MIMO Routing with QoS Provisioning

Athanasios Gkelias[1], Federico Boccardi[2], Chi Harold Liu[1], Kin K. Leung[1]

[1] *Imperial College London,* [2] *Bell-Labs, Alcatel-Lucent*
Email: a.gkelias@imperial.ac.uk

Abstract— Multiple-Input-Multiple-Output (MIMO) transceivers can take advantage of random channel fading and multi-path delay spread for multiplying transfer rates (multiplexing gain), improve transmission quality/reliability (diversity gain) at no cost of extra spectrum, or allow the spatial separation of the signals of different transceivers in a multi-user scenario, thus providing a multiple-access gain.

While multiple antenna techniques have been widely analyzed for single link transmission scenarios or from the medium access control (MAC) perspective, their usage and impact on network layer and more specifically their interaction with routing has not received much research attention.

In this paper we investigate the interaction between multiple antenna techniques and routing. A novel routing scheme that exploits the multi-user gain capability of MIMO by controlling the number of ongoing transmission in order to increase the end-to-end throughput and guarantee the required quality-of-service (QoS) constraints is presented and analyzed.

Index Terms— Ad Hoc, QoS, MIMO, Routing

I. INTRODUCTION

Uncoordinated wireless multi-hop networks such as ad-hoc, mesh and sensor networks gained increased research interest over the last decade. This is because they are easy to deploy, robust to node failures and promising to increase the overall network capacity. On the other hand most of such networks and especially wireless mesh networks are expected to serve an increased amount of data traffic with various Quality-of-Service (QoS) constraints. However, the lack of central control and the multi-hop, interference limited nature of the network makes the QoS provisioning an extremely difficult and challenging task. Therefore, it is of paramount importance the introduction of some new technology enablers, such as multiple-input multiple-output (MIMO) antennas and novel, optimized cross-layer algorithms.

MIMO links have been proven to provide high spectral efficiency in rich multi-path environments through multiple spatial channels without additional bandwidth requirements. This increase in spectral efficiency is obtained for a single link with no external interference. In a wireless mesh environment though, there will be channel reuse and therefore co-channel interference from other adjacent transmissions. Recent research results have shown that co-channel interference can seriously degrade the overall capacity when MIMO channels are used in a cellular system [1].

Moreover, it has been proven that for flat Rayleigh fading channels, with independent fading coefficients for each path, it is possible to achieve higher capacity by reducing the number of MIMO streams. More specifically, for a system with N receive antennas, M transmit streams and K interfering streams, all the M transmit streams can be isolated and decoded successfully as long as $M + K \geq N$. A group of M antenna elements will be used for data reception while the remaining $N - M$ elements are used to null out the interfering streams. The best performance is achieved when all the degrees of freedom of the MIMO channel are used, i.e., $M = N - K$.

However, if the incoming streams are more than the receiver antenna elements (i.e., $M + K < N$), it may be not be possible for the receiver to decode any of the desired signal streams if the excess streams degrade the overall SINR below a threshold. It must be noted here that if the interfering (K) streams are far weaker than the desired (M) streams, it may be still possible to decode the desired streams (even if $M + K < N$) given that the SINR is above the required threshold.

A lot of work has been performed on MIMO techniques; however, most of it is concentrated on single link transmissions without interference while there are some limited references on multiple access techniques with MIMO channels [1]-[2] (for a detailed review on multiple access schemes with multiple antennas see [3]). Moreover, the research community interest over the last decades regarding wireless routing has been only concentrated on omnidirectional transmissions. To the best of our knowledge, the interaction between multiples antennas and QoS routing has not been adequately investigated. Some of the most well known QoS routing protocols for ad hoc networks include QoS Ad hoc On-demand Distance Vector (AODV) [4], Bandwidth Routing (BR) [5], On-demand QoS Routing (OQR) [6], On-demand Link-state Multi-path QoS Routing (OLMQR) [7]. However, none of them considers the effect of MIMO antennas and more specifically the impact of stream control and interference cancelation techniques on the Network Layer.

Our contribution: In this framework we present a novel QoS routing algorithm that exploits some main MIMO properties such as multiplexing gain and interference cancelation, but also the benefits of a space-division-multiple-access (SDMA) scheme. In this way, we are able not only to increase the overall network throughput but also to optimally distribute the data traffic over the network topology in a way that we can guarantee the required long-term QoS throughput to the underlying applications. This becomes feasible by controlling the available degrees of freedom for each wireless MIMO transceiver so that it has enough resources to cancel out the interference generated by adjacent transmissions.

The rest of the paper is organized as follows. The main con-

cepts of the proposed cross-layer framework are demonstrated in Section II and the corresponding subsections. The performance analysis of the routing algorithm and some numerical results are presented in Section III. Finally conclusions are drawn in Section IV.

II. ALGORITHM DESCRIPTION

The main challenge of this framework is that the routing algorithm needs long term knowledge of the channel throughput in order to guarantee the QoS requirements during the whole lifetime of a data session. However the capacity of a link is highly affected by the interference generated from adjacent links transmissions. This interference dynamically changes over the time as a function of number of antenna elements used by the interfering transmitters, the network data traffic and the chosen routes in the network.

Since predicting the interference and (hence the channel throughput) is a difficult task, in the proposed scheme we try to control the interference. A traditional approach would be to separate transmission in space and time such that no more than one transmitters operate around each receiver at the same time. However, recent works have shown that the actual channel capacity in a multiple access scenario can be achieved if multiple users are allowed to transmit simultaneously rather than using the channel resources in an orthogonal way [8]. More specifically in the case of MIMO transceivers multiple transmissions can take place in independent streams given that the number of data streams plus interfering streams are less or equal to degrees of freedom at the receiver end.

Therefore our cross-layer scheme needs to allow simultaneous transmissions take place while at the same time be able to predict and control the interference before it is actually generated. In this way the routing algorithm will be able to establish end-to-end routes that can guarantee the required QoS throughput throughout the whole duration of the underlying data sessions. In order to achieve this the proposed scheme goes through the following steps:
1. Differentiate strong from weak interfering streams.
2. Control the number of strong interfering streams while treat the weak ones as gaussian noise.
3. Estimate the expected link throughput without considering any interference.
4. Include a resource reservation margin ξ to account for the allowing interference.
5. Use a routing algorithm with a novel utility function to guarantee the end-to-end QoS throughput requirements for the whole duration of each data session.

These functionalities are described in more details in the following:

A. Stream Control

Each wireless transceiver classifies its neighbor nodes into three groups based on the average received signal-to-noise-ratio (SNR): Data nodes (i.e., nodes that are within transmission range), strong interfering nodes and weak interfering nodes. Let L_i be the set of all nodes able to generate strong

interference to node i, and let D_i be the set of all nodes that are within transmission range of node i. It is clear that $D_i \subseteq L_i$.

Our scheme controls the number of streams allowed for a given transmission from node i to node j by making sure that: first, the receiver j has enough available degrees of freedom for reception of node i's streams, and second, all the remaining surrounding receivers have enough degrees of freedom to cancel out the interfering streams generated by the transmitter i. In other words, let us denote as $N_{0,i}$ and $M_{0,j}$ the number of overall antenna elements at the transmitter i and receiver j respectively. Moreover, $N_{t,i}(t)$ and $M_{r,j}(t)$, $t = 1, 2, ..., T$ (where T is the number of slots per MAC frame) represents the number of antenna elements used in slot t for ongoing transmissions and receptions at nodes i and j respectively. Finally, N_{max} represents the maximum number of streams that are allowed to simultaneously operate in a given area. The number of interfering streams that node i can handle in slot t is calculated as

$$W_i(t) = N_{max} - M_{r,i}(t) - \sum_j N_{t,j}(t), \forall j \in L_i. \quad (1)$$

The maximum interference that a node n is allowed to generated is limited by the $\min_i \{W_i\}, \forall i \in L_n$ of its surrounding nodes. Therefore the maximum number of antenna elements that node n can use for transmission is given by:

$$N_n(t) = min\{N_{0,n} - N_{t,n}(t), \min_i\{W_i(t)\}\}, \forall i \in L_n. \quad (2)$$

Similarly, at the receiver side, the maximum number of antenna elements that node m can use for reception is given by:

$$M_m(t) = min\{M_{0,m}(t) - M_{r,m}(t), W_m(t)\}. \quad (3)$$

All the interference coming from nodes that do not belong to the set L_m will be treated by node m as Gaussian noise.

B. Capacity Estimation

Let us model the channel between the m-th antenna, $m = 1, \ldots, M$, of the i-th node and the n-th antenna, $n = 1, \ldots, N$, of the j-th node as follows:

$$h_{n,m}^{(j,i)} = \beta_{n,m}^{(j,i)} \sqrt{\left[d^{(j,i)}/d_0\right]^\delta \rho} \quad (4)$$

where $\beta_{n,m}^{(j,i)}$ is the Rayleigh fading, $d^{(j,i)}$ is the distance between the i-th and j-th node, d_0 is the reference distance, δ is the pathloss coefficient and ρ is the reference SNR defined as the SNR measured at the reference distance assuming a single transmit antenna transmitting at full power, accounting only for the pathloss. Note the distance between the M antennas of the i-th node and the N antennas of the j-th node is assumed to be the same.

The received signal at the jth node can be written as follows

$$\mathbf{y}_j = \mathbf{H}^{(j,i)}\mathbf{x}_i + \sum_{k=1}^{K_i} \mathbf{H}^{(j,k)}\mathbf{x}_k + \mathbf{n}_j \quad (5)$$

where $\left[\mathbf{H}^{(j,i)}\right]_{n,m} = h_{n,m}^{(j,i)}$, $\mathbf{x}_i$ is the transmitted signal vector at the ith node, $\sum_{k=1}^{N_i} \mathbf{H}^{(j,k)}\mathbf{x}_k$ is the term due to the K_i

interfering links and $\mathbf{n}_j$ is the term due to the thermal noise. We assume that the channel state information (CSI) is not available at the transmit side, whereas at the receiver the CSI is perfectly known. Future extensions will assume transmit CSI available by means of channel reciprocity between transmitter and receiver.

During the link throughput estimation phase we assume that each transmitter estimates the ergodic open-loop capacity [9] without considering the interference from other links. The open-loop capacity of the (i, j) link (without considering the interference term) is given by

$$C = \log_2 \det\left(\mathbf{I}_N + \frac{1}{M}\mathbf{H}^{(j,i)}\mathbf{H}^{(j,i)^H}\right) \qquad (6)$$

We use the following lower bound for the ergodic open-loop capacity [10]

$$E\{C(N, M)\} \geq \alpha \log_2\left(1 + \frac{\rho}{M}\exp\left(\frac{1}{\alpha}\sum_{l=1}^{\alpha}\sum_{p=1}^{\beta-l}\frac{1}{p} - \gamma\right)\right) \qquad (7)$$

where $\alpha = \min(M, N)$, $\beta = \max(M, N)$ and $\gamma \approx 0.57721566$ is Euler's constant. Therefore, by using (2) and (3) the estimated link $\ell_{n,m}$ throughput can be expressed as $E\{C(N_n(t), M_m(t))\}$.

In order to account for the weak interfering streams a resource reservation margin $\xi \geq 1$ is introduced. Our scheme will reserve resources such that for each link $\ell_{n,m}$

$$\sum_{t=1}^{T} E\{C(N_n(t), M_m(t))\} = \xi T_{QoS}. \qquad (8)$$

Capacity Estimation Signaling: The channel estimation is based on pilot symbols. Each node has to broadcast periodically (every T_p) a pilot symbol. All the surrounding nodes must listen for that pilot symbol for the channel estimation. Each pilot symbol transmission should not interfere with other pilot transmissions and moreover, each receiver should know the ID of the node that transmits the pilot symbol. This implies that all nodes within two-hop area must have dedicated and unique slots for their pilot symbol transmissions.

In order to achieve this, we introduce a synchronization phase that will be repeated periodically (every T_s, where $T_s >> T_p$) that the nodes will reserve a unique slot for their pilot symbol transmission. All their neighbors will be aware of the slots and the ID of the transmitting node. In that contention phase, each node transmits a request packet with its ID then backs off and waits for nodes replies. The reply can be only negative, i.e., only if any of the neighbor nodes is unable to resolve the ID replies with a negative acknowledgment packet (this can be just a busy tone since more than one nodes may send a negative acknowledgment) that implies that a collision took place. If there is no reply from the surrounding nodes, the node assumes that it is the unique transmitter and reserves the pilot slot corresponds to the synchronization slot.

This can be used as discovery phase in ad hoc networks where nodes are listening for the IDs of new nodes. If a node does not receive any ID packet it can assume that it is

disconnected. Nevertheless it has to continue to periodically send its request packet that may be heard by any new node. All the one-hop neighbors are synchronized to transmit their pilot symbols in different slots. In this case the interference of the adjacent nodes is not included.

C. Actual Throughput

Assuming now that controlled transmissions take place and that minimum square error estimation with successive interference cancelation (MMSE-SIC) receivers are used, each MIMO receiver should be able to achieve data rates that are not limited by the interference [8]. Using the procedure proposed in [11], the actual mutual information for a MIMO link ℓ_{ij} with interference known at the receiver side is expressed as

$$I = \log_2 \det\left(\mathbf{I}_N + \frac{1}{M}\mathbf{H}^{(j,i)}\mathbf{H}^{(j,i)^H}\mathbf{R}_j^{-1}\right) \qquad (9)$$

where $\mathbf{R}_j$ represents the whitening matrix

$$\mathbf{R}_j = \sum_{k=1}^{N_i}\mathbf{H}^{(j,k)}\mathbf{H}^{(j,k)^H} + \mathbf{I}_N \qquad (10)$$

A QoS data session is assumed to be successful if $I \geq T_{QoS}$ throughput the route and for the whole duration of the session.

D. Routing Algorithm Description

The proposed algorithm is based on the min-max concept. In other words, it tries to find the route with the maximum throughput bottleneck from source to sink. The bottleneck must satisfy the throughput QoS requirements. Unlike the single antenna case or the full MIMO case where all the antenna elements are used for transmission/reception, the achievable link throughput and bottleneck now depend on the number of streams used for transmission/reception and interference cancellation. The routing algorithm can be summarized in the following steps:

1. The source s initiates the route discovery by broadcasting a packet with the QoS throughput requirements and the available antenna elements $N_s(t)$, $t = 1, 2, ..., T$ for transmission and the destination node. Note here that a node may not target a specific destination node but any possible gateway.

2. *Throughput Estimation*: All nodes $j \in D_s$ perform capacity estimation based on the available transmitting $N_s(t)$ and receiving $M_j(t)$ streams in each time slot t and estimate the achievable link throughput as $T_{sj} = \sum_{t=1}^{T} E\{C(N_s(t), M_j(t))\}$. If the QoS requirements are satisfied in link ℓ_{sj}, node j broadcasts again the packet to all its neighbors $k \in D_j$. The same procedure is repeated until a packet reaches the destination node or expires. Note here that in the QoS requirements the resource reservation margin is included.

Moreover, receivers have to check if new bottlenecks in the route occur. Generally, if T_i^k is the bottleneck of route k up to the node i and link ℓ_{ij} is the next hop in the route, the receiver j has to update the route k bottleneck as $T_j^k = min\{T_i^k, T_{ij}\}$. Finally, in order to avoid loops the packet contains an ID-list of all the nodes it has traversed in the route to that point. The receiver will discard the packet if its own ID exists in this list.

3. *Resource Reservation*: Before a node forwards a route discovery packet it performs temporal reservation of streams per time slot. Let $NT_i^k = [N_i(1), N_i(2), ..., N_i(T)]$ and $MR_j^k = [M_j(1), M_j(2), ..., M_j(T)]$ be the temporal resource reservation vectors for transmission and reception respectively in the link ℓ_{ij}. Both nodes i and j have to broadcast this information to all the nodes belong to the L_i and L_j sets respectively. Note that this reservation is temporal and will be used only in the calculations of the following hops of the specific route k. All the nodes belong to the L_i and L_j sets will recalculate their available resources based on the equations (2) and (3).

4. The final destination node (e.g., a gateway), upon reception of a request packet, performs the last link capacity estimation to check if the QoS requirements are satisfied. If they do, it waits until the expiration time of the packet, in case other request packets arrive from the same source through different routes. Then it chooses the one with the greater bottleneck link and replies backwards through this route with the serial number of the route.

5. Each node i throughout the chosen route uses the NT_i^k and MR_i^k vectors to temporarily reserve the required resources and broadcasts this information to all the $j \in L_i$ neighbors.

III. NUMERICAL RESULTS

In order to assess the performance of the proposed algorithm a slotted, event-driven simulation platform has been developed. A number of wireless MIMO transceiver and gateways are randomly and independently deployed on a rectangular two-dimensional space. Each MIMO transceiver can independently generate poisson distributed data sessions of constant length, certain throughput QoS requirement and randomly choose any of the gateways. Upon arrival of a new data session the source node initiates a route discovery phase to find the optimum route to its assigned gateway according to the proposed algorithm. We assume here that the route discovery phase of different sources do not overlap in time. After the end-to-end route has been established, nodes constantly check the link throughput of each hop. If the throughput QoS requirement is not satisfied in any of the links throughout the route, the session is dropped and all the resources reserved by this route are released.

In our simulation the data transmission range for each receiver i has been set equal to the strong interference range (i.e., $L_i = D_i$) and includes all the nodes with SNR $\geq 10dB$ from the given receiver. All generated sessions have the same QoS throughput requirement of $T_{QoS} = 1Mbps$. Finally, the maximum number of hops from source to destination has been set to 5, the MAC frame has a duration of 100ms that comprises 40 slots ($T = 40$) and the channel bandwidth is 20MHz.

Figure 1 depicts the performance of proposed framework as the number of successful QoS data sessions versus the offered traffic for different values of the resource reservation margin ξ. The antenna elements per MIMO transceiver and the maximum number of simultaneous streams have been set to $N_0 = 6$ and $N_{max} = 6$ respectively. It can be observed
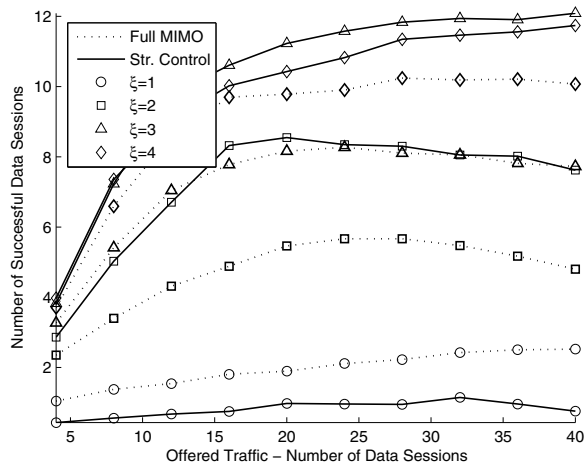

Fig. 1. Number of successful QoS data sessions versus offered traffic for different values of the resource reservation margin ξ, with and without stream control ($N_0 = N_{max} = 6$).

that our routing algorithm with end-to-end stream control (SC) can provide 20% higher throughput as compared to the case where the MIMO transceivers are using all their antenna elements for transmission/reception (referred to as "full MIMO"). Moreover, the maximum network throughput is achieved for resource reservation margin $\xi = 3$ for the SC case while for "full MIMO" transmission a $\xi = 4$ is required. For both cases the network throughput saturates for more than 30 incoming sessions (for the given scenario).
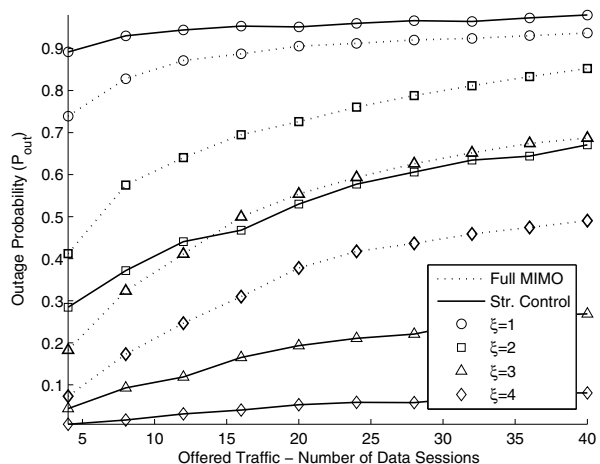

Fig. 2. Outage probability (P_{out}) of admitted QoS data sessions versus offered traffic for different values of the resource reservation margin ξ, with and without stream control ($N_0 = N_{max} = 6$).

In figure 2, the outage probability of the admitted data sessions is demonstrated as a function of the offered traffic and for different values of ξ. Graph shows again that the proposed algorithm outperform the "full MIMO" by being able to provide less than 10% outage even for high traffic conditions. It is interesting here to observe though that the minimum outage probability is obtained for $\xi = 4$ as compared
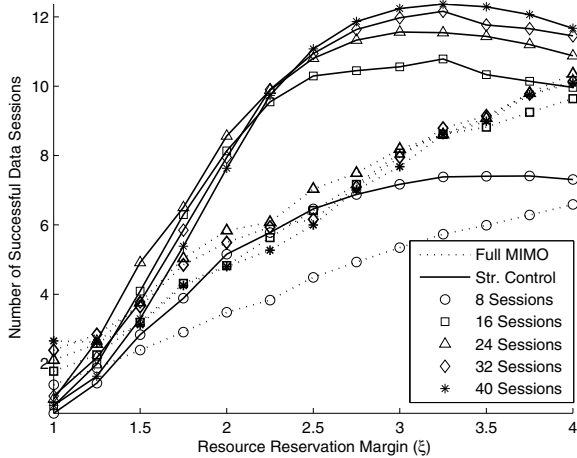
to $\xi = 3$ for maximum network throughput.

Fig. 3. Number of successful QoS data sessions versus the resource reservation margin ξ for different values of offered traffic (number of incoming sessions), with and without stream control ($N_0 = N_{max} = 6$).

Figure 3 demonstrates the dependance of network throughput on the resource reservation margin ξ for different values of incoming data traffic. The optimum value for the specific scenario it appears to be $\xi = 3.25$ for the SC case. However, the network throughput for the "full MIMO" case increases linearly as more resources are allocated to each session. It can be observed again that the network throughput saturates when the incoming traffic exceeds the 30 data sessions.
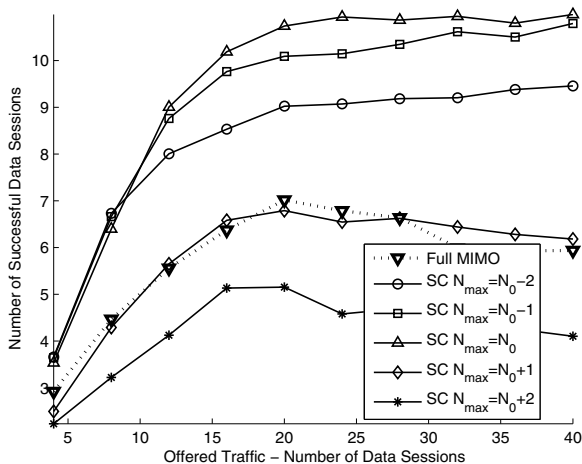

Fig. 4. Number of successful QoS data sessions versus offered traffic for different values of the maximum number of allowed streams per receiver (N_{max}), with and without stream control ($\xi = 2.5$).

In figure 4, we demonstrate the effect of maximum allowed number of simultaneous streams around each MIMO receiver (N_{max}) on the network throughput. The resource reservation margin has been set to $\xi = 2.5$. We can observe that the network throughput is maximized when the number of data streams plus interfering streams is equal to the number of receiving streams (receiver's antenna elements), i.e., $N_{max} =$

N_0. Moreover, it can be seen that $N_{max} < N_0$ gives much better performance than $N_{max} > N_0$. However, this may change if the we modify the L_i set (for instance by reducing or increasing the threshold between weak and strong interferers). How this threshold affects the overall network performance and how to calibrate and optimize such a network will be part of our future work.

IV. CONCLUSIONS

A novel QoS routing algorithm has been presented that exploits the multiplexing gain and interference cancelation properties of MIMO antennas. The proposed algorithms performs end-to-end stream control for each route such that more than one MIMO transceivers can operate in the same area at the same time, while each MIMO receiver has enough streams to cancel out the interference generated by any adjacent transmission. Simulation results showed that the proposed cross-layer framework can guarantee long-term QoS while at the same time significantly increases the overall network throughput as compared to the conventional orthogonal in space and time coordination.

REFERENCES

[1] M. F. Damirkol, M. A. Ingram, "Stream Control in Networks with Interfering MIMO Links," in *Proc. of WCNC 2003*, vol 1, 16-20 March 2003, pp. 343 - 348.
[2] K. Sundaresan, R. Sivakumar, M.A. Ingram, T.-Y. Chang, "Medium Access Control in Ad Hoc Networks with MIMO Links: Optimization Considerations and Algorithms," *IEEE Trans. on Mobile Computing*, pp. 350-365,Vol. 3, No. 4, Oct.-Dec. 2004.
[3] A. Gkelias and K. K. Leung, "Multiple Antenna Techniques for Wireless Mesh Networks," *Wireless Mesh Networks: Architectures, Protocols, and Applications,* Spinger Science, 2007.
[4] C.E. Perkins, E.M. Royers, S.R. Das, "Quality of service for ad hoc on-demand distance vector routing (work in progress)," IETF Internet Draft, draft-ietf-manet-aodvqos-00.txt, July 2000.
[5] C.R. Lin, J. Liu, "QoS routing in ad hoc wireless networks," *IEEE J. Sel. Areas Commun.*, vol 17, no. 8, Aug. 1999, pp. 1426-1438.
[6] C.R. Lin, "On-demand QoS routing in multihop mobile networks," in *Proc. of IEEE INFOCOM 2001*, vol 3, April 2001, pp. 1735-1744.
[7] Y. Chen, Y. Tseng, J. Sheu, P. Kuo, "On-demand, link-state, multi-path QoS routing in a wireless mobile ad-hoc network," in *Proc. of European Wireless 2002*, Feb. 2002, pp. 135-141.
[8] D. Tse and P. Viswanath, *"Fundamentals of Wireless Communications,"* Cambridge University Press 2005.
[9] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications.* Cambridge University Press, 2003.
[10] Ö. Oyman, R. Nabar, H. Bölcskei and A. J. Paulraj, "Tight lower bounds on the ergodic capacity of Rayleigh fading MIMO channels," in *IEEE Global Communication Conference (GLOBECOM)*, (Taipei, Taiwan), Nov. 2002.
[11] R. S. Blum, "MIMO Capacity with Interference," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 793–801, June 2003.