On the Capacity of *p*-Persistent CSMA

Marek Miśkowicz

Department of Electronics, AGH University of Science and Technology, Kraków, Poland

Summary
The CSMA-based algorithms constitute a heart of contemporary wireless media access control technology. In last decades, various enhancements have been introduced to basic CSMA schemes, first of all, in order to support collision avoidance. At the same time, the performance analysis of the *p*-persistent CSMA has gained a renewed interest recently since the behavior of many CSMA protocols might be studied by the *p*-persistent model. The capacity of the *p*-CSMA is an important performance measure. It establishes the best-case channel utilization that may be obtained in the *p*-persistent CSMA with varying *p*. The goal of the paper is to compare the maximum channel utilization of the *p*-persistent CSMA with the corresponding measure for the predictive *p*-persistent CSMA implemented in MAC sublayer of LonTalk/EIA-709.1 protocol used for communication between intelligent sensors and actuators in LonWorks technology. A comparison of both CSMA schemes shows that the predictive *p*-persistent CSMA throughput is close to the capacity of the *p*-persistent CSMA only for small network sizes. If the number of contenders is greater, the throughput of the predictive *p*-CSMA is much smaller than *p*-CSMA capacity.

Key words:
Media access control, Carrier sense multiple access protocols, Performance evaluation

1. Introduction

Although carrier sense multiple access (CSMA) protocols were invented in the 70s, they are still used in modern networking due to the inherent flexibility of random access systems. One of generic and widely used CSMA algorithms is the *p*-persistent CSMA protocol. As is well-known, a node contending for the shared channel according to the *p*-CSMA, transmits with probability *p* if the channel is idle, and defers transmission with the probability (1-*p*) if the channel is busy [1].
The CSMA-based algorithms constitute a heart of contemporary wireless media access control technology. In last decades, various enhancements have been introduced to basic CSMA algorithms in order to support *collision avoidance*, which is the priority task of random access protocols for wireless networks.
At the same time, the performance analysis of the *p*-persistent CSMA has gained a renewed interest recently since the behavior of many CSMA protocols, for example, IEEE 802.11, LonTalk/EIA-709.1 and Sift protocol, might be studied by the *p*-persistent model [2-6].
In general, MAC protocols for wireless networking need high bandwidth utilization because the wireless networks deliver much lower bandwidth than wired networks. The fraction of channel bandwidth used for successful transmissions gives a main indication of the overhead introduced by the media access protocol to hold its coordination task among the nodes.
The channel utilization in the *p*-persistent CSMA is strongly affected by the *p* value, which represents the *persistence level* of the protocol. In particular, large *p* values cause excessive collisions, while small *p* values degrade the bandwidth utilization forcing the channel to be idle. To keep the bandwidth utilization on the satisfactory level, a tradeoff between large and small values is necessary.
A given persistence level, *p*, maximizes the channel utilization only for a preselected number of contending nodes. If a number of contenders is unknown *a priori* or varies in time, the *p* value cannot be set optimally, and consequently the performance of *p*-persistent CSMA may be considerably degraded. Therefore, the CSMA-based protocols with *collision avoidance* (CSMA/CA), which are the enhancements of the pure *p*-persistent CSMA, try to adapt to the number of contending nodes. The modification of *p* value is usually accomplished by decreasing *p* in case of collisions, and by increase of *p* after each successful transmission.
One of important adaptive CSMA algorithms is the *predictive p-persistent CSMA* where the probability *p* is variable and dynamically adjusted to the expected traffic load [7]. This protocol has been designed for sensor/control networking where the traffic produced by sensing devices might be bursty. The predictive *p*-persistent CSMA is commercially implemented in MAC sublayer of LonTalk protocol registered as ANSI/EIA 709.1 and ENV 13154-2 standards and exploited in Local Operating Networks (LonWorks) technology for communication between intelligent sensors and actuators [8,9]. The performance analyses of the predictive *p*-persistent CSMA has been studied using either simulation [10] and analytical methods [3,4].
The goal of the paper is to compare the maximum channel utilization of the *p*-persistent CSMA with the corresponding measure for the predictive *p*-persistent

Manuscript received November 5, 2007.
Manuscript revised November 20, 2007.

CSMA. Although the *p*-CSMA has been extensively studied in past decades for various traffic models (see e.g., [1,2,10]), the simple comparison of that results with performance analyses of the predictive *p*-persistent CSMA is difficult or even impossible due to the use of different analytical methods and network model assumptions. Therefore, to keep consistency of the analysis, in the present paper we adapt the performance evaluation method developed in [3] for the predictive *p*-persistent CSMA in order to estimate the channel capacity of the *p*-persistent CSMA.

2. Protocol Model and Performance Measures

2.1 Protocol Model

We assume a slotted-CSMA algorithm where the time axis is split into segments, called *contention slots*, whose duration is equal to β_2. All the nodes are forced to start transmission only at the beginning of a slot. Even if the nodes are ready to transmit in the middle of a contention slot, they have to wait until the slot finishes, and a new one begins. When two packets conflict, they will overlap completely rather than partially, which greatly reduces the probability of collision and provides an increase of channel efficiency. The slotted-CSMA has been derived from the slotted-ALOHA protocol invented in the middle of the 70s.

The algorithm operates in the following way. A node attempting to transmit monitors the state of the channel. If the channel is busy, the node continues sensing. When the node detects no transmission during the β_1 period, delays a random number of time slots of β_2 duration. In CSMA-based schemes, this time is called the *backoff*. Both β_1 and β_2 are configurable time constants determined by Physical Layer parameters such as the propagation delay defined by the media length or a distance between the nodes, the detection and turn-around delay within the MAC sublayer.

If the channel is still idle when the random backoff expires, the node transmits. Otherwise, the node receives incoming packet and competes for the channel access again. If more than one node choose the same slot number, and where that slot has the lowest number selected by any node with a packet to send, then a *collision* happens. All the packets involved in a collision are corrupted.

The backoff time is expressed as a pseudorandom number of time slots β_2 drawn from the uniform distribution between 0 and *W*, where *W* is the *size of the contention window*. Since each slot is selected by a node with equal probability, the CSMA protocol with a number of *W* contention slots is (1/*W*)-persistent CSMA protocol.

2.2 Network Model

We assume that a network stays at the *saturation status*, that is, the system consists of a constant number of contending nodes that always have packets ready to send. Thus, there are no idle packet cycles in the channel access. We assume also that there are no store-and-forward routers in the analyzed networked system. Next, we suppose that all the packets sent via the channel are of a constant length denoted by *Pkt Length*.

As was stated in the introduction, in order to compare the channel capacity, we provide the *p*-persistent CSMA protocol model consistent with that of the predictive *p*-persistent CSMA.

2.3 Performance Measures Definitions

The network *throughput* is defined as the fraction of time used for successful transmissions of packets in the channel [12]. We express the throughput as a percentage of the channel bit rate. Therefore, we use the terms the “throughput” and the “channel utilization” interchangeably in the paper although the former is often expressed in the literature by the number of packets successfully transmitted per a unit of time.

The maximum achievable throughput (or the channel utilization) is called the *channel capacity* and is found by maximizing the throughput with respect to the offered load [12]. The channel utilization/throughput is maximized by balancing the time wasted in collisions with the time spent for listening to the channel [11].

Alternatively, the capacity might be found by minimizing the ratio of the packet length to the mean time between consecutive successful transmissions in the channel [11].

The protocol capacity is one of primary measures characterizing MAC protocols. For an ideal MAC protocol, the capacity equals one. Furthermore, the channel capacity depends on several network parameters, first of all, on the number of active nodes and their contribution to the offered load.

3. Throughput Evaluation

The throughput of the 1/*W*-persistent CSMA might be found analytically as follows [3] :

$$Throughput(n,W) = \frac{PktLength}{\left(\frac{1}{p_{succ}^{(W)}(n)} - 1\right)\tau_{coll}^{(W)}(n) + \tau_{succ}^{(W)}(n)} \, . \quad (1)$$

where

$$p_{succ}^{(W)}(n) = n\sum_{s=1}^{W}\frac{1}{W}\left(\frac{W-s}{W}\right)^{n-1} \quad (2)$$

is the probability of a successful transmission of a packet if a number of n nodes contends for a shared channel and the contention window consists of a number of W slots.

The probability $p_{succ}^{(W)}(n)$ is expressed as the sum multipled by the number of contenders of the following probabilities calculated for each one from $1,...,W$ slots :

(i) the probability that a winner selects a certain slot $s, s = 1,...,W$, which equals to $1/W$,

(ii) the probability that all the other $(n-1)$ nodes draw one from $(W-s)$ later slots, which equals to $((W-s)/W)^{n-1}$.

Further, $\tau_{succ}^{(W)}(n)$, $\tau_{coll}^{(W)}(n)$ denote the mean lengths of a successful and unsuccessful packet cycles, respectively, given by the following formulas :

$$\tau_{succ}^{(W)}(n) = \beta_1 + [d_{succ}^{(W)}(n) - 1]\beta_2 + PktLength \quad (3)$$

$$\tau_{coll}^{(W)}(n) = \beta_1 + [d_{coll}^{(W)}(n) - 1]\beta_2 + PktLength \quad (4)$$

where $d_{succ}^{(W)}(n)$, $d_{coll}^{(W)}(n)$ represent the mean slot numbers when the successful transmission starts or the collision occurs, accordingly, and may be found as follows :

$$d_{succ}^{(W)}(n) = \frac{\sum_{s=1}^{W}(W-s)^{n-1}s}{\sum_{s=1}^{W}(W-s)^{n-1}} \quad (5)$$

$$d_{coll}^{(W)}(n) = \frac{1}{W^{n-1}}\sum_{s=1}^{W}s^{n-1} \quad (6)$$

The expression (6) is defined for $n \geq 2$ since the collision may happen if more than one node compete for the channel access.

It can be proved that $d_{succ}^{(W)}(n) < d_{coll}^{(W)}(n)$ for any number of nodes, $n; n \geq 2$, which is intuitively clear since the collisions are more likely in later slots. Moreover :

$$\lim_{n\to\infty} d_{succ}^{(W)}(n) = \lim_{n\to\infty} d_{coll}^{(W)}(n) = 1 \quad (7)$$

as follows from (5) and (6).

We assume that the interpacket space is $\beta_1 = 4[\text{bits}]$ long, the contention slot width equals $\beta_2 = 2[\text{bits}]$, the packet length is constant and equal to $PktLength = 96[\text{bits}]$. Note that packets transmitted through the sensor or control network are relatively short. See [3] for derivations of the analytical formulas presented above.

4. Numerical Results

4.1 Throughput vs. Number of Nodes

The numerical results of the throughput for selected contention window sizes (W=32, W=80, and W=160 slots) according to the analytical procedure presented in Sect. 3 are listed in Table 1. The corresponding plots are presented in Fig. 1. The throughput versus the number of active contenders for a certain contention window size is a function that has a single maximum. The number of active nodes corresponding to the channel capacity represents the *optimal offered load* for a given protocol persistence level $1/W$.

Table 1: Channel utilization versus number of contenders for selected sizes of the contention window

Number of nodes, n	***Throughput W=32***	***Throughput W=80***	***Throughput W=160***
5	0.808	0.740	0.620
10	0.779	0.793	0.726
20	0.675	0.792	0.789
50	0.393	0.675	0.776

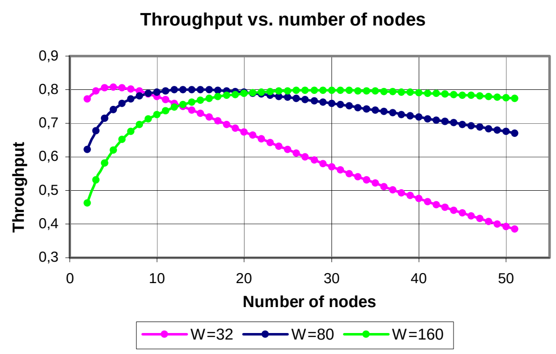

Fig. 1 Channel utilization versus the number of contenders for selected sizes of the contention window.

Two factors influence the CSMA protocol performance: the collisions, and the waste of bandwidth to randomize the uncoordinated channel access. The optimal channel utilization of the p-persistent CSMA is characterized by the balancing between duration of collisions and idle times [11]. The long time intervals between successful transmissions that occur for the *p*-persistent CSMA with small *p* are the effect of the high number of empty slots preceding each transmission. Then, the probability that two or more stations start transmitting simultaneously is negligible. On the other hand, the use of large *p* values yields excessive collisions between relatively rare successful transmissions. The throughput maximization corresponds to such a *p* value for which both effects balance.

As expected, the probability of a successful transmission $p_{succ}^{(W)}(n)$ given by the formula (2) is large if the number of contenders *n* is small. On the other hand, as follows from the formula (5) the mean slot number when the successful transmission starts, $d_{succ}^{(W)}(n)$, is also relatively large for small *n*. Thus, for the small (suboptimal) number of contenders the dominant component of the bandwidth watse is a certain number of contention slots that are wasted in order to avoid collisions. The analysis of the formula (5) shows that $d_{succ}^{(W)}(n)$ is a decreasing function of *n*. Thus, $d_{succ}^{(W)}(n)$ for $n=1$ reaches its maximum equal to $\max_{n\geq 1}[d_{succ}^{(W)}(n)] = d_{succ}^{(W)}(n=1) = (W+1)/2$. Roughly speaking, the maximum average number of slots lost due to contention is equal nearly to the half of the contention window size ($W/2$).

If the contention is high (i.e. many nodes want to transmit), the dominant component of the bandwidth waste are collisions. The fraction of bandwidth used for a randomization of the channel access is then insignificant. It is because, for a large number of contenders, $p_{succ}^{(W)}(n)$ is strongly decreased due to excessive collisions (see the formula (2)), and $d_{succ}^{(W)}(n)$ approaches asymptotically one (see the formula (7)). In other words, a node that wins a contention has to select some early slot, even close to the first one.

4.2 Channel Capacity vs. Number of Nodes

In Table 2 and the corresponding Fig. 2, the channel capacity versus the number of active nodes is presented. These results are obtained by finding the optimal number of nodes ($n_{opt}^{(W)}$) that maximizes the channel throughput for selected sizes of a contention window according to the formula (1). As expected, the channel capacity is a decreasing function of the number of active nodes due to the increase of the probability of collision. However, a decrease of the channel capacity is small and varies from 0.82 for two contenders to almost 0.8 for dozens of active nodes. Summing up, the capacity of the *p*-persistent CSMA with a minor error might be assumed to be constant and approximated by 0.8 for packet lengths in the order of ten bytes typically used in sensor/control networking. This result is an important quantitative reference for the evaluation of variable-window CSMA protocols because the channel capacity plot (Fig. 2) shows the best-case channel utilization that may be obtained in the *p*-CSMA with varying *p*.

Table 2: Channel capacity versus number of contenders

Optimal number of nodes, n_{opt}	***Channel capacity***
2	0.8205
5	0.8082
11	0.819
20	0.7992
59	0.7969
119	0.7963

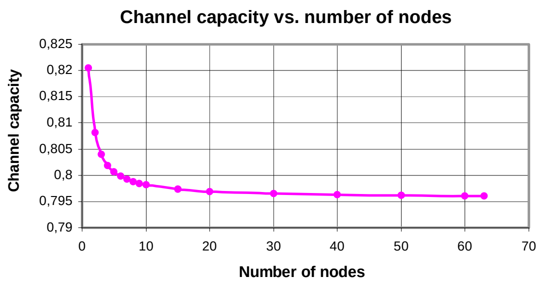

Fig. 2 Channel capacity versus the number of contenders.

4.3 Optimal Contention Window Size

The next interesting issue is the answer to a question how to control the size of the contention window in order to keep the channel throughput close to its maximum shown in Fig. 2. This problem is of a significant interest in the variable-window CSMA protocols with collision avoidance where the size of the contention window is dynamically adjusted to the current channel workload (e.g., LonTalk/EIA-709.1, IEEE 802.11). The aim of such CSMA schemes is to provide the effective adjustment of the window size as a response for varying load conditions that are often bursty especially in event-driven networked sensor/control systems [9,13].

The formula for the optimal number of contention slots W_{opt} with a given number of contenders cannot be derived explicitly from the formula (1) since W_{opt} is the implicit discrete function of the number of contending nodes *n*. We have found the relationship $W_{opt}(n)$ numerically. The corresponding results are shown in Table 3 and Figure 3.

As follows from the analysis of the provided results, the size of the optimal contention window grows almost linearly with growing number of contenders. In particular, the rate of an increase of the optimal size of the contention window equals nearly 5 slots per a node.

Table 3: Optimal size of contention window versus number of nodes

Number of nodes	*Optimal number of slots, W_{opt}*
2	13
5	29
10	56
20	109
30	162

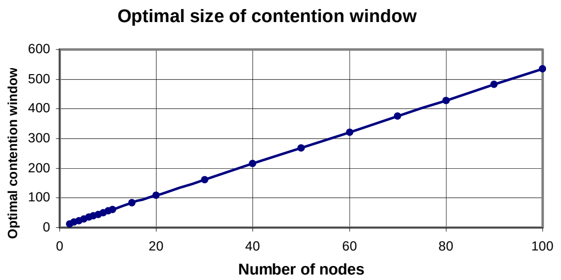

Fig. 3 The optimal size of the contention window versus the number of active nodes.

5. Throughput of Predictive *p*-CSMA vs. Capacity of *p*-CSMA

As was stated in the introduction, the goal of the investigation of the *p*-persistent CSMA capacity was to compare it with the throughput of the predictive CSMA implemented at the MAC sublayer of LonTalk/EIA-709.1 protocol for networked control systems. The plot of the throughput versus the number of active nodes for the predictive *p*-persistent CSMA are presented in Fig. 4. These results are cited from Ref. [3] and have been obtained using the analytical approach based on Markov chains. Since the performance of the predictive CSMA (unlike the *p*-persistent CSMA one) depends on the structure of the traffic transmitted via the channel, several load scenarios are considered in Fig. 4. See [3] for a detailed specification of a particular scenario.
As follows from Fig. 4, for small networks (up to 10 nodes), the throughput equals about 0.8 for any acknowledged service load scenario. For larger networks, the throughput decreases, but establishes next at the constant level for a network containing more than 100 nodes. This measure represents the *sustained throughput* and constitutes the worst-case throughput if the prediction built in the protocol is *effective*, i.e. the contention widow is not limited by its maximum size equal to 1008 slots. The sustained throughput depends on the load scenario, and equals about [3] :

(i) 0.48 if all the messages are unacknowledged,
(ii) 0.63 if all the messages are acknowledged and unicast,
(iii) 0.72 if all the messages are acknowledged and multicast addressed to two recipients.

For networks with a number of nodes greater than about one thousand, the throughput of the predictive CSMA is consistent with the throughput of the (1/1008)-persistent CSMA due to a limitation of the contention window size.
A comparison of the predictive *p*-CSMA throughput and capacity of the *p*-persistent CSMA shows that the former is close to the latter only for networks containing up to 10 nodes. If the number of active nodes is greater, the throughput of the predictive *p*-CSMA is smaller than *p*-CSMA capacity shown in Fig. 2, and ranges from 0.48 to 0.72 depending on the load scenario.
As a matter of fact, it seems that for some network scenarios the throughput of the predictive CSMA is not far from the *p*-persistent CSMA capacity equal to 0.8. However, the analysis reported in [14] shows that the high channel utilization in the predictive *p*-persistent CSMA is obtained at the cost of minimization of the fraction of bandwidth devoted to a transmission of messages carrying application data. In other words, if the throughput of the predictive *p*-CSMA becomes high, most of packets transmitted through the channel are acknowledgements.

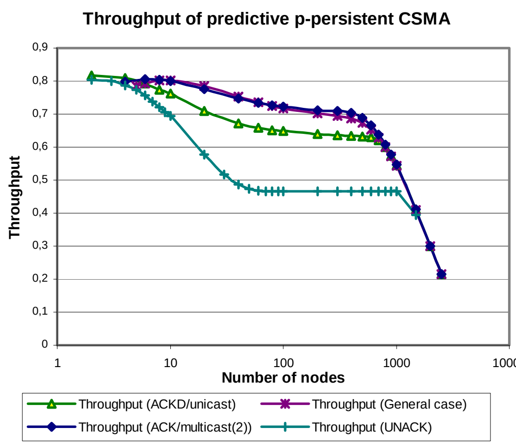

Fig. 5 The throughput of the predictive p-persistent CSMA for various load scenarios [3].

6. Conclusion

The capacity of the slotted-CSMA protocol with a constant contention window has been studied in the paper. The saturation conditions as a traffic model have been supposed.

Summing up, the capacity of the *p*-persistent CSMA with a minor error might be assumed to be constant and approximated by 0.8 for packet lengths typically used in sensor/control networking. This result is an important quantitative reference for the evaluation of variable-window CSMA protocols because the channel capacity plot shows the best-case channel utilization that may be obtained in the *p*-CSMA with varying *p*.
The capacity of the *p*-persistent CSMA has been further compared with the channel utilization of the predictive *p*-persistent CSMA which is an important MAC scheme used in control networked systems. A comparison of the predictive *p*-persistent CSMA throughput and capacity of the *p*-persistent CSMA shows that the former is close to the latter only for small networks containing up to 10 nodes. If the number of active nodes is greater, the throughput of the predictive *p*-CSMA is smaller and ranges from 0.48 to 0.72 depending on the load scenario. Moreover, the relatively high channel utilization (e.g. close to 0.7) of the predictive *p*-CSMA is obtained at the cost of minimization of the fraction of bandwidth devoted to a transmission of messages carrying application data.

References

[1] L. Kleinrock and F.A. Tobagi, “Carrier sense multiple-access modes and their throughput-delay characteristics,” IEEE Trans. Comm., vol.23, pp.1400-1416, 1983.
[2] F. Cali, M. Conti and E. Gregori, “Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit,” IEEE/ACM Trans. on Netw., vol.8, pp.785-799, 2000.
[3] M. Miśkowicz, “A generalized analytic approach to the evaluation of predictive *p*-CSMA/CD saturation performance,” Proc. IEEE Intern. Workshop on Factory Communication Systems WFCS’2006, pp.342-352, 2006.
[4] P. Buchholz and J. Plonnigs, “Analytical analysis of access-schemes of the CSMA type,” Proc. IEEE Intern. Workshop on Factory Communication Systems, WFCS’2004, pp.127-136, 2004.
[5] K. Jamieson, H. Balakrishnan and Y. Tay, “Sift: A MAC protocol for event-driven wireless sensor networks,” MIT, LCS-TR-894, 2003.
[6] Y.C. Tay, Kyle Jamieson, and Hari Balakrishnan, “Collision-Minimizing CSMA and its Applications to Wireless Sensor Networks,” IEEE J. Select. Areas Commun., vol.22, pp.1048-57, 2004.
[7] LonTalk Protocol Specification, Version 3.0, Echelon Corporation, 1995.
[8] D. Dietrich, D. Loy and H.-J. Schweinzer, Open Control Networks: LonWorks/EIA 709 Technology, Kluwer Academic Publishers, 2001.
[9] M. Miśkowicz, R. Golański, “LON technology in wireless sensor networking applications,” Sensors, vol. 6, pp.30-48, 2006.
[10] M. Miśkowicz, M. Sapor, W. Latawiec and M. Zych, “Performance analysis of predictive *p*-persistent CSMA protocol for control networks,” Proc. IEEE Intern. Workshop on Factory Communication Systems WFCS’2002, pp.249-256, 2002.
[11] R Bruno, M Conti, E Gregori, “Optimal capacity of *p*-persistent CSMA protocols,” IEEE Commun. Lett., vol. 7, pp.139-141, 2003.
[12] J. Yeh, “Simulation of local computer networks – a case study ,” Computer Networks, vol. 3, pp.401-417, 1979.
[13] M. Miśkowicz, “Send-on-delta concept: an event-based data reporting strategy,” Sensors, vol. 6, pp.49-63, 2006.
[14] M. Miśkowicz, “Limits of performance of adaptive variable-window CSMA,” Proc. IFAC International Conference on Fieldbuses & Networks in Industrial & Embedded Systems FeT’2007, pp.239-244, 2007.

Marek Miśkowicz reveived his M.Sc. and Ph.D. degrees in Electronic Engineering respectively in 1995 and 2004 from AGH University of Science and Technology (AGH-UST), Krakow, Poland. Currently, he is an Assistant Professor at the Department of Electronics, AGH-UST. His main research interest have been focused on level-crossing sampling, asynchronous analog-to-digital conversion, and performance modeling of networked sensor and control systems. He has published more than 30 papers in international journals and conference proceedings.