Throughput and Value Maximization in Wireless Packet Scheduling under Energy and Time Constraints *

Fan Zhang and Samuel T. Chanson
Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Email: {zhangfan, chanson}@cs.ust.hk

Abstract

Energy-efficient data transmission is an important issue for mobile communication devices. In this paper, we study throughput and value-based packet scheduling over a wireless link under time and energy constraints. A mobile node is required to deliver multiple data streams to multiple receivers. These data streams may have different importance and incur different energy costs. In this research, point-to-point wireless data transmission is studied. The objective is to develop schemes that selectively transmit data streams at certain power levels so that the data throughput or the value of data delivery is maximized without violating the energy and time constraints. Exploiting the special properties of the throughput maximization problem, we propose an iterative algorithm that yields the optimal transmission schedule. The value maximization problem is formulated as a convex optimization problem and the optimal results can be obtained by solving the Kuhn-Tucker conditions. When time and transmission power are allocated at discrete slots and levels, respectively, a dynamic programming approach is proposed to find the optimal value.

1 Introduction

Mobile and wireless devices are becoming increasingly popular in the workplace and at home. Wireless LANs and sensor networks are widely deployed. With increasing processor computation power and wireless transmission rate, the energy consumption of wireless devices is growing rapidly. As batteries are the primary energy source of mobile devices, increasing the energy-efficiency of computing and communications is of critical importance.

Energy-efficient computing systems have been extensively studied in the literature. For example, processor voltage scaling has been introduced to trade processing speed for lower energy consumption [1–4]. A similar technique has been developed for wireless communications. With proper wireless channel

^{*}The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKUST6178/03E)

coding or modulation schemes, the energy for transmitting a given amount of data can be reduced substantially by lowering the transmission power and increasing the transmission time [5–7].

Energy-efficient scheduling of packets with timing constraints over wireless links has received much attention recently. The objective is usually to minimize energy consumption given a fixed amount of data [5, 6, 8, 9] or to maximize throughput under a given energy constraint [9]. It is known that with proper channel coding the energy required for transmitting any given amount of data is strictly convex and monotonically decreases with transmission time [5]. In other words, sending data over a longer interval saves energy. Prabhakar et al. have studied the energy minimization problem where all packets have a common deadline [5], and they have further investigated the multiple receiver case where transmitting to different receivers incurs different energy costs [6]. In a recent work, it was found that more energy can be saved by exploiting the energy recovery effect of batteries [8]. Energy-efficient packet scheduling under a fading channel is more complicated. As the channel states change with time, the current and future channel states must be considered when allocating transmission power [9, 10].

We consider an ad hoc sensor network where the communication medium (frequency band) is shared among the nodes by time-division multiplexing (TDMA). To resolve collision and minimize energy wastage, medium access control can be achieved by implementing periodic reservation protocols such as the Demand Assigned Multiple Access (DAMA) protocol or using static schemes such as I-EDF which assigns fixed amount of time to each node periodically [11]. In both cases, a node will receive a finite number of time slots, during which it can transmit data, in each fixed interval. As the lifetime of a node is constrained by the battery capacity, only a limited amount of energy should be consumed in each interval (or period). For solar-powered sensor nodes, this energy budget can be the amount of solar energy collected in each period. The node should selectively transmit the data accumulated in its transmit buffer within the time and energy constraints. As data streams destined to different receiving nodes may consume different amount of energy, the energy cost must be considered in maximizing data throughput. Furthermore, if data streams have different priorities and are of different importance (or value), it is also desirable to take the values of the streams into consideration and maximize the total value. Recently, Fu et al. [9] have studied the throughput and value maximization problem for energy constrained data transmission in satellite systems. However, they assumed that the value from successful transmission changed randomly with time and were independent of the streams. The value maximization problem for real-time tasks has been investigated by Rusu et al. [12], but the iterative algorithm they proposed does not give the optimal solution.

In this paper, we study throughput and value maximization problems for point-to-point data transmission over wireless links. Given the value functions and energy cost functions for a set of incoming data streams on the transmitting node, the objective is to maximize the throughput or the total value from successful data transmission under time and energy constraints. Two types of decisions have to be made during scheduling: data selection and power control. For the throughput maximization problem, we analyze the properties of optimal transmission schedules and propose an optimal iterative algorithm exploiting these properties. The algorithm has low complexity and can be implemented efficiently. The value maximization problem can be solved using the Kuhn-Tucker conditions if time and power can be adjusted continuously. When time is slotted and transmission power is adjustable at discrete levels only, we propose a dynamic programming approach to compute the optimal schedule. We found that the optimal value achievable with discrete time and power allocation was close to that for the continuous case. In designing mobile low-power devices, it may be worthwhile to use slotted time and discrete power levels so as to simplify the design of wireless transmitters without sacrificing much performance. Furthermore, we show that the dual problem, that is, energy minimization under time and value constraints can be solved using a similar method.

The rest of the paper is organized as follows: Section 2 introduces the system model. In Section 3 we present an optimal iterative algorithm for the throughput maximization problem. Value-based packet scheduling under time and energy constraints is studied in Section 4. Section 5 addresses the energy minimization problem under time and value constraints. Finally, Section 6 concludes this paper.

2 System model

2.1 Data model

We consider an ad hoc network environment where a single node is transmitting to multiple receivers over a wireless channel. The data to be transmitted are stored in the buffer before transmission. These data are either generated by the local sensor or forwarded from remote nodes, so the buffer may contain a number of different streams with different priorities or values. At any time, the transmitter can only communicate with one receiver. We assume in each period L, T units of time are allocated to the transmitter, where $0 \le T < L$ depending on the MAC layer protocol. Furthermore, let E be the energy budget in each period. The problem is to selectively transmit data to maximize throughput or the transmission value within the time constraint T and energy constraint E. Without loss of generality, we assume there are n data streams destined to m receiving nodes waiting to be transmitted during interval (0, T].

The system receives an increment in value for each successful data transmission. Transmission of partial data is acceptable but is of lower value. Practical examples include transmitting a lower resolution picture instead of a full size one, and selectively sending only part of the sensor data collected. The value received from stream k follows the value function $V_k(s_k)$ where s_k is the number of units of data transmitted. Following common practice, we make the following assumptions about the value function $V_k(s_k)$:

- 1. $V_k(s_k)$ is non-negative.
- 2. $V_k(s_k)$ is continuous and monotonically increases with s_k .
- 3. $V_k(s_k)$ is differentiable anywhere and is concave.

$$R'_{k}(P_{k}) \cdot t_{k} - \mu \cdot t_{k} + \lambda^{1}_{k} \cdot R'_{k}(P_{k}) \cdot t_{k} - \lambda^{2}_{k} \cdot R'_{k}(P_{k}) \cdot t_{k} + \lambda^{3}_{k} - \lambda^{4}_{k} = 0, \qquad (2)$$

$$R_k(P_k) - \mu \cdot P_k - \nu + \lambda_k^1 \cdot R_k(P_k) - \lambda_k^2 \cdot R_k(P_k) = 0, \qquad (3)$$

$$\lambda_k^1 \cdot (d_k - R_k(P_k) \cdot t_k) = 0, \qquad (4)$$

$$\lambda_k^2 \cdot (R_k(P_k) \cdot t_k - D_k) = 0, \qquad (5)$$

$$\lambda_k^3 \cdot P_k = 0, \tag{6}$$

$$\lambda_k^4 \cdot (P_k - P_{\max}) = 0, \qquad (7)$$

For each data stream k, at least d_k units of data must be transmitted. Furthermore, the amount of data transmitted (s_k) for stream k cannot exceed its total amount of data D_k . Throughout this paper, we assume $D_k > d_k > 0$.

2.2 Power model

Using proper channel coding can effectively reduce the energy consumption in wireless communications. Take the Additive White Gaussian Noise (AWGN) channel model for example. With optimal channel coding, the maximum channel capacity is:

$$C = \frac{1}{2}\log_2\left(1 + \frac{P'}{N}\right) \quad \text{bits/transmission},\tag{1}$$

where P' is the received signal power and N is the noise power. Let B be the channel bandwidth, the relationship between transmission rate R and received power P' can be described as:

$$P' = N \cdot \left(2^{\frac{2R}{B}} - 1\right).$$

Considering power attenuation, we have:

$$P = A \cdot P' = A \cdot N \cdot \left(2^{\frac{2R}{B}} - 1\right),$$

where P is the transmission power and A is the *attenuation factor* for the transmitter-receiver pair. The maximum transmission power is denoted as P_{max} . It is easy to see that the required transmission power P is strictly increasing and strictly convex in transmission rate R. When suboptimal channel coding is deployed, the power function still bears the same characteristics [5]. We further assume the power function P(R) is continuous and has continuous first derivatives. The inverse function of the power function, namely the rate function R(P) is assumed to be continuous, strictly increasing and strictly concave in the transmission power.

Note that the power functions (and the rate functions) intersect with each other only at point (0,0) unless the two functions are identical. Consider two receivers that are at different distances from the transmitter. With the same transmission power, the node further away n_F will always have a lower transmission rate. Moreover, to maintain the same transmission rate for the two receivers, a higher transmission power is required to transmit reliably to n_F .

3 Throughput maximization under time/energy constraints

3.1 Problem formulation

Since the receivers are at different distances from the transmitter, the power attenuation at the receivers is generally different from one another. As a result, the power required to maintain reliable transmission at a given speed varies from stream to stream. We denote the power function of data stream k by:

$$P_k(R_k) = A_k \cdot N \cdot \left(2^{\frac{2R_k}{B}} - 1\right)$$

If two streams are destined to the same receiving node, their power functions are identical. It is not necessary to distinguish streams destined to the same receiver in the throughput maximization problem. Therefore, without loss of generality, we assume the n data streams are destined to n distinct receiving nodes. Note that the coefficient A_k may differ depending on the distance between the transmitter and the receiver. The channel states and receiving nodes are assumed to be static within (0, T], so the power attenuation factors $A_k(1 \le k \le n)$ do not change over the period. Also, N and B are constant and are the same for all streams.

In the optimal schedule, the transmission rate throughout the transmission interval of any single data stream is constant [5]. The transmission power for stream k is denoted by P_k . Given t_k units of time, the data throughput of the stream is $s_k = R_k(P_k) \cdot t_k$, where $R_k(P_k)$ is the rate function for stream k. If all streams can transmit their D_k within the time and energy constraints, the problem can be solved using the MoveRightExpress algorithm [6] which will give the minimum energy solution. In the following, we only consider the case in which not all the streams can transmit D_k units of data. In this case, the given time and energy will be fully utilized in the optimal solution. The throughput maximization problem is therefore formulated as follows:

$$\begin{array}{ll} \text{maximize} & \sum_{k=1}^{n} R_k(P_k) \cdot t_k \\ \text{subject to} & \sum_{k=1}^{n} P_k \cdot t_k - E = 0, \\ & \sum_{k=1}^{n} t_k - T = 0, \\ & d_k \leq R_k(P_k) \cdot t_k \leq D_k, \; \forall k, 1 \leq k \leq n \\ & 0 \leq P_k \leq P_{\max}, \; \; \forall k, 1 \leq k \leq n. \end{array}$$

Applying the method of Lagrange and Kuhn-Tucker conditions, we obtain the necessary conditions for the optimal solution: equations (2) — (7) for all $k, 1 \leq k \leq n$, where $\mu, \nu, \lambda_k^1, \lambda_k^2, \lambda_k^3, \lambda_k^4$ are generalized Lagrange multipliers and are non-negative constants [13]. The optimal solution can be obtained by solving equations (2) — (7). However, when the number of streams is large, solving the equations can be very complicated and time-consuming. Fortunately, the optimal solution for the throughput optimization problem possesses a few special properties. In the following, we shall exploit these properties and compute the optimal transmission schedule in a much simpler way.

3.2 Properties of the optimal solution

Consider the streams that are transmitted at power P_k where $0 < P_k < P_{\text{max}}$ in the optimal schedule. According to equations (6) and (7), $\lambda_k^3 = \lambda_k^4 = 0$, so equations (2) and (3) become (as $d_k > 0$, t_k must be positive):

$$R'_{k}(P_{k}) + \lambda_{k}^{1}R'_{k}(P_{k}) - \lambda_{k}^{2}R'_{k}(P_{k}) = \mu, \qquad (8)$$

$$R_k(P_k) - \mu P_k + \lambda_k^1 R_k(P_k) - \lambda_k^2 R_k(P_k) = \nu, \qquad (9)$$

respectively.

We consider the following two cases:

1) $\mu = 0$. This case is further divided into two subcases:

1a) At least one stream is transmitted at power P_k , $0 < P_k < P_{\text{max}}$. According to equation (8), we have:

$$(1 + \lambda_k^1 - \lambda_k^2) \cdot R'_k(P_k) = 0.$$

As $R_k(P_k)$ is strictly increasing, $R'(P_k) \neq 0$. Therefore, $(1 + \lambda_k^1 - \lambda_k^2) = 0$. As $\lambda_k^1, \lambda_k^2 \geq 0, \lambda_k^2$ must be non-zero. From equation (5), we have $s_k = R_k(P_k) \cdot t_k = D_k$. Furthermore, from equation (9) we get:

$$\nu = (1 + \lambda_k^1 - \lambda_k^2) \cdot R_k(P_k) - \mu \cdot P_k = 0.$$

For all other streams, that is, $P_i = P_{\text{max}}$ (as $d_i > 0$, P_i cannot be 0), equation (3) becomes

$$(1 + \lambda_i^1 - \lambda_i^2) \cdot R_i(P_i) = 0.$$

As $P_i = P_{\text{max}} > 0$, $(1 + \lambda_i^1 - \lambda_i^2) = 0$, which indicates $\lambda_i^2 > 0$. From equation (5), we will get $s_i = R_i(P_i) \cdot t_i = D_i$.

In summary, we have $s_k = D_k$, $(\forall k, 1 \le k \le n)$. As mentioned previously, the optimal transmission schedule in this case can be obtained using the MoveRightExpress algorithm [6].

1b) All streams are transmitted at power P_{max} . In this case, $P_{\text{max}} \cdot T = E$. The problem is simplified to a time allocation problem. The optimal solution can be obtained as follows: Every stream is first allocated time such that $d_k(1 \le k \le n)$ units of data can be transmitted. The remaining time will be allocated to the streams by their transmission rates $R_k(P_{\text{max}})$ in descending order. That is, time will be allocated to the stream with the highest transmission rate first until it has enough time to transmit all its D_k units of data. Then the stream with the next highest transmission rate will proceed, and so on. 2) $\mu \ne 0$.



Figure 1: Tangent lines of the rate functions cross the x-axis at the same point.

We first consider the streams transmitted at power P_k where $0 < P_k < P_{\text{max}}$. Dividing both sides of equation (9) by μ , we get:

$$\frac{(1+\lambda_k^1-\lambda_k^2)\cdot R_k(P_k)}{(1+\lambda_k^1-\lambda_k^2)\cdot R'_k(P_k)} - P_k = \frac{\nu}{\mu} \Rightarrow \frac{R_k(P_k)}{R'_k(P_k)} - P_k = \frac{\nu}{\mu}.$$
(10)

Equation (10) reveals a special relationship of the transmission powers in the optimal solution. We use an example to demonstrate this relationship. In the example, three streams are transmitted. Their transmission rate functions are plotted in Figure 1. P_1 , P_2 , $P_3 < P_{\text{max}}$ are the transmission powers in the optimal solution and (P_1, R_1) , (P_2, R_2) and (P_3, R_3) are called the *solution points*. We draw tangent lines to the three rate function curves as shown at their solution points. The slopes of the three lines are $R'_1(P_1)$, $R'_2(P_2)$ and $R'_3(P_3)$, respectively. As $R'_k(P_k) > 0$ ($1 \le k \le 3$), all the lines cross the x-axis. Let the x-coordinates of the cross points be A_1 , A_2 and A_3 , respectively. It is not difficult to see from the figure that:

$$P_k - A_k = R_k / R'_k (P_k), \quad (1 \le k \le 3).$$

From equation (10), we have:

$$A_1 = A_2 = A_3 = -\nu/\mu,$$

that is, all the tangent lines cross the x-axis at the same point. We shall call this point the *control point*, denoted by A.

If the transmission power of a stream is P_{max} in the optimal solution, equations (2) and (3) become:

$$R'_{k}(P_{k})t_{k} - \mu t_{k} + \lambda_{k}^{1}R'_{k}(P_{k})t_{k} - \lambda_{k}^{2}R'_{k}(P_{k})t_{k} - \lambda_{k}^{4} = 0,$$

$$R_{k}(P_{k}) - \mu P_{k} - \nu + \lambda_{k}^{1}R_{k}(P_{k}) - \lambda_{k}^{2}R_{k}(P_{k}) = 0,$$

respectively. Based on these two equations, we have:

$$\frac{R_k}{R'_k(P_k)} - P_k = \frac{\nu t_k - \lambda_k^4 P_k}{\mu t_k + \lambda_k^4} \le \frac{\nu}{\mu}.$$
(11)



Figure 2: If the power value of the tangent point exceeds P_{max} , then P_{max} is the solution.

This means if we draw the tangent line at point $(P_{\text{max}}, R_k(P_{\text{max}}))$, the line will cross the x-axis at or to the right of the control point. In other words, if we draw a tangent line to the rate function curve from the control point $(-\nu/\mu, 0)$, the x-coordinate of the tangent point will be larger than or equal to P_{max} (see streams 2 and 3 in Figure 2 for example).

The two special cases when $\mu = 0$ can be easily identified and tackled with existing techniques. In the rest of this subsection, we focus on the more general case when $\mu \neq 0$.

Property 1 For any optimal transmission schedule, there exists a control point $A(-\nu/\mu, 0)$ on the xaxis such that the optimal transmission power and rate values for all streams are on the tangent lines from this point to the corresponding rate function curves. However, if the power value of a tangent point exceeds P_{max} , the optimal transmission power for the corresponding stream is P_{max} .

Property 1 reveals some characteristics of the transmission powers in the optimal transmission schedule. Next we shall study the relation between the amounts of data transmitted for different streams.

We draw lines L_k $(1 \le k \le n)$ between the control point $A(\nu/\mu, 0)$ and every optimal solution point (P_k, R_k) . If the data transmitted in stream k is more than d_k but less than D_k , then by equation (3) the slope of L_k is:

$$\frac{R_k}{\nu/\mu + P_k} = \frac{\mu P_k + \nu}{\nu/\mu + P_k} = \mu.$$

If stream *i* transmits D_i units of data, the slope of L_i is:

$$\frac{R_i}{\nu/\mu + P_i} = \frac{\mu P_i + \nu + \lambda_i^2 R_i(P_i)}{\nu/\mu + P_i} = \mu + \frac{\mu \lambda_i^2 R_i(P_i)}{\nu + \mu P_i} \ge \mu.$$

Similarly, if stream j transmits d_j units of data in the optimal solution, the slope of L_j is:

$$\frac{R_j}{\nu/\mu + P_j} = \frac{\mu P_j + \nu - \lambda_j^1 R_j(P_j)}{\nu/\mu + P_j} = \mu - \frac{\mu \lambda_j^1 R_j(P_j)}{\nu + \mu P_j} \le \mu.$$

As a stream must fall into one and only one of the above categories, we have the following property:



Figure 3: The order of the slopes does not change with the control point.

Property 2 Given the optimal solution, draw lines from the control point to all the solution points (P_k, R_k) $(1 \le k \le n)$. If the transmitted data volume for stream k is larger than d_k and smaller than D_k , the slope of L_k is equal to μ . If the slope of L_i is larger than μ , in order to obtain the optimal schedule, stream i should transmit D_i units of data. Otherwise, if the slope of L_j is smaller than μ , then stream j should transmit d_j units of data.

As the rate functions of all streams intersect only at point (0,0) (see Section 2.2), the ordering of the slopes of L_k is equal to the ordering of $R'_k(0)$ $(1 \le k \le n)$ and does not change with the control point. For example, in Figure 3, the slope of stream 1 is always larger than those of the other two streams despite the position change of the control point. This claim still holds if some streams are transmitted at P_{max} . According to Property 2, a stream is allowed to transmit more than its minimum data volume d_k only if all the streams with larger slopes have transmitted all their data. Based on these observations, an iterative algorithm is presented in the next section that yields the optimal transmission schedule.

3.3 An iterative algorithm

We propose an iterative algorithm to tackle the throughput maximization problem. This algorithm does not address the special cases when $T \cdot P_{\max} \leq E$ or $s_k = D_k(\forall k, 1 \leq k \leq n)$, both of which can be solved using the methods described in Section 3.2. In the following discussion, we assume at least one stream will be transmitted at a power level lower than P_{\max} and not all streams transmit D_k units of data in the optimal solution.

By the two properties presented in Section 3.2, given the energy and time constraints, there exists a control point on the x-axis corresponding to the optimal transmission schedule that maximizes data throughput. In fact, given a control point and a time interval of length T, it is also possible to find the energy E such that the control point corresponds to the optimal solution of the (E, T) constrained throughput maximization problem.

Suppose the streams are sorted in decreasing order of $R'_k(0)$. We use function **Energy**(a,T) to

calculate the required energy given the time constraint T and the control point (a, 0). The procedure is as follows. Given the control point, we can determine the transmission power of each stream by Property 1. Let them be P_1, P_2, \ldots, P_n corresponding to streams 1 to n, respectively. We first allocate time to each stream such that every stream can transmit $d_k(1 \le k \le n)$ units of data. The remaining time is allocated to the streams in decreasing order of $R'_k(0)$. Therefore, a stream will receive additional time only if all streams before it have been allocated adequate time to transmit all their data.

Theorem 1 Suppose the control point (a, 0) and time constraint T are given. If the energy computed by **Energy**(a, T) is E, then (a, 0) is the control point corresponding to the optimal solution of the throughput maximization problem with energy constraint E and time constraint T.

Proof: We prove the claim by contradiction. Suppose the control point corresponding to the (E,T) constrained throughput maximization problem is (b,0) instead of (a,0). As the optimal solution must satisfy the two properties presented in Section 3.2, we consider the following two cases:

1) a < b. As (b, 0) corresponds to the optimal solution, according to the two properties listed in Section 3.2, all streams should transmit more (or at least the same amount of) data compared with the solution corresponding to (a, 0). However, since b > a, the transmission rate of at least one stream is decreased, so it takes longer time to complete data transmission. Note that the solution corresponding to the control point (a, 0) fully utilizes the available time T, so the solution for (b, 0) would violate the time constraint.

2) a > b. Similar to the first case, all streams should transmit more (or at least the same amount of) data compared with the solution corresponding to (a, 0). Furthermore, the transmission rate of at least one stream is increased in the schedule for point (b, 0). As the power function is strictly convex in the transmission rate, the energy consumption of this solution will exceed the energy consumed in the solution corresponding to (a, 0) which is E. Therefore the energy constraint is violated.

By Theorem 1, if the time constraint is fixed, there is a correspondence between the control point and the energy constraint. In the following we show that it is a one-to-one mapping.

Theorem 2 Suppose the data streams and the time constraint are given. The energy computed by function **Energy()** strictly increases as the control point moves towards the left. Similarly, the energy computed by function **Energy()** strictly decreases as the control point moves towards the right.

Proof: As the transmission power of a stream is on the tangent line from the control point to the rate function, when the control point moves to the left, the transmission power of each stream will increase unless it has already reached P_{max} . As the time constraint is fixed, by Property 2, the throughput of at least one stream will increase while the throughput of the other streams will remain the same. Due to the increasing data throughput and transmission power, the total energy consumption would increase. The same reasoning can be applied to the case when the control point moves to the right.

Suppose the energy constraint E and time constraint T are given. To find the optimal solution, we just need to find the control point of this (E, T) constrained problem. We first use the MoveRightExpress

```
Iterative Algorithm:

Input: E, CPTL, CPTR, \varepsilon

Output: CPT

energy = 0;

while |energy - E| > \varepsilon

CPT = (CPTR + CPTL)/2;

energy = \mathbf{Energy}(CPT, T);

if energy < E

CPTR = CPT;

else

CPTL = CPT;

end if

end while
```

Figure 4: The iterative algorithm.

algorithm [6] to compute the minimum energy solutions for the two extreme cases when all streams transmit d_k or $D_k(1 \le k \le n)$ units of data, respectively. Let the corresponding control points¹ be (CPTR, 0) and (CPTL, 0), respectively, where $|CPTR| \le |CPTL|$. The energy consumption in the two cases are denoted by E_d and E_D , respectively. If $E_d > E$, there is no valid solution. If $E_D \le E$ or $E_d = E$, the solution generated by MoveRightExpress is the optimal solution. If $E_d < E < E_D$, the control point corresponding to the (E, T) constrained throughput maximization problem must fall between (CPTR, 0) and (CPTL, 0). The optimal control point can then be found by the iterative algorithm in Figure 4

The algorithm utilizes a bisection method and is guaranteed to converge. If in an iteration, the required energy corresponding to the current CPT exceeds the given energy constraint E, the optimal control point must fall to the right of the current CPT; otherwise the optimal control point must fall to the left. The algorithm terminates when the difference between the energy corresponding to CPT and the energy constraint E is smaller than ε , the termination threshold. After the algorithm terminates, the transmission schedule can be formed easily based on the computed control point (CPT, 0). Note that although the *energy* may not be exactly equal to E, the difference is usually negligible when ε is sufficiently small. Otherwise, the schedule can be adjusted to satisfy the energy constraint as follows. Suppose the streams are sorted in decreasing order of $R'_k(0)$. If *energy* < E, the unused (E - energy) units of energy can be allocated to the first stream k which transmits less than D_k units of data, until it has sufficient energy to transmit D_k units of data or its transmission power reaches P_{max} . The remaining energy can be allocated to the subsequent streams in the same way. If *energy* > E, then (*energy* - E) units of energy should be deducted from the last stream k that transmits more than d_k units of data, until it transmits only d_k units of data. The remaining energy can be deducted from the preceding

¹The optimal solution of the energy minimization problem with given data and time constraints also satisfies Property 1 (see appendix). The control point can be found in the same way.



Figure 5: Optimal data throughput under time and energy constraint (5 receivers, time constraint = 300ms).

streams in the same way.

3.4 Time analysis

As stated previously, time is assigned to the transmitting nodes by the MAC protocol at the beginning of each period L. In order to guarantee successful transmission of the minimum data requirement $(\sum_{i=1}^{n} d_i)$, a certain amount of time can be reserved by the transmitting node instead of obtained by contention. Given the energy budget, calculating the minimum time requirement T^* for the transmitter is analogous to the energy minimization problem under time constraints, and can be solved in a way similar to the MoveRight algorithm [5]. When the data streams of a transmitting node change, the required time T^* can be recalculated and the reservations can be adjusted accordingly. After all the reservations have been made, the remaining bandwidth can be allocated dynamically by bidding or contention periodically.

Consider an individual node. Stream i is entitled to d_i units of data in each period L. If the actual data rate of stream i is below or equal to d_i in each period, then the delay experienced by the stream is at most $L + T^*$, where T^* is the reserved transmission time of the node. The worst case occurs when the data arrive at the beginning of the transmission in the current period, and are eventually transmitted at the end of the transmission in the next period. The out-of-band data will be served in a best-effort manner if the transmitter is allocated extra time besides the reserved transmission time.

3.5 Numerical results

Data transmission in an AWGN channel with 1MHz bandwidth was simulated. The transmission power was continuously adjustable between 0 to 100. Transmitting at power i ($0 \le i \le 100$) would consume i units of energy every millisecond. The signal attenuation factors and the noises at the receivers are generated such that the signal-to-noise ratios (SNR) at the receivers were uniformly between 18db and 30db when the transmission power was 1. Based on formula (1), the corresponding channel capacities were 3 to 5 bits/transmission. The transmission rates were therefore 3 to 5 Mbits/second when the transmission power was 1.

We have conducted simulations with 5 and 10 data streams, destined to 5 and 10 distinct receivers, respectively. The lower bound of the data transmitted for each stream was uniformly distributed between 100Kbits to 400Kbits. The upper bound of a stream was uniformly distributed between 1 to 2 times its lower bound. For the 5-receiver case, the time constraint was set at 300ms. The minimum energy required to transmit the data lower bound and upper bound, computed using the MoveRightExpress algorithm, were 358 and 2241 units, respectively. Therefore, if the energy constraint was lower than 358 units, no feasible solution could exist. On the other hand, if the energy constraint was higher than 2241 units, all data could be transmitted. We therefore varied the energy constraint in our simulation between 400 and 2200 units. The results for the 5-receiver case is shown in Figure 5, the optimal data throughput was a concave function of the energy constraint. The results of the 10-receiver case had the same characteristics. The computation overhead of the iterative algorithm is low. For example, in the 5-receiver case, it took less than 10 milliseconds to find a feasible schedule with a throughput within 0.1% of the optimal value on a 900MHz Athlon processor.

4 Value maximization under time/energy constraints

4.1 Problem formulation

Maximizing the transmission value is much more complicated than maximizing the data throughput. The value of a stream may not be a linear function of the data transmitted. In fact, even for the case where the value functions are linear, that is, $V_k = v_k \cdot R_k \cdot t_k$ where v_k is the value per data unit, because the $v_k \cdot R_k(P_k)$ functions of different streams may intersect at more than one point, it is impossible to maintain a fixed order of the streams based on their slopes as presented in Section 3.2. Therefore, the iterative algorithm given in the last section is not applicable.

Suppose s_k units of data are transmitted for stream k using t_k units of time. The optimization problem is formulated as follows:

maximize
$$\sum_{k=1}^{n} V_{k}(s_{k})$$
subject to
$$\sum_{k=1}^{n} P_{k}(\frac{s_{k}}{t_{k}}) \cdot t_{k} - E \leq 0,$$

$$\sum_{k=1}^{n} t_{k} - T \leq 0,$$

$$s_{k}/R_{k}(P_{\max}) - t_{k} \leq 0, \forall k, 1 \leq k \leq n,$$

$$d_{k} - s_{k} \leq 0, \quad \forall k, 1 \leq k \leq n,$$

$$s_{k} - D_{k} < 0, \quad \forall k, 1 \leq k \leq n,$$
(12)

The objective function $\sum_{k=1}^{n} V_k(s_k)$ is the sum of concave functions, so it is also concave. As $P_k(x)$

is strictly convex, we can prove that the first constraint is convex (see Appendix). The other constraints are linear functions. With a concave objective function and convex constraint functions, the Kuhn-Tucker conditions are both sufficient and necessary for the global maximum to exist [13]. The optimal solution can therefore be obtained by solving the Kuhn-Tucker conditions.

4.2 The discrete case

Due to the large number of constraint functions, it is sometimes computationally prohibitive to find a solution by solving the Kuhn-Tucker conditions. More importantly, the value function and power function of a data stream may not be continuous. For a data stream there may exist several discrete versions with different values from which the transmitter can choose to deliver. On the other hand, the transmission power of a mobile transmitter may be adjustable only at a few discrete levels.

Assume each data stream k contains up to M versions whose sizes are $s_k^1, s_k^2, \ldots, s_k^M$ with corresponding values $v_k^1, v_k^2, \ldots, v_k^M$. The transmission rate is adjustable in N levels $r_k^1, r_k^2, \ldots, r_k^N$ corresponding to the N transmission power levels. Dynamic programming is used to find the optimal solution.

The dynamic programming approach is formulated as follows. Sort the available data streams in any order. The value maximization problem is divided into n stages. At stage k, the amount of data transmitted for stream k and the corresponding transmission power is determined. Let t_k denote the transmission time and r_k be the transmission rate. We use $V_k[T_k, E_k]$ to denote the optimal obtainable value by transmitting data from streams $k, k + 1, \ldots, n$ given residue time T_k and available energy E_k . The optimal total value that can be obtained from all the streams is:

$$V_1[T_1, E_1] = V_1[T, E].$$

The optimal value function for each stage k is:

$$V_k[T_k, E_k] = \max_{0 \le i \le M; \ 0 \le j \le N} \left\{ V_{k+1}[T_k - t_{i,j}, E_k - P_k(r_k^j) \cdot t_{i,j}] + v_k^i \right\}, \quad 1 \le k \le n-1,$$

where $t_{i,j} = s_k^i / r_k^j$ is the required time to transmit s_k^i units of data at rate r_k^j . The maximization is taken over the size of the transmitted data s_k^i and the transmission rate r_k^j .

The value function at the final stage, that is, the boundary condition is:

$$V_n[T_n, E_n] = \max_{0 \le i \le M; \ 0 \le j \le N} \left\{ v_n^i \mid t_{i,j} \le T_n \text{ and} P_n(r_n^j) \cdot t_{i,j} \le E_n, \text{ where } t_{i,j} = s_n^i / r_n^j \right\}.$$

The data for transmission is chosen subject to the time and energy constraints (T_n, E_n) . The optimal total value can be obtained by iteratively computing V_n, V_{n-1} up to V_1 .

The computation complexity can be derived as follows. For any given (k, T_k, E_k) , the time complexity of computing the optimal $V_k[T_k, E_k]$ is O(MN). For each stage k, as there are at most TE different system states, the time complexity is multiplied by TE. The time complexity for all n stages is $O(n \cdot TEMN)$, linear with the number of data streams. As the interval (0,T] under investigation is usually short, the overall computation cost is low.



Figure 6: Optimal value with varying time and energy.

Note that in the dynamic programming formulation, no assumption is made regarding the convexity or concavity of the value and power functions. Furthermore, this approach can easily be modified to solve the throughput maximization problem with discrete time and power.

4.3 Numerical results

We have built simulators to evaluate the performance of the proposed algorithms. The experimental results for the value maximization problem are presented in this section.

The simulator simulated scheduling 20 data streams destined to 5 different receivers. For each stream, the scheduler could choose to transmit one out of four versions — a basic version and three extended versions. The sizes of the extended versions were two, three, and four times that of the basic version, respectively. The sizes of the basic versions were uniformly distributed between 5kb and 10kb. The values of different versions were proportional to their sizes. The value of the basic version was uniformly distributed between 5 and 10.

The AWGN channel with 1MHz bandwidth was simulated. The transmission power was adjustable at 10 discrete levels. Transmitting at power level i ($1 \le i \le 10$) would consume i units of energy in every time slot. The receiving hosts were generated in the same way as in Section 3.5 with comparable signal-to-noise ratios. Each time slot was 1ms.

We have conducted simulations with T = 200, 150 and 100 time slots. The available energy was varied between 300 and 60 units. For each time/energy constraint setting, 100 distinct sets of data streams were generated and the average achieved value was taken. The results are shown in Figure 6. Given the same amount of energy, a higher value can be obtained when the time is increased. As the usable energy decreases, the value drops. Furthermore, as the number of energy units decreases to below the number of time slots, the value curve drops sharply. In this case, as the lowest power level is 1, not all available time slots can be used for transmission. For example, if 70 units of energy are available, then at most 70 time slots can be used for data transmission. The remaining time slots have to be left idle. In Figure 6, the three curves eventually coincide as the energy decreases. As energy becomes the



Figure 7: Comparison of continuous and discrete time/power allocation.

bottleneck, the number of available time slots would not make a difference.

The dynamic programming approach is designed for the case when time is slotted and transmission power levels are discrete. If both time and power can be allocated continuously, a higher value can usually be obtained. To study the impact on performance due to the discreteness assumption, we simulated the continuous case using the optimization toolbox of Matlab where the transmission power was continuously adjustable between 1 and 10. As shown in Figure 7, the optimal values in the continuous case are only slightly higher than those in the discrete case.

5 Energy minimization under time/value Constraints

5.1 Problem formulation

In the previous section, the value maximization problem under energy and time constraints was investigated. In the following, we study how to minimize energy consumption subject to time and value constraints.

This problem is analogous to the value maximization problem. Let s_k and t_k be the transmitted data size and transmission time for data stream k respectively, and V be the value that should be achieved. The energy minimization problem can be formulated as follows.

The objective is to minimize:

$$\sum_{k=1}^{n} P_k(\frac{s_k}{t_k}) \cdot t_k.$$

The value constraint can be written as:

$$V - \sum_{k=1}^{n} V_k(s_k) \le 0.$$

Assume the value function $V_k(s_k)$ is concave, $V - \sum_{k=1}^n V_k(s_k)$ is convex in s_k . The other constraints are the same as in (12) and are linear. As both the objective function and the constraint functions are convex, the Kuhn-Tucker conditions are sufficient and necessary for a global minimum to exist.

5.2 The discrete case

Similar to its dual problem, if time and power are allocated in discrete slots and levels, dynamic programming can be used to solve the energy minimization problem.

Assume data stream k contains up to M versions and the transmission power is adjustable in N levels. The dynamic programming problem can be formulated as follows. Sort the data streams in any order. The problem is divided into n stages. At stage k, the amount of data to be transmitted for stream k and the transmission time are determined. We use $E_k[T_k, V_k]$ to denote the minimum energy required to obtain a total value of V_k from streams $k, k + 1, \ldots, n$ within time T_k . The minimum total energy required is:

$$E_1[T_1, V_1] = E_1[T, V].$$

The minimum energy function for each stage k is:

$$E_k[T_k, V_k] = \min_{0 \le i \le M; \ 0 \le j \le N} \left\{ E_{k+1}[T_k - t_{i,j}, V_k - v_k^i] + P_k(r_k^j) \cdot t_{i,j} \right\}, \quad 1 \le k \le n-1,$$

where $t_{i,j} = s_k^i / r_k^j$ is the required time interval to transmit s_k^i units of data at rate r_k^j . The energy minimization is taken over all possible data versions and transmission rates.

The energy function at the final stage is:

$$E_n[T_n, V_n] = \min_{0 \le i \le M; \ 0 \le j \le N} \Big\{ P_n(r_n^j) \cdot t_{i,j} \mid t_{i,j} \le T_n \text{ and } v_n^i \ge V_n, \text{ where } t_{i,j} = s_n^i / r_n^j \Big\}.$$

The minimum energy can be obtained by iteratively solving $E_n, E_{n-1}, \ldots, E_1$. The computation complexity is similar to that of the value maximization problem, i.e., $O(n \cdot TEMN)$.

6 Conclusion

In this paper, throughput and value-based wireless packet scheduling under time and energy constraints has been studied. As transmitting data streams to different receivers incurs different amounts of energy and produces different values, two schemes were proposed, one for maximizing the data throughput and the other for the total value of successful transmissions. The throughput maximization problem is solved using an iterative algorithm, exploiting the special characteristics of optimal transmission schedules. The value maximization problem is formulated as a convex optimization problem. The optimal solution can be found by solving the corresponding Kuhn-Tucker conditions. Furthermore, for the case when time and power are allocated at discrete slots and levels respectively, we have developed a dynamic programming approach that computes the optimal transmission schedule efficiently. Simulation results have shown that allocating time and power at discrete levels only slightly affects performance adversely compared to the continuous case.

We have also studied the dual problem of the value maximization problem, that is, minimizing energy while achieving a given value within a given time constraint. Similarly, the Kuhn-Tucker conditions can be used in the continuous case while the dynamic programming approach can be applied in finding the optimal solution in the discrete case.

7 Acknowledgement

The authors of this paper would like to thank Prof. Chang-Gun Lee and the anonymous reviewers for their valuable comments which greatly helped to improve the quality of the paper.

References

- F. Zhang and S. T. Chanson. Processor voltage scheduling for real-time tasks with non-preemptible sections. In *Proceedings of IEEE Real-Time Systems Symposium (RTSS)*, pages 235–245, December 2002.
- [2] H. Aydin, R. Melhem, D. Mosse, and P. Mejia-Alvarez. Dynamic and aggressive scheduling techniques for power-aware real-time systems. In *Proceedings of IEEE Real-Time Systems Symposium* (*RTSS*), pages 95–105, December 2001.
- [3] P. Pillai and K. G. Shin. Real-time dynamic voltage scaling for low-power embedded operating systems. In *Proceedings of the 18th ACM Symposium on Operating Systems Principles*, pages 89–102, 2001.
- [4] I. Hong, D. Kirovski, G. Qu, M. Potkonjak, and M. B. Srivastava. Power optimization of variablevoltage core-based systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits* and Systems, 8(12):1702–1714, 1999.
- [5] B. Prabhakar, E.U. Biyikoglu, and A. El Gamal. Energy-efficient transmission over a wireless link via lazy packet scheduling. In *INFOCOM'01*, pages 386–394, April 2001.
- [6] A. El Gamal, C. Nair, B. Prabhakar, E.U. Biyikoglu, and S. Zahedi. Energy-efficient scheduling of packet transmissions over wireless networks. In *INFOCOM'02*, pages 1773–1782, June 2002.
- [7] C. Schurgers, V. Raghunathan, and M.B. Srivastava. Power management for energy-aware communication systems. ACM Transactions on Embedded Computing Systems, 2(3):431–447, August 2003.
- [8] P. Nuggehalli, V. Srinivasan, and R.R. Rao. Delay constrained energy efficient transmission strategies for wireless devices. In *INFOCOM'02*, pages 1765–1772, June 2002.
- [9] A. Fu, E. Modiano, and J. N. Tsitsiklis. Optimal energy allocation for delay-Constrained data transmission over a time-varying channel. In *INFOCOM'03*, pages 1095–1105, April 2003.
- [10] M. Goyal, A. Kumar, and V. Sharma. Power constrained and delay optimal policies for scheduling transmission over a fading channel. In *INFOCOM'03*, pages 311–320, April 2003.
- [11] Marco Caccamo, Lynn Y. Zhang, Lui Sha, and Giorgio Buttazzo. An implicit prioritized access protocol for wireless sensor networks. In *Proceedings of IEEE Real-Time Systems Symposium* (*RTSS*), pages 39–48, December 2002.

- [12] C. Rusu, R. Melhem, and D. Mosse. Maximizing rewards for real-time applications with energy constraints. To appear in ACM Transactions on Embedded Computing Systems, 2003.
- [13] L.R. Foulds. Optimization Techniques: An Introduction. Springer-Verlag, 1981.

Appendix

Claim: The optimal solution of the energy minimization problem with given data and time constraints satisfies Property 1 in Section 3.2.

Proof: The energy minimization problem can be formulated as follows:

minimize
$$\sum_{k=1}^{n} P_k \cdot t_k$$

subject to
$$\sum_{k=1}^{n} t_k - T = 0,$$
$$R_k(P_k) \cdot t_k = s_k, \ \forall k, 1 \le k \le n,$$
$$0 \le P_k \le P_{\max}, \ \forall k, 1 \le k \le n,$$
(13)

where $s_k(1 \le k \le n)$ are the data to be transmitted. Using the Lagrange method and Kuhn-Tucker conditions, we get the following necessary conditions for the optimal solution:

$$t_k + \lambda_k^1 \cdot R'_k(P_k) \cdot t_k - \lambda_k^2 + \lambda_k^3 = 0$$
$$P_k + \mu + \lambda_k^1 \cdot R_k(P_k) = 0$$
$$\lambda_k^2 \cdot P_k = 0$$
$$\lambda_k^3 \cdot (P_k - P_{\max}) = 0.$$

Similar to Section 3.2, we can get:

$$R_k(P_k)/R'_k(P_k) - P_k = \mu, \quad (0 < P_k < P_{\max}), \text{ and}$$

 $R_k(P_k)/R'_k(P_k) - P_k \leq \mu, \quad (P_k = P_{\max}).$

Therefore the optimal solution satisfies Property 1 and the control point is $(-\mu, 0)$.

Claim: $\sum_{k=1}^{n} P_k(\frac{s_k}{t_k}) \cdot t_k - E$ is convex in s_k and t_k .

Proof: In order to prove $\sum_{k=1}^{n} P_k(\frac{s_k}{t_k}) \cdot t_k - E$ is convex, it is sufficient to prove each individual $P_k(\frac{s_k}{t_k}) \cdot t_k$ (k = 1, ..., n) is convex. For simplicity, we rewrite the above function as:

$$f(x,y) = P\left(\frac{x}{y}\right) \cdot y.$$

We prove f(x, y) is convex by evaluating the corresponding Hessian matrix. Calculating the second derivatives of f(x, y), we get its Hessian matrix:

$$H = \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{y}P''\left(\frac{x}{y}\right) & -\frac{x}{y^2}P''\left(\frac{x}{y}\right) \\ -\frac{x}{y^2}P''\left(\frac{x}{y}\right) & \frac{x^2}{y^3}P''\left(\frac{x}{y}\right) \end{pmatrix}.$$

The leading principal minor of the Hessian is $f_{xx}(x, y) > 0$ and the determinant |H| = 0, so the Hessian matrix is positive semidefinite. Therefore, f(x, y) is convex.