Multipacket Reception in Random Access Wireless Networks: From Signal Processing to Optimal Medium Access Control

Lang Tong, Qing Zhao, and Gokhan Mergen, Cornell University

This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and the Army Research Office under Grant ARO-DAAB19-00-1-0507.

ABSTRACT

Recently, there has been considerable interest in the idea of cross-layer design of wireless networks. This is motivated by the need to provide a greater level of adaptivity to variations of wireless channels. This article examines one aspect of the interaction between the physical and medium access control layers. In particular, we consider the impact of signal processing techniques that enable multipacket reception on the throughput and design of random access protocols.

INTRODUCTION

Traditionally, the medium access control (MAC) layer is designed with minimum input from the physical layer and using simple collision models. Most conventional random access protocols assume that the channel is noiseless, and the failure of reception is caused by collisions among users; packets transmitted at the same time are destroyed, and retransmissions must be made later. The basic approach to improving performance has been "resolving" collisions by limiting the transmissions of users. One way is to randomize retransmissions as in Aloha; another is to split successively the set of users until collisions are resolved [1].

The advent of sophisticated signal processing has changed many of the underlying assumptions made by conventional MAC techniques. In code-division multiple access (CDMA), for example, one of the basic premises of multiuser detection [2] is that signals from different users should be estimated jointly, which makes it possible for the node to receive multiple packets simultaneously. The use of antenna arrays also makes it possible to have multipacket receptions.

What are the impacts of these advances at the physical layer on the performance and design of MAC protocols? If there is a high probability that simultaneously transmitted packets can be received correctly, should the MAC "encourage," rather than limit, transmissions of users? We first consider receiver multipacket reception (MPR) capability at the physical layer. Possibilities of obtaining receiver MPR at the modulation level through space-time processing are discussed. Impacts of receiver MPR on the network throughput are considered. The design of MAC protocols that can take advantage of the MPR property of the network is also a topic.

MPR NODES

Users in a wireless network share a common medium, and their transmissions may interfere with one another. An objective of receiver design is to extract, in some optimal way, signals of interest from interference and noise. If a node of the network is capable of correctly receiving signals from multiple transmitters, the node is an MPR node.

DIVERSITIES AND MPR

Receiver MPR depends on the ability of separating signals transmitted simultaneously from different users, and the key to signal separation is exploiting transmission and reception diversity. At the modulation level, signals can be separated by designing waveforms judiciously. A well known example is CDMA, where each user transmits with a specific signature waveform: the spreading code. Temporal and spectral diversities of spread spectrum signals can be exploited by the receiver for signal separation.

Spatial diversity is an additional avenue for signal separation. Employing transmitting and receiving antenna arrays with properly designed space-time codes can significantly increase the rate of reliable communications, allowing the separation of multiple users transmitting at the same time, using the same modulation and access protocol [3].

Geographical locations of users provide network diversities in signal propagation. This has recently been exploited for signal-processing-based collision resolution [4]. Also related is spread Aloha [5] where, by subsampling the output of the matched filter, a receiver can separate signals from users transmitting asynchronously or with different propagation delays.

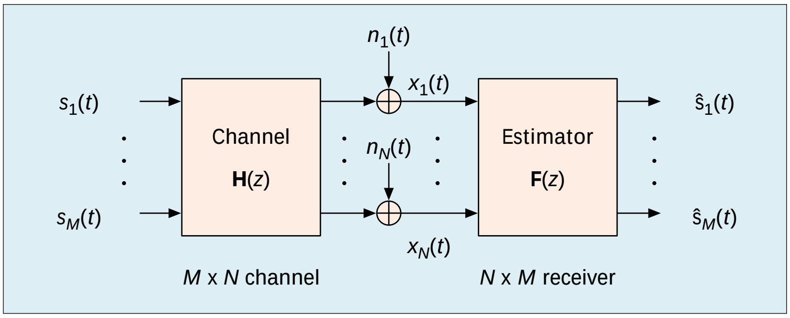

Figure 1. *A general model for multiuser communications and receiver MPR.*

Signal Processing for MPR

A general model of a multiuser system that includes spatial, temporal, and code diversities is the multi-input and multi-output (MIMO) channel shown in Fig. 1. Here $s_i(t)$ are transmitted signals from M users, and $x_i(t)$ are received signals from antenna array elements or virtual receivers of temporal processing [6, vol. 1, ch. 8]. The channel impulse response $\mathbf{H}(z)$ depends on the form of modulation, the transmission protocol, and the configuration of transceiver antenna arrays.

The basic signal separation problem is to design an estimator such that multiple sources are extracted in some optimal fashion. Although optimal estimators are nonlinear in general, to reduce implementation cost, one is often restricted to an MIMO linear filter with finite impulse response $\mathbf{F}(z)$.The design of $\mathbf{F}(z)$depends on knowledge of the channel $\mathbf{H}(z)$ and the format of transmission.

It is unrealistic to assume that the receiver knows the channel response $\mathbf{H}(z)$ in a wireless mobile network. It is then necessary to "train" the receiver by introducing pilot or training symbols in the data stream. Knowing the training symbols from user i, a linear estimator for that user can be designed based on, for example, the least squares criterion. The least squares optimization, when there is a sufficient amount of training, can be implemented adaptively, offering the ability to track one or a group of users. Furthermore, the receiver only needs to know the training symbols from node i in order to design the optimal receiver for that node, and there is no requirement of synchronization among users. The performance of the receiver, however, does depend on the presence of other nodes and the level of interference. As users drop in and out of the network and the channel varies with time, training needs to be done repeatedly.

The use of training significantly simplifies the problem of receiver design for MPR. However, it has several practical and theoretical drawbacks. For example, the training symbols and their locations in a data packet must be known to the receiver. This may be possible for networks with scheduled transmissions but may not be practical for random access networks. The overhead associated with training may also be too excessive. It is therefore desirable to develop self-adaptive algorithms that are able to track users without relying on training.

There has been considerable research in blind and semi-blind signal separation in recent years (see the survey of papers in [7]). Without a sufficient number of training symbols, the key to signal separation is to utilize the structure of the channel and characteristics of the input sources. For example, communication signals often have the constant modulus property, which enables the separation of multiple sources by minimizing the signal dispersion using the constant modulus algorithm (CMA) [6, 8, 9]. The finite alphabet property of communication signals may also be exploited for signal separation [6]. The statistical dependency among sources is another condition that leads to a number of effective source separation algorithms [7]. In a transmitter-oriented CDMA system, code information can be exploited for signal separation [6]. The diversity of the propagation channel from each transmitter to the receiver provides yet another possibility for packet separation. In [4], simultaneously transmitted packets are separated according to the duration of their channel impulse responses. For related topics in this growing field of research, readers are referred to [6].

Networks with MPR Nodes

To characterize the performance of a network with MPR nodes, MPR needs to be modeled at the node level. For slotted networks, Ghez, Verdu, and Schwartz provided a convenient model [10] where MPR capability of a node is modeled by the MPR matrix

$$\mathbf{R} = \begin{pmatrix} R_{1,0} & R_{1,1} & & \\ R_{2,0} & R_{2,1} & R_{2,2} & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad (1)$$

where $R_{n,k}$ is the conditional probability that k packets are correctly received given that n packets are transmitted. The weakest MPR is the conventional collision channel $\mathbf{R}_0$, and the strongest MPR is $\mathbf{R}_1$ that models perfect packet separation:

$$\mathbf{R}_0 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}, \mathbf{R}_1 = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In between, the MPR channel matrix can take various forms as a function of the channel conditions and signal separation algorithms.

Local Throughput

Local throughput is measured by the number of successfully received packets averaged over time

A more flexible approach is needed for MPR channels because the protocol should allow the optimal number of users to transmit. This implies that the set of users to access the channel should be enlarged if there were not enough users holding packets in the previous slot and shrunk if too many users attempted to transmit.

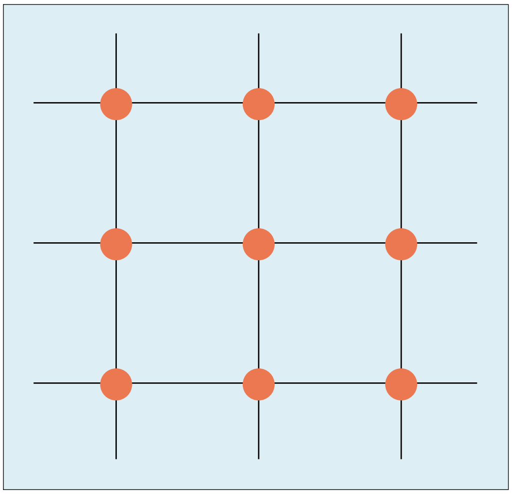

Figure 2. *The Manhattan Network.*

at network nodes. This of course does not accurately reflect the network throughput since the packets successfully received at a node of the network are not necessarily those that successfully reach their final destinations. On the other hand, local throughput often directly affects the network throughput.

For a network with slotted transmissions, conditional on the number i of transmitting neighbors, the local throughput of a node with MPR matrix **R** is given by

$$\eta_i \triangleq \sum_{j=0}^{i} jR_{i,j},$$

Hence, the local throughput is upper bounded by

$$\eta_i \triangleq \sum_{j=0}^{i} jR_{i,j}, \tag{2}$$

$$\eta = \sup_i ni. \tag{3}$$

If the network has a finite number of nodes and is fully connected, the maximum throughput given in Eq. 3 can indeed be achieved in heavy traffic by random access protocols outlined later (for details see [10]). For a fully connected network with an infinite number of users, under certain conditions on the MPR matrix **R**, the maximum stable throughput for an Aloha-like protocol was shown [11] to be

$$\eta_{\text{Aloha}} = \sup_{\lambda \geq 0} e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^i}{i!} \eta i. \tag{4}$$

Intuitively, $_{\text{Aloha}}$ is the average number of successfully received packets maximized over the rate of Poisson arrivals.

End-to-End Throughput

For cellular systems, there is a one-to-one correspondence between receiver MPR at the base station and network MPR because all traffic goes through the base station, and all packets received by the base station are intended for it. This, however, is not true in general for ad hoc networks. Even if a node successfully receives multiple packets, some of these packets may not be intended for that node. To evaluate network throughput, one must convert the receiver MPR to the network MPR [4].

Unfortunately, the MPR model at the node level cannot accurately describe multihop ad hoc networks. Issues beyond MAC (e.g., routing) must be considered in the throughput evaluation. However, insights can be gained by examining networks with regular structures as in [12]. An informative example is the rectangular grid, the so-called Manhattan Network, shown in Fig. 2, where each node has four neighbors. Although the receiver MPR at the node level is well defined, the network MPR cannot be defined by a single matrix.

For a network of size N, it can be shown [13] that the maximum achievable end-to-end throughput is given by

$$\eta = \max_{i=1,\ldots,4} \frac{2\eta i}{i+1} \sqrt{N} \tag{5}$$

with the maximum gain of 60 percent over the throughput (÷N) of networks without MPR. For slotted Aloha without MPR, using a similar analysis as in [12], it can be shown that the maximum throughput is 0.16 ÷n, which is 1/6 of the maximum throughput of a non-MPR network.

MAC Protocols for MPR

MPR offers the potential of improving network performance. At the same time, it presents several new challenges. The outcome of a particular slot in the conventional collision channel can be a success, collision, or no transmission. In contrast, there is a higher level of uncertainty (hence a greater amount of information) associated with the outcome of a particular slot for networks with MPR. Specifically, the successful reception of a packet at an MPR node does not imply that only one neighbor transmitted. To improve the network throughput, we are no longer restricted to splitting users in order to resolve collisions. We outline next two approaches that exploit the MPR capability of cellular networks.

An Optimal MAC for MPR Channels

The key to maximizing throughput is to grant an appropriate subset of users access to the MPR channel. For the conventional collision channel, this can be accomplished by splitting users in the event of collision. A more flexible approach is necessary for MPR channels because the protocol should allow the optimal number of users to transmit. This implies that the set of users to access the channel should be enlarged if there were not enough users holding packets in the previous slot and shrunk if too many users attempted to transmit. Ideally, N_o — the number that maximizes $_i$ — users should be allowed to transmit in order to achieve the maximum throughput. Unfortunately, this is not always possible because the number of users holding packets is a random variable not known to the receiver. One should extract information from the joint distribution of the states of all users.

The Multi-Queue Service Room (MQSR)

protocol [10] is designed explicitly for general MPR channels. The protocol is designed to accommodate groups of users with different delay requirements. Here we consider the case when there is only one group of users with the same delay requirement. As shown in Fig. 3, users are queued, waiting to enter a service room where transmissions are allowed. The division of users into those inside and those outside the service room allows decomposition of the joint distribution of the user states so that this joint distribution can be updated effectively. To allow the flexibility to enlarge and shrink the set of users accessing the channel, the service room is divided into the access and waiting rooms. Only users in the access room are allowed to transmit. If there are too many users in the access room, the last users entering the access room are pushed back into the waiting room. If there are too few users in the access room, on the other hand, users in the waiting room, and users outside the service room if necessary, are allowed to enter the access room. The design of the optimal number of users entering the access room is based on the maximization of the network throughput for each slot.

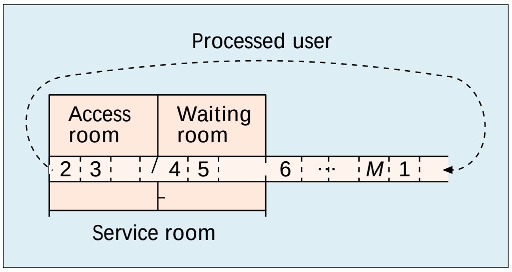

Figure 3. *The basic structure of the service room protocol.*

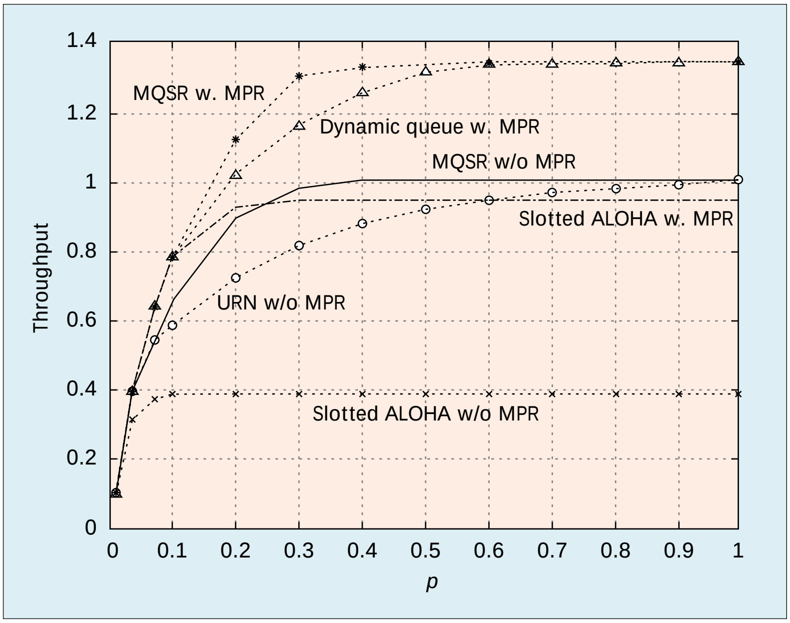

Figure 4. *Throughput comparison. Ten users are present, each with probability.*

A Suboptimal MAC for MPR Channels

The service room protocol optimally exploits all available information for the efficient utilization of MPR channels. The drawback is its high computational cost due to updates of the joint distribution of all users' states. Analogous to the dynamic tree protocol [14] for the conventional channel and much simpler than the service room protocol, the dynamic queue protocol [15] has the property that it approaches the maximum throughput in heavy traffic and has low latency in light traffic. Different from the tree protocol, however, is the way it determines the set of users gaining access to the channel.

The basic idea is to partition the time axis into transmission periods (TPs) where the ith TP is dedicated to the transmission of the packets generated in the $(i - 1)$th TP. Based on the probability that a user has a packet and the MPR matrix, the protocol determines the optimal access set for this TP by minimizing the length of the transmission period. It turns out that this minimization can be carried out by computing the absorbing time of a finite state discrete time Markov chain. Furthermore, the optimal size of the contention class can be computed offline, and the protocol can be implemented by a simple lookup table.

Signal Processing vs. MAC

To gain insights into the roles of signal processing and MAC, as an illustration, we compare first the performance of the optimal protocol with those of the URN and slotted Aloha protocols for a fully connected network with 10 users and a central controller. This example intentionally favors the two conventional protocols since the URN protocol [16] assumes the knowledge of the number of nodes with packets, and the Aloha used in the comparison was implemented using the optimal retransmission probability. In Fig. 4 we see that, for the conventional collision channel without MPR, three protocols behave similarly when the traffic is light. As the traffic load increases, the throughput of the optimal protocol quickly reaches the maximum achievable throughput of 1, whereas the slotted Aloha remains at around 0.4. The URN protocol has the same performance in both light and heavy traffic but lags in the mid-range of the traffic load. Note that the gain of throughput from around 0.4 to 1 in Fig. 4 is due to the optimal MAC protocol without MPR. If the receiver MPR is introduced using, in this example, the signal-processing-based collision resolution technique in [4], another 30 percent gain can be achieved by the optimal protocol. This gain comes from the receiver MPR. The throughput of the Aloha protocol with MPR is twice that of the conventional collision channel.

Conclusions

In this article we have considered potential impacts of receiver MPR at the physical layer on the performance and design of MAC protocols. Cross-layer design is a methodology that requires further investigation, and issues involved are broad and deep. Is it simple enough to implement? Does it scale? Is it robust? A critical element in cross-layer design is choosing an appropriate set of parameters that serve as

A critical element in cross-layer design is choosing an appropriate set of parameters that serve as agents carrying information between layers, parameters that are simple, but not too simple, so that the network can be designed to be adaptive to channel variations, but at the same time, robust to modeling errors.

agents carrying information between layers, parameters that are simple, but not too simple, so that the network can be designed to be adaptive to channel variations, but at the same time robust to modeling errors.

REFERENCES

[1] D. P. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, 1992.

[2] S. Verdu, *Multiuser Detection*, Cambridge Univ. Press, 1998.

[3] A. Paulraj and B. Ng, "Space-time Modems for Wireless Personal Communications," *IEEE Pers. Commun.*, vol. 5, no. 1, Feb 1998, pp. 56–48.

[4] Q. Zhao and L. Tong, "Semi-blind Collision Resolution in Random Access Ad Hoc Wireless Networks," *IEEE Trans. Sig. Proc.*, vol. 48, no. 9, Sept. 2000.

[5] N. Abramson, "Multiple Access in Wireless Digital Networks," *Proc. IEEE*, vol. 82, no. 9, Sept. 1994, pp. 1360–70.

[6] G. Giannakis *et al.*, *Signal Processing Advances in Wireless Communications: Vol. I Trends in Channel Estimation and Equalization, Vol. II Trends in Single- and Multi-User Systems*, Englewood Cliffs, NJ: PTR Prentice-Hall, 2001.

[7] R. Liu and L. Tong Ed., "Special Issue on Blind System Identification and Estimation," *IEEE Proc.*, vol. 86, no. 10, Oct. 1998.

[8] J. R. Treichler and B. G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. ASSP*, vol. ASSP-31, no. 2, Apr. 1983, pp. 459–72.

[9] A. van der Veen, "Algebraic Methods for Deterministic Blind Beamforming," *IEEE Proc.*, vol. 86, no. 10, Oct. 1998, pp. 1987–2008.

[10] Q. Zhao and L. Tong, " A Multi-Queue Service Room MAC Protocol for Wireless Network with Multipacket Reception," submitted to *IEEE/ACM Trans. Net.*, Jan. 2001; http://www.ee.cornell.edu/~ltong/

[11] S. Ghez, S. Verdu, and S. Schwartz, "Optimal Decentralized Control in the Random Access Multipacket Channel," *IEEE Trans. Auto. Control*, vol. 34, no. 11, Nov. 1989, pp. 1153–63.

[12] J. A. Silvester and L. Kleinrock, "On the Capacity of Multihop Slotted ALOHA Networks with Regular Structure," *IEEE Trans. Commun.*, vol. 31, no. 8, Aug.–Mar. 1983, pp. 974–82.

[13] G. Mergen and L. Tong, "Multipacket Reception in Multihop Ad-Hoc Networks," to appear, *Allerton Conf. Commun., Control, and Comp.*, 2001.

[14] J. I. Capetanakis, "Generalized TDMA: The Multi-Accessing Tree Protocol," *IEEE Trans. Commun.,* vol. 27, no. 10, Oct. 1979, pp.1476–84.

[15] Q. Zhao and L. Tong, " A Dynamic Queue MAC Protocol for Random Access Channels with Multipacket Reception," submitted to *IEEE/ACM Trans. Net.*, June 2001; http://www.ee.cornell.edu/~ltong

[16] L. Kleinrock and Y. Yemini, "An Optimal Adaptive Scheme for Multiple Access Broadcast Communication," *Proc. Int'l. Conf. Commun.*, June 1978, pp. 7.2.1–7.2.5.

ADDITIONAL READING

[1] M. K. Tsatsanis, R. Zhang, and S. Banerjee, "Network Assisted Diversity for Random Access Wireless Networks," *IEEE Trans. Sig. Proc.*, SP-48, no. 3, Mar. 2000, pp. 702–11.

BIOGRAPHIES

LANG TONG [S'87,M'91] (ltong@ee.cornell.edu) received a B.E. degree from Tsinghua University, Beijing, China, in 1985, and M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Indiana. He was a postdoctoral research affiliate at the Information Systems Laboratory, Stanford University in 1991. Currently, he is an associate professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He received a Young Investigator Award from the Office of Naval Research in 1996, and the Outstanding Young Author Award from the IEEE Circuits and Systems Society. His areas of interest include statistical signal processing, adaptive receiver design for communication systems, signal processing for communication networks, and information theory.

QING ZHAO is currently a Ph.D. student at the School of Electrical and Computer Engineering, Cornell University. She received a B.S. degree in electrical engineering in 1994 from Sichuan University, Chengdu, China, and an M.S. degree in 1997 from Fudan University, Shanghai, China. Her research interests lie in the intersection of signal processing, wireless communications, and communication networks. She has been working on adaptive receiver design for wireless communication systems, signal processing for communication networks, and optimal medium access protocol design for multiaccess communication networks. She received the IEEE Signal Processing Society 2000 Young Author Best Paper Award.

GOKHAN MERGEN is currently a Ph.D. student at the School of Electrical and Computer Engineering, Cornell University. He received B.S. and B.A degrees in electrical and electronics engineering and mathematics, respectively, in 2000 from Middle East Technical University, Turkey. His research interests lie in the areas of statistical signal processing, wireless communications, and communication networks.