 The definitive version of this paper has been published in Fourth International Conference on 3G Mobile Communication Technologies (3G 2003), 2003, 10.1049\/cp:20030376.

WEB TRAFFIC PERFORMANCE IN WIRELESS LAN HOT SPOTS

K Heck

University of Würzburg, Germany

ABSTRACT

Wireless Internet access gains more and more attraction with the ongoing introduction of Wireless LAN hot spots. In this article the interaction of typical Web traffic and the Wireless LAN Medium Access Control protocol is evaluated. The focus lies on the user-experienced quality received in different scenarios and varying cell loads. We also include the effects of the RTS/CTS and fragmentation mechanisms, which were introduced to overcome performance degradation problems that are specific for wireless environments, such as the hidden node problem. The results provide insights on the advantages and deficiencies of these mechanisms and allow conclusions about the user-experienced performance of Web users in hot spot environments. We thereby build the foundation for the capacity planning of WLAN hot spots as part of the future 4th generation mobile networks.

INTRODUCTION

Discussions about 4th generation mobile networks highlight the future importance of the IEEE 802.11 Wireless LAN (WLAN) standard. WLAN hot spots pop up all across the world with an increasing pace, while multi-mode devices are being developed to bring together heterogeneous network technologies such as UMTS and Wireless LAN. The bandwidth of up to 54 Mbps supported in hot spot environments encourages ISPs to provide high-speed Internet access to wireless users.

Various performance studies of the Wireless LAN Medium Access Control (MAC) protocol can be found in the literature (see Bruno, Conti and Gregori (1) or Köpsel, Ebert and Wolisz (2)). These publications, however, focus on the properties of the MAC protocol itself, such as the maximum achievable throughput or fairness, but they ignore application-specific influences, which are of great importance to the subjective quality experienced by single users, e.g. in differing cell-load situations.

In this article we study the effect of the number of concurrently active Web users on the system performance and the user-experienced quality of service. This does not only include the question of how many Web users can be served with adequate quality within a single cell, but it also shows the effect of mechanisms that were defined to overcome problems found solely in the wireless environment, such as the hidden node problem, and how these extensions affect the system performance.

The goal of our studies is to provide Wireless Internet Service Providers (WISP) with a better understanding of the capability of their WLAN infrastructure. We draw conclusions about the realistic capacity of single WLAN cells, which allows a better planning of WISPs' Internet access networks.

This paper is organized as follows. The next section will provide an overview of the simulation model for the WLAN MAC protocol and user source traffic model. This is followed by the numerical results of the simulation studies. Finally, a conclusion of the main results is provided and an outlook on future work is given.

SIMULATION MODEL

In order to perform simulation studies of WLAN hot spots, three main entities have to be modeled. First, there is the underlying WLAN MAC protocol together with a set of protocol extensions. Then, the user behavior has to be specified in the form of a source traffic model. In our case, only Web traffic is analyzed, such that a single Web source traffic model will be used. Finally, the simulation environment has to be chosen. Typical scenarios are used to account for the various possibilities.

Each of these three individual tasks will be explained in the following.

Wireless LAN Medium Access Control Protocol

The fundamental access method in Wireless LAN is provided by the Distributed Coordination Function (DCF). It allows the automatic medium sharing between attached clients through the use of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) and a random back-off time following a busy medium condition. Basically, each client is only allowed to transmit its data, if the medium is found idle prior to the transmission attempt. A successful transmission is immediately acknowledged by the recipient issuing an ACK packet. If the ACK is not

received within a predefined ACK timeout interval, the sender automatically schedules the retransmission of the packet.

This basic mechanism is very similar to the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol also known as the Ethernet protocol. However, in the case of a wireless scenario, some specific properties of the channel have to be accounted for. Due to the limited transmission power and coverage range of WLAN hot spots, the rapidly changing radio conditions will lead to situations where clients can not receive each other's signals, while they can still communicate with their associated access point. These two clients are said to be in a hidden node relationship. In this case, the carrier sensing mechanism can not prevent the clients to simultaneously transmit data to the access point, which will cause collisions and a degradation of the system performance.

Thus, a refinement of the basic access protocol becomes necessary under these circumstances to minimize collisions. The WLAN protocol defines the RTS/CTS (Request To Send / Clear To Send) mechanism. Here, the transmitting and receiving clients exchange short control frames prior to the data transmission. These frames contain medium reservation information. All clients receiving the RTS or CTS frame will defer their own transmission for the specified amount of time, which is stored in the Network Allocation Vector (NAV).

The procedure is depicted in Figure 1. The source node transmits an RTS packet to indicate its intention to send a data packet. The destination node answers an RTS request by issuing a CTS packet. All other stations within the reception range of either the originating node or the destination node take notice of the medium reservation and delay their own transmission attempts for at least the amount of time specified in the NAV counter. This mechanism ensures that nodes, which are hidden from the source, are informed of the ongoing transmission.

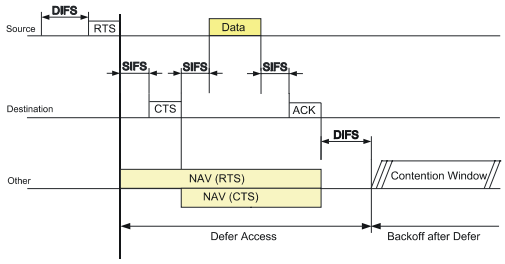

Figure 1: RTS/CTS mechanism and NAV counter

The current version of the Wireless LAN standard does not specify a Forward Error Correction (FEC) mechanism. A single bit in error causes the recipient to drop the packet. In this case the source node performs an automatic retransmission. However, due to the erroneous nature of the wireless channel, a large amount of packets can be disturbed and dropped. The problem gets even worse the larger the data packets are. In order to overcome this problem, a fragmentation mechanism has been introduced. It partitions the data packets into smaller fragments and transmits the fragments one after another as shown in Figure 2.

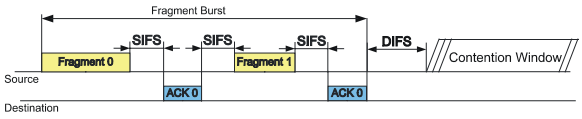

Figure 2: WLAN fragmentation mechanism

Dropped fragments are automatically retransmitted. On the downside this causes more overhead, however, the lower fragment error probability compared to the packet error probability increases the system reliability and can lead to a higher overall performance in certain situations. We consider both, the RTS/CTS and the fragmentation mechanism in our simulation model.

Web Source Traffic Model

The source traffic model that determines the behavior of the Web users basically follows the model presented by Staehle, Leibnitz and Tran-Gia (4). It is based on real measurements of network traffic created by users surfing the WWW and thus allows to directly map our results to real hot spots.

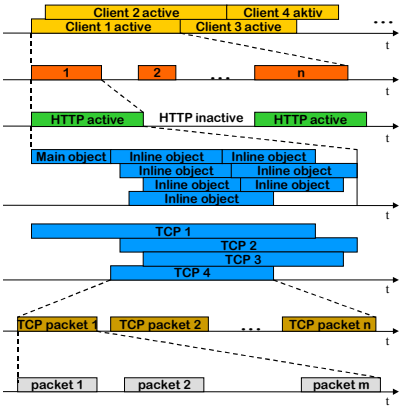

Figure 3: HTTP source traffic model

The overall model is shown in Figure 3. On the top layer it defines activity phases for each user. An activity phase consists of a number of Web sessions. Within a Web session a user has HTTP active and HTTP inactive phases. In the active phase the actual download is performed, while the inactive phases

describe the time while the user is viewing the downloaded page. Depending on the HTTP version, up to four simultaneous TCP connection are used to download a single Web page, which consists of a single main object and a number of inline objects, e.g. images or applets.
In our simulations the users are assumed to be always active. Therefore, our results relate to the number of users that are simultaneously browsing the Web. We do not consider inactive users, as we are interested in evaluating worst-case scenarios for capacity planning.

Simulation Scenarios

Two different simulation scenarios are considered. In the first scenario, all involved clients are in reception range of each other as shown in Figure 4. The second scenario, on the other hand, defines two distinct groups of users which are hidden from each other, as indicated in Figure 5.

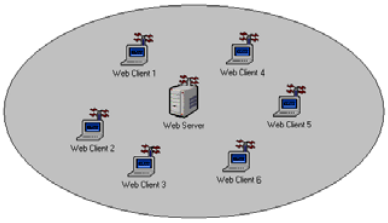

Figure 4: WLAN simulation scenario

For both of these simulation scenarios, the effect of the RTS/CTS and fragmentation mechanisms is evaluated. Comparing the results yields the effect of the hidden nodes. We also varied the maximum bandwidth of the nodes from 1 Mbps up to 11 Mbps in order to compare the capability of the different WLAN configurations.

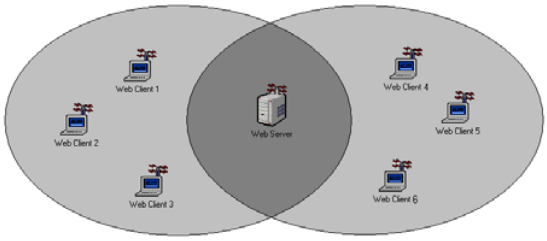

Figure 5: Hidden node simulation scenario

RESULTS

The simulation scenarios are compared in terms of the average Web page download time, which is a subjective measure for the user-perceived quality of service. According to the Web source traffic model, the average page size greatly varies. Therefore, the important measure is the relative behavior of the average page download time rather than the absolute value. A large number of users in this kind of simulation results in an increase in the delay experienced by each user. This causes TCP retransmissions once a certain delay limit is reached. Any further increase of the cell load will cause TCP to exceed its maximum number of retransmissions, such that the TCP connection is dropped and the page download is cancelled. Such dropped downloads do not contribute to our statistics. Therefore, the number of possible TCP retransmissions is set to unlimited.
The RTS threshold is set to 256 bytes, which means that if packets larger than the threshold have to be send, an RTS packet is issued prior to the data transmission. The WLAN standard allows fragmentation thresholds in the range of 256 up to 1024 bytes. Only packets larger than the threshold are fragmented. In our simulations a fragmentation threshold of 256 bytes is used.
Figure 6 shows the results for the 1 Mbps scenario. The two solid lines represent the results for pure CSMA/CA and a single group of clients (no hidden nodes) and the case with two groups of users (with hidden nodes). It can be seen that the average page download time increases from approximately 0.6 seconds for the 10 client case to more than 10 seconds for 100 clients. This increase corresponds to a factor of more than 16, which means that a user in the 100 client case experiences a page download time of 16 times longer than for 10 clients. Such an increase is not acceptable for the users, which means that the maximum number of Web users in the 1 Mbps scenario should not exceed 40 clients. Comparing the two curves yields the degradation of the system performance due to the hidden nodes. The gray line for the hidden node case is only about 3 percent above the black curve. Thus, the hidden nodes have a small but noticeable effect.

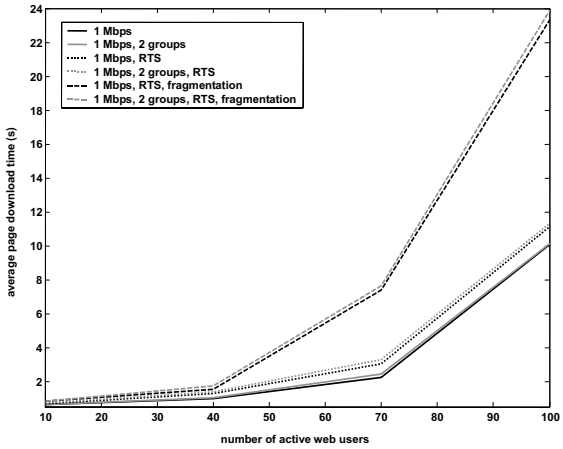

Figure 6: WLAN performance at 1 Mbps

The dotted lines in Figure 6 show the results for the case that the RTS/CTS mechanism is activated. The average page download time for this case is always found to be about 10 percent above the scenario without RTS/CTS. This is true for the one and two groups scenarios. As we have discussed earlier, the RTS/CTS mechanism lowers the number of collisions in the 2 groups scenario (gray curves). However, it produces more overhead in our cases than can be gained by decreasing the collisions.

Finally, the dashed lines correspond to the cases with additional fragmentation of packets larger than 256 bytes. In our case, the wireless channel was assumed to be free of errors. The results therefore display the overhead introduced by fragmentation. It can be easily seen that the page download times are by far greater and that the fragmentation overhead has a major effect on the overall system performance.

The results for the 2 Mbps scenario can be seen in Figure 7. The increase of the page download time for the solid lines reaches a factor of 3 from 10 clients to 100 clients. The hidden nodes again results in a system degradation of about 3 percent. The RTS/CTS mechanism overhead causes an increase of the average page download time of approximately 10 percent. As for the 1 Mbps scenario, the RTS/CTS mechanism does not improve the system capacity. A maximum number of 70 Web clients seems possible. However, the fragmentation overhead leads to an explosion of the page download times.

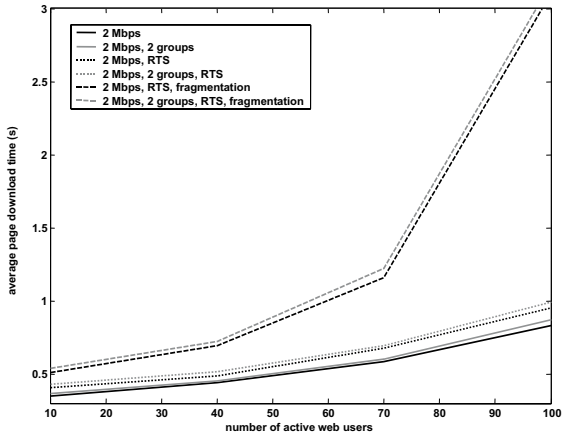

Figure 7: WLAN performance at 2 Mbps

The results for the WLAN cell with a maximum bandwidth of 5.5 Mbps are shown in Figure 8. The overall tendencies found for the 1 and 2 Mbps cases can be recognized as well. Two groups of users that are hidden from one another cause a system degradation of about 3 percent. The RTS/CTS mechanism raises the average page download time by approximately 10 percent. The overhead of the RTS/CTS mechanism is larger than the system degradation caused by the hidden node problem. Again, fragmentation causes a huge overhead.

However, the solid lines allow the conclusion, that the system can handle 100 to 140 clients with appropriate quality. The average page download time does not even double from the 10 clients case to the 100 clients case. This also applies to the RTS/CTS case. The performance for 100 to 140 clients is still in a range that can satisfy the users' demands. Fragmentation on the other hand, should not be deployed if more than 70 users are located in one WLAN cell.

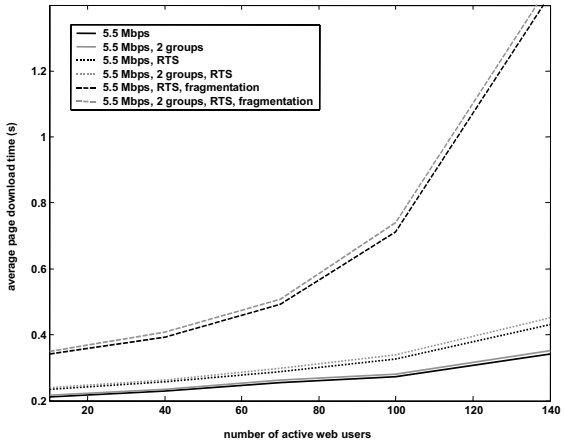

Figure 8: WLAN performance at 5.5 Mbps

Similar results can be found for the case where the maximum data rate is set to 11 Mbps as displayed in Figure 9. The hidden nodes (gray lines) cause a performance degradation of no more than 3 percent compared to the case without hidden nodes (black lines). The RTS/CTS mechanism overhead again reaches about 10 percent. Therefore, RTS/CTS does not improve the overall performance, but leads to a further increase of the average page download times. The situation changes drastically, once the fragmentation mechanism is activated. The page download times explode and the WLAN cell can not handle more than 40 clients appropriately.

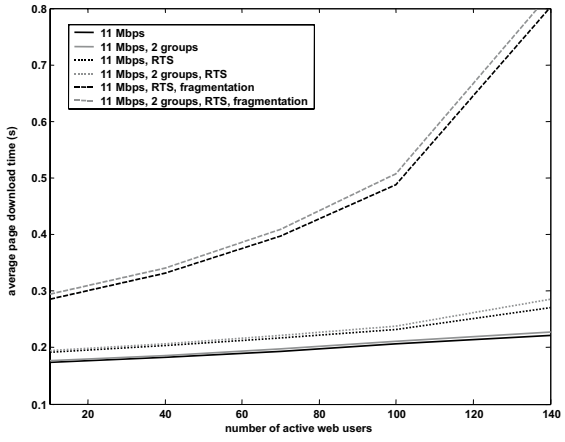

Figure 9: WLAN performance at 11 Mbps

Nevertheless, the results in Figure 9 show that in the 11 Mbps case the system can easily handle the 140 clients as long as fragmentation is not used. The average page download time for the 140 client case is less than 20 percent above the 10 client case. Considering the fact, that most currently available access points often support less than 140 simultaneously attached clients, we can conclude that the performance of the system is still good enough to satisfy the Web users' demands even in environments with high cell loads.

CONCLUSIONS AND OUTLOOK

In this paper we discussed simulation studies performed on a single Wireless LAN cell. In order to account for application-specific influences, the users were modeled using a realistic Web source traffic model. We studied the effect of the number of concurrently active Web users on the system performance and the user-experienced quality of service. The measure used to quantify the level of service was the average page download time. Using this realistic approach allows to directly map our results to real WLAN hot spots.

The wireless nature of the medium necessitates a set of extensions to the basic CSMA/CA medium access control mechanism. On the one hand, there is the hidden node problem that causes performance degradation. The standard specifies the RTS/CTS (Request to Send / Clear to Send) mechanism to overcome this deficiency. On the other hand, the wireless medium frequently exhibits large bit error probabilities. However, the WLAN standard does not incorporate a Forward Error Correction code, but recommends the use of fragmentation. Large data packets are partitioned in smaller fragments with lower packet error probabilities. We considered both of these extensions in our simulations. The wireless channel was assumed to be free of errors. Therefore, our results show the pure overhead caused by the different mechanisms.

Two simulation scenarios were defined. One with a single group of users all within the reception range of each other. The other scenario defined two groups of users that are hidden from one another. Results were presented for a set of varying transmission bandwidths, ranging from 1 Mbps up to 11 Mbps.

We have seen that two groups of hidden nodes cause an increase of the average page download time of approximately 3 percent. The RTS/CTS mechanism causes a system degradation of about 10 percent, such that it does not solve the hidden node problem in our scenarios. The overhead of the fragmentation mechanism was shown to be extremely large. In most of the cases it causes the average page download time to increase excessively.

Using our results, we can derive the maximum allowable number of concurrently active Web users within a single Wireless LAN cell. For the case of a maximum bandwidth of 1 Mbps, the number of users should not exceed 40. If the data rate is 2 Mbps, we can allow a maximum number of Web users in the range of 40 to 70. In the 5.5 Mbps case, 100 to 140 simultaneous Web users were shown to be feasible, while in the 11 Mbps case, the limit is more the maximum number of clients that can associate to a single access point rather than the increase in page download time.

The RTS/CTS mechanism could not increase the system capacity even in the case of two hidden node groups. It does not seem realistic, that there are many more hidden nodes in a single Wireless LAN cell, such that RTS/CTS should always be deactivated. Fragmentation has a major effect on the system capacity. It heavily degrades the system performance. However, high bit error probabilities were not considered. It is left to future work to evaluate the fragmentation mechanism in such cases.

ACKNOWLEDGEMENTS

The author would like to thank Tom Wirth and Rastin Pries for their contributions to this paper.

REFERENCES

1. Bruno R, Conti M and Gregori E, 2002, "IEEE 802.11 Optimal Performances: RTS/CTS Mechanism vs. Basic Access", Proc. of the 13th IEEE Intl. Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)

2. Köpsel A, Ebert J and Wolisz A, 2000, "A Performance Comparison of Point and Distributed Coordination Function of an IEEE 802.11 WLAN in the Presence of Real-Time Requirements" Proc. of the 7th Intl. Workshop on Mobile Multimedia Communications (MoMuC)

3. IEEE, 1999, "IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification", ISO/IEC 8802-11

4. Staehle D, Leibnitz K and Tran-Gia P, 2001, "Source Traffic Modeling of Wireless Applications", International Journal of Electronics and Communications, Vol. 55