent, global feedback.*

Congestion algorithms	Methods of control	Time of action	Feedback information
Adaptive admission control [18]	Using the round-trip delay between the source and destination as a congestion control parameter to enforce an adaptive admission rate-control for each flow entering in the [illegible]work.	Upon receiving the feedback information of the virtual delay, a single control parameter β is derived, with β 1 indicating low congestion and β 0 high congestion. The time interval between packet insertion into the network is controlled by β: the larger the β, the shorter the inter-packet gap is, and vice versa.	Special sampling packets are periodically sent from the source to the destination to calculate the virtual delay that is piggybacked to the source.
Q-bit scheme [42]	Besides the congestion bit (C-bit) in the binary feedback scheme, an additional bit called Q-bit is added in each packet, which indicates queuing condition of packets in a switching node. Also, instead of using window-based load control, this scheme uses rate-based load control.	The Q-bit is set whenever a packet is forced to wait in a queue at a switching node. The source increases or decreases the rate of flow entering the network, based on the states of C-bit and Q-bit.	The states of congestion indication bit and the Q-bit in each packet.
Loss load curves [47]	Each switching node in the network monitors local traffic load and provides a loss load curve as feedback to the senders at the sources of traffic. A raw throughput curve can be derived at the source using the loss load curve.	Senders at the entering point of the network have the responsibility to choose their own transmission rate based on the tradeoff between throughput and packet loss, which is obtained from the feedback information of the loss load curve. This mechanism shows less dependence on round-trip delay and the number of hops transversed, and therefore achieves quicker response time than the slow start scheme.	Information of the loss load curve of each traffic is feedbacked to the source.

■ **Figure 4b.** *Closed loop control with persistent, global feedback (continued from Table 4a).*

Control algorithms	Methods of control	Time of action	Feedback information
Hop by hop control [29]	Service rates of individual circuits (connections) or aggregates at each switching node are dynamically adjusted, using feedback information provided by the neighboring switches.	Each switching node monitors and collects buffer occupancy information for each connection per outgoing link, which is sent to each nearest upstream switching node. The service rate of a connection is calculated from feedbacks in terms of the change in the number of waiting packets and the net inflow of packets.	Buffer occupancy information.

Open Loop with Destination Control

Congestion control algorithms under this category perform control operations at the destination end without any knowledge of feedback. They include the isarithmetic control policy [5], packet discarding [44], and selective packet discarding schemes [48]. Comparisons of these algorithms are summarized in Table 2.

Closed Loop Control with Implicit Feedback

Congestion control algorithms under this category realize closed loop control through certain feedback information between destination and source (global). However, feedbacks are not based on any specific messages or explicit actions regarding the traffic conditions in the network. Algorithms under this category include the slow start scheme of Jacobson [19], the timeout-based congestion control scheme [20], and the Tri-S scheme [46]. Table 3 presents features of these algorithms.

Closed Loop Control with Persistent, Global Feedback

Many of the existing congestion control algorithms fall into this category, in which the feedback information regarding the state of the network traffic is constantly (periodically) present between the destination and the source ends. Examples of such algorithms are the binary feedback scheme [38], the adaptive admission congestion control scheme [18], the congestion control algorithm in BBN network [40], the adaptive admission control scheme [14], the Q-bit control scheme [42], and the Loss load curves algorithm [47]. Table 4 lists the major functions of these algorithms.

Closed Loop Control with Persistent, Local Feedback

The difference between algorithms in this category and the previous category is that the feedback information is propagated between immediate neighbors instead of sending it all the way from the destination to the source. There is only one algorithm in this category: the hop by hop control scheme [29] (Table 5).

Closed Loop Control with Responsive, Global Feedback

Congestion control algorithms in this category generate feedback information in response to the traffic conditions in the network, such as when the queue length in a switch raises beyond certain limit. The feedback information is sent between the destination and the source. The source quench [37] or algorithms such as the choke packet scheme [21, 44], rate-based congestion control [4], and the dynamic time windows algorithm [31] are some examples (Table 6).

Closed Loop Control with Responsive, Local Feedback

The single algorithm in this category is the source quench scheme [44], which is a closed loop control scheme with feedbacks only being generated in response to congestive conditions in the network and being sent to the upsteam neighbors of the traffic. The Table 7 gives a brief summary of the algorithm.

Classification and Comparisons of Congestion Control Algorithms of Frame-Relay and ATM Networks

Rapid advances in telecommunication and computer technologies have lead to the emergence and evolution of many new techniques for computer communications and networking. Among them, the ISDN frame-relay and ATM networks are probably two of the most influential developments that will shape up the future of data communication and computer networking. With the growing concern of congestion control in frame-relay and ATM networks, a variety of congestion control strategies have been proposed. Although frame-relay and ATM networks are also packet switching networks, the congestion algorithms for these networks are generally different, due to their specific characteristics, from the algorithms discussed in the previous section. With the increasing number of new congestion control algorithms reported for frame-relay and ATM networks in recent years, a thorough study and comparison of these schemes is beyond the scope of this article. In this section, we apply our taxonomy to some of the congestion control algorithms of frame-relay and ATM networks to show how these algorithm can fit to various categories of the taxonomy.

Congestion Control Schemes of Frame-Relay Network

The motivation behind the ISDN frame-relay service is to fully take advantage of the high speed

Control algorithms	Methods of control	Time of action	Feedback information
Choke packet scheme [45]	A source reduces its traffic sent to a particular destination by some percentage when it receives choke packets. Sources ignore repeating choke packets for a fixed interval of time. If no further choke packets arrive after a certain time, the source will again increase the traffic.	Choke packets are sent to the source when the utilization of an output link at each switching node exceeds above a threshold (say 75 percent), the output link enters a warning state. If a newly arriving packet enters a link in a warning state, a choke packet is sent back to the source.	Choke packets.
Rate-based control [4]	Sources monitor the incoming traffic to each destination. New packets to a destination are rejected if the rate exceeds available capacity. The source adjusts the transmission rate by increasing or decreasing the inter-packet gap for packets to the affected destination when receiving rate control messages.	Rate control messages are generated if the average number of buffered packets (the queue length) exceeds some thresholds at a switching node. The messages are broadcast to all nodes in the network.	Rate control messages reporting the amount of capacity available on the link, the overload factor, and the destination of the offending traffic.
Dynamic time windows [31]	Users reserve a traffic rate based on the time window, a time interval over which the source's average rate is computed. Source throughput is not modified, but instead is maintained by controlling the variance of the traffic rate within the time window.	Each virtual circuit (VC) has a time window associated with it, specified at setup time and varied dynamically for the life of VC. When new users join or existing ones leave the network, feedback information will be sent to the source to adjust the interval sizes of the time windows.	Special fields in data packets signaling the changes in network traffic.

■ **Figure 6.** *Closed loop control with responsive, global feedback.*

and increasing quality of modern digital transmission technologies and therefore to minimize transit delay and maximize throughput. This goal is achieved by simplifying and streamlining the low-level protocol of packet switching. In frame-relay networking, each network node (relay node) only conducts the core procedure of LAPD, i.e., only to check the valid frame-check-sequence (FCS), and address fields of the incoming frames. The frames are either relayed to the destination via their virtual circuit, or simply discarded because of errors or the condition at the switching node. No retransmission and flow control are performed in the network relaying nodes, which only operates between the LAPD end points on an end-to-end basis.

Because of the limited means available to the frame handler in controlling the flow of frames between relaying nodes, special measures have been considered by CCITT and ANSI to provide support for congestion control in the frame-relay networking. Two bits in the address field of each frame are included: one is Backward Explicit Congestion Notification (BECN), and the other is Forward Explicit Congestion Notification (FECN) [43]. The frame relaying nodes has choices to set either of these bits to frames of some logical connections (virtual circuits) when they detect congestion conditions. The end nodes (users) then can take some proper actions in response to these notifications. Based on our new taxonomy, this scheme of congestion control in frame relay can be classified as close-loop control with explicit, responsive, global feedback. There are many other congestion control strategies for frame-relay networks that fall into the categories of open loop control, implicit feedback, and etc. [8, 15]. Table 8 lists the classification of these schemes.

Congestion Control Schemes of ATM Network

Another main development in shaping up the future high-speed (gigabit) networking is the emergence of Broadband ISDN (B-ISDN) and ATM. With its so-called cell switching and the support of virtual path (VP) and virtual circuit (VC), ATM can provide a wide variety of traffic and diverse services, including real-time multimedia (data, voice, and video) applications. Because of its efficiency and flexibility, ATM is considered the most promising transfer technique for the implementation of B-ISDN and for the future of high-speed wide- and local-area networks.

In ATM networks, all data is transmitted in small, fixed-size packets. Due to the high-speed transfer rate (in the range of hundreds to thousands of Mb/s) and rather short cell length (53 bytes), the ratio of propagation delay to cell transmission time and the ratio of processing time to cell transmission time of ATM networks will increase significantly more than that of existing networks. This leads to a shift in the network's performance bottleneck from channel transmission speed (in most existing networks) to propagation delay of the channel and the processing speed at the network switching nodes [1]. Therefore, a major issue in the flow and congestion control of ATM networks is how to handle the conditions of a large number of cells being in transit (or "data in flight" [35]) between two ATM switching nodes. As a result, many congestion

Control algorithms	Methods of control	Time of action	Feedback information
Source quench scheme [44]	When a source receives source quench messages, it will reduce its transmission rate. The rate will be restored gradually without receiving source quench messages for a certain time period.	When the buffer pool at a switching node is filling up, further incoming traffic will be dropped and a source quench message is sent to the upstream nodes for every discarded packet.	Source quench messages.

■ **Table 7.** *Closed loop control with responsive, local feedback.*

Control algorithms	Classification	Methods of control	Feedback information
Window-based control scheme [8]	Open-Loop with source control	The protocol reserves a number of buffers for the entire window size per virtual circuit (determined at call setup time). The call setup control limits the number of virtual circuits created. Implicit priority is given to delay-sensitive short frames.	No feedback required.
Adaptive window scheme [3]	Closed-loop with implicit feedback	The window size is reduced upon the receipt of a REJECT frame from the destination, and the size is increased after *n* successful delivery of frames. Different disciplines in adjusting the window size, including reduce one by one, reduce to w_{min}, and reduce by a factor of αw, lead to different performance outcomes.	REJECT frames from the destination due to the receipt of out-of sequences frames.

■ **Table 8.** *Classification of frame-relay congestion control schemes.*

control algorithms of existing packet switching networks would not work appropriately for ATM networks. With the growing concern of traffic and congestion control on ATM networks, new congestion control schemes targeted to ATM networks have been proposed or reported in recent years [1]. Many of these algorithms fall into the categories of open-loop control or closed-loop control with local feedback to cope with the issue of increased ratio between propagation delay and transmission time. Table 9 classifies some major congestion control schemes of ATM networks based on the taxonomy.

Conclusions

In this article we propose a new taxonomy for congestion control algorithms in packet switching networks based on control theory. This taxonomy provides a coherent framework for comparative study of the existing algorithms and hints at the development of new congestion control strategies. Generally, all open loop schemes have a continuous activation feature and are based on certain admission mechanisms that intend to stabilize the traffic arrival process at the source ends. Due to the lack of global information, open loop schemes are usually not robust enough and therefore cannot guard the network against all possible traffic patterns. However, the open loop schemes can act much faster than the closed loop schemes, specifically in the case of high channel propagation delay found in ATM networks. Meanwhile, it is only possible for closed loop schemes to explicitly distribute indications of resource utilization and traffic conditions throughout the network, and to allow traffic sources to respond rapidly and precisely to the onset of situations that could lead to congestion in the network. Issues of which performance measures should be used as indications for traffic conditions in the network, how this information could promptly be disseminated throughout the network and effectively be used for congestion control remain major challenges in the design and development of future congestion control strategies for computer networks.

Acknowledgment

The authors would like to thank Run-Sheng Yang of ZhongShang University, China, for the discussions and his comments on the earlier version of the article. Thanks are also extended to the anonymous reviewers and the editors who have provided valuable criticisms and suggestions for the improvement of the article.

References

[1] J. Bae and T. Suda, "Survey of Traffic Control Schemes and Protocols in ATM Networks," *Proc. of the IEEE*, Feb. 1991.
[2] A. Berger and A. Eckberg, "A B-ISDN/ATM Traffic Descriptor, and Its Use in Traffic and Congestion Controls," *Proc. IEEE GLOBECOM '91*, 1991, pp. 266-270.
[3] K. Chen and K. Rege, "A Comparative Performance Study of Various Congestion Controls for ISDN Frame-Relay Networks," *Proc. IEEE INFOCOM '89*, April 1989, pp. 674-675.
[4] D. Comer and R. Yavatkar, "A Rate-Based Congestion Avoidance and Control Scheme for Packet Switched Networks," *Proc. of 10th ICDCS*, IEEE, 1990, pp. 390-397.

Control algorithms	Classification	Methods of control	Feedback information
Tagging-based control scheme [9, 17, 26]	Open-loop with destination control	Two kind of flows are admitted at each source: guaranteed and excessive, with the later indicated by a special tag (traffic-violation tagging). If congestion is encountered, the tagged frames can be selectively discarded.	No feedback required.
Service-based control scheme [15]	Open-loop with source control	The network provides two different services, express service and first-class service. The former is for real-time applications and not subject to any traffic control in the network. The source admission control (per VC) limits the maximum reserved bandwidth. The first-class service subjects to choking/relieving in case of congestion conditions.	No feedback required.
Credit-based control scheme [23]	Closed-loop with local feedback	For each VC link, data cells can only be sent after receive of credit cells from the receiver. Therefore, the flow control is managed per VC, link-by-link.	Circulation of credits on the link.
Rate-based control scheme [25, 49]	Closed-loop with global feedback	Window size (or input traffic rate) is adjusted based on the receipt of acknowledgment or explicit congestion indication information from the destination. Various disciplines can be used for the reduction/increment of the window sizes.	End-to-end acknowledgment (implicit) or congestion indication (explicit).
Balanced control scheme [16]	Closed-loop with persistent, global feedback	The source traffic is admitted according to both average and peak rate, which are decided at the link setup time. The rates are adjusted based on the traffic condition of the network, which is monitored through the continuous feedback between the source and destination.	Feedback cells are periodically sent from end-to-end.

■ Figure 9. *Classification of ATM congestion control schemes.*

[5] D. Davies, "The Control of Congestion in Packet Switch Networks," *IEEE Trans. on Commun.*, COM 20 no. 3, June 1972.

[6] A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm," *Proc. Symp. on Communication Architectures and Protocols*, ACM SIGCOMM '89, Sept. 1989.

[7] Digital Equipment Corporation, DECnet Digital Network Architecture [phase IV] General Description, Order AA-N149A-TC, Digital Equipment Corporation, 1982.

[8] B. Doshi and H. Nguyen, "Congestion Control in ISDN Frame-Relay Networks," *ATT Technical Journal*, Nov./Dec. 1988, pp. 35-46.

[9] A. Eckberg, D. Luan, and D. Llucantoni, "Meeting the Challenge: Congestion and Flow Control Strategies for Broadband Information Transport," *Proc. IEEE GLOBECOM '89*, 1989, pp. 1769-1773.

[10] G. Finn, "A Connectionless Congestion Control Algorithm," *Computer Commun. Review*, vol. 19, no. 5, Oct. 1989, pp. 12-31.

[11] S. Gelostani, "Congestion-Free Communications in High-Speed Packet Networks," *IEEE Trans. on Commun.*, vol. 39, no. 12, Dec. 1991.

[12] F. George and G. Young, "SNA Flow Control: Architecture and Implementation," *IBM System Journal*, vol. 21, no. 2, 1982, pp. 179-210.

[13] M. Gerla and L. Kleinrock, "Flow Control: A Comparative Survey," *IEEE Trans. on Commun.*, vol. 28, no. 4, Apr. 1980, pp. 553-574.

[14] M. Gerla and L. Kleinrock, "Congestion Control in Interconnected LANs," *IEEE Network*, vol. 2, no. 1, Jan. 1988.

[15] A. Gersht and K. Lee, "A Congestion Control Framework for ATM Networks," *Proc. IEEE INFOCOM '89*, April 1989, pp. 701-710.

[16] Y. Gong and I. Akyildiz, "Dynamic Traffic Control Using Feedback and Traffic Prediction in ATM Networks," *Proc. IEEE INFOCOM*, 1944, pp. 91-98.

[17] A. Gravey, P. Boyer, and G. Hebuterne, Tagging versus Strict Rate Envorcement in ATM Networks, *Proc. IEEE GLOBECOM '91*, 1991, pp. 271-275.

[18] Z. Haas, "Adaptive Admission Congestion Control," *ACM SIGCOMM '91*, 1991.

[19] V. Jacobson, "Congestion Avoidance and Control," *Proc. of SIGCOMM '88*, 1988.

[20] R. Jain, "A Time-based Congestion Control Scheme for Window Flow-controlled Networks," *IEEE JSAC*, SAC-4, no. 7.

[21] R. Jain and K. Ramakrishnan, "Congestion Avoidance in Computer Networks with a Connectionless Network Layer: Concepts, Goals, and Methodology," *Proc. of Computer Networking Symposium*, April 1988.

[22] S. Keshav, "A Control-Theoretic Approach to Flow Control," *Proc. SIGCOMM '91*, Sept. 1991.

[23] H. Kung, T. Blackwell, and A. Chapman, "Credit-Based Flow Control for ATM Networks: Credit Update Protocol, Adaptive Credit Allocation, and Statistical Multiplexing," *Proc. SIGCOMM '94*, London, UK, Aug. 31 -Sept. 2, 1994, pp. 101-114.

[24] S. Lam and M. Reiser, "Congestion Control of Store-and-Forward Network by Input Buffer Limits Analysis," *IEEE Trans. on Commun.*, COM-27, no. 1, Jan. 1979.

[25] W. Leland, "Window-Based Congestion Management in

Broadband ATM Networks: the Performance of Three Access-Control Policies, Proceedings," *IEEE GLOBECOM '89*, 1989, pp. 1794-1800.
[26] B. Makrucki, "On the Performance of Submitting Excess Traffic to ATM Networks," *Proc. IEEE GLOBECOM '91*, 1991, pp. 281-287.
[27] A. Mankin, "Random Drop Congestion Control," *Proc. of SIGCOMM '90*, 20, no. 4, Sept. 1990.
[28] A. Mankin and K. Ramakrishnan, "Gateway Congestion Control Survey," RFC-1254, Network Working Group, Aug. 1991.
[29] P. Mishra, H. Kanakia, "A Hop by Hop Rate-based Congestion Control Scheme," *Proc. SIGCOMM '92*, Oct. 1992.
[30] A. Mukherjee, L. Landwebber, T. Faber, "Dynamic Time Windows and Generalized Virtual Clock: Combined Closed-loop/Open-Loop Congestion Control," *IEEE INFOCOMM '92*.
[31] A. Mukherjee, L. Landwebber, and T. Faber, "Dynamic Time Windows: Packet Admission Control with Feedback," *Proc. SIGCOMM '92*, Oct. 1992.
[32] U. Mukherji, "A Schedule-Based Approach for Flow-Control in Data Communication Networks," Ph.d. Thesis, Massachusetts Institute of Technology, Feb. 1986.
[33] J. Nagle, "On Packet Switches with Infinite Storage," *IEEE Trans. on Commun.*, vol. 35, no. 4, April 1987, pp. 435-438.
[34] K. Park, "Warp Control: a Dynamically Stable Congestion Protocol and Analysis," *Proc. SIGCOMM '93*, Sept. 1993.
[35] C. Partridge, Gigabit Networking, (Addison-Weseley Publishing Co., 1994).
[36] J. B. Postel, C. A. Sunshine, and D. Cohen, "The ARPA Internet Protocol," *Computer Networks*, vol. 5, no. 4, 1981, pp. 171-261.
[37] W. Prue, and J. Postel, "Something a Host Could Do with Source Quench," RFC-1016, Network Working Group, July 1987.
[38] K. Ramakrishnan and R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks," *ACM Trans. on Computer Systems*, vol. 8, no. 2, May 1990, pp. 158-181.
[39] K. Ramakrishnan, R. Jain, and D. Chiu, "A Selective Binary Feedback Scheme for General Topologies," Networks, Tech. Report, DEC-TR-510, Digital Equipment, Corporation, 1991.
[40] J. Robinson, D. Friedman, and M. Steenstrup, "Congestion Control in BBN Packet-Switched Networks," *ACM Computer Commun. Review*, Jan. 1990, pp. 76-90.
[41] A. Romanow and S. Floyd, "Dynamics of TCP Traffic over ATM Networks," *Proc. SIGCOMM '94*, London, UK, Aug. 31 -Sept. 2, 1994, pp. 79-88.
[42] O. Rose, "The Q-bit Scheme," *ACM Computer Commun. Review*, April 1992, pp. 29-42.
[43] W. Stallings, Networking Standards: a Guide to OSI, ISDN, LAN, and MAN Standards, (Addison-Wesley Publishing Co., Reading, Mass., 1993).
[44] A. Tanenbaum, Computer Networks, (Prentice-Hall, Englewood Cliffs, NJ, 1981).
[45] R. Varakulsiripunth, N. Shiratori, and S. Noguchi, "A Congestion-Control Policy on the Internetwork Gateway," *Computer Networks and ISDN Systems*, 1986, pp. 43-58.
[46] Z. Wang and J. Crowcroft, "A New Congestion Control Scheme: Slow Start and Search [Tri-S]," *Computer Commun. Review*, vol. 21, no. 1, Jan. 1991, pp. 32-43.
[47] C. Williamson, "Optimizing File Transfer Response Time Using the Loss-Load Curve, Congestion Control Mechanism" *Proc. SIGCOMM '93*, Sept. 1993.
[48] N. Yin, S. Li, and T. Stern, "Congestion Control for Packet Voice by Selective Packet Discarding," *IEEE Trans. on Commun.*, vol. 38, no. 5, May 1990, pp. 674-683.
[49] N. Yin and M. Hluchyj, On Closed-Loop Rate Control for ATM Cell Relay Networks, *Proc. IEEE INFOCOM*, 1944, pp. 99-108.
[50] H. Zhang and S. Keshav, "Comparison of Rate-Based Service Disciplines," *Proc. SIGCOMM '91*, Sept. 1991.
[51] L. Zhang, "Congestion Control in Packet-Switched Computer Networks," *Proc. of 2nd Int'l Conf. on Computers and Applications*, 1987, pp. 273-280.
[52] L. Zhang, "VirtualClock: A New Traffic Control Algorithm for Packet Switching Networks" *Proc. SIGCOMM '90*, Sept. 1990, pp. 19-29.

Biographies

CUI-QING YANG received an undergraduate degree in computer engineering from Nanjing Institute of Technology, China, and M.S. and Ph.D. degrees in computer sciences from the University of Wisconsin-Madison. He is an associate professor of Computer Sciences at the University of North Texas, Denton, Texas. His research interests include distributed and parallel systems, operating systems, data communication and computer networks. He is a member of the IEEE Computer Society and ACM SIGCOMM. His e-mail address is: cqyang@cs.unt.edu.

ALAPATI V. S. REDDY currently works at BNR, Richardson, Texas. for the Cellular Digital Packet Data Group (CDPD). He received a Bachelor's of Technology in computer science in 1992 from Regional Engineering College, Warangal, India, and a Master's in computer science in 1995 from the University of North Texas. His research interests include scheduling algorithms for multimedia communications, congestion control, and routing algorithms for computer networks. His e-mail address is: alapati@bnr.ca.

ATM is considered the most promising transfer technique for the implementation of B-ISDN and for the future of high-speed wide- and local-area networks.