

Exploiting Idle Communication Power to Improve Wireless Network Performance and Energy Efficiency

Lei Guo¹, Xiaoning Ding¹, Haining Wang², Qun Li², Songqing Chen³, Xiaodong Zhang¹

¹Ohio State University
Columbus, OH 43210, USA
{lguo,dingxn,zhang}@cse.ohio-state.edu

²College of William and Mary
Williamsburg, VA 23187, USA
{hnw,liqun}@cs.wm.edu

³George Mason University
Fairfax, VA 22030, USA
sqchen@cs.gmu.edu

Abstract—As a family of wireless local area network (WLAN) protocols between physical layer and higher-layer protocols, IEEE 802.11 has to accommodate the features and requirements of both ends. However, current practice has addressed the problems separately and is far from being satisfactory. On the one end, due to varying channel conditions, WLANs have to provide multiple data channel rates to support various bit error rates. A low channel rate station not only suffers low throughput itself, but also significantly degrades the throughput of other stations. On the other end, TCP is not energy efficient running on 802.11. This is because a wireless network interface (WNI) has to stay awake to generate timely acknowledgments during a TCP session, and hence, the energy consumed during idle awake time is wasted for channel listening.

In this paper, considering the needs of both ends, we utilize the idle communication power of the WNI to improve the throughput and energy efficiency of stations in WLANs supporting multiple channel rates. We characterize the energy efficiency as energy per bit, instead of energy per second. Based on modeling and analysis, we propose a data forwarding mechanism and an energy-aware channel allocation mechanism. In such a system, a high channel rate station relays data frames between its neighboring stations with low channel rates and Access Point, improving their throughput and energy efficiency. Different from traditional relaying approaches, our scheme compensates for the energy consumption for data forwarding. The forwarding station gets additional channel access time from its beneficiaries, leading to the increase of its own throughput without compromising its energy efficiency. We implement a prototype of our proposed system and evaluate it through extensive experiments. Our results show significant performance improvements for both low and high channel rate stations.

I. INTRODUCTION

Mobile devices are usually driven by battery power. Due to limited battery capacity, it is essential to reduce power consumption of mobile devices without degrading their performance. In mobile communications, wireless network interfaces (WNIs) consume a significant portion of energy. For instance, the energy consumed by WNIs can account for more than 50% of the energy consumption in handheld computers and up to 10% in laptop computers [4], [18]. As shown in [26], the energy consumption of WNIs is dominated by the idle time of WNIs, instead of the amount of transferred data. To save energy in wireless devices, the basic principle is to put the WNI into sleep mode when it is idle, e.g., IEEE 802.11 power saving mechanism [22]. Nonetheless, due to

the overhead of **mode switches** of WNIs and lagged data reception, frequent waking up and sleeping of WNIs may result in serious performance degradation and may even increase overall energy consumption in mobile devices [4], [16]. Furthermore, to improve throughput and reduce response time to clients, WNIs should always stay awake in TCP sessions to generate timely TCP acknowledgments [16]. The attempt to sleep, which induces a delay in the generation of ACKs, will adversely affect TCP throughput. In other words, during an ongoing TCP session, the WNI has to be always active. Thus, a significant portion of power is wasted on channel listening, which we call the *idle communication power* of a station.

In addition to battery power, mobile devices are very susceptible to physical signal quality degradations such as fading, attenuation, and interference. Due to varying channel conditions, wireless local area networks (WLANs) have to provide multiple data channel rates to support various signal qualities, such as IEEE 802.11a (6-54 Mbps, 8 levels) and IEEE 802.11b (1-11 Mbps, 4 levels). The basic IEEE 802.11 channel access method guarantees an equal channel access probability in the long-term to all stations. Since a low channel rate station takes a much longer time to receive or transmit one data frame, it occupies a longer channel access time and penalizes stations with high channel rates. Therefore, low channel rate stations not only suffer low throughput themselves, but also significantly degrade throughput of other stations, and thus the entire WLAN [9]. To address this performance anomaly in multi-rate WLANs, a time-based fairness channel access method [27] has been proposed, in which each station equally occupies the channel with other stations, regardless of channel rates. However, while the time-based scheme protects high channel rate stations from dramatic performance degradation, it aggravates the throughput of stations with low channel rates.

In this paper, we explore how to utilize idle communication power to improve throughput and energy efficiency of mobile stations in multi-rate WLANs. Instead of simply measuring the energy consumed on WNIs per second, we characterize the energy efficiency in wireless communications as energy per bit, which reflects the actual performance demands that users care about. Since the WNI of a mobile station with a high channel rate is idle in most time, it can forward data frames as a proxy between its neighboring stations with low

channel rates and Access Point. Due to proximity, the channel rates between the proxy and its clients are much higher than those between the clients and the Access Point. Therefore, with the introduction of proxy forwarding mechanism, we can significantly improve the network performance and energy efficiency of mobile clients. Because the proxy consumes additional energy for data forwarding, we propose an energy-aware channel allocation to reward the proxy with longer channel occupancy time from its clients, resulting in the increase of its own throughput without loss of its energy efficiency. **Under such an incentive mechanism, the forwarding service is profitable and thus becomes a kind of resource that stations want to compete for, which is significantly different from previous multi-hop routing algorithms in ad hoc networks. To ensure the fairness of this competition, we propose an auction-based mechanism for proxy selection.**

Based on a mathematical model we propose, we analyze performance gains of proxies and clients in WLANs with the support of multi-hop relay and channel time allocation. The analytical results give theoretical bounds of performance gains under different circumstances. According to the theoretical guideline, we elaborate the system design, which consists of three components: (1) a proxy selection algorithm to choose relay stations for low channel rate stations; (2) a multi-hop forwarding algorithm to provide reliable communication at the MAC layer and coordinate intermediate stations along a forwarding path; (3) a token-based, energy-aware channel allocation algorithm to provide channel occupancy time compensation to forwarding stations under time-based fairness and max-min fairness.

To evaluate the proposed system, we implement a prototype and conduct extensive experiments on our testbed. Our experimental results show that integrating the proxy forwarding and energy-aware channel allocation schemes, high channel rate mobile stations (proxies) not only significantly improve the network performance and energy efficiency of low channel rate mobile stations (clients), but also increase their own throughput and the aggregate throughput of the entire WLAN, without compromising their energy efficiency. Compared with time-based fairness scheduling, the client and proxy throughput can be improved by 138% and by 23% respectively, and the aggregate throughput of the entire WLAN can be improved by 79%.

The remainder of this paper is organized as follows. Section II surveys related work. Section III discusses our motivation. Section IV describes our system model and performance metrics. Section V presents the proxy forwarding and energy-aware token rewarding mechanisms. Section VI details our system design. We implement a prototype of the proposed system and validate its efficacy on our testbed in Section VII, and make concluding remarks in Section VIII.

II. RELATED WORK

Most current WLANs support multiple channel rates for mobile stations with different signal qualities. In outdoor WLANs, radio signal strength attenuates rapidly when the

propagation distance increases. Studies [15], [27] have shown that rate diversity is prevalent in many indoor WLANs and exists even in a small room, because of the signal strength diversity caused by noise, interference, multi-path, and user mobility. In [27], the authors also showed that wireless channels are often saturated due to channel contention among different users.

Heusse et al. [9] identified a performance anomaly in 802.11b that supports four different channel rates. A mobile station transmitting at 1 Mbps degrades the throughput of stations with high channel rates (e.g., 11 Mbps) down below 1 Mbps. The main reason is because a mobile station with lower channel rate takes much longer time to transmit or receive a data frame, and hence, it occupies much more channel time than higher channel rate stations. To address this anomaly, Tan and Gutttag [27] proposed a time-based fairness scheduling algorithm in multi-rate WLANs. In their algorithm, channel access time is equally allocated among all stations with different channel rates. Thus, high channel rate stations are shielded from throughput degradation, but the performance of low channel rate stations is decreased.

IEEE 802.11 supports a power saving mechanism [22]. When a mobile station has no communication workload, it may switch to power saving mode and notify the Access Point to buffer incoming data for it during its sleeping time. **In 802.11 WLANs**, the Access Point periodically broadcasts beacon messages so that mobile stations can synchronize their clocks. **In each beacon message**, the Access Point also transmits a *traffic indication map*, which contains a list of stations that have data frames buffered at the Access Point. A mobile station in power saving mode periodically wakes up and listens to the beacon message. If there are data frames buffered at the Access Point for it, the station polls the Access Point, and then the Access Point transmits the data frames to the station. Afterwards the station returns to sleep mode again.

IEEE 802.11 power saving mode may significantly degrade the network performance in TCP [16] or RPC [4]. This is because it increases the round-trip-time (RTT) to the beacon interval (about 100 ms), which is much greater than a typical end-to-end RTT over the Internet. In [16], the authors demonstrated the performance degradation of Web access caused by power saving mode, and proposed a bounded slowdown protocol to resolve the problem by adapting the sleep and awake durations based on the degree of network activities. Anand et al. [4] have shown the performance degradation of UDP-based RPC calls caused by power saving mode, and presented a self-tuning power management approach to adapting the behaviors of a mobile station to the access patterns and intents of its applications. Note that these solutions are orthogonal to our idle communication power exploitation, and hence, can be integrated with our proposed schemes for better network performance and power savings.

Exploiting spatial reuse in cellular networks, Hsieh and Sivakumar [12] have studied multi-hop ad hoc models to improve network throughput and reduce energy consumption for stations with poor signal qualities. However, spatial reuse

is infeasible in WLANs because of the channel overlapping problem. In [21], a unified cellular and ad-hoc network architecture has been presented, using both a 3G cellular network interface and a 802.11 network interface. In [29], a relay-enabled MAC protocol is proposed for ad hoc networks. In [19], the authors proposed a multi-hop WLAN architecture and demonstrated its benefits to wireless clients. However, none of these solutions can provide effective incentive mechanisms to encourage stations to relay data for other stations. In contrast, our approach quantitatively compensate proxy stations by rewarding them with additional channel occupancy time to improve their own throughput without compromising energy efficiency.

To improve TCP performance in the wireless WAN environment, instead of WLAN, a proxy-based TCP PRISM [14] has been proposed, in which multi-homed mobile stations utilize the diversity of wireless WAN connections for masking adverse effects upon network performance. Besides the closely related work mentioned above, extensive research has been conducted to investigate network performance and power consumption in mobile systems, such as [18], [25], and [26]. Moreover, a variety of transport-layer enhancements have also been developed to improve networking performance over wireless links, such as [5], [6], [7], [10], [11], [13], [17], [20], [24], and [28].

III. MOTIVATION AND RATIONALE

In this section, we describe the rationale behind the proposed data forwarding and channel access time compensation in an IEEE 802.11 WLAN. We have two observations. First, a mobile station involved in a TCP session has to stay awake to generate timely acknowledgments (ACKs) upon data arrivals. Channel listening consumes a considerable amount of energy. Second, a mobile station with a low channel rate will impact other stations with high channel rates and plummet the aggregate throughput of the entire WLAN.

One solution to these problems would be to recruit mobile stations with high channel rates to harvest their idle time and forward data frames for the mobile stations with low channel rates. A low channel rate station improves its throughput via a much higher channel-rate path. However, a high channel rate station has to consume extra energy on forwarding data frames for the low channel rate stations, which it may be unwilling to do. Instead of forcing the high channel rate stations to sacrifice their energy for data relay, the low channel rate stations should compensate a certain amount of time slots to the high channel rate stations, and hence, at least the high channel rate stations will not be penalized by being helpful.

Initially, each mobile station should be assigned the same amount of channel access time for data communication, following the rule of time-based fairness [27]. In our scheme, the allotted time for each station can be traded for higher throughput. Indeed a mobile station always desires a high throughput and low energy consumption. A mobile station can improve its throughput either by obtaining more time slots for its own communication or by increasing the channel rate at

which its data are transmitted. If a high channel rate station gets extra time slots from low channel rate stations, while a low channel rate station increases its data transmission rate through a high channel-rate path composed of proxies; then it will be a *win-win* scenario.

The efficiency of energy utilization needs a little more careful consideration. Energy consumption can be expressed as energy consumed per unit time, or energy consumed per data volume. A fixed data transmission rate for a mobile station gives us an illusion that a user cares about energy consumption per unit time, which is not always true. We believe the user actually cares more about how much energy consumed for a certain amount of data communicated, because the WNI can be put into sleep mode or turned off when it has no communication workload. We define the *energy utility* of the WNI of a mobile station as the average number of effective bits transmitted/received per unit energy when the power is on. Thus, the best way to save energy is to reduce the energy cost of every effective bit or increase the energy utility, and turn off the WNI when the communication session terminates.

Both throughput and energy efficiency are important for TCP communications. To encourage a high channel rate station to relay data for a low channel rate station, its energy utility should not be reduced. A WNI can work in three modes with different power consumption levels: transmission, receiving/listening, and sleep mode. The power consumption of transmission mode is usually much higher than that of receiving/listening mode¹. Thus, the energy utility metric of a high channel rate station will become lower if it relays for a low channel rate station without any compensation. However, if the low channel rate station contributes a fraction of its allotted time slots to the proxy station, the proxy station can use these bonus time slots for its own communication, leading to the increase of its throughput and the decrease of its WNI working time. As a result, although the proxy station spends extra energy for the data forwarding service, its energy utility can remain intact or even increase.

IV. SYSTEM MODEL AND PERFORMANCE METRICS

We now focus on network description and basic notations before we proceed to discuss protocol design. The WLAN in consideration is composed of an *Access Point (AP)*, S_0 , and n ($n \geq 2$) *mobile stations* (denoted as *station* in the remainder of this paper), S_1, S_2, \dots, S_n . The radio channel is shared by all stations and the Access Point. Two stations S_i and S_j can communicate with each other at a channel rate $R_{i,j}$ ($i \neq j$ and $0 \leq i, j \leq n$). Specifically, each station S_i ($1 \leq i \leq n$) can communicate with the AP with channel rate $R_{0,i}$, and we denote $R_{0,i}$ as R_i for simplicity.

Let P_t be the power consumption (energy per second) of a station in the transmission mode, and P_r be the power consumption of a station in the listening or data receiving mode. Assume $P_t = \alpha P_r$ ($\alpha > 1$). We further assume the

¹For example, the typical power consumption of the Cisco Aironet 350 series WNI is 450 mA at transmission mode, 270 mA at receiving/listening mode, and 15 mA at sleep mode (all under 5V DC) [1].

TABLE I
SYMBOLS AND NOTATIONS

symbol	meaning and unit
$P(S_i)$	power consumption of station S_i (Joule/sec)
$T(S_i)$	throughput of station S_i (bit/sec)
$E(S_i)$	energy utility of station S_i (bit/Joule)
$R_{i,j}$	the channel rate between station S_i and S_j (bit/sec)
f_i	the fraction of outgoing traffic in S_i 's workload
t_i	the allocated time of S_i
$x_{j,k}^i$	the fraction of channel time during which the traffic of S_i is forwarded between S_j and S_k
y_j^i	the fraction of channel time that S_i rewards S_j
$U(S_i)$	utilization of allocated time of station S_i
$g_T^0(S_i)$	the throughput gain when S_i has no clients
$g_E^0(S_i)$	the energy utility gain when S_i has no clients
$g_T(S_i)$	the throughput gain of S_i
$g_E(S_i)$	the energy utility gain of S_i

fraction of allocated channel occupancy time of station S_i is t_i , in which the fraction for data transmitting is f_i ($0 \leq f_i \leq 1$). In time-based fairness scheduling [27], each station is assigned the same fraction of channel time. Thus, $t_i = \Delta t = \frac{1}{n}$ ($1 \leq i \leq n$), and we also have the bound $0 < t_i \leq \frac{1}{2}$.

Based on the energy consumption of a station S_i per unit time, $P(S_i)$, we define two performance metrics for a wireless station as follows:

- *Throughput*, $T(S_i)$, the number of effective bits a station transmits and receives per unit time;²
- *Energy utility*, $E(S_i)$, the average number of effective bits per unit energy. That is, $E(S_i) = \frac{T(S_i)}{P(S_i)}$.

According to the assumptions of our model, we have

$$\begin{cases} P(S_i) &= P_t t_i f_i + P_r (1 - t_i f_i) \\ &= P_r (1 + (\alpha - 1) t_i f_i), \\ T(S_i) &= R_i t_i, \\ E(S_i) &= \frac{1}{(\alpha - 1) f_i + \frac{1}{P_r}}, \end{cases} \quad (\text{IV.1})$$

where $1 \leq i \leq n$.

We can compare the original time-based fairness scheme with our new scheme by considering the performance gain in terms of throughput and energy utility:

$$\begin{cases} g_T(S_i) &= \frac{T'(S_i)}{T(S_i)}, \\ g_E(S_i) &= \frac{E'(S_i)}{E(S_i)}, \end{cases} \quad (\text{IV.2})$$

where $P(S_i)$, $T(S_i)$, $E(S_i)$ and $P'(S_i)$, $T'(S_i)$, $E'(S_i)$, are the power consumption, throughput, and energy utility of a station S_i before and after a forwarding service it provides/receives, respectively. Table I lists the notations that will be used in the following analysis.

V. CHANNEL ALLOCATION FOR MULTI-HOP FORWARDING

In this section we will analyze how much time a low rate station has to offer the high rate station for the forwarding service so that the latter will not be penalized. We analyze a simple one-hop case first, and then extend the one-hop relay to the general case of multi-hop relay.

²The bits for MAC level retransmission and the forwarding data for other stations are not counted as effective bits.

A. Channel Occupancy Time Allocation

Assuming time-based fairness scheduling is enabled, each station is assigned an equal fraction of channel time in units of time slot. In such a WLAN, a station that can communicate with AP at a high channel rate can work as the *proxy station* for stations that can only communicate with AP at low channel rates, if they can communicate with each other at high channel rates. The time slots used for the forwarding service should come from the time slots that the *client stations* are allotted based on time-based fairness. Meanwhile, since transmitting data for clients consumes its energy, the proxy station should be rewarded additional time slots from its client stations for compensation. We define the fraction of channel time that a client S_q rewards its proxy S_p to keep the energy utility of S_p unchanged as the *cost price* (or valuation) of the proxy's forwarding service to the client, denoted as $cost(p, q)$.

We define the fraction of channel time that a station is assigned by the time-based fairness scheduling as the *assigned time* of the station, and the fraction of channel time that a station can use for its own communication as the *effective time* of the station. We also define the fraction of channel time that a client rewards each of its proxies as its *rewarding time* to the proxy or the *rewarded time* of that proxy. The effective time of a proxy is its assigned time plus all rewarded time from its clients. The effective time of a client is its assigned time subtracting the fraction of channel time it rewards its proxies and the fraction of channel time for its data relaying (transmitting or receiving) between the AP and the immediate proxy of the client (*relaying time*).

We further define the sum of a station's assigned time according to time-based fairness scheduling and its rewarded time getting from its clients as the *allocated time* of the station, since it can use these times for its own communication and the rewarding to its proxies. Therefore, we define the *utilization of the allocated time* of a station S_i , $U(S_i)$, as the ratio of its effective time to its allocated time.

B. Performance Gain Analysis for One-hop Relay

First, we consider one client and one proxy for simplicity. Assume client station S_q is relayed by proxy station S_p . The assigned time of S_q should be divided into 3 pieces:

$$t_q = \Delta t = x_{0,p} + x_{p,q} + y_p^q, \quad (\text{V.3})$$

where $x_{0,p}$ is the fraction of channel time used for data relaying between AP (S_0) and the proxy station S_p (relaying time), $x_{p,q}$ is the fraction of channel time that the client station S_q is transmitting/receiving data to the proxy station (effective time), and y_p^q is the fraction of channel time that the client station compensates S_p (rewarding time). The utilization of S_q 's allocated time is $U(S_q) = \frac{x_{p,q}}{\Delta t}$.

The effective time of S_p is

$$t_p' = t_p + y_p^q = \Delta t + y_p^q, \quad (\text{V.4})$$

where t_p is its assigned time and y_p^q is its rewarded time from client S_q . The utilization of S_p 's allocated time is 1 since it

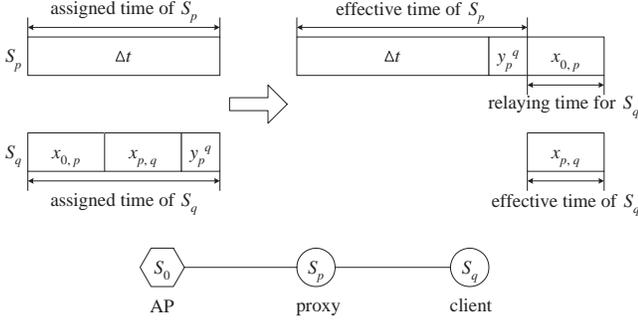


Fig. 1. Channel time allocation

can use all its assigned time and rewarded time for its own communication. Figure 1 shows the channel time allocation in one-hop proxy forwarding.

Lemma 1: In one-hop forwarding, the allocated time utilization, rewarding time, throughput gain and energy utility gain of a client S_q when it pays the cost price to its proxy S_p for the forwarding service are

$$\begin{cases} U(S_q) &= \frac{R_{0,p}}{R_{0,p}+R_{p,q}+(\alpha-1)\Delta t[f_q R_{p,q}+(1-f_q)R_{0,p}]}, \\ y_p^q &= (\Delta t)^2 U(S_q) R_{p,q} (\alpha-1) \left(\frac{f_q}{R_{0,p}} + \frac{1-f_q}{R_{p,q}} \right), \\ g_T(S_q) &= \frac{R_{p,q}}{R_{0,q}} U(S_q), \\ g_E(S_q) &= \frac{R_{p,q}}{R_{0,q}} U(S_q) \frac{(\alpha-1)\Delta t f_q + 1}{U(S_q)(\alpha-1)\Delta t f_q + 1}. \end{cases}$$

Proof: Two constraints dictate how much time a low rate station has to offer to a high rate station: (1) every client station allocates sufficient time for the transmission and forwarding of its data; (2) the energy utility of the high rate station is unchanged.

First, we have

$$T'(S_q) = x_{0,p} R_{0,p} = x_{p,q} R_{p,q}, \quad (\text{V.5})$$

which means the flow rate in each hop along the forwarding path of client S_q is equal.

Second, the energy utility of the proxy is unchanged, that is, the cost price of S_p serving S_q is the rewarding time of S_q to keep the energy utility of S_p unchanged

$$E(S_p) = E'(S_p). \quad (\text{V.6})$$

Equation IV.1 gives the power consumption, throughput and energy utility of S_p when it has no clients. Denote the power consumption, throughput and energy utility of S_p when S_p serves client S_q as $P'(S_p)$, $T'(S_p)$, and $E'(S_p)$, respectively, we have

$$\begin{cases} P'(S_p) &= P_r(1 + (\alpha-1)t_p^f), \\ T'(S_p) &= R(S_p)(\Delta t + y_p^q), \\ E'(S_p) &= \frac{P'(S_p)}{T'(S_p)}, \end{cases} \quad (\text{V.7})$$

where $t_p^f = f_p(\Delta t + y_p^q) + f_q x_{0,p} + (1-f_q)x_{p,q}$ is the total time of proxy S_p used for data transmission. In t_p^f , $f_p(\Delta t + y_p^q)$ is the time that S_p transmits its own upstream workload to AP, $f_q x_{0,p}$ is the time that S_p forwards the upstream workload of

S_q to AP, and $x_{p,q}$ is the time that S_p forwards the downstream workload of S_q to S_q .

Resolving Equation V.3, V.5, and V.6, we have

$$\begin{cases} T'(S_q) &= \frac{\Delta t}{\frac{1}{R_{0,p}} + \frac{1}{R_{p,q}} + (\alpha-1)\Delta t \left(\frac{f_q}{R_{0,p}} + \frac{1-f_q}{R_{p,q}} \right)}, \\ U(S_q) &= \frac{x_{p,q}}{\Delta t} = \frac{T'(S_q)}{R_{p,q}\Delta t} \\ &= \frac{R_{p,q}}{R_{0,p}+R_{p,q}+(\alpha-1)\Delta t[(1-f_q)R_{p,q}+f_q R_{0,p}]}, \\ y_p^q &= t_p(\alpha-1)(f_q x_{0,p} + (1-f_q)x_{p,q}) \\ &= \Delta t(\alpha-1)T'(S_q) \left(\frac{f_q}{R_{0,p}} + \frac{1-f_q}{R_{p,q}} \right), \end{cases} \quad (\text{V.8})$$

where $y_p^q = \text{cost}(p, q)$. According to Equation IV.2, for client station S_q , we have

$$\begin{cases} g_T(S_q) &= \frac{R_{p,q}}{R_{0,q}} U(S_q), \\ g_E(S_q) &= \frac{R_{p,q}}{R_{0,q}} U(S_q) \frac{(\alpha-1)\Delta t f_q + 1}{U(S_q)(\alpha-1)\Delta t f_q + 1}. \end{cases} \quad (\text{V.9})$$

$U(S_q)$ and $g_T(S_q)$ increase with the increase in the number of stations (the decrease of Δt in the WLAN). We have $U(S_q) < \frac{R_{p,q}}{R_{0,p}+R_{p,q}}$ and $g_T(S_q) < \frac{R_{p,q}}{R_{0,q}} \frac{R_{p,q}}{R_{0,p}+R_{p,q}}$. Relaying is only useful when the throughput gain $g_T(S_q) > 1$. Since $U(S_q) < 1$, $f_q \geq 0$, by examining Equation V.9, we have $g_E(S_q) \geq g_T(S_q)$. That is, relaying can always increase the energy utility of client station as long as its throughput can be improved.

For a special case when $R_{0,p} = R_{p,q}$, we have

$$\begin{cases} y_p^q &= \frac{(\alpha-1)\Delta t^2}{2+(\alpha-1)\Delta t}, \quad 0 < y_p^q \leq \frac{1}{2(\alpha+3)}, \\ T'(S_q) &= \frac{\Delta t R_{0,p}}{2+(\alpha-1)\Delta t}, \quad 0 < T'(S_q) \leq \frac{R_{0,p}}{\alpha+3}, \\ U(S_q) &= \frac{1}{2+(\alpha-1)\Delta t}, \quad \frac{2}{\alpha+3} \leq U(S_q) < \frac{1}{2}, \\ g_T(S_q) &= \frac{1}{2+(\alpha-1)\Delta t} \frac{R_{0,p}}{R_{0,q}}, \\ &\frac{2}{\alpha+3} \frac{R_{0,p}}{R_{0,q}} \leq g_T(S_q) < \frac{1}{2} \frac{R_{0,p}}{R_{0,q}}, \\ g_E(S_q) &= \frac{1+(\alpha-1)\Delta t f_q}{2+(\alpha-1)\Delta t(1+f_q)} \frac{R_{0,p}}{R_{0,q}}. \end{cases} \quad (\text{V.10})$$

A proxy station S_p can serve multiple clients at the same time, and these client stations may have different channel rates and different data transmitting/receiving ratios. **We have the following lemma.**

Lemma 2: Assume station S_p provides forwarding services to k client stations, $S_{q_1}, S_{q_2}, \dots, S_{q_k}$ ($k > 1$), and these client stations independently contribute their rewarding time to S_p to keep the energy utility of S_p unchanged, we have

$$\begin{cases} U(S_p) &= 1, \\ g_T(S_p) &= 1 + (\alpha-1) \sum_{i=1}^k T'(S_{q_i}) \left(\frac{f_{q_i}}{R_{0,p}} + \frac{1-f_{q_i}}{R_{p,q_i}} \right), \\ g_E(S_p) &= 1, \end{cases}$$

where $T'(S_{q_i})$ is the throughput of client S_{q_i} ($1 \leq i \leq k$), with the forwarding service of S_p .

Proof: It is easy to see that $U(S_p) = 1$ and $g_E(S_p) = 1$. Since each client rewards S_p independently, similar to the last

formula in Equation V.8, we have

$$\begin{aligned} \frac{1}{t_p} &= \frac{(\alpha-1)(f_{q_1}x_{0,p}+(1-f_{q_1})x_{p,q_1})}{y_p^{q_1}} \\ &= \dots \\ &= \frac{(\alpha-1)(f_{q_k}x_{0,p}+(1-f_{q_k})x_{p,q_k})}{y_p^{q_k}}, \\ \frac{1}{t_p} &= \frac{1+(\alpha-1)\sum_{i=1}^k[(f_{q_i}x_{0,p}+(1-f_{q_i})x_{p,q_i})]}{t_p+\sum_{i=1}^k y_p^{q_i}}. \end{aligned}$$

The effective time of S_p is $t'_p = t_p + \sum_{i=1}^k y_p^{q_i}$. Thus, we have

$$\begin{aligned} g_T(S_p) &= \frac{T'(S_p)}{T(S_p)} = \frac{t'_p}{t_p} = 1 + \frac{\sum_{i=1}^k y_p^{q_i}}{t_p} \\ &= 1 + (\alpha-1)\Delta t \sum_{i=1}^k T'(S_{q_i}) \left(\frac{f_{q_i}}{R_{0,p}} + \frac{1-f_{q_i}}{R_{p,q_i}} \right). \end{aligned}$$

In case $R_{0,p} = R_{p,q_i}$ ($1 \leq i \leq k$), we have

$$g_T(S_p) = 1 + (\alpha-1) \frac{k\Delta t}{2 + (\alpha-1)\Delta t}. \quad (\text{V.11})$$

Since $k\Delta t = \frac{k}{n} < 1$, $g_T(S_p)$ is bounded by

$$1 < g_T(S_p) < \frac{\alpha+1}{2}. \quad (\text{V.12})$$

C. A Generic Analysis for Channel Allocation in Multi-hop Forwarding

A station S_i that is relayed by other stations can still work as the proxy for stations of lower channel rates, and gets rewarded time from its clients. However, only a fraction of its rewarded time can be used for its own communication, since S_i also needs to reward its relaying stations. We consider the relay chain $S_0 \rightarrow S_1 \rightarrow \dots \rightarrow S_{i-1} \rightarrow S_i$ starting from the AP (S_0). In order for S_1 to relay data for S_2 , S_1 has to keep its energy utility unchanged. After S_1 decides to relay data for S_2 , S_2 will have a higher energy utility than that before it is relayed. S_2 would like to keep this new energy utility unchanged when it decides to relay for S_3 , and so on. The following Lemma describes the performance gain of a station in such scenarios. The proof basically formalizes the above process.

Denote the throughput gain and energy utility gain when S_i has no clients as $g_T^0(S_i)$ and $g_E^0(S_i)$, respectively. We have the following lemma.

Lemma 3: Assume each station has at most one immediate relaying station in the WLAN, and each station rewards its relaying stations independently to keep their energy utilities unchanged. For station S_i that is relayed by $i-1$ ($i \geq 1$) stations along the path $S_0 \rightarrow S_1 \rightarrow \dots \rightarrow S_{i-1} \rightarrow S_i$, and S_i has m_i indirect or direct clients ($S_{q_1}, S_{q_2}, \dots, S_{q_{m_i}}$), we have

$$\begin{cases} g_T^0(S_i) &= \frac{R_{i-1,i} U(S_i)}{R_{0,i}} \\ g_E^0(S_i) &= \frac{R_{i-1,i} U(S_i)}{R_{0,i}} \frac{(\alpha-1)\Delta t f_i + 1}{U(S_i)(\alpha-1)\Delta t f_i + 1}, \end{cases}$$

where $U(S_i) = \frac{1}{1 + R_{i-1,i} \sum_{j=1}^{i-1} [\frac{1}{R_{j-1,j}} + (\alpha-1)\Delta t (\frac{f_j}{R_{j-1,j}} + \frac{1-f_j}{R_{j,j+1}})]}$, and

$$\begin{cases} g_T(S_i) &= g_T^0(S_i) \left(1 + \frac{\sum_{j=1}^{m_i} y_i^{q_j}}{\Delta t} \right) & i \geq 1, \\ g_E(S_i) &= g_E^0(S_i) & i \geq 1, \end{cases}$$

where $y_i^{q_j} = \Delta t(\alpha-1)T'(S_{q_j}) \left(\frac{f_{q_j}}{R_{i-1,i}} + \frac{1-f_{q_j}}{R_{i,i_j}} \right)$, $T'(S_{q_j})$ is the throughput of S_{q_j} when it is forwarded by S_i , and S_{i_j} is the next hop station of S_i to reach S_{q_j} .

Proof: For station S_i ($i > 1$) that is relayed by stations S_1, \dots, S_{i-1} , we have

$$t_i = (x_{0,1}^i + \dots + x_{i-2,i-1}^i) + x_{i-1,i}^i + (y_1^i + \dots + y_{i-1}^i). \quad (\text{V.13})$$

The flow rate of S_i 's own traffic in each hop along the forwarding path is equal, so we have

$$T'(S_i) = x_{0,1}^i R_{0,1} = \dots = x_{i-2,i-1}^i R_{i-2,i-1} = x_{i-1,i}^i R_{i-1,i}. \quad (\text{V.14})$$

For a relaying station of S_i , S_j ($0 < j < i$), when S_j has no clients, we have

$$\begin{cases} P(S_j) &= P_t \Delta t f_j U(S_j) + P_r (1 - \Delta t f_j) \\ &= P_r [1 + (\alpha-1)\Delta t f_j U(S_j)], \\ T(S_j) &= R(S_j) \Delta t U(S_j), \end{cases} \quad (\text{V.15})$$

where $U(S_j) = 1$ when S_j has no proxy ($j = 1$), and $U(S_j) < 1$ when S_j is relayed by other stations ($1 < j < i$). When S_j serves station S_{j+1}, \dots, S_i , we have

$$\begin{cases} P'(S_j) &= P_r [1 + (\alpha-1)t_j^f], \\ T'(S_j) &= R(S_j) (\Delta t + \sum_{l=j+1}^i y_l^j) U(S_j), \end{cases} \quad (\text{V.16})$$

where $t_j^f = f_j (\Delta t + \sum_{l=j+1}^i y_l^j) U(S_j) + \sum_{l=j+1}^i f_l x_{j-1,j}^l + \sum_{l=j+1}^i (1-f_l) x_{j,j+1}^l$. In t_j^f , $f_j (\Delta t + \sum_{l=j+1}^i y_l^j) U(S_j)$ is the time used by S_j to transmit its own workload to S_{j-1} , $f_l x_{j-1,j}^l$ is the time used by S_j to transmit the upstream workload of S_l to S_{j-1} , and $(1-f_l) x_{j,j+1}^l$ is the time used by S_j to transmit the downstream workload of S_l to S_{j+1} .

Considering the energy utility of S_j , we have

$$\begin{cases} E(S_j) &= \frac{R(S_j)}{P_r} \frac{\Delta t U(S_j)}{1 + (\alpha-1)\Delta t f_j U(S_j)}, \\ E'(S_j) &= \frac{R(S_j)}{P_r} \frac{(\Delta t + \sum_{l=j+1}^i y_l^j) U(S_j)}{1 + (\alpha-1)t_j^f}. \end{cases}$$

The energy utility of S_j should be unchanged, that is, $E(S_j) = E'(S_j)$. By substituting $E(S_j)$ and $E'(S_j)$, we have

$$\begin{aligned} (\alpha-1)f_j + \frac{1}{\Delta t U(S_j)} &= \\ (\alpha-1)f_j + \frac{1 + (\alpha-1) \sum_{l=j+1}^i (f_l x_{j-1,j}^l + (1-f_l) x_{j,j+1}^l)}{(\Delta t + \sum_{l=j+1}^i y_l^j) U(S_j)} &, \end{aligned}$$

Simplifying the equation, we have

$$\frac{1}{\Delta t} = \frac{(\alpha-1) \sum_{l=j+1}^i (f_l x_{j-1,j}^l + (1-f_l) x_{j,j+1}^l)}{\sum_{l=j+1}^i y_l^j}.$$

Since each station S_l ($j+1 \leq l \leq i$) rewards time slots to S_j independently, it is easy to prove

$$\frac{1}{\Delta t} = \frac{(\alpha-1)(f_l x_{j-1,j}^l + (1-f_l) x_{j,j+1}^l)}{y_l^j}.$$

Thus, we have

$$\begin{aligned} y_l^j &= \Delta t(\alpha-1)(f_l x_{j-1,j}^l + (1-f_l) x_{j,j+1}^l) \\ &= \Delta t(\alpha-1)T'(S_l) \left(\frac{f_l}{R_{j-1,j}} + \frac{1-f_l}{R_{j,j+1}} \right), \end{aligned} \quad (\text{V.17})$$

where $T'(S_l)$ is the throughput of S_l when it is served by S_j and $T'(S_l) = R_{l-1,l} \times t_l U(S_l)$, where $U(S_l)$ is the allocated time utilization of S_l .

When S_i has no clients, we have $t_i = \Delta t$. Considering Equation V.13, V.14, and V.17, for station S_i , we have

$$U(S_i) = \frac{T'(S_i)}{R_{i-1,i} t_i} = \frac{1}{1 + R_{i-1,i} \sum_{j=1}^{i-1} \left[\frac{1}{R_{j-1,j}} + (\alpha-1) \Delta t \left(\frac{f_j}{R_{j-1,j}} + \frac{1-f_j}{R_{j,j+1}} \right) \right]}. \quad (\text{V.18})$$

Accordingly, we get

$$\begin{cases} g_T^0(S_i) &= \frac{T'(S_i)}{T(S_i)} = \frac{R_{i-1,i} t_i U(S_i)}{R_{0,i} \Delta t} = \frac{R_{i-1,i}}{R_{0,i}} U(S_i), \\ g_E^0(S_i) &= \frac{E'(S_i)}{E(S_i)} = \frac{R_{i-1,i} U(S_i) P(S_i)}{R_{0,i} P(S_i)} \\ &= \frac{R_{i-1,i}}{R_{0,i}} U(S_i) \frac{(\alpha-1) \Delta t f_i + 1}{U(S_i) (\alpha-1) \Delta t f_i + 1}. \end{cases} \quad (\text{V.19})$$

When S_i has m_i clients $S_{q_1}, \dots, S_{q_{m_i}}$, since each client rewards S_i time slots independently, the new throughput is $T''(S_i) = U(S_i) R_{i-1,i} (\Delta t + \sum_{j=1}^{m_i} y_i^{q_j})$. The performance gain becomes

$$\begin{cases} g_T(S_i) &= \frac{T''(S_i)}{T(S_i)} = g_T^0(S_i) \left(1 + \frac{\sum_{j=1}^{m_i} y_i^{q_j}}{\Delta t} \right) & i \geq 1, \\ g_E(S_i) &= g_E^0(S_i) & i \geq 1, \end{cases}$$

where $y_i^{q_j}$ follows Equation V.17. ■

The above analysis of the channel allocation for one-hop and multi-hop forwarding shows the performance gains of low channel rate clients and high channel rate proxies. Specifically, we show that a high rate station will increase its throughput even with its energy utility unchanged, providing a strong incentive for the forwarding service.

VI. SYSTEM DESIGN

In this section, we describe the system design of the multi-hop forwarding service. The proposed system consists of three major components: a proxy selection algorithm, a token-based energy-aware channel scheduling algorithm, and a multi-hop forwarding algorithm. The *proxy selection algorithm* runs on the AP, choosing relay proxies for stations with low channel rates. The *energy-aware channel scheduling algorithm* also runs on the AP, arbitrating channel time allocation and ensuring time-based and max-min fairness among stations. The *multi-hop forwarding algorithm* is a distributed algorithm running on both the AP and each station, in order to coordinate intermediate stations along the forwarding path and provide reliable communication at the MAC layer. The three algorithms work together to enable the data forwarding among stations in a WLAN.

As shown in Figure 2, stations in the WLAN are organized into a tree rooted at the AP for the multi-hop forwarding service. Each non-root node of the tree represents a station, and the weight of each edge represents the channel rate between two nodes. The AP (root) maintains the topology and the weights of edges of the forwarding tree. Each station maintains the information about its children and predecessors, and the weight of each edge along the path. Note that the height of

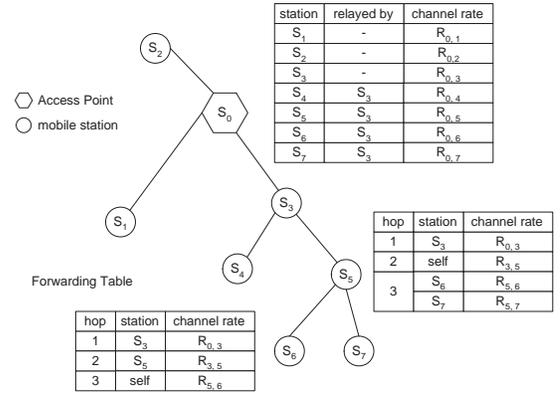


Fig. 2. Multi-hop forwarding structure

the forwarding tree should be small (typically two or three in 802.11b). The main reason for this is that the data forwarding along each hop requires the occupancy of channel resources (spatial reuse is difficult in WLANs). With the increase in the number of forwarding hops, the improvement of a client's throughput decreases rapidly. Moreover, due to the possible mobility of stations, it is much easier to maintain a short tree than a tall one.

A. Proxy Selection and Association

With the channel time compensation, the forwarding service is profitable and thus becomes a kind of resource that stations want to compete for. This is significantly different from previous multi-hop routing algorithms in ad hoc networks. To ensure the fairness of this competition, we propose an auction-based mechanism for proxy selection.

Our proxy selection algorithm runs on the AP, which works as the auctioneer. When a station S_q is working at a low channel rate to communicate with the AP, it broadcasts a sequence of *SFP* (search for proxy) messages with different channel rates. Upon receiving an SFP, each high channel rate station computes the expected throughput gain it can provide to S_q and the cost price based on Lemma 3, then bids for the forwarding service with the cost price.

Upon receiving an SFP, the AP collects the bids from all bidders within the bidding time, and then selects the station that can provide the largest throughput gain for S_q as the proxy. A client would always like to pay less and get more, while a proxy would always like to being paid more and serve less. In our mechanism design, the *dominant strategy* for all bidders—the “best” strategy they can expect—should be to bid with the cost price of their services. The *second price sealed bid* auction rule [23] can be applied to provide such a dominant strategy and finish the auction in one bidding round. In this mechanism, station S_q will pay the proxy at the price of the bidder who offers the second largest throughput gain (please refer to our technical report [8] for a detailed description of the auction mechanism).

When the proxy is selected, the AP sends (or piggybacks) the MAC address of the proxy and the corresponding price to S_q . Then S_q sends a *RFR* (request for relay) message to the

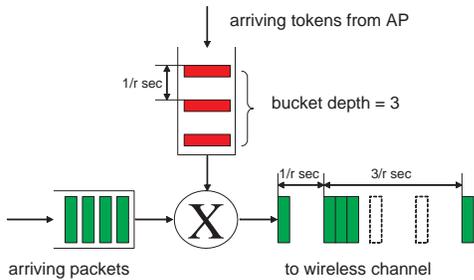


Fig. 3. Token bucket: the AP distributes tokens in a rate r (one round per $1/r$ seconds)

proxy, and the proxy acknowledges the request and reports to the AP to commit the proxy association. When the client does not need data forwarding any longer, it sends a notification to the AP directly through the low-rate channel to cancel the forwarding service.

Many high channel rate stations may compete with each other to get more rewarded time slots to improve their own throughput. The AP needs to balance the profits among proxy candidates that can provide the same forwarding services in a WLAN. For example, if two stations can provide the same throughput gain for S_q , the AP should favor the station with less throughput as the proxy of S_q . Other factors, such as the history of activity and mobility of the proxy candidates, may also be taken into consideration for proxy selection.

B. Channel Allocation and Scheduling

The channel scheduling and the forwarding coordination can be easily implemented in 802.11 WLANs under PCF (point coordination function) with polling MAC mechanism. However, most 802.11 commercial products only support DCF (distributed coordination function) MAC control. In what follows, we describe our system design for 802.11 WLAN under the DCF MAC mechanism.

In the proposed system, the channel is allocated in units of time slot, same as the unit of station's back-off time for PHY medium access ($50 \mu\text{s}$ for FHSS and $20 \mu\text{s}$ for DSSS). As shown in Figure 3, the time slot allocation is performed by the AP based on the *token bucket model*. Each station is assigned a certain number of *tokens* for use in channel contention. A station competes for channel only when it has available tokens. At regular intervals, the AP evenly distributes tokens to each station, ensuring time-based fairness. When the bucket of a station is full, the overflowing tokens are returned to AP, and are re-distributed equally to other stations to ensure max-min fairness. The token bucket shapes the frame transmission of a station at a constant rate in the long run, while allowing bursty frame transmission of a station in the short term. The tokens can be distributed individually or be piggybacked within the data/control frames to stations.

A station transmits data frames only when it has enough tokens. Similarly, the AP buffers data frames for stations without tokens, and postpones their data transmission to the next round of time-slot allocation. Thus, the number of tokens a station holds determines whether it is qualified for channel

competition. Meanwhile, channel contention is fair for those stations with tokens. Therefore, the channel occupancy time of a station is dependent on the token allocation scheme in the long term, although it is non-deterministic in the short term.

We use the similar method as that in [27] to compute the channel occupancy time of a station. For each station, a token counter is maintained at this station and the AP, respectively. Upon receiving/sending a data frame from/to the AP, the station deducts the corresponding tokens from its token counter. At the same time, the AP deducts the same number of tokens of that station. **In 802.11 protocol, the number of retries that a sender transmits a frame is included in the frame header that the receiver gets. However, current hardware does not return the number of retries when the frame is successfully transmitted. Thus, the sender cannot accurately compute the number of tokens used for data transmission, and the two counters may be inconsistent. To minimize this effect, the receiver piggybacks the number of tokens that are used for the last data transmission of its peer in the data frame, and the peer adjusts its token counter accordingly.**

To simplify token management, a proxy station does not maintain token counters for its clients. Once a client associates to the proxy, the tokens, including those that the client should reward its proxy and those that are used to receive/forward data frames for the client, are delivered to the proxy directly by the AP during the token distribution. Correspondingly, the same number of rewarding tokens is deducted from the token counter of the client by the AP. Once the client cancels the forwarding service, the proxy automatically suspends the data forwarding at the next round of token distribution, because the AP will no longer convey the client's rewarding tokens.

C. Multi-Hop Forwarding

1) *Basic Mechanism*: To support multi-hop forwarding, each data frame is appended with two fields indicating the source and destination MAC addresses of the frame, respectively. Each station maintains a forwarding table as shown in Figure 2. Upon receiving a data frame, the station compares the destination MAC address with its own MAC address. If they are different, the station looks up the MAC address for the next-hop station in the forwarding table. Then it modifies the destination address of the frame header and forwards it to the next-hop station. The forwarding table also records the uplink channel rates of the station's predecessors, in order to compute the cost price for the forwarding service it provides its clients, and throughput gains its clients can achieve.

2) *Forwarding Path Maintenance*: The channel rates along the forwarding path and between the client and the AP may change with the mobility of stations or signal instability. Furthermore, the forwarding path may even be broken, due to hardware failure, signal error (or interference), and the mobility of proxy stations. To cope with the possible change of channel rates, each client periodically re-evaluates the forwarding service. If the service quality is significantly degraded, it re-broadcasts SFPs for a new proxy.

3) *Power Management in Multi-hop Forwarding*: Most existing power saving solutions [4], [16] utilize heuristic algorithms to adapt the sleep behavior of a WNI with its network activities. When a station has no network traffic, it will still be up for a while before it goes to sleep based on the prediction of its network activity. The station may also change its waking period adaptively to save the energy consumption on beacon listening. In our scheme, each station has the flexibility to set its own power saving policy.

Any station that wants to sleep needs to send a request to the AP, so that the AP can buffer the incoming data frames for it. The sleep request of a client is sent to the AP at a low channel rate directly. When a proxy decides to switch to power saving mode, it notifies its immediate clients first. If any child of the proxy has clients, the notification will be propagated recursively. Upon receiving the ACK from all its clients, the proxy sends a request to the AP, and shifts to power saving mode. Then, its clients (immediate or non-immediate) search for new proxies.

VII. IMPLEMENTATION AND EXPERIMENTAL EVALUATION

This section presents a prototype implementation of our proposed scheme and its experimental evaluation. Our purpose is twofold: (1) to demonstrate that our data forwarding mechanism is feasible under the framework of the current IEEE 802.11 protocol; and (2) to validate its efficacy in significantly improving the throughput and energy utility for stations in the WLAN.

A. Prototype Implementation

We have implemented a prototype of the proposed scheme and evaluated it on our testbed, which includes an Access Point and six mobile stations. The AP is a desktop PC equipped with a NetGear MA311 802.11b PCI wireless adaptor running Linux kernel 2.4.20. The mobile stations are six HP laptop computers, each equipped with a NetGear MA401 802.11b PCMCIA wireless adaptor running Linux kernel 2.4.20. One of the six works as the proxy, the others work as the clients. All wireless adaptors in the AP and mobile stations use the Intersil Prism2 chipset.

We have modified the HostAP Linux driver for Prism2/2.5/3 [2] as the driver of our Access Point. The AP maintains the forwarding structure for each station associated to it, as described in Section VI. The bidding time for proxy selection is set to 50 ms and the token distribution interval is set to 100 ms. Each token denotes 20 μ s channel occupancy time. To implement the token distribution, the HostAP driver maintains the number of available tokens owned by each mobile station that is currently associated to the AP. In each round of the token distribution, the HostAP driver first evenly allocates tokens based on the number of stations, then transfers the rewarding tokens from each client to its proxy based on their service agreement.

We have also modified the ORiNOCO Linux driver 0.15rc2 for wireless cards [3] as the driver of our proxy and client

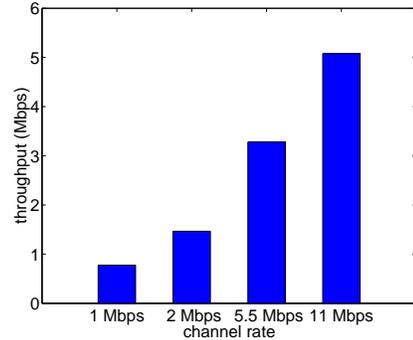


Fig. 4. The effective bandwidth of 802.11b WLAN under different channel rates

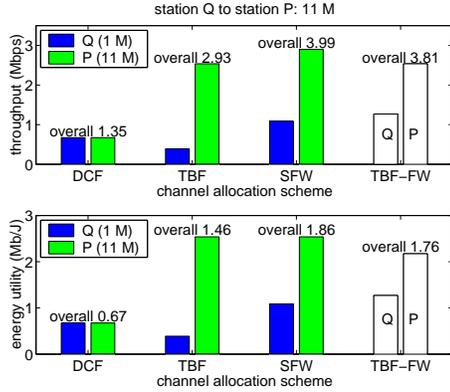
stations. Inside the driver, we have implemented a simple multi-hop forwarding protocol.

B. Experimental Evaluation

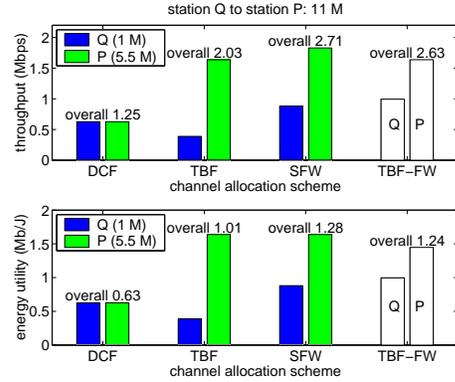
To evaluate the implemented prototype, we have conducted extensive experiments on our testbed with respect to FTP-like and Web-like workload, respectively. Due to page limitations, we only present the results of FTP-like workload (please refer to our technical report [8] for a detailed evaluation of Web-like workload).

1) *Performance Baseline Measurement*: The ideal channel rate of IEEE 802.11 WLAN cannot be achieved in reality, due to the overhead of control frames, inter-frame spaces, physical and MAC layer headers, channel contention, and possible data losses. Therefore, we first measure the effective throughput of a WNI as the baseline for performance comparisons. In this evaluation, we first set up a small 802.11b WLAN that consists of an AP and a mobile station. A large file (about 1 GB) is downloaded from the AP to the station to measure the user level throughput under different channel rates. Figure 4 shows the effective bandwidth of the 802.11b WLAN under channel rates of 1 Mbps, 2 Mbps, 5.5 Mbps, and 11 Mbps, respectively. The higher the channel rate is, the less efficient the channel utilization is. The reason is that all physical layer headers are transmitted at the lowest channel rate in the 802.11b protocol, in order to ensure that all stations in a WLAN can listen to the channel for collision avoidance. However, the diversity of user level throughput under different channel rates is still large enough to benefit low channel rate users through data forwarding. In WLANs supporting more levels of channel rates such as 802.11a, multi-hop data forwarding would have greater potential to improve the system performance.

2) *Experiments on FTP-like Workload*: We implement four channel allocation protocols as listed in Table II and compare their throughput and energy utility with a FTP-like data transmission workload. In these schemes, DCF denotes the normal DCF MAC in a 802.11b WLAN, and TBF denotes the time-based fairness channel contention mechanism proposed in [27]. SFW denotes our proposed mechanism, meaning *selfish multi-hop forwarding*, in which the client pays the cost price for the forwarding service (since there is only one proxy). In order to show the advantage



(a) Channel rates between stations: Q-AP, 1 M; P-AP, 11 M; Q-P, 11 M



(b) Channel rates between stations: Q-AP, 1 M; P-AP, 5.5 M; Q-P, 11 M

Fig. 5. The throughput and energy utility of stations under different channel allocation schemes (1 proxy and 1 client)

TABLE II
CHANNEL ALLOCATION SCHEME

Scheme	Scheme Description
DCF	802.11 DCF MAC (without data forwarding)
TBF	time-based fairness scheduling (without data forwarding)
SFW	selfish forwarding under TBF scheduling
TBF-FW	data forwarding under TBF scheduling

of our proposed channel time compensation mechanism, we also implement data forwarding under time-based fairness for comparison, called TBF-FW. In this mechanism, each station is assigned equal channel time to ensure time-based fairness, and the proxy voluntarily forwards data for its clients using the channel time of its clients, without any time slot rewarded. Note that this is a *phantom* mechanism just for comparison, neither proposed nor implemented before.

In the experiments, we simultaneously download a large file from the HostAP machine to the proxy and client stations, respectively. The throughput is measured by recording the data volume transferred between each client and its proxy (or the proxy and the AP) under different channel allocation schemes. The energy consumption on data transmitting is computed as the product of the data transmission time of physical frames and the power consumption of the wireless card in the transmitting mode (provided by the manufacturer). The energy consumption on receiving/listening is computed in a similar way.

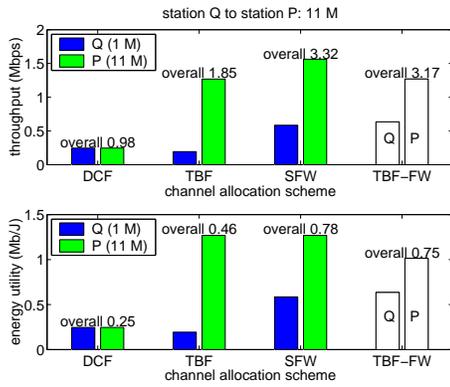
We conduct experiments for the one-hop forwarding case, where the WLAN consists of 1 AP, 1 proxy (denoted by P), and multiple clients (denoted by Q) varying from 1 to 5. Assuming all clients have the same channel rate, there are eight possible combinations for the data forwarding service:

- the channel rate is 1 M or 2 M between Q-AP, 11 M between P-AP, and 11 M between Q-P;
- the channel rate is 1 M or 2 M between Q-AP, 5.5 M between P-AP, and 11 M between Q-P;
- the channel rate is 1 M or 2 M between Q-AP, 11 M between P-AP, and 5.5 M between Q-P;

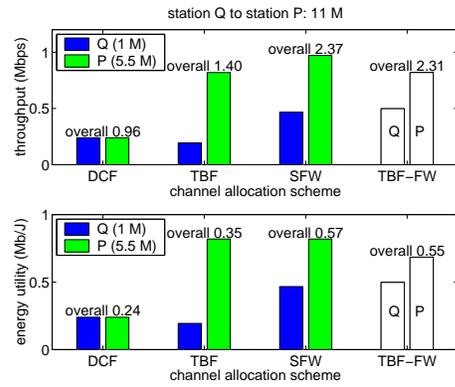
- the channel rate is 1 M or 2 M between Q-AP, 5.5 M between P-AP, and 5.5 M between Q-P.

Each experiment is repeated three times. Figures 5, 6, and 7 show the performance in a WLAN with 1 AP, 1 proxy, and 1, 3, and 5 clients, respectively. Due to page limitations, only part of results are presented (other results are similar). In the figures, the number on the top of the bar group denotes the overall throughput (in Mbps) or the overall energy utility (in Mb per Joule) of the proxy and client stations in the WLAN. We also present the corresponding performance of DCF and TBF for comparisons. The performance of phantom TBF-FW is presented as white bars.

The results can be summarized as follows. SFW has the highest overall performance for both throughput and energy utility, while DCF has the worst overall performance. By enforcing time-based fairness, TBF improves the performance of high channel rate stations but decreases the performance of low channel rate stations. TBF-FW improves the throughput of low channel stations (clients) by data forwarding, but significantly decreases the energy utility of the forwarding station (proxy), which the proxy is unwilling to do. **So the phantom TBF-FW scheme, though simple and effective for low channel rate stations, is not likely feasible in practice. In contrast,** in our proposed forwarding scheme, the proxy gets additional channel time compensation from its clients, resulting in the improvement of its own throughput without decreasing its energy utility, **and thus providing a strong incentive for the forwarding service.** The client stations sacrifice a few channel time tokens for the forwarding service, but the overhead is minor. For example, as shown in Figure 6(a), the client throughput of SFW is 138% higher than that of DCF, more than 2 times over that of TBF, and about 93% of that of TBF-FW, while the proxy throughput of SFW is more than 5 times over that of DCF, and 23% higher than those of TBF and TBF-FW. The proxy energy utility of SFW is more than 4 times over that of DCF, and is same as that of TBF. On the other hand, compared with SFW, the proxy energy utility of TBF-FW is 20% lower than

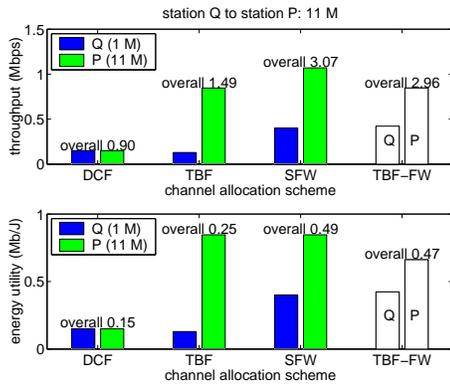


(a) Channel rates between stations: Q-AP, 1 M; P-AP, 11 M; Q-P, 11 M

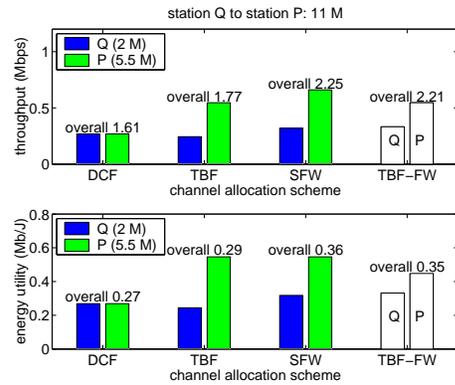


(b) Channel rates between stations: Q-AP, 1 M; P-AP, 5.5 M; Q-P, 11 M

Fig. 6. The throughput and energy utility of stations under different channel allocation schemes (1 proxy and 3 clients)



(a) Channel rates between stations: Q-AP, 1 M; P-AP, 5.5 M; Q-P, 11 M



(b) Channel rates between stations: Q-AP, 2 M; P-AP, 5.5 M; Q-P, 11 M

Fig. 7. The throughput and energy utility of stations under different channel allocation schemes (1 proxy and 5 clients)

that of TBF without any throughput improvement for the forwarding service. Furthermore, with our proposed SFW, the overall performance in the WLAN is also better than that of TBF-FW. These results evidence that our proposed scheme not only provides a strong incentive for data forwarding, but also balances the trade off between the performance of individual stations and the entire WLAN.

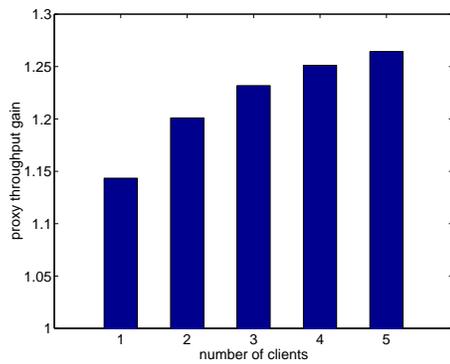
Figure 8(a) shows the proxy throughput gain in SFW (the proxy throughput of SFW over that of TBF) with the increasing number of clients in the WLAN. In this experiment, the proxy (working at 11 Mbps channel rate with the AP) serves all other stations (working at 1 Mbps with the AP and 11 Mbps with the proxy) in the same WLAN. The throughput gain of the proxy increases with the number of clients it serves. With channel time compensation, even in 1 client and 1 proxy case, the proxy throughput can still be improved by 14% over TBF. Figure 8(b) shows the proxy energy utility gain in TBF-FW (the proxy energy utility of TBF-FW over that of TBF) in the same circumstances as above. The energy utility gain is less than 1, meaning it is smaller than that in TBF. Figure 8(b) also indicates that in TBF-FW, the proxy may have to consume more than 22% energy for its clients, which could

prevent the proxy from providing such service.

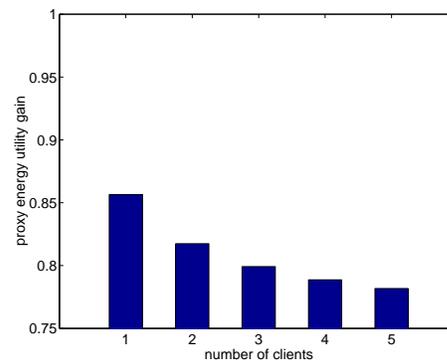
VIII. CONCLUSION

In this paper, we aim to (1) address the throughput degradation induced by low channel rate stations in a WLAN, and (2) utilize the inevitable energy waste in channel listening during a communication session. We characterize energy efficiency as energy per bit, instead of energy per second. Utilizing idle communication power, we present a data forwarding mechanism and an energy-aware token rewarding scheme to supplement the IEEE 802.11 protocols. In data forwarding, a high channel rate station forwards data for a low channel rate station, resulting in a significant improvement of its throughput. To give high channel rate stations an incentive to be proxies, we design an energy-aware token rewarding scheme, in which low channel rate stations compensate for proxies with additional time slots. Thus, a proxy can also improve its own throughput without compromising its energy efficiency.

We have presented a mathematical model to guide the protocol design, and have proposed algorithms for proxy selection, channel allocation and scheduling, and data forwarding in



(a) Improvement of proxy throughput gain in SFW



(b) Reduction of proxy energy utility gain in TBF-FW

Fig. 8. Performance gain of proxy with different number of clients

the IEEE 802.11 WLAN. To evaluate our proposed scheme, we have implemented a prototype of the proposed system by modifying the HostAP driver running on a Linux PC serving as an Access Point, and ORiNOCO Linux driver for wireless cards. We have conducted a set of experiments on our testbed. The experimental results show that the proposed data forwarding and channel access time compensation schemes significantly improve the system performance of the entire WLAN.

ACKNOWLEDGMENT

We would like to thank William L. Bynum for his constructive comments and suggestions. This work is partially supported by the National Science Foundation under grants xxx.

REFERENCES

- [1] <http://www.cisco.com/en/US/products/hw/wireless/>.
- [2] Host AP driver for intersil prism2/2.5/3, hostapd, and WPA supplicant. <http://hostap.epitest.fi/>.
- [3] The Linux ORiNOCO driver. <http://www.nongnu.org/orinoco/>.
- [4] M. Anand, E. Nightingale, and J. Flinn. Self-tuning wireless network power management. In *Proc. of ACM MOBICOM*, San Diego, CA, Sept. 2003.
- [5] A. Bakre and B. Badrinath. I-TCP: Indirect TCP for mobile hosts. In *Proc. of IEEE ICDCS*, Vancouver, Canada, May - June 1995.
- [6] H. Balakrishnan, S. Seshan, and R. Katz. Improving reliable transport and handoff performance in cellular wireless networks. *ACM Wireless Networks*, 1(4):469 – 481, 1995.
- [7] K. Brown and S. Singh. M-TCP: TCP for mobile cellular networks. *ACM SIGCOMM Computer Communication Review*, 27(5):19 – 43, 1997.
- [8] L. Guo, X. Ding, H. Wang, Q. Li, S. Chen, and X. Zhang. Exploiting idle communication power to improve wireless network performance and energy efficiency. Technical Report, <http://www.cse.ohio-state.edu/~lguo/idle-comm-power-techreport.pdf>.
- [9] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. Performance anomaly of 802.11b. In *Proc. of IEEE INFOCOM*, San Francisco, CA, May 2003.
- [10] G. Holland, N. H. Vaidya, and P. Bahl. A rate-adaptive MAC protocol for multi-hop wireless networks. In *Proc. of ACM MOBICOM*, Rome, Italy, July 2001.
- [11] H. Hsieh, K. Kim, Y. Zhu, and R. Sivakumar. A receiver-centric transport protocol for mobile hosts with heterogeneous wireless interfaces. In *Proc. of ACM MOBICOM*, San Diego, CA, Sept. 2003.
- [12] H. Hsieh and R. Sivakumar. On using the ad-hoc network model in cellular packet data networks. In *Proc. of MOBIHOC*, Lausanne, Switzerland, June 2002.
- [13] H. Kim and J. Hou. Improving protocol capacity with model-based frame scheduling in IEEE 802.11-operated WLANs. In *Proc. of ACM MOBICOM*, San Diego, CA, Sept. 2003.
- [14] K. Kim and K. G. Shin. Improving TCP performance over wireless networks with collaborative multi-homed mobile hosts. In *Proc. of ACM MOBISYS*, Seattle, WA, June 2005.
- [15] D. Kotz, C. Newport, and C. Elliott. The mistaken axioms of wireless-network research. Technical Report TR2003-467, Computer Science, Dartmouth College, July 2003.
- [16] R. Krashinsky and H. Balakrishnan. Minimizing energy for wireless Web access with bounded slowdown. In *Proc. of ACM MOBICOM*, Atlanta, GA, Sept. 2002.
- [17] R. Kravets and P. Krishnan. Power management techniques for mobile communication. In *Proc. of ACM MOBICOM*, Dallas, TX, 1998.
- [18] R. Kravets and P. Krishnan. Application-driven power management for mobile communication. *Wireless Networks*, 6(4):263–277, 2000.
- [19] S. Lee, S. Banerjee, and B. Bhaattacharjee. The case for a multi-hop wireless local area network. In *Proc. of IEEE INFOCOM*, Hong Kong, China, Mar. 2004.
- [20] S. Lu, V. Bharghavan, and R. Srikant. Fair scheduling in wireless packet networks. *IEEE/ACM Transactions on Networking*, 7(4):473–489, 1999.
- [21] H. Luo, R. Ramjee, P. Sinha, L. Li, and S. Lu. UCAN: a unified cellular and ad-hoc network architecture. In *Proc. of ACM MOBICOM*, San Diego, CA, Sept. 2003.
- [22] L. S. C. of the IEEE Computer Society. Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. Technical Report IEEE Std 802.11, the Institute of Electrical and Electronics Engineers, Inc., 1999.
- [23] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press, Cambridge, MA, 1994.
- [24] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. Opportunistic media access for multirate ad hoc networks. In *Proc. of ACM MOBICOM*, Atlanta, GA, Sept. 2002.
- [25] E. Shih, P. Bahl, and M. Sinclair. Wake on wireless: An event driven energy saving strategy for battery operated devices. In *Proc. of ACM MOBICOM*, Atlanta, GA, Sept. 2002.
- [26] M. Stemm and R. Katz. Measuring and reducing energy consumption of network interfaces in hand-held devices. *IEICE Transactions on Fundamentals of Electronics, Communications, and Computer Science, Special Issue on Mobile Computing*, E 80-B(8):1125–1131, 1997.
- [27] G. Tan and J. Gutttag. Time-based fairness improves performance in multi-rate WLANs. In *Proc. of USENIX Annual Technical Conference*, Boston, MA, June 2004.
- [28] N. H. Vaidya, P. Bahl, and S. Gupta. Distributed fair scheduling in a wireless LAN. In *Proc. of ACM MOBICOM*, Boston, MA, Aug. 2000.
- [29] H. Zhu and G. Cao. rDCF: A relay-enabled medium access control protocol for wireless ad hoc networks. In *Proc. of IEEE INFOCOM*, Miami, FL, Mar. 2005.