Multi-user Diversity for IEEE 802.11 Infrastructure Wireless LAN

Sung Won Kim

School of Electrical Engineering and Computer Science, Yeungnam University,
Gyeongsangbuk-do, 712-749, Korea
ksw@ieee.org

Abstract. To realize high data rate wireless communication systems, much attention is being payed to multi-user diversity due to large bandwidth availability. Multi-user diversity based opportunistic scheduling is a modern view communication over fading wireless channels, whereby, unlike rate adaptation based schemes, channel variations are exploited rather than mitigated. This paper proposes a multi-user diversity scheme for IEEE 802.11 infrastructure wireless LAN to enhance the throughput. Numerical investigations show the throughput superiority of the scheme over IEEE 802.11 standard and other method.

1 Introduction

The transmission medium used by wireless data networks is inherently time-varying due to e.g. multipath propagation, user mobility, and non-stationary clutter. Also, the wireless resource is scarce and expensive, requiring optimized usage to maximize the throughput (spectral efficiency). Achieving overall throughput maximization requires scheduler to momentarily postpone scheduling packets to a node with poor link quality until its link hits near its peak. Opportunistic scheduling, used to extract multi-user diversity gain, was first proposed in [1] and then extended to many wireless communication systems [2][3]. An opportunistic scheduling algorithm that exploits the inherent multi-user diversity has been implemented as the standard algorithm in the third-generation cellular system IS-856 [4] (also known as high data rate, HDR). To enable the opportunistic multi-user communications, timely channel information of each link is required for an effective scheduling. Feedback of predicted link quality of each active node that is required in opportunistic scheduling is usually integrated into wireless systems. Usually, each receiver measures the received signal-to-noise ratio (SNR) on the channel and then feeds it back to the transmitter.

When it comes down to wireless local area networks (WLANs), it is difficult to utilize the multi-user diversity. The access point (AP) cannot track the channel fluctuations of each link because of the single shared medium and the distributed Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) Medium Access Control (MAC) protocol. Wang et al. [5] presented the opportunistic packet scheduling method for WLANs. The key mechanism of the method is

E. Sha et al. (Eds.): EUC 2006, LNCS 4096, pp. 214–223, 2006.

the use of multicast RTS (Request-To-Send) and priority-based CTS (Clear-To-Send) to probe the channel status information. Since their method requires the modification of RTS and CTS in the standard, the scheme cannot be directly applied into widely deployed IEEE 802.11 typed WLANs.

On the other hand, this form of the multi-user wireless system produces asymmetric traffic loads where most of the traffic loads converge into APs. For example, Internet access or mobile computing uses transmission control protocol (TCP) or user datagram protocol (UDP) in which the offered traffic load is strongly biased toward the downlink (from AP to nodes) against the uplink (from nodes to AP) or the direct link (from nodes to nodes). Thus, these traffic flows for the downlink are completely blocked due to the CSMA/CA MAC protocol in distributed environments.

To alleviate this downlink bottleneck problem, some resource allocation algorithms between the uplink and the downlink are proposed in [6]–[8]. In [6], the authors observe a significant unfairness between the uplink and the downlink flows when DCF is employed in a WLAN. The reason is that in a WLAN with N nodes there are N uplink CSMA/CA instances contending with only one downlink CSMA/CA instance. Thus, when the downlink has much more offered traffic load than that of the uplink, the downlink becomes bottleneck of the system capacity and much more APs should be deployed to accommodate such nodes. The TCP fairness issue between the uplink and the downlink in WLANs has been studied in [7]. The authors are interested in a solution that results in uplink and downlink TCP flows having an equal share of the wireless bandwidth. Because this solution operates on the TCP layer, it is not effective when there exist traffic flows other than TCP. In [8][9], we proposed FAIR that is a dynamic resource allocation method between the uplink and the downlink. FAIR estimates the utilization ratio between the uplink and the downlink to determine the AP access method. FAIR does not consider the throughput improvement by using the multi-user diversity. Moreover, the parameter estimation method proposed in FAIR is not stable.

To mitigate the bottleneck problem in the downlink and to increase the throughput in WLANs, we propose a MAC protocol that exploits the multi-user diversity. The remainder of this paper is organized as follows. The next section presents system model. Section 3 describes the proposed method. In Section 4, we investigate the enhancement of the proposed method with some numerical results. Finally, the paper is concluded in Section 5.

2 System Model

2.1 Infrastructure WLAN

MAC protocol in the IEEE 802.11 standard [10] consists of two coordination functions: mandatory Distributed Coordination Function (DCF) and optional Point Coordination Function (PCF). In DCF, a set of wireless nodes communicates with each other using a contention-based channel access method, CSMA/CA. CSMA/CA is known for its inherent fairness between nodes and robustness. It is quite effective in supporting symmetric traffic loads in ad hoc networks where the traffic loads between nodes are similar.

DCF achieves automatic medium sharing between compatible nodes through the use of CSMA/CA. Before initiating a transmission, a node senses the channel to determine whether or not another node is transmitting. If the medium is sensed idle for a specified time interval, called the distributed interframe space (DIFS), the node is allowed to transmit. If the medium is sensed busy, the transmission is deferred until the ongoing transmission terminates.

If two or more nodes find that the channel is idle at the same time, a collision occurs. In order to reduce the probability of such collisions, a node has to perform a backoff procedure before starting a transmission. The duration of this backoff is determined by the Contention Window (CW) size which is initially set to CW_{min}. The CW value is used to randomly choose the number of slot times in the range of [0, $CW-1$], which is used for backoff duration. In case of an unsuccessful transmission, the CW value is updated to $CW \times 2$ while it does not exceed CW_{max}. This will guarantee that in case of a collision, the probability of another collision at the time of next transmission attempt is further decreased.

A transmitter and receiver pair exchanges short RTS and CTS control packets prior to the actual data transmission to avoid the collision of data packets. An acknowledgement (ACK) packet will be sent by the receiver upon successful reception of a data packet. It is only after receiving an ACK packet correctly that the transmitter assumes successful delivery of the corresponding data packet. Short InterFrame Space (SIFS), which is smaller than DIFS, is a time interval between RTS, CTS, data packet, and ACK. Using this small gap between transmissions within the packet exchange sequence prevents other nodes from attempting to use the medium. As a consequence, it gives priority to completion of the ongoing packet exchange sequence.

Fig. 1 illustrates the system model of an infrastructure WLAN. The AP plays the important role for relaying the traffic between the mobile nodes (wireless

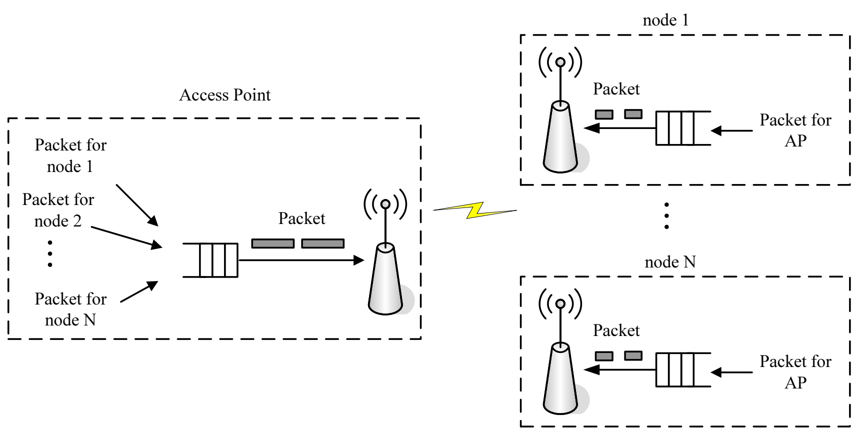

Fig. 1. System model for an infrastructure WLAN

stations) and the wired network which results in asymmetric traffic load between the AP and single mobile node in the infrastructure WLANs. Although PCF is designed for infrastructure networks, the problem is that currently most of the WLAN cards do not support the PCF mode. With DCF mode, the CSMA/CA mechanism makes the AP and mobile nodes have the same priority to access the medium. This leads to the significant unfair WLAN bandwidth distribution between uplink and downlink flows.

2.2 Rate Adaptation

The auto-rate fallback (ARF) protocol for IEEE 802.11 has been presented in [11]. Specifically, if the ACKs for two consecutive data packets are not received by the sender, the sender reduces the transmission rate to the next lower data rate and starts a timer. When the timer expires or ten consecutive ACKs are received, the transmission rate is raised to the next higher data rate and the timer is canceled. However, if an ACK is not received for the immediately next data packet, the rate is lowered again and the timer is restarted. The ARF protocol is simple and easy to incorporate into the IEEE 802.11. However, as pointed out in [12], it is purely heuristic and cannot react quickly when the wireless channel conditions fluctuate.

In the above algorithm, the rate adaptation is performed at the sender. However, it is the receiver that can perceive the channel quality, and thus determine the transmission rate more precisely. Observing this, the authors in [13] have presented a receiver-based auto-rate (RBAR) protocol assuming that the RTS/CTS mechanism is there. The basic idea of RBAR is as follows. First, the receiver estimates the wireless channel quality using a sample of the SNR of the received RTS, then selects an appropriate transmission rate for the data packet, and piggybacks the chosen rate in the responding CTS packet. Then, the sender transmits the data packet at the rate advertised by the CTS. The simulation results in [13] show that the RBAR protocol can adapt to the channel conditions more quickly and in a more precise manner than does the ARF protocol, and thus it improves the performance greatly.

3 Proposed MAC Protocol

3.1 Downlink Channel Access

Each node can directly communicate only with the AP (uplink or downlink), since we focus on AP-coordinated infrastructure WLANs. The AP manages the downlink packet queues for each node as shown in Fig. 1. We propose that the AP determines the downlink channel access method according to the operation mode, that is *normal mode* and *opportunistic mode*. In normal mode, nodes and AP use the DCF mechanism with RTS/CTS handshaking, where each node should wait for DIFS and backoff window time after previous ACK packet.

Let N be the number of active nodes except AP. Then the probability that the successful packet transmission is performed by node n is given as

$$P_n = \frac{1}{N+1}, \quad \text{for } n = 1, 2, ..N. \tag{1}$$

The same probability applies to the AP. Let Γ be the maximum available system throughput. Then, the system throughput allocated to the downlink, Γ_d, and the uplink, Γ_u, are given as

$$\Gamma_d = \Gamma \times P_n = \Gamma \frac{1}{N+1}, \tag{2}$$

$$\Gamma_u = \Gamma \times \sum_{n=1}^{N} P_n = \Gamma \frac{N}{N+1}, \tag{3}$$

where the packet size is assumed to be the same. The ratio between the uplink throughput and the downlink throughput is given as

$$\frac{\Gamma_d}{\Gamma_u} = \left(\frac{\Gamma}{N+1}\right) \Big/ \left(\frac{\Gamma N}{N+1}\right) = \frac{1}{N}. \tag{4}$$

Thus, in DCF, the allocated downlink throughput decreases as the number of nodes increases because the system throughput is shared equally between nodes. This method is not efficient when the traffic load is asymmetric between the uplink and the downlink such as TCP and UDP. Even in the case of symmetric traffic load, the downlink traffic in DCF gets less throughput than that of the uplink and this causes the increased delay of the downlink traffic. To solve this problem, the opportunistic mode is used in the AP.

In the opportunistic mode, the AP waits only for SIFS interval instead of DIFS and backoff interval. By shorting the interval period, the AP can access the channel without collision because all other nodes should wait at least DIFS period which is longer than SIFS period. By using the opportunistic mode, more throughput can be allocated to the downlink.

For the change of the operation mode, the AP has counters for the uplink and the downlink, denoted by ST_u and ST_d, respectively. The counter values increase by one whenever there is a successful packet transmission in the uplink and the downlink, respectively. When $ST_d \geq ST_u$, which means the accumulated number of the successful downlink packet transmission is equal to or larger than that of the uplink, the operation mode of the AP is set to the normal mode. On the contrary, when $ST_d < ST_u$, the operation mode of the AP is changed to the opportunistic mode to allocate more throughput to the downlink. The two counters, ST_u and ST_d, also run in the opportunistic mode and the operation mode will be changed to the normal mode as soon as it becomes $ST_d \geq ST_u$. The mode change algorithm is illustrated in Fig. 2.

3.2 Packet Scheduling Algorithm

In the normal mode, the packet scheduling algorithm adopts the first-in first-out (FIFO) algorithm. In the opportunistic mode, the AP schedules the packet based on the channel quality. The link with better channel quality is given higher

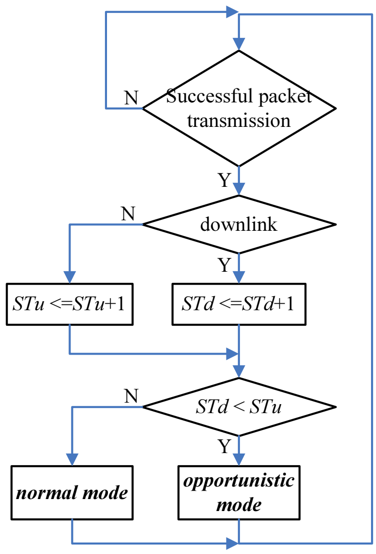

Fig. 2. Mode change algorithm

priority in packet scheduling. In order to track the latest channel quality, it is necessary to send the control packet to the node. However, this method will increase the overhead and need the modification of the IEEE 802.11 standard. Our design goal is that the scheduling method can be implemented without the modification of the nodes already deployed in the system. Thus, we propose that the AP updates the channel quality of each link after every successful packet transmissions. The channel quality is reported from the physical layer of AP by measuring the SNR of the received packets, e.g. CTS and ACK packets for the downlink traffic. This estimation of the channel quality may not be the timely information. However, the estimation error is in the acceptable range as will be shown in the next section. Moreover, the proposed method can be implemented without the modification of the deployed nodes.

The AP lists all the communication links according to the estimated channel quality. When the AP is in the opportunistic mode, the link that recorded the best channel quality in the previous successful packet transmission is given the first chance to transmit the packet in the queue. When there is no packet in the queue for that link, the next communication link in the list is given the second chance to transmit the packet.

One of the problems in the previous opportunistic scheduling method is the unfairness between the nodes [5]. The node that has the better channel quality gets more throughput and this may lead to the starvation problems for other nodes. However, in our method, opportunistic scheduling is compromised with FIFO scheduling and this alleviates the unfairness problem.

4 Numerical Results

We evaluate the performance of the proposed method by computer simulations. The IEEE 802.11 DCF and FAIR in [8] are compared with the proposed method. The parameter values used to obtain numerical results of the simulation runs are based on the IEEE 802.11b direct sequence spread spectrum (DSSS) standard [10]. To reflect the fact that the surrounding environmental clutter may be significantly different for each pair of communication nodes with the same distance separation, we use the log-normal shadowing channel model [14].

We assume that all nodes except the AP are randomly distributed in the circle area with diameter 150 meters and move randomly at speed 0.1 m/sec. The AP is located at the center of the area. To evaluate the maximum performance, traffic load is saturated in each nodes and the destination addresses of the packets are the AP. In the AP, there are N connections, each for one node, and packets are generated for each connections with the same distrubution as those in each nodes. To make an asymmetric traffic load condition between uplink and downlink, the size of the downlink and uplink packets are 1024 and 64 bytes, respectively. The number of node N is set to 25. The effects of the uplink packet size and the number of nodes on the performance are also evaluated by the simulation.

In FAIR, the system resource is allocates based on the dynamic estimation of the number of nodes and FIFO scheduling algorithm is used. Simulation results of the dynamic update method of the number of downlink and uplink nodes in FAIR are shown in Fig. 3. The ideal value for the downlink and uplink ratio is one because the number of downlink flows and that of uplink flows are the same in the simulation. However, the estimated values are different for three simulation runs, R_1, R_2, and R_3. Thus, the throughput allocation ratio between uplink and downlink may be far from the ideal value in some cases.

The time wasted by the packet collision of the proposed method is compared with those of DCF and FAIR in Fig. 4. The proposed method is denoted by MUD

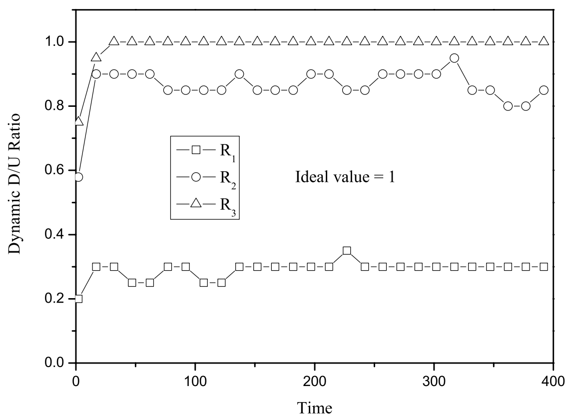

Fig. 3. Dynamic parameter estimation of Down/Up ratio in FAIR

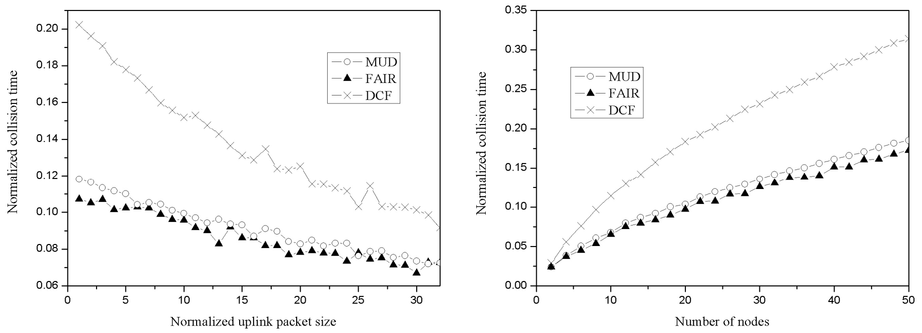

Fig. 4. Normalized collision time as a function of normalized uplink packet size and number of nodes

in the figure. The collision time is normalized to the total simulation time and the uplink packet size is normalized to 64 bytes. The collision time decreases as the uplink packet size increases because the data packet length per a transmission increases. The probability of the packet collision increases as the number of nodes increases. FAIR and MUD show less collision time than DCF because they provide the access method without the collision. FAIR shows less collision time than MUD. It is because MUD utilizes the multi-user diversity during the opportunistic mode which increases the throughput of the opportunistic mode. Thus, more time for the channel access can be allocated to the normal mode in MUD. This will be shown again in the next figure.

The channel access number of the opportunistic mode divided by total channel access number in the proposed method is compared with FAIR in Fig. 5. There is not the opportunistic mode in FAIR and the similar concept, called downlink compensation access, is compared in the figure. The opportunistic mode ratio does not change by the uplink packet size because the access method is changed by the number of channel access. As the number of nodes increases, uplink flows can easily get more throughput as explained in (4). Thus, more opportunistic mode is required to compensate the unfairness. Note that the opportunistic mode ratio is not 0.5 because the AP transmits packets both in normal mode and opportunistic mode. The opportunistic mode ratio of MUD is less than that of FAIR because of the same reason in Fig. 4.

The system throughput of the proposed method is compared with those of DCF and FAIR in Fig. 6. The system throughput increases as the uplink packet size increases because of the reduced overhead per a transmission. The system throughput of DCF decreases as the number of nodes increases because of the increased collisions. However, the system throughput of FAIR and MUD are not changed by the number of nodes because the opportunistic mode ratio is controlled by the number of nodes as shown in Fig. 5. The system throughput of MUD is larger than that of FAIR because of the multi-user diversity gain during the opportunistic mode.

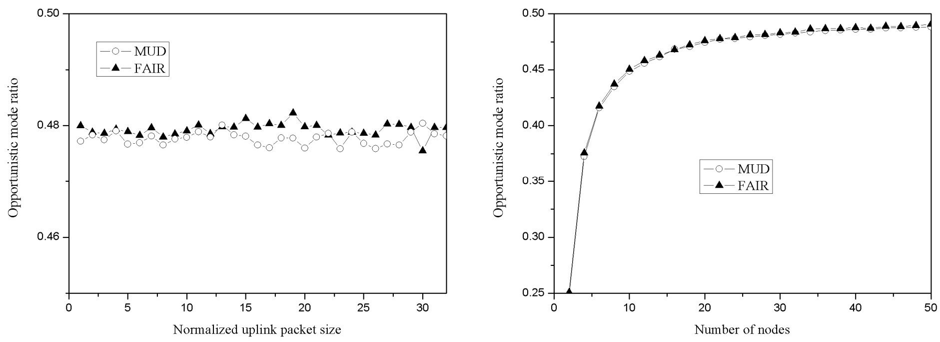

Fig. 5. Opportunistic mode ratio as a function of normalized uplink packet size and number of nodes

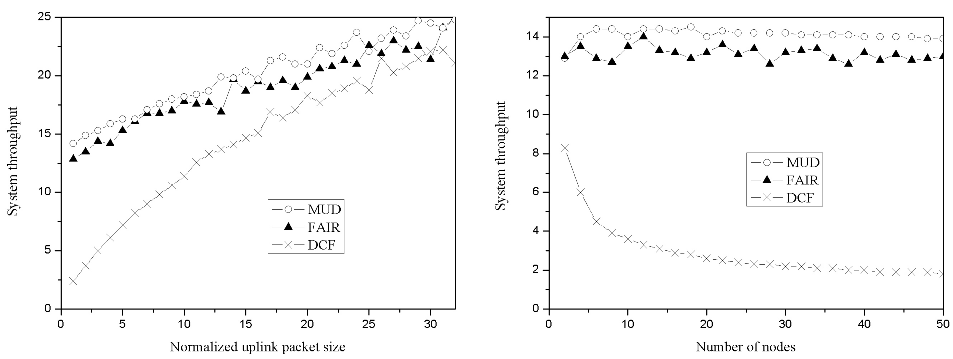

Fig. 6. System throughput as a function of normalized uplink packet size and number of nodes

5 Conclusion

We have proposed a multi-user diversity method to enhance the system throughput of the IEEE 802.11 DCF. Moreover, the proposed method alleviates the throughput unbalance between uplink and downlink. The proposed method can be implemented without the modification of the IEEE 802.11 standard for nodes that are widely deployed.

The efficiency of the proposed system has been demonstrated by computer simulation. The results show that the proposed method enhances the system throughput for asymmetric traffic load. This, in turn, reduces the blocking probability of multimedia data packets in the proposed systems compared with that in the IEEE 802.11 DCF where most of bandwidth is occupied by the uplink.

References

1. Knopp, R., Humblet, P.A.: Information capacity and power control in single cell multiuser communications. In: Proc. IEEE ICC 1995. (1995) 331–335
2. Ajib, W., Haccoun, D.: An overview of scheduling algoriths in MIMO-based fourth-generation wireless systems. IEEE Network **19** (2005) 43–48
3. Gyasi-Agyei, A.: Multiuser diversity based opportunistic scheduling for wireless data networks. IEEE Commun. Lett. **9** (2005) 670–672
4. IS-856: CDMA 2000 standard: High rate packet data air interface specification. (2000)
5. Wang, J., Zhai, H., Fang, Y.: Opportunistic packet scheduling and media access control for wireless LANs and multi-hop ad hoc networks. In: Proc. IEEE WCNC'04, Atlanta, Georgia (2004) 1234–1239
6. Grilo, A., Nunes, M.: Performance evaluation of IEEE 802.11e. In: Proc. IEEE PIMRC'02, Lisboa, Portugal (2002)
7. Pilosof, S., Ramjee, R., Raz, D., Shavitt, Y., Sinha, P.: Understanding TCP fairness over wireless LAN. In: Proc. IEEE Infocom'03, San Francisco, CA, USA (2003)
8. Kim, S.W., Kim, B., Fang, Y.: Downlink and uplink resource allocation in IEEE 802.11 wireless LANs. IEEE Trans. Veh. Technol. **54** (2005) 320–327
9. Kim, S.W., Kim, B.: Reource allocation based on traffic load over relayed wireless access networks. In: Proc. ICESS'05, Xian, China (2005) 441–451
10. IEEE Std 802.11b-1999: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band. (1999)
11. Kamerman, A., Monteban, L.: WaveLAN-II: A high-performance wireless LAN for the unlicensed band. Bell Labs Tech. J. **2** (1997) 118–133
12. Qiao, D., Choi, S., Shin, K.G.: Goodput analysis and link adaptation for IEEE 802.11a wireless LANs. IEEE Trans. Mob. Comput. **1** (2002) 278–292
13. Holland, G., Vaidya, N., Bahl, P.: A rate-adaptive MAC protocol for multi-hop wireless networks. In: Proc. IEEE/ACM MOBICOM'01, Boston, MA, USA (2001) 236–251
14. T. S. Rappaport: Wireless communications: principles and practices, 2nd Ed. Prentice Hall (2002)