This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2005 proceedings.

Energy-Aware Resource Allocation in WLAN Mobile Devices*

Junsung Kim†, Minsu Shin, Sachin Lal Shrestha and Song Chong
quasar@lge.com
†Mobile Communication Technology Lab., LG Electronics, Anyang, Korea
(msshin, sachinls, song)@netsys.kaist.ac.kr
Network Systems Lab. Dept. of EECS KAIST, Taejon, Korea

***Abstract*— This paper focuses on low-power usages of mobile devices in WLAN (Wireless Local Area Network) environments. Recently, the researchers have been concentrating on power issues for long battery life. However, many of them consider only the MAC (Medium Access Control) protocol. In this paper, we develop a resource allocation algorithm for low-power consumption without considering lower layer protocol. This algorithm provides fairness, efficiency and stability by using an optimization framework in mobile stations. We verify the efficiency of our algorithm by simulation using Intel $Centrino^{TM}$ parameters, which shows the efficiency of our algorithm.**

I. Introduction

WLANs, also known as 802.11 [6], are high-speed wireless networks of increasingly popular wireless devices. They behave like a traditional Ethernet without wires. Recently, WLAN hot spots, which provide a wireless Internet environment, have proliferated. For using wireless devices, power control for long battery life is essential. Without it, a node executes and transmits all packets at maximum power. Therefore, the method of low-power consumption is being studied in many fields: low-power consuming CPU technology [4], low-power techniques in wireless transceiver [2], interconnection networks [1], low-power MAC protocol design [8], power control issues in wireless communication, battery technology, etc.

However, advances in battery technology and low-power circuit design alone will not result in the best performance [10]. Neither will the appearance of the most efficient low-power MAC protocol. For example, $Centrino^{TM}$ succeeded in $Sonoma^{TM}$ achieving low-power consumption by an efficient hardware combination of the Pentium M, 855 Chipset family and a Wireless Pro Network Connection. Further, there is no algorithm that manages overall flow rates and associated applications while also considering battery power. At the system level of a wireless device as one node, there are three key concerns: Computational power resources, Link capacity resources, and Power resources. In this paper, we show that effective control of packet flow and application processing can also manage power resources well. We develop a resource allocation algorithm to save power and use mobile devices efficiently.

This research was supported by the Ministry of Information and Communication, Korea, under the grant for BrOMA-ITRC program supervised by IITA.

II. Power Model

In this section, we describe the power model that we consider. We first provide a brief overview of the CPU power model and then describe the transceiver power model.

A. CPU Power Model

For a given working point of core voltage V and frequency F, well over 90% of the total power dissipation for CMOS microprocessors can be approximated with the following equation [4]:

$$P = \alpha \cdot C \cdot V^2 \cdot F \tag{1}$$

where α is the activity factor, P is the power consumption and C is the effective capacitance for a given design. Reducing any of the terms in this equation will lower the overall processor power consumption and extend battery life. In Eq. (1), the frequency term represents potential areas of savings if the processor can match its active operating level to the performance requirements of the application. To achieve this goal, companies such as Intel, AMD, and Transmeta use processor architecture that manages dynamically the frequency and voltage levels discretely at runtime to meet the needs of a given application. In this paper, we assume for the analysis that processor architecture can dynamically manage frequency continuously.

B. Wireless Transceiver Power Model

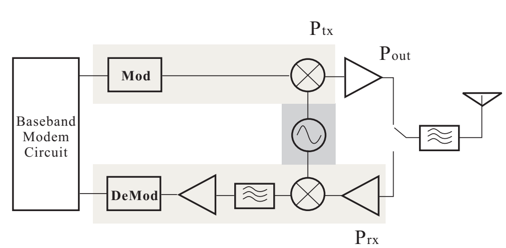


Fig. 1. Block Diagram of a Wireless Transceiver.

The wireless transceiver [2] is composed of many modules, such as the VCO(Voltage Controlled Oscillator), frequency

 0-7803-9415-1/05/$20.00 © 2005 IEEE

synthesizer, and mixers. To analyze the radio power consumption for the resource allocation, these modules are categorized into three components: transmitter, output power amplifier and receiver, as shown in Figure 1. We assume that there is no variation in power consumption by the baseband modem circuit, according to packet streaming. Note that the transmitter will be regarded as the modulator part of the radio (i.e., mixer, frequency synthesizer) and excludes the output power amplifier stage. The output amplifier stage is decoupled from the transmitter because its power consumption is determined primarily by the 802.11 standard and channel environment. The average power consumption of the radio can be described by the following equation:

$$\begin{aligned} Power &= N_{tx}[P_{tx}(T_{on-tx}+T_{start})P_{out}T_{on-tx}] \\ &\quad +N_{rx}P_{rx}(T_{on-rx}+T_{start}) \end{aligned} \quad (2)$$

where $N_{tx/rx}$ is the average number of packets that a transceiver receives/transmits during one time unit, $P_{tx/rx}$ is the power consumption of the transmitter/receiver, P_{out} is the output transmit power, $T_{on-tx/rx}$ is the actual data transmission/reception time, and T_{start} is the start-up time of the transceiver. We assume that the start-up times of transmitter and receiver are the same. For an environment monitoring application, $N_{tx/rx}$ depends on the sum of the rates of the sources(applications). Also note that $T_{on-tx} = \frac{L}{R}$, where L is the average length of transmitted packets in bits and R is the data rate in bits per second.

C. Merged Power Model

In the previous two sections, the CPU and wireless transceiver power model were described. To achieve efficient power control we need a power model that contains properties of both the CPU and wireless transceiver.

In Eq. (1), it is assumed that the frequency can only be controlled continuously. Therefore, the equation is

$$P = k \cdot F \quad (3)$$

where k is a constant that denotes αCV^2.

Eq. (2) is the power model of a wireless transceiver. Because we consider only transmit power and on-time in this paper, terms about reception power and $N_{tx}P_{tx}T_{start}$ can be omitted. N_{tx} denotes the number of packets transmitted. It can be expressed as $\frac{\sum x}{L}$, where $\sum x$ is the aggregated rates in bits per second and L is the average packet length in bits. The simplified equation is as follows:

$$\begin{aligned} Power &= N_{tx}(P_{tx}+P_{out})T_{on-tx} = \frac{\sum x}{L}(P_{tx}+P_{out})\frac{L}{C_B} \\ &= \frac{\sum x}{C_B}(P_{tx}+P_{out}) \end{aligned} \quad (4)$$

where C_B is the transmission capacity of that node. The form of the merged power model is simply the summation of Eq. (3) and Eq. (4). Therefore,

$$P_{total} = k \cdot F + \frac{\sum x}{C_B}(P_{tx}+P_{out}) \quad (5)$$

Next, we develop an algorithm to guarantee longer battery lifetime or to save power by using Eq. (5) and by introducing the power boundary P_{max}.

III. Formulation of the Problem

We formulate the problem that we consider in 802.11b WLAN infrastructure mode. For the analysis, it is assumed that a static network topology that has an AP(Access Point) and a node n, and node n is always in active mode. Before we progress to the formulation of the problem, we simplify further the power model, Eq. (5). Consider a WLAN network that consists of nodes that run a set A of applications given by $A = \{1, 2, \cdots, N\}$ with transmission capacity ϵC_{BW}, where ϵ is the factor that reflects a change in transmission capacity in 802.11b WLAN according to the current SNR(Signal-to-Noise Ratio). Generally, since the SNR changes as the distance between a mobile node and the AP varies, we can regard data link rates as a function of the distance. We can see the relation in Figure 2 [9]. Accordingly, ϵ can have four discrete values, since 802.11b WLAN standards define four transmission bit rates(1Mbps, 2Mbps, 5.5Mbps, 11Mbps). In addition, we assume that ϵ does not change frequently.

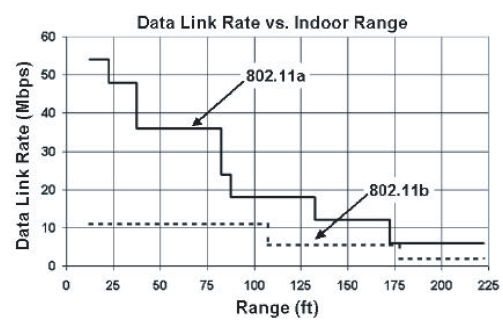


Fig. 2. The relation between data link rate and indoor range

Let w_a [3] denote the per-bit processing demand in unit of cycles per bit of application a. The rate r_a satisfies $m_a \le r_a \le M_a$ in active mode, where m_a and M_a are the minimum and maximum transmission rates of application a, respectively. Then, $w_a r_a$ denotes the quantity of frequency demand during the unit time, and the sum of these values constitutes the total demand for CPU power.

In Section 2, we assumed that processor frequency can be controlled continuously. According to this assumption, we can change Eq. (3) like this:

$$P_{CPU} = k \cdot F = k \cdot \sum_{a \in A} w_a r_a \quad (6)$$

Eq. (4) should be redefined following current notation. Then,

$$P_{Radio} = (P_{tx}+P_{out})\frac{\sum_{a \in A} r_a}{\epsilon C_{BW}} \quad (7)$$

Therefore, P_{Total}, Eq. (5), is expressed as follows by using

Eq. (6) and Eq. (7).

$$P_{Total} = P_{CPU} + P_{Radio} = \sum_{a \in A}(kw_a + \frac{P_{tx} + P_{out}}{\epsilon C_{BW}})r_a = \sum_{a \in A} p_a r_a \quad (8)$$

where p_a has properties similar to w_a. Let p_a denote the per-bit power demand in units of J/sec(watt) of application a, and $p_a = kw_a + \frac{P_{tx}+P_{out}}{\epsilon C_{BW}}$. Therefore, note that the per-bit power demand can differ depending on applications, protocols, and transmitted bits.

A. Primal Problem

Consider a node that has a processing capacity C_P, a transmission capacity ϵC_{BW}, and a power limit for guaranteed lifetime P_{max}. The units of C_P, C_{BW}, and P_{max} are cycles/sec, bits/sec, and Joule/sec, respectively. Let $I_a = [m_a, M_a]$ denote the range in which application source rate r_a must lie. U_a $R_+ \to R$ is a utility function. We associate each application $a \in A$ with U_a. The application a attains a utility $U_a(r_a)$ when it transmit at rate r_a that satisfies $m_a \leq r_a \leq M_a$. We assume that in the interval $I_a = [m_a, M_a]$, the utility functions U_a are increasing, strictly concave, and twice continuously differentiable, and the curvatures of U_a are bounded away from zero on I_a, i.e., $-U_a^{''}(r_a) \geq 1/\alpha_a > 0$ for all $r_a \in I_a$ where $U_a^{''}(r_a)$ denotes the second derivatives of U_a with respect to r_a.

Our objective is to find a rate vector $r = [r_1, \cdots, r_a]^T$ such that

$$\mathbf{P} : \quad \max_{r_a \in I_a} \sum_{a \in A} U_a(r_a) \quad (9)$$

$$\text{subject to} \quad \sum_{a \in A} w_a r_a \leq C_P \quad (10)$$

$$\sum_{a \in A} r_a \leq \epsilon C_{BW} \quad (11)$$

$$\sum_{a \in A} p_a r_a \leq P_{max} \quad (12)$$

The processing capacity constraint, Eq. (10), means that the aggregate processing load (in units of cycles per seconds) of all applications does not exceed the capacity C_P. The transmission capacity constraint Eq. (11) has a meaning similar to that of the processing capacity constraint. The power limit Eq. (12) guarantees the lifetime of mobile devices. Using the assumptions above, there exists a unique maximizer that is the optimal solution for the primary problem because the objective function is strictly concave, and continuous, and the feasible solution set is closed and bounded.

B. Dual Problem

Because the application rates r_a are decoupled by the applications, solving the primal problem directly is difficult in real hardware. Therefore, to solve the problem by a simple method, we use its dual problem in this section. Define the Lagrangian

$$L(r, \lambda, \mu, \nu) = \sum_{a \in A}(U_a(r_a) - r_a(w_a\lambda + \mu + p_a\nu)) + \lambda C_P + \mu\epsilon C_{BW} + \nu P_{max} \quad (13)$$

where λ, μ, ν are Lagrangian multipliers for the processing capacity constraint, the link capacity constraint, and the power limit constraint, respectively, and $p = [\lambda, \mu, \nu]$ denotes the Lagrangian multiplier vector.

The objective function of the dual problem is

$$D(p) = \max_{r_a \in I_a} L(r, p) = \sum_{a \in A} B_a(p^a) + \lambda C_P + \mu\epsilon C_{BW} + \nu P_{max}$$

where

$$B_a(p^a) = \max_{r_a \in I_a} U_a(r_a) - r_a p^a \quad (14)$$

$$p^a = w_a\lambda + \mu + p_a\nu \quad (15)$$

and the dual problem is

$$\mathbf{D:} \quad \min_{p \geq 0} D(p). \quad (16)$$

If we interpret λ, μ, ν as the price per unit processing capacity, the price per unit bandwidth capacity, and the price per managed power, respectively, then p^a is the total cost that an application a must pay when it transmits. Unit bandwidth transmission of application a induces the processing cost $w_a\lambda$ and the bandwidth cost μ, in addition to the power cost $p_a\nu$. Therefore, $r_a p^a$ represents the total cost to application a when it transmits at rate r_a, and $B_a(p^a)$ represents the maximum benefit that application a can achieve at the given price p^a.

For each p^a, a unique maximizer, denoted by $r_a(p^a)$, exists in maximization Eq. (14) since U_a is strictly concave, and is given by

$$r_a(p^a) = [U_a'^{-1}(p^a)]_{m_a}^{M_a} \quad (17)$$

where $U_a'^{-1}$ is the inverse of U_a', that exists over the range $[U_a'(M_a), U_a'(m_a)]$, since U_a is strictly concave and U_a' is continuous. In fact, $r_a(p^a)$ is the demand function in microeconomics. Let $r(p) = [r_1(p^1), r_2(p^2), \cdots, r_A(p^A)]^T$. By duality theory [5], there exists a dual optimal price vector $p^* \geq 0$ such that $r(p^*)$ is primal optimal. Hence, once we obtain p^* by solving the dual problem Eq. (16), the primal optimal source rates r_a^* can be computed by individual applications a separately by using their own demand function in Eq. (17). In this paper, we use a logarithmic function for all utility functions, because $\log x$ satisfies all the assumptions made above.

IV. Iterative Algorithm

In order to solve the dual problem Eq. (16), we use the gradient projection method [5] where prices are adjusted in the opposite direction to the gradient $\nabla D(p)$ as follows:

$$p_i(t+1) = \left[p_i(t) - \gamma \frac{\partial}{\partial p_i} D(p(t))\right]^+ \quad (18)$$

 0-7803-9415-1/05/$20.00 © 2005 IEEE

where γ is a positive step size and $[z]^+ = \max\{z, 0\}$. Since U_a is strictly concave, $D(p)$ is continuously differentiable with partial derivatives, so we obtain the following price update rules for processor, bandwidth, and power.

$$\lambda(t+1) = \left[\lambda(t) - \gamma\left(C_P - \sum_{a\in A} w_a r_a(p^a(t))\right)\right]^+ \quad (19)$$

$$\mu(t+1) = \left[\mu(t) - \gamma\left(C_{BW} - \sum_{a\in A} r_a(p^a(t))\right)\right]^+ \quad (20)$$

$$\nu(t+1) = \left[\nu(t) - \gamma\left(P_{max} - \sum_{a\in A} p_a r_a(p^a(t))\right)\right]^+ \quad (21)$$

This appears to be consistent with the law of supply and demand. That is, for example, if the processing demand, $\sum_{a\in A} w_a r_a(p^a(t))$, at the processor exceeds the processing supply C_P, the processing price, $\lambda(t)$ rises. As with the application algorithm Eq. (17), the price update algorithm suggests treating the processor, the transceiver, the battery and the applications a as processors in a simple computation system to solve the dual problem. In each iteration, applications a individually solve (17) and communicate their results $r_a(p^a)$ to the processor, transceiver, and battery on their own hardware. Then the processor, transceiver, and battery update their prices, according to the price update algorithm, and communicate the new prices to applications a, and the iteration repeats. This algorithmic model is practical if there is an arbiter to make updates at the applications and the other applications synchronize at times $t = 1, 2, \cdots$.

V. System Properties

Fairness and efficiency are two objectives of resource allocation. Fairness means that no users are penalized severely and efficiency means that the resource should be utilized fully. The context of this paper is different from the traditional bandwidth allocation problem; we consider three resources. We first need to determine what should be the allocations that are efficient and fair in this case. First, we define fairness.

Definition 5.1: A feasible rate vector r is said to be bandwidth-proportionally fair if, for any other feasible vector r', the aggregate of proportional changes in terms of bandwidth usage is non-positive $\sum_{a\in A} \frac{r'_a - r_a}{r_a} \le 0$

Definition 5.2: A feasible rate vector r is said to be processor(power)-proportionally fair if, for any other feasible vector r', the aggregate of proportional changes in terms of processor(power) usage is non-positive $\sum_{a\in A} \frac{w_a r'_a - w_a r_a}{w_a r_a} \le 0$ $(\sum_{a\in A} \frac{p_a r'_a - p_a r_a}{p_a r_a} \le 0)$

Theorem 5.1: If all utility functions are logarithmic, $U_a(r_a) = \log r_a$, $a \in A$, the primal optimal solution r^* is bandwidth-proportional fair, processor-proportional fair, and power-fair.

Definition 5.3: An allocation $r = (r_1, r_2, \cdots, r_n)$ is bandwidth-fair if all r_is are the same. Similarly, an allocation r is processor(power)-fair if all $w_i r_i (p_i r_i)$s are the same.

A. Characterization of Optimal Solutions

From the optimization theory, the problem has unique x^*. If we let λ^*, μ^*, and ν^* be the optimal dual variables corresponding to Eq. (10), Eq. (11), and Eq. (12), respectively, then $U'_a(r^*_a) = w_a\lambda^* + \mu^* + p_a\nu^*$, for all a. The values of λ^* μ^* ν^* can be one of the seven cases shown in Table I.

TABLE I
Characterization of Optimal Solutions

Case	Properties of optimal solutions
$\lambda^* > 0$, $\mu^* = 0$, $\nu^* = 0$	CPU fair point, if feasible
$\lambda^* = 0$, $\mu^* > 0$, $\nu^* = 0$	BW fair point, if feasible
$\lambda^* = 0$, $\mu^* = 0$, $\nu^* > 0$	Power fair point, if feasible
$\lambda^* > 0$, $\mu^* > 0$, $\nu^* = 0$	BW, CPU full utilization point, if feasible
$\lambda^* > 0$, $\mu^* = 0$, $\nu^* > 0$	CPU Throughput maximization point, if feasible
$\lambda^* = 0$, $\mu^* > 0$, $\nu^* > 0$	BW Throughput maximization point, if feasible
$\lambda^* > 0$, $\mu^* > 0$, $\nu^* > 0$	The other case

B. Fairness and Efficiency Bound

The proposed algorithm manages the power to control both processing and bandwidth resources simultaneously. It is different from the conventional network flow control algorithm, which controls only bandwidth capacity. The existing DVS(Dynamic Voltage Scaling) algorithm controls CPU clock frequency and core voltage according to CPU load without considering network flows. However, recently developed WLAN devices require an efficient algorithm that controls both resources with limited power. Here, we show how to compare the conventional algorithm with the proposed algorithm. Fairness and throughput are just for bandwidth criteria. In the case of the proposed algorithm, when CPU resources are limited, fairness of CPU usage accompanies a reduction in bandwidth fairness. However, this reduction in fairness and throughput of bandwidth lie within certain bounds, given by the following theorems.

Theorem 5.2: For $\forall i, j$, if $w_i > w_j$ and $p_i > p_j$ for $i > j$, the fairness bound is

$$1 \ge \frac{r_i}{r_j} \ge \min(\frac{w_j}{w_i}, \frac{p_j}{p_i}) \quad (22)$$

Theorem 5.3: The efficiency bound is

$$\min(\frac{C_P}{nC_{BW-max}} \sum \frac{1}{w_i}, \frac{P_{max}}{nC_{BW-max}} \sum \frac{1}{p_i}) \le \frac{\sum r_i}{C_{BW}} \le 1 \quad (23)$$

where n denotes the number of elements of applications, and C_{BW-max} denotes the maximum value of variable C_{BW}.

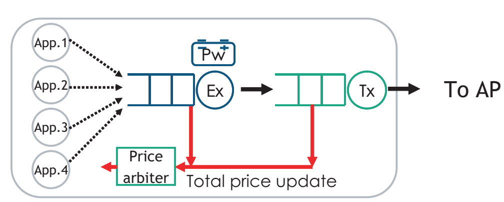


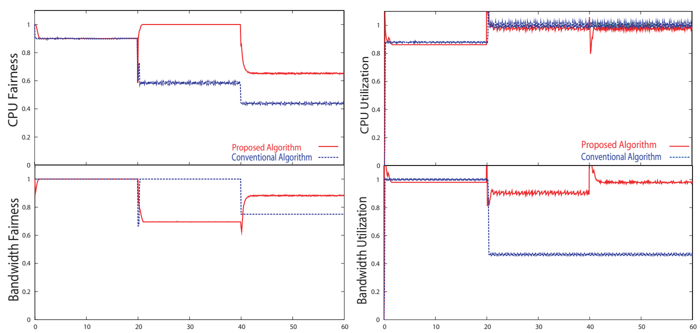
Fig. 3. Proposed System Model

 0-7803-9415-1/05/$20.00 © 2005 IEEE

VI. SIMULATION

In this section, we compare the proposed algorithm with the conventional algorithm, considering only bandwidth resources such as wired networks, because this kind of research has previously not been conducted. This study was done using *ns-2* ver.2.27. Our simulation network consists of one node and one Access Point. Figure.3 shows the system model. We assume that the node uses the INTEL *Centrino Mobile Technology* parameters, which are drawn from the INTEL data sheet on the INTEL website, and the proposed algorithm is applied to this node. From [4], the value of k is 1.5312×10^{-8}. From [7], we know that $P_{tx} + P_{out} = 1.6W$. The WLAN card uses IEEE 802.11b. We assume that the processing boundary of applications using the network capacity is 600MHz.

In this simulation, we define the fairness index as a simulation value, which gives intuitive fairness values. Greater fairness is indicated by the convergence to 1 of the fairness index. The processing demands are different, and those values are in proportion(1,2,4,8). The power demands are determined by using $p_a = kw_a + \frac{P_{tx}+P_{out}}{\epsilon C_{BW}}$. According to the change of link capacity every 20ms, the figures show that the solutions converge well, and that the converged values exist in different optimization ranges.



(a) Fairness (b) Utilization

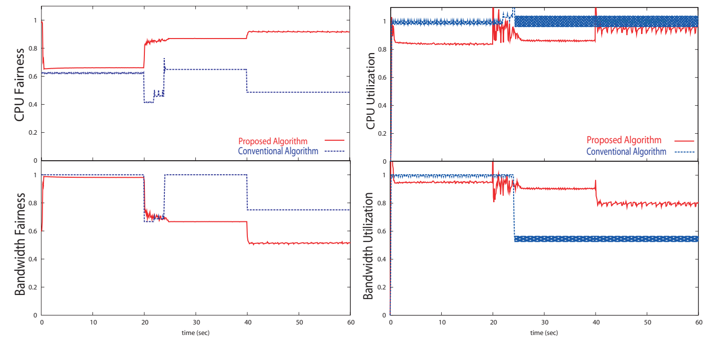
Fig. 4. Without Power Managed Case

Figure.4 represents the fairness and the resource utilization. In this case, there is no power limit. During the first 20ms, bandwidth fairness is achieved, so the fairness index and utilization of bandwidth converge to 1. During the next 20ms, CPU fairness is achieved, so the fairness index and utilization of CPU converge to 1. During the last period, full bandwidth and CPU utilization is achieved, so we obtain maximum utilization of available resources. We lose a little fairness in bandwidth compared to the conventional algorithm, instead.

Figure.5 represents the case in which there exists a power limit. Therefore, the algorithm finds the bandwidth or CPU maximum utilization point under the designated power limit. Figure.6 shows a well-managed power line. Therefore, we can manage the lifetime by setting a power limit.

VII. CONCLUSION AND FURTHER WORK

In this paper, we described an optimization approach to the resource allocation of processor, bandwidth, and power in



(a) Fairness (b) Utilization

Fig. 5. With Power Managed Case

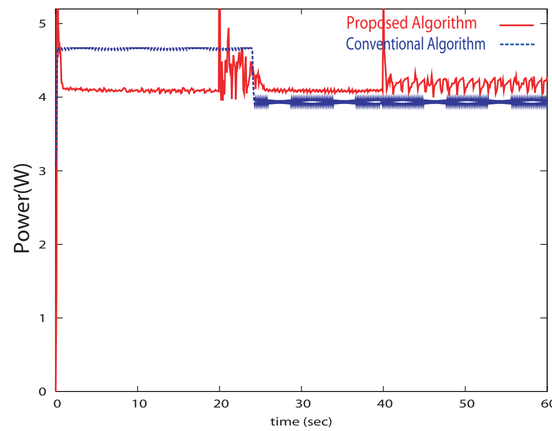


Fig. 6. Power Consumption

WLAN mobile devices. For the proper management of power, we constructed a model for WLAN mobile devices such that a predetermined value for the lifetime of the devices will be guaranteed. In addition to managing and saving power for long battery life, we can use this mechanism to design hardware parameters or predict lifetime. Recently, many devices, such as PDAs, 3G cellular phones, and Smart Phones, have been developed that use many applications that use air connections simultaneously. This algorithm can be applied to any of these devices. With a few modifications, this algorithm also can be applied easily to the ad hoc mode of WLAN.

REFERENCES

[1] Li Shang, etc., *Dynamic voltage scaling with links for power optimization of interconnection networks*, HPCA9, 2003.
[2] SeongHwan Cho, *Energy Efficient RF Communication Systems for Wireless Microsensors*, Ph.D. Thesis, MIT, 2002.
[3] Minsu Shin, *Bandwidth allocation with processing constraint*, in Prep.
[4] Simcha Gochman, etc., *The Intel Pentium M Processor: Microarchitecture and Performance*, Intel Technology Journal, Intel Press, 2003.
[5] S.H. Low and D.E. Lapsley, *Optimization Flow Control, I: Basic Algorithm and Convergence*, IEEE/ACM TON, 1999.
[6] IEEE 802.11b Standard.
[7] Intel PRO/Wireless 2100 LAN MiniPCI Adapters Types 3B page.
[8] E.S Jung and N.H. Vaidya, *A Power Control MAC protocol for ad hoc networks*, ACM MOBICOMM'02, 2002.
[9] James, C. Chen, *Measured Performance of 5 Ghz 802.11a Wireless LAN Systems*, Atheros, 2001.
[10] Jason Flinn and M. Satyanarayanan, *Energy-aware adaptation for mobile applications*, 1999.

 0-7803-9415-1/05/$20.00 © 2005 IEEE