Novel Packet-Level Resource Allocation with Effective QoS Provisioning for Wireless Mesh Networks

Ho Ting Cheng, *Student Member, IEEE*, and Weihua Zhuang, *Fellow, IEEE*

***Abstract*—Joint power-subcarrier-time resource allocation is imperative for wireless mesh networks due to the necessity of packet scheduling for quality-of-service (QoS) provisioning, multi-channel communications, and opportunistic power allocation. In this work, we propose an efficient intra-cluster packet-level resource allocation approach. Our approach takes power allocation, subcarrier allocation, packet scheduling, and QoS support into account. The proposed approach combines the merits of a Karush-Kuhn-Tucker (KKT)-driven approach and a genetic algorithm (GA)-based approach. It is shown to achieve a desired balance between time complexity and system performance. Bounds for the throughputs obtained by real-time and non-real-time traffic are also derived analytically.**

***Index Terms*—Genetic algorithm (GA), Karush-Kuhn-Tucker (KKT), quality-of-service (QoS) provisioning, resource allocation, wireless mesh network (WMN).**

I. Introduction

Wireless mesh networking has emerged as a promising technology for future broadband wireless access [1]. Wireless mesh networks (WMNs) generally comprise gateways, mesh routers, and mesh clients, organized in a three-tier hierarchical architecture [1]. Recently, wireless mesh networking has been attracting significant attention from industry and academia. This networking paradigm provides a viable and economical solution for both peer-to-peer applications and Internet access.

In the literature, node clustering is shown to be indispensable for networking stability and system throughput increase in a large wireless network [2], [3]. In a multi-channel environment, channel negotiation can be performed among clusters instead of individual nodes. The notion of frequency reuse, therefore, can be facilitated and the system capacity can be elevated. In addition, collision-free scheduling is feasible within a cluster, facilitating quality-of-service (QoS) provisioning. In this work, we focus on resource allocation within a cluster for the mesh backbone with an orthogonal frequency division multiplexing (OFDM) physical layer.

In order to effectively support multimedia services (e.g., voice, video, and data) with QoS assurance in the medium access control (MAC) layer, bandwidth reservation and hence packet scheduling is imperative. With the OFDM physical layer, subcarrier allocation is necessary to grant diverse transmission rates for various multimedia applications. By assigning different subcarriers to different nodes, simultaneous transmissions are fostered in a cooperative manner so as to increase system capacity. Since different subcarriers experience different channel fading characteristics, power allocation with respect to channel conditions is proven to be crucial. As a consequence, joint power-subcarrier-time resource allocation is vital.

The contribution of our work is two-fold.

- We propose an efficient intra-cluster packet-level resource allocation approach with effective QoS provisioning, Combined-KKT-GA, taking power allocation, subcarrier allocation, and packet scheduling into consideration. The approach combines the merits of a Karush-Kuhn-Tucker (KKT)-driven approach and a genetic algorithm (GA)-based approach.
- Our proposed approach is shown to achieve good system performance in terms of throughput and packet dropping rate. The time complexity of Combined-KKT-GA is low, resulting in a preferred candidate for practical implementation. We also analytically derive the bounds for the throughputs obtained by real-time and non-real-time traffic, validating our simulation results.

II. Related Work

The problem of subcarrier-bit-power allocation for OFDM systems has been researched extensively with respect to the physical layer [4]–[10]. Due to the hardness of the problem, the joint optimization problem is usually decoupled into two subproblems [4]. In [4], the objective function is to maximize the total achievable rate. However, with a Hungarian approach, the complexity is at least on the order of $O(N^3)$, where N is the number of subcarriers. QoS demands and fairness constraints are taken into account in [5], where heuristic schemes are proposed for convex optimization problems (e.g., the signal-to-noise-ratio (SNR) maximization problem). In [6], a heuristic resource allocation scheme is proposed for the uplink orthogonal frequency division multiple access (OFDMA) systems, taking fairness and time-varying fading into consideration. In [7], a two-level resource allocation scheme for downlink OFDMA systems is proposed to achieve Nash bargaining fairness. A Lagrangian-based approach is employed in [8] to solve the total transmit power minimization problem, but the high complexity of the proposed algorithm impedes its practical implementation. In order to reduce the computational complexity, KKT interpretations are popularly employed in designing resource allocation algorithms (e.g., [9], [10] for utility maximization in uplink OFDMA/WiMAX systems). In the preceding work, packet scheduling, however, is not

Manuscript received June 5 2008; revised August 10 2008; accepted September 15 2008. The associate editor coordinating the review of this letter and approving it for publication was Y. Fang.

The authors are with the Centre for Wireless Communications, Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1 (e-mail: {htcheng, wzhuang}@bbcr.uwaterloo.ca).

This work was supported in part by the Natural Science and Engineering Research Council (NSERC) of Canada.

Digital Object Identifier 10.1109/TWC.2009.080739

addressed properly. QoS provisioning is often neglected in the MAC layer. Thus, applying these existing resource allocation schemes directly to the WMNs with heterogeneous packet-level QoS requirements can be ineffective in provisioning QoS. In addition, those schemes derived from the theory of convex optimization may not be efficient or even applicable to non-convex optimization problems.

Regarding the problem of power-subcarrier-time resource allocation with QoS assurance, a KKT-driven resource allocation algorithm is proposed in [11], effectively provisioning QoS at the packet level (e.g., packet dropping rates). This simple KKT-driven algorithm is shown to outperform a random scheme. Unlike some other KKT-driven approaches (e.g., [10]), the proposed KKT-driven approach considers all three resource dimensions (i.e., power, subcarrier, and time) individually in succession. Thus, the throughput performance can be quite far from optimal. Genetic algorithms (GAs) are commonly used as search algorithms for near-optimal solutions [12]. GA-based resource allocation can incorporate all three resource dimensions simultaneously, potentially achieving better throughput performance and provisioning QoS more effectively. Yet, GAs are usually computationally expensive due to a long convergence time. Our solution proposed in this paper, Combined-KKT-GA, combines the features of a KKT-driven approach (i.e., low complexity) and a GA-based approach (i.e., near-optimal performance). This novel Combined-KKT-GA approach is demonstrated as an efficient and effective packet-level resource allocation scheme for WMNs with QoS support.

III. System Model

We consider a WMN divided into a number of clusters (see Fig. 1). We assume that the mesh routers are stationary and the channel gains can be estimated accurately. The cluster structure and hence the clusters can be changed at times for the sake of traffic load distribution variations; addressing this issue, however, is beyond the scope of this research work. One node is selected as a clusterhead in a cluster. The main responsibility of the clusterhead is to provide timing information and perform resource allocation for the active connections in the cluster. Time is partitioned into frames. Each frame is further divided into a beacon slot, a control slot, and a number of DATA slots. The beacon is used to provide timing and cluster information, and broadcast scheduling decisions for the DATA slots. In the control slot, the clusterhead collects the requests from its clustermembers, and announces the resource allocation in the subsequent beacon. Each mesh router is equipped with one omni-directional transceiver, so that it cannot transmit and receive at the same time. A mesh router can be a sender, relay, or receiver at different times. Some mesh routers are directly connected to a gateway, whereas the others are scattered around, forming a multi-hop network. With the help of multi-channel OFDM technology, each router can choose a set of subcarriers for DATA transmissions and/or receptions. In order to focus on intra-cluster resource allocation, we make the following assumptions for presentation clarity: 1) Effective call admission control is in place, such that the QoS requirements of the admitted calls can be satisfied; 2) There is a clustering algorithm in place, so that a set of channels allocated in the cluster of interest can be determined [3]; 3) The co-channel interference follows a Gaussian distribution.

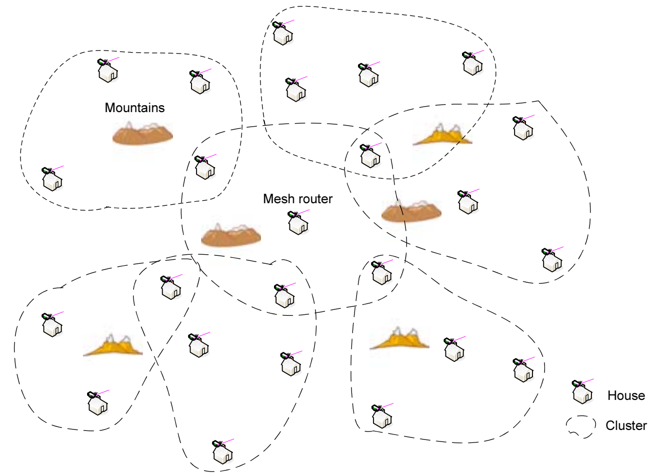

Fig. 1. An illustration of a clustered WMN, where stationary wireless routers are mounted on rooftops of a residential area.

IV. Joint Power-Subcarrier-Time Resource Allocation

There are various system constraints associated with the joint power-subcarrier-time resource allocation problem. The sum of the transmit power of each node on the allocated subcarriers is bounded by a maximum power level. Here, we consider the case where each subcarrier can only be allocated to one transmission link in a cluster. Different traffic types require different packet transmission rates. In our problem formulation, we take the instantaneous rate requirements of different traffic types (e.g., voice, video, and data) in the current frame, if any, into account. QoS provisioning is to be handled by MAC-layer packet scheduling (discussed later). Given the allocation of transmit power, subcarriers, and timeslots, the achievable transmission rate is computed using the Shannon capacity formula. With the above constraints, we employ the well-known utility maximization framework to abstract the objective. The objective function can be generalized to optimize system throughput, fairness, or tradeoffs among several system performance metrics (e.g., a tradeoff between throughput and fairness [13]). In this work, the objective function is chosen to maximize the system throughput.

Problem Formulation: Let M, N, and L denote the number of active links in a cluster, the number of subcarriers available in a cluster, and the number of timeslots (i.e., DATA slots) in a frame, respectively. Consider the following resource allocation optimization problem

$$\max_{\mathbf{a},\mathbf{p}} \left\{ \sum_{m=1}^{M} U_m(R_m(\mathbf{a},\mathbf{p})) \right\} \quad (1)$$

$$\text{subject to } R_m(\mathbf{a},\mathbf{p}) \geq R_m^d, \forall_m \quad (2)$$

$$p_{m,n}^l \geq 0, \forall_{m,n,l} \quad (3)$$

$$\sum_{n=1}^{N} p_{m,n}^l = P_m^{\max}, \forall_{m,l} \quad (4)$$

$$\sum_{m=1}^{M} a_{m,n}^{l} = 1, \forall_{n,l} \tag{5}$$

$$a_{m,n}^{l} \in \{0,1\}, \forall_{m,n,l} \tag{6}$$

where $U_m(.)$ is the utility function of the m^{th} link, $R_m(\mathbf{a},\mathbf{p}) = \sum_{l=1}^{L}\sum_{n=1}^{N} a_{m,n}^{l} \ln\left(1 + g_{mm,n}^{l} p_{m,n}^{l}\right)$ represents the actual aggregate transmission rate of the m^{th} link over a frame with $\mathbf{a} = [a_{m,n}^{l}]_{M\times N\times L}$, $\mathbf{p} = [p_{m,n}^{l}]_{M\times N\times L}$, and $g_{mk,n}^{l} = \varphi G_{mk,n}/(I_n + \eta)$, $a_{m,n}^{l}$ is the indicator of allocating the n^{th} subcarrier to the m^{th} link on the l^{th} timeslot, $p_{m,n}^{l}$ is the transmit power over the n^{th} subcarrier of the m^{th} link's transmitter on the l^{th} timeslot, R_m^d is the instantaneous transmission rate demand of the m^{th} link in the current frame, and $P_m^{\max}$ is the maximum power constraint of the m^{th} link's transmitter. $G_{mk,n}$ is the channel gain from the k^{th} link's transmitter to the m^{th} link's receiver over the n^{th} subcarrier, I_n is the interference power at the n^{th} subcarrier, φ is a BER measure, and η is the background noise power. The optimization problem is NP-hard as shown in Appendix-A.

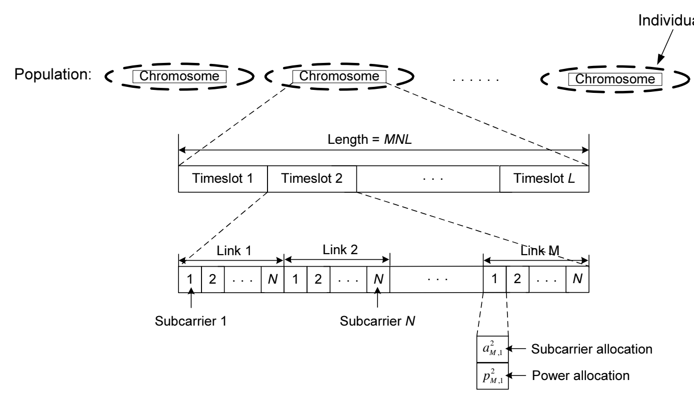

Fig. 2. A genetic representation of a set of feasible solutions to the optimization problem.

V. Combined-KKT-GA Resource Allocation Approach

We first study two resource allocation approaches: 1) Cheng's KKT-driven resource allocation (proposed in [11]) and 2) GA-based resource allocation. The Cheng's KKT-driven algorithm first uses uniform power allocation over all the subcarriers, allocates subcarriers based on the optimal subcarrier allocation criterion [11], re-allocates the subcarriers until all the system constraints are met, and finally performs water-filling for power re-allocation. The time complexity of the Cheng's KKT-driven algorithm is on the order of $O(LMN)$. Despite low complexity, this approach is only suboptimal, because it considers three resource dimensions (i.e., power, subcarrier, and time) individually in succession.

A GA-based resource allocation scheme starts with an initial population of individuals characterized by *chromosomes*, and then improves the quality of the population through evolution. In each generation, three operations are carried out one by one to yield a new population: 1) selection; 2) crossover; and 3) mutation. A chromosome length of MNL is chosen for our GA-based resource allocation scheme. Each individual represents a joint power-subcarrier-time resource allocation solution. To form an initial population, all subcarriers in a frame are allocated randomly to the links in such a way that every individual is a feasible solution to the optimization problem. The fitness function is chosen to be the objective function given in (1). Optimal power allocation is employed in the fitness evaluation of an individual. A genetic representation of a set of feasible solutions is depicted in Fig. 2. For the selection operation, we consider a well-known roulette wheel selection operator [12]. An individual is selected with a probability proportional to its fitness value. Genes for crossover are chosen according to a uniform distribution, whereas only one gene in a chromosome is randomly chosen for mutation. The recommended crossover and mutation probabilities are 0.7 and 0.01, respectively [12]. Based on our observations, the preferred value for the maximum number of iterations, T, and that for the population size, S, are $25,000$ and 100, respectively. The time complexity of the GA-based resource allocation scheme is on the order of $O(TSLMN)$.

Since the GA-based approach considers all three resource dimensions simultaneously, it can plausibly achieve better system performance than the Cheng's KKT-driven approach. However, in practice, running the GA-based resource allocation algorithm with $S = 100$ and $T = 25,000$ on a mesh router is not preferred due to its high computational cost. To strive for a desired balance of system performance and complexity, we propose a scheme named *Combined-KKT-GA* by combining two aforementioned approaches. By studying the performance comparison of the Cheng's KKT-driven scheme and the GA-based scheme with $S = 100$, we have two observations: 1) If the GA-based scheme performs better than the Cheng's KKT-driven one after a few iterations (e.g., 10 iterations), then the former one will give a better resource allocation solution eventually; 2) If the GA-based scheme cannot prevail after a small number of iterations, it will be very likely to take several thousands of iterations to win out. Since the Cheng's resource allocation solution will likely be locally optimal, the rationale for the above observations is mainly due to the quality of the initial population generated in the GA-based algorithm (i.e., the initial set of the feasible solutions) and the structure of the feasible region of our optimization problem (i.e., the solution space governed by the system constraints). If the initial population is residing at a favorable *neighborhood* of the solution space where the path to reaching the global optimum is easy, the GA-based scheme can achieve a better solution in a few iterations with ease. In contrast, if the initial population is residing at a disadvantageous neighborhood of the solution space and/or the solution space contains many various local optimums in between the global optimum and the starting point of search, the GA-based scheme is expected to undergo lots of iterations before reaching the globally optimal point. Since efficiency is one of the most important goals in resource allocation for future WMNs, we propose to combine the Cheng's KKT-driven and GA-based resource allocation schemes as follows.

Step 1: Obtain a resource allocation solution using the Cheng's KKT-driven approach.
Step 2: Obtain a resource allocation solution using the GA-based approach with $S = 100$ and $T = 10$.
Step 3: Choose the better solution out of the above two solutions.

Therefore, on average, the performance of this novel approach, Combined-KKT-GA, is expected to be lower bounded and upper bounded by that of the Cheng's KKT-driven scheme and that of the GA-based scheme, respectively. Note that our proposed approach resembles the notion of *SoftMAC* [14]. SoftMAC selects the best MAC protocol from the protocol pool based on various system parameters such as the channel condition and the packet delay. In our case, the solution of the proposed Combined-KKT-GA approach is highly contingent on the quality of the initially generated population in the GA-based algorithm and the solution space of the optimization problem. How to devise a SoftMAC-like selection strategy is a vital research issue which is, however, left for further research.

VI. Performance Evaluation

We consider a cluster with a number of wireless nodes randomly located in a 1km x 1km coverage area. We assume that the routing is predetermined so that the transmission source and destination pair of an incoming packet is known in advance. Simulation parameters are chosen as follows: $\eta \sim N(0, 10^{-12}\text{W})$, $I_n \sim N(0, 10^{-10}\text{W})$, $\forall_n$, $P_m^{max} = 8\text{mW}$, $\forall_m$, $U_m(R_m(\mathbf{a}, \mathbf{p})) = R_m(\mathbf{a}, \mathbf{p})$, $\forall_m$, $\varphi = 1$, $N = 100$ and $L = 4$. The maximum transmission rate of each subcarrier is set to be 1Mbps. We adopt the channel model suggested in [15]. We employ the traffic models for voice, video, and data, the MAC protocol, and call admission control as described in [11]. Voice traffic and video traffic are generated according to a two-state ON-OFF model. For voice traffic, in the ON state, a fixed-size packet arrives at a constant rate, whereas in the OFF state, no packet is generated. The duration of an ON period and that of an OFF period both follow an exponential distribution. For video traffic, different incoming packets in the ON state have different packet sizes, generating a variable-rate traffic [16]. We consider two data traffic models: 1) data traffic is the background traffic and available anytime; and 2) data packet arrivals follow a Poisson process with mean rate 50 packets/second, where the packet size follows a Weibull distribution (i.e., Weibull(2,2)). Data traffic is assigned the lowest priority and does not have any rate requirements. To resemble mixed traffic types in the mesh backbone, we assume that there are a voice source, a video source, and a data source residing at every node. There are 1 beacon slot, 1 control slot, and 4 DATA slot(s) in a frame. The duration of a timeslot is 5ms, and hence the duration of a frame is 30ms. The polling is done in every 100ms. Both the polling and the beacon packet transmissions are assumed to be error-free. To facilitate QoS provisioning, resources allocated to a particular node are reserved for its packet transmissions until the next polling. For real-time traffic, the higher the packet dropping rate a traffic flow experiences, the higher the priority of the packets associated with that flow. A clusterhead serves those higher-priority packets first, effectively capping the packet dropping rate. The maximum number of admitted voice calls and that of admitted video calls are 13 and 12, respectively. We perform the simulations for 1,000 runs and average the results, where each simulation run sustains 10,000 frames.

A. Simulation

We carry out simulations to compare our proposed scheme, Combined-KKT-GA, with a random scheme (in [11]), the Cheng's KKT-driven scheme (in [11]), the Shi's KKT-driven scheme (in [9]), the Gao's scheme (in [6]), and a GA-based scheme with $S = 100$ and $T = 25,000$. We evaluate system performance in terms of throughput and packet dropping rate. We study the impact of the number of links on the system performance for all six approaches. Under the assumptions of no packet dropping for the real-time traffic and perfect statistical multiplexing for all the traffic, we derive the upper bound for the throughputs obtained for all the traffic (see Appendix-B).

Case 1 with background data traffic: Fig. 3 shows the throughput performance for background data traffic versus the number of links. The standard deviations of the results are also plotted for reference. Clearly, the GA-based approach outperforms the other three approaches. This result is justifiable because the GA-based method basically takes all the three resource dimensions (i.e., power, subcarrier, and time) into consideration simultaneously. As expected, the throughput performance of Combined-KKT-GA is somewhat between that of the Cheng's KKT-driven approach and that of the GA-based approach. The Shi's KKT-driven approach is better than the Cheng's KKT-driven one because of different subcarrier re-allocation strategies. The former aims at maximizing the system throughput, whereas the latter focuses on QoS satisfaction of the admitted calls. Despite optimal power allocation in place, the Gao's scheme employs suboptimal subcarrier allocation and hence performs worse than the Cheng's KKT-driven approach. Nonetheless, the proposed Combined-KKT-GA approach achieves higher throughput than the Shi's KKT-driven, Gao's, and random approaches. Since more network resources are reserved for the real-time traffic with the number of links, the total throughput of the background data traffic in our proposed algorithm drops. On the other hand, due to the call admission control in place, no more than 13 voice sources and 12 video sources can be admitted into the system. Hence, from $M = 15$ onward, the total throughputs level off. The slight fluctuation of the curves is partly because of the random arrivals of voice and video packets. The upper bound for the throughput obtained by the data traffic is also plotted for reference. For resource utilization, we observe that the Combined-KKT-GA approach utilizes the resources better than the random scheme (which has less than 10% efficiency), the Gao's approach (about 45% from $M = 15$ onward), the Cheng's KKT-driven approach (about 50% from $M = 15$ onward), and the Shi's KKT-driven approach (about 55% from $M = 20$ onward). In fact, the utilization achieved by our proposed approach is close to that achieved by the GA-based approach. For the Combined-KKT-GA algorithm, the resource utilization drops from about 95% to about 65% when the number of links increases from $M = 5$ to 20,

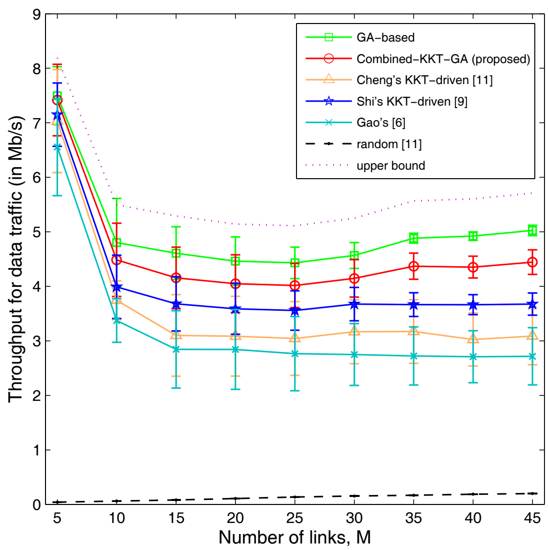

Fig. 3. Comparison of the data throughput performance of the proposed scheme (Combined-KKT-GA), that of the GA-based scheme, that of the Cheng's KKT-driven scheme [11], that of the Shi's KKT-driven scheme [9], that of the Gao's scheme [6], and that of the random scheme [11] vs. the number of links for background data traffic.

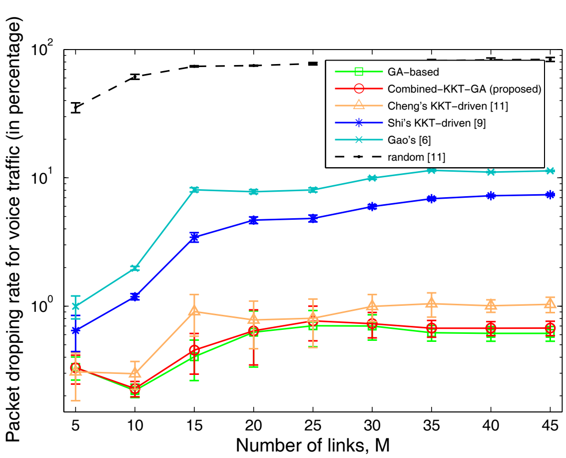

Fig. 4. Comparison of the voice packet dropping rates of the proposed scheme (Combined-KKT-GA), that of the GA-based scheme, that of the Cheng's KKT-driven scheme [11], that of the Shi's KKT-driven scheme [9], that of the Gao's scheme [6], and that of the random scheme [11] vs. the number of links for background data traffic.

and the utilization stays around 60% from $M = 20$ onward. On the other hand, the resource utilization achieved by the GA-based approach drops from 97% to about 65% along with the number of links. Notice that the reduction in the utilization is mainly ascribed to the bandwidth reservation for the real-time traffic. The simulation results confirm the fact that QoS provisioning and resource utilization maximization are conflicting with each other [13]. In Fig. 4, the voice packet dropping rates are depicted. The packet dropping rates for voice traffic are capped in the Combined-KKT-GA (about 0.7%), the KKT-driven (about 1%), and the GA-based (about 0.6%) approaches. From $M = 15$ onward, the packet dropping rates are about 6% and 11% in the Shi's KKT-driven approach and the Gao's approach, respectively. Suboptimal subcarrier allocation in the Gao's approach leads to the incapability of provisioning QoS effectively. On the other hand, in the Shi's KKT-driven approach, the subcarrier re-allocation step aims at the system-wise throughput improvement at the cost of link-wise (or call-wise) QoS satisfaction, resulting in poor QoS provisioning performance. In the random scheme, the voice packet dropping rates increase as M increases. After all, the Shi's KKT-driven, Gao's, and random approaches are not designed for effective packet-level QoS provisioning and, therefore, they cannot provide the same level of QoS assurance offered by our proposed approach. We observe a similar trend for the packet dropping rates for video traffic.

Case 2 with bursty data traffic: Here, we study the case where data packet arrivals are modeled by a Poisson process. Fig. 5 shows the system throughput obtained in serving three traffic types versus the number of links. The curve of data capacity defined by the maximum achievable system throughput without QoS constraints is also plotted for reference. The trend of the system throughput performance in all approaches except the random one increases with the number of links. The GA-based scheme performs the best among all schemes. The performance of the Combined-KKT-GA approach is almost the average of that of the Cheng's KKT-driven approach and that of the GA-based approach. As expected, due to the optimal subcarrier allocation, the Cheng's KKT-driven approach performs better than the Gao's approach. However, the Cheng's KKT-driven approach attains almost the same system throughput as the Shi's KKT-driven one. The reason is that the data traffic load is low and hence the effect of the throughput-sensitive subcarrier re-allocation step in the Shi's approach is not significant. In fact, there are two key factors determining the trend of the curves, namely traffic load and resource reservation for the real-time traffic. Since the network is not saturated, the higher the traffic load, the higher the throughput. The curves go up when M increases from 5 to 15, as there is a higher traffic load. From $M = 15$ onward, the influence of the resource reservation for QoS provisioning becomes more significant. Thus, the increment of system throughputs is smaller. The data capacity increases with the number of links because of the multi-user diversity. It is noteworthy that in Fig. 5, there is an obvious performance gap between the system throughputs obtained from all the approaches and the data capacity. This gap is ascribed to the low traffic load and resource reservation. We observe that both the system throughput and resource utilization for the proposed algorithms are improved when the traffic load increases. The upper bound for the system throughput obtained by analysis is also plotted for reference. For the dropping rates of the real-time traffic, we observe that the results are similar to Fig. 4. Thus, it can be concluded that the data traffic model has little influence on the performance of the real-time traffic, for the data packets are assigned the lowest priority.

Simulation results show that the proposed Combined-KKT-GA approach can effectively provision QoS for real-time traffic and achieve satisfactory system throughput performance. This novel approach is also of low computational complexity, leading to a viable candidate for practical implementation. The characteristics of the aforementioned resource allocation approaches are summarized in Table I.

TABLE I
THE CHARACTERISTICS OF DIFFERENT RESOURCE ALLOCATION APPROACHES.

Resource Allocation Approach	Complexity	Throughput Performance	Effective QoS Provisioning	Practicality
Exhaustive Search	$O(M^{NL})$	Optimal	Yes	No
GA-based	$O(TSLMN)$	Close to optimal*	Yes	Yes*
Combined-KKT-GA (proposed)	$O(CLMN)$**	Good	Yes	Yes
Cheng's KKT-driven [11]	$O(LMN)$	Average	Yes	Yes
Shi's KKT-driven [9]	$O(LM^2N)$	Good	No	Yes
Gao's [6]	$O(LMN)$	Average	No	Yes
Random [11]	$O(N)$	Poor	No	Yes

Note *: The results are contingent on the number of iterations used in the GA-based resource allocation scheme.

Note **: The value of C is far less than that of TS.

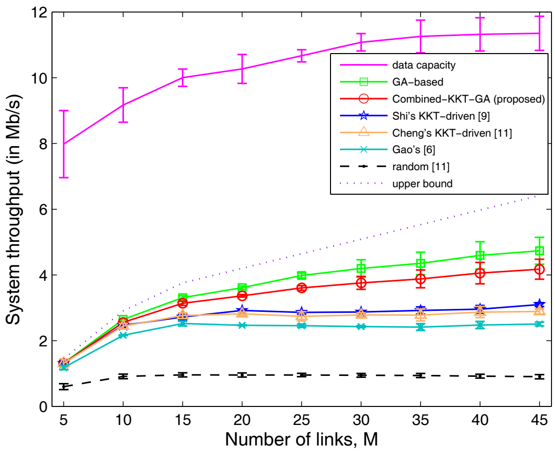

Fig. 5. System throughputs of the proposed scheme (Combined-KKT-GA), the GA-based scheme, the Cheng's KKT-driven scheme [11], the Shi's KKT-driven scheme [9], the Gao's scheme [6], the random scheme [11], and the data capacity vs. the number of links for bursty data traffic.

VII. CONCLUSION

A novel QoS-aware packet-level resource allocation approach, Combined-KKT-GA, is proposed for the joint power-subcarrier-time intra-cluster resource allocation problem in WMNs. By combining the merits of the KKT-driven and the GA-based approaches, the newly proposed Combined-KKT-GA approach is shown to effectively facilitate QoS provisioning at the packet level, and achieve a desired balance between system performance and computational complexity. Furthermore, bounds for the throughputs obtained by real-time and non-real-time traffic are derived, validating our simulation results.

APPENDIX

A. Complexity of the Optimization Problem

Proposition 1: The optimization problem is an NP-hard problem.

Proof: We prove the NP-hardness by reducing the NP-complete *number partitioning problem* (denoted by PARTITION) to the optimization problem. We consider the special case of the optimization problem by fixing the power allocation and letting $M = 2$, $r^l_{1,n} = r^l_{2,n} = r^l_n$, $\forall n, l$, and $R^d_1 = R^d_2 = \sum_{n=1}^{N} \sum_{l=1}^{L} \frac{r^l_n}{2}$, and the "size" of the n^{th} subcarrier at the l^{th} timeslot is r^l_n. Since each subcarrier can only be allocated to one link at any timeslot (i.e., $a^l_{1,n} + a^l_{2,n} = 1$, $\forall n, l$), the solution to this special-case problem is exactly the same as that of the PARTITION problem. In other words, the PARTITION problem can be *polynomially transformed* into the special-case problem, and vice versa. Since the PARTITION problem is NP-complete, the special-case problem is also NP-complete. Further, the above special-case problem can be generalized by the optimization problem. Thus, the optimization problem is NP-hard. ■

B. Upper Bound for Throughput

Denote ρ_1 and ρ_2 as the voice activity factor and the video activity factor, respectively. The upper bound for the throughput obtained for each of the traffic is given as follows.

- Throughput obtained for voice traffic, T_{voice}: $T_{voice} = \rho_1 N_{voice} R_{voice}/4$, where N_{voice} and R_{voice} are the number of voice sources in the system and the constant data rate of voice traffic, respectively.
- Throughput obtained for video traffic, T_{video} : $T_{video} = \rho_2 N_{video} R_{video}$, where N_{video} and R_{video} are the number of video sources in the system and the mean data rate of video traffic, respectively.
- Throughput obtained for background data traffic, T_{data_1}: $T_{data_1} = \sum_m R_m - T_{voice} - T_{video}$, where R_m is the achieved data rate for the m^{th} link.
- Throughput obtained for bursty data traffic, T_{data_2}: $T_{data_2} = \lambda N_{data} \sqrt{\pi}$, where λ and N_{data} are the packet arrival rate and the number of data sources in the system, respectively.

REFERENCES

[1] H. T. Cheng, H. Jiang, and W. Zhuang, "Distributed medium access control for wireless mesh networks," *Wireless Commun. Mobile Computing*, vol. 6, no. 6, pp. 845-864, Sept. 2006.

[2] J. Yu and P. Chong, "A survey of clustering schemes for mobile ad hoc networks," *IEEE Commun. Surveys Tutorials*, vol. 7, no. 1, pp. 32-48, First Qtr. 2005.

[3] H. T. Cheng and W. Zhuang, "Pareto optimal resource management for wireless mesh networks with QoS assurance: joint node clustering and subcarrier allocation," *IEEE Trans. Wireless Commun.*, in press.

[4] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *Proc. IEEE Globecom'00*, vol. 1, 2000, pp. 103-107.

[5] D. Julian, M. Chiang, D. O'Neill, and S. Boyd, "QoS and fairness constrained convex optimization of resource allocation for wireless cellular and ad hoc networks," in *Proc. IEEE INFOCOM'02*, vol. 2, June 2002, pp. 477-486.

[6] L. Gao and S. Cui, "Efficient subcarrier, power, and rate allocation with fairness consideration for OFDMA uplink," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1507-1511, May 2008.

[7] X. Song, Z. He, K. Niu, and W. Wu, "A hierarchical resource allocation for OFDMA distributed wireless communication systems," in *Proc. IEEE Globecom'07*, pp. 5195-5199.

[8] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[9] J. Shi and A. Hu, "Maximum utility-based resource allocation algorithm in the IEEE 802.16 OFDMA system," in *Proc. IEEE Globecom'08*, pp. 311-316.

[10] C. Y. Ng and C. W. Sung, "Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1667-1675, May 2008.

[11] H. T. Cheng and W. Zhuang, "Joint power-frequency-time resource allocation in clustered wireless mesh networks," *IEEE Network*, vol. 22, no. 1, pp. 45-51, Jan./Feb. 2008.

[12] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA: The MIT Press, 1998.

[13] H. T. Cheng and W. Zhuang, "An optimization framework for balancing throughput and fairness in wireless networks with QoS support," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 584-593, Feb. 2008.

[14] M. Neufeld *et al.*, "SoftMAC–flexible wireless research platform," in *Proc. ACM SIGCOMMM'05*, pp. 1-5, Nov. 2005.

[15] IEEE 802.16 WG, "Channel models for fixed wireless applications," 2003. [Online]. Available: http://www.ieee802.org/16

[16] D. P. Heyman, "The GBAR source model for VBR videoconferences," *IEEE/ACM Trans. Networking*, vol. 5, no. 4, pp. 554-560, Aug. 1997.