A Novel Real-Time MAC Protocol Exploiting Spatial and Temporal Channel Diversity in Wireless Industrial Networks

Kavitha Balasubramanian, G. S. Anil Kumar, G. Manimaran and Z. Wang

Dept. of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011, USA
{kavitha,anil,gmani,zhengdao}@iastate.edu

Abstract. Wireless technology is increasingly finding its way into industrial communication because of the tremendous advantages it is capable of offering. However, the high bit error rate characteristics of wireless channel due to conditions like attenuation, noise, channel fading and interference seriously impact the timeliness and guarantee that need to be provided for real-time traffic. Existing wired protocols including the popular PROFIBUS perform unfavorably when extended or adapted to the wireless context. Other wireless protocols proposed either do not adapt well to erroneous channel conditions or do not provide real-time guarantees. In this paper, we present a novel real-time MAC (Medium Access Control) protocol that is specifically tailored to the message characteristics and requirements of the industrial environments. The protocol exploits both the spatial and temporal diversity of the wireless channel to effectively schedule real-time messages in the presence of bursty channel error conditions. Simulation results show that the proposed protocol achieves much better loss rate compared to baseline protocols under bursty channel conditions.

1 Introduction

The term industrial communication denotes the interaction between various classes of devices in setups such as production control, control of chemical plants, air control, communication systems in cars, planes and trains, power station control and so on. The applications in these setups are very complex, therefore their functionality needs to be distributed to a number of systems or devices, which communicate with each other. In this paper, we are concerned mainly with the traffic generated on a network operating at the device level of factory communication systems which includes various controllers, sensors and actuators.

Industrial networks differ significantly from traditional LANs due to special requirements of their applications like the need for hard timing and bandwidth guarantees and supporting priorities. Predictable inter-task communication is extremely critical in such industrial real-time systems because unpredictable delays in the delivery of messages can affect the completion time of the tasks participating in message communication, resulting in deadline misses and eventually performance losses, halts/resets of manufacturing pipelines or defects in products. Several wired protocols like PROFIBUS are being used in industries and factories that meet such stringent timing requirements.

Recently, the growing popularity of wireless communication in numerous fields has led to its increased dependability, performance improvement and cost reduction. Hence wireless networks are beginning to represent a viable choice for industrial applications because they can offer several attractive features like reduced cost of cabling, ease of configuration and maintenance, extended mechanical freedom and mobility and preventing losses arising due to potential damage of cabling caused by mechanical moving parts, high temperatures and other hostile conditions. Thus, it is very likely in the near future, there will be a proliferation of wireless implementations of factory communication systems.

In spite of having such clear benefits, wireless technology has its own drawbacks arising due to the unreliable characteristics of the wireless medium which makes it, in its current state, unsuitable for supporting real-time communication. Effects due to fading, interference from other users and shadowing from objects degrade the channel performance. In addition, distance dependent path loss and co-/adjacent channel interference influence the channel. Hence the wave propagation environment (number of propagation paths, their respective losses) and its time varying nature (moving people, moving machines and metal surfaces) play a dominant role in constituting channel characteristics [1]. Also due to heavy obstruction, the wireless medium of industrial environments are known to suffer more serious large-scale path loss and fading than other indoor environments [2]. Consequently, the wireless link exhibits both bit errors and packet losses (change in bit values in a packets data part) which vary strongly over time and tend to occur in bursts.

Since wireless networks are substantially different from their wired counterparts with respect to the channel conditions, technologies developed for wired networks cannot be directly adopted. In most wired network models for real-time systems, the communication links are assumed to have a fixed capacity over time. This assumption may be invalid in wireless environments, where link capacities can be temporarily degraded due to fading, attenuation, and path blockage [1]. In addition, existing wireless standards such as IEEE 802.15.1 (Bluetooth) and IEEE 802.15.4 (Zigbee) also provide no mechanisms for supporting real-time messages. Hence, there arises a need to design and develop special MAC protocols and techniques which take both the channel characteristics and the hard real-time requirements of the messages into account. In the next section, we present the related work in this area.

2 Related work and motivation

A number of measurement studies[1, 3, 4]reveal the time-variable and high error rates of the wireless channel. Results published by Willig et. al.[1] indicate that the popular Gilbert Elliot model with some modifications is a useful tool for simulating bit errors on a wireless link, which we use in the present work.

Several proposals have been made that extend the wired protocols used for industrial communication over to a wireless medium. In [5], the authors explore the use of IEEE 802.11 for industrial communication by analyzing the possibility of implementing protocols based on master-slave architecture of traditional field buses on a IEEE 802.11 PHY. In [6], the adaptive-intervals MAC protocol has been proposed that uses a polling-based approach combined with group testing feature for improving the delay in low load conditions. In [7], the authors discuss different architectures that make use of a spread spectrum repeater to integrate distant wireless stations with a wired segment.

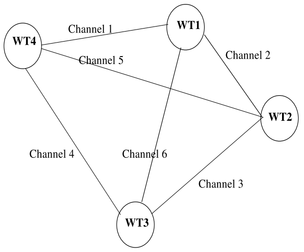
In addition, many MAC protocols and schemes have been proposed to increase the reliability offered by wireless links. In [8] and [9], the authors make use of channel conditions while making packet dispatching decisions. However, the traffic considered in [9] is best effort. In [8], a technique that estimates the channel state beforehand and uses a centralized priority queue based scheduling mechanism is proposed. However, accurate estimation techniques that predict the exact future channel state is unfeasible. In [10], the authors investigate schemes to support combined scheduling of periodic and aperiodic real-time traffic over master-slave Bluetooth networks. In [2], the authors explore the use of Direct sequence spread spectrum(DSSS) CDMA technology to build Industrial Control Wireless LAN with enhanced robustness. In [11] and [12], the authors introduce the concept of antenna redundancy and compare it with modifications made to the Automatic Repeat Request (ARQ) protocol. The ARQ schemes proposed do not work well at high error rates and antenna redundancy requires additional hardware in all communicating devices if any-to-any communication need to be implemented.

The rest of the paper is organized as follows. In Section 3, details about the system model are provided following which we introduce our basic framework in section 4. In section 5, we explain in detail about the Exchange Protocol and present the findings of the simulation studies in section 6. We conclude in section 7 providing directions for our future work.

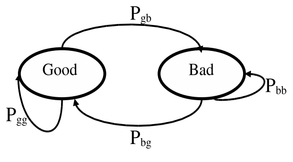
3 Network and channel model

We study a single-hop industrial environment consisting predominantly of real-time periodic message with occasional aperiodic messages/alarms being generated due to faulty or abnormal outcome of some process which require higher priority. The communication medium is wireless characterized by high bit error rate due to phenomena like noise, attenuation, fading and interference. We assume that messages destination is a node in the single hop.

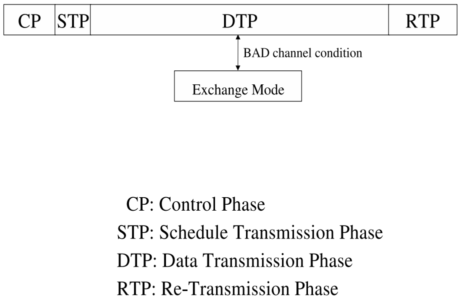
The bursty error characteristics of the wireless environment in a typical industrial setup can be captured by the Discrete-Time Gilbert-Elliot Channel Model [1, 13, 14]. Time in the super-frame is divided into slots and the model works with slotted time where state transitions happen at the end of each slot. The



(a) A typical Industrial wireless network



(b) Gilbert-Elliot model for each channel



CP: Control Phase
STP: Schedule Transmission Phase
DTP: Data Transmission Phase
RTP: Re-Transmission Phase

(c) Super Frame Organization

Fig. 1. Channel model and super-frame format

state space of the Gilbert-Elliot model contains the following two states: GOOD and BAD. When in the GOOD state, no bit errors occur in the data sent in the corresponding slot. Hence the transmission succeeds when done in an exclusive manner. On the other hand, when the channel is in BAD state certain bit errors occur in the received data unit and the data transmission is considered erroneous since an Acknowledgement (ACK) is not received. Fig. 1(b) shows the state diagram along with the transition probabilities. We assume that each channel between a given source-destination pair is statistically independent. In Fig. 1(a), each solid line between two wireless nodes represent an independent channel over which the Gilbert-Elliot channel model is applied.

4 Basic framework

The medium is shared by all the wireless nodes and transmissions follow a super-frame structure that repeats itself. The super-frame is divided into slots and each message would occupy several slots. In a slot, a sender is able to transmit a unit of the message and receive the corresponding acknowledgment(ACK). The absence of an ACK indicates that the channel between the source an destination is in a bad state and the unit is marked for re-transmission.

The basic framework consists of a centralized scheduler that collects all the messages available in the system before every super-frame. The scheduler then prepares a schedule that is followed by all nodes in the system. To facilitate such an approach, every super-frame is divided into the following four phases (see Fig. 1(c)):

- **Control phase (CP):** All the messages in the system are sent to the central scheduler which performs an admission test and constructs a non-overlapping transmission schedule for the admitted messages. The admission test checks if the super-frame has enough free slots to accommodate the next message and its recurring instances (for periodic messages only) before its deadline. Consider a periodic message of size M_i occupying N_i slots of the super-frame. Let N_{data} denote the number of slots of the data transmission phase, $N_{admitted}$ denote the number of slots of the super-frame occupied by already admitted messages; $N_{transfer}$ denote the number of transfer slots and N_{exchg} denote the number of exchange slots (more details about the usage of these slots are provided in Section 5). The admission test checks if

  $$N_i \leq N_{data} - N_{admitted} - N_{transfer} - N_{exchg} \tag{1}$$

  If the above condition is satisfied, the message is admitted to the system and the scheduler reserves N_i slots exclusively for the message; else the message is rejected. However, aperiodic messages are always admitted into the system by removing an instance of the periodic message, since they require higher priority.
- **Schedule transmission phase (STP):** The central scheduler broadcasts the above constructed schedule to all the nodes in the network.
- **Data transmission phase (DTP):** Each wireless node begins its transmission in its scheduled slot. We assume that all the messages that need to be transmitted during the data phase become ready at the beginning of this phase and every message needs to complete before the end of the super-frame. In spite of allocating enough time slots in an exclusive manner, not all messages will reach the destination without errors because of the erroneous channel condition. Therefore, some messages might miss the deadline. The number of deadline misses will depend on the exact data transmission protocol. We present two basic schemes here which would be used for transmitting messages in this phase. However, our main contribution is the Exchange protocol which we present in Section 5 and compare it against the following two basic schemes.

  In Time Division Multiplexing with Variable number of Retransmissions (**TDMVR**), when the channel is underloaded, all the unutilized slots towards the end of the super frame are used for re-transmission. In Time Division Multiplexing with Constant number of Retransmissions (**TDMCR**), the schedule is formed in such a way that all the unutilized slots are equally distributed between the transmitting nodes. Although these schemes enable full utilization of the channel in case of of an underloaded system by increasing the attempts available for existing message transmissions, they do not adapt to the bursty error conditions of the channel. The exchange protocol presented in the next section adapts to the channel conditions thereby decreasing the number of deadline misses and increasing the effective system utilization.
- **Re-Transmission phase (RTP):** All wireless nodes which could not successfully transmit all their messages during the DTP in the first attempt

contend for channel access (CSMA) and employ a backoff algorithm on collision. A fixed percentage of slots in DTP is allocated for re-transmission. At the end of the superframe, the slots that were unable to be successfully transmitted are declared as deadline misses.

5 Slot exchange protocol

We now present the Slot exchange protocol that comes into effect during the DTP as shown in Fig. 1(c). The exchange protocol dynamically adapts to adverse channel conditions and enables effective scheduling of real-time messages in addition to preserving the schedulability guarantees provided to existing messages. Schedulability guarantee implies the fact that when a message is admitted into the system, it is given a certain number of slots (as is occupied by the message) exclusively for data transmission. The scheme caches on two characteristic features - spatial and temporal diversity of the wireless channel; temporal diversity signifies the fact that when a channel is the bad state, it would eventually move to the good state and spatial diversity indicates the condition that if one channel is in bad state, it is possible that a different channel would be in good state.

5.1 Basic idea and illustrative example

During the DTP, each wireless node begins its transmission in its scheduled slot. When a channel between a source destination pair is bad, transmissions begin to fail. During this state, the Exchange protocol is used that works around the occurrences of error bursts. The primary intuition behind the scheme is to postpone the transmission on a channel in a bad state to a later time and schedule transmissions on a channel in a good state with the hope that the channel in the bad state would change into good state in the meantime. This protocol forms its basis on the wireless channel characteristic of correlated packet losses i.e. on a channel which is erroneous, a single packet loss would be followed by back-to-back packet losses. Hence the exchange protocol takes advantage of this characteristic feature to perform efficient scheduling of real-time messages.

Consider a simple network with three wireless nodes shown in Fig 2(a). Let the messages that need to be transmitted be: 12, 23 and 13 where the first number indicates the source and second number indicates the destination. Figure 2(b) shows the channel condition variation with time. The shaded slots indicate that the channel is in bad state. The original schedule given by the central scheduler is shown in Fig. 2(c) and the schedule of the basic schemes is given in Fig. 2(d) which would lead to 6 slots being unsuccessful.

In the exchange protocol, once a node (exchange initiator), notices that its channel to the destination is in bad state, it exchanges its slots (as many as possible) with a different node(exchanged sender). As a result of the exchange, the exchange initiator performs its transmissions in the slots of the exchanged sender and vice versa. This basic idea is depicted in Fig. 2(e), where the exchange initiator, node 2 exchanges its 6 slots with the exchanged sender 1. The final schedule due to the exchange is shown in Fig. 2(f) where only 1 slot is unsuccessful.

Several different heuristics can be applied for a choice of the exchanged sender, based on channel correlation, estimation of the burst length and priority. In this paper, for simplicity, we use the next node in the trasnmission schedule which has a message to transmit for exchange.

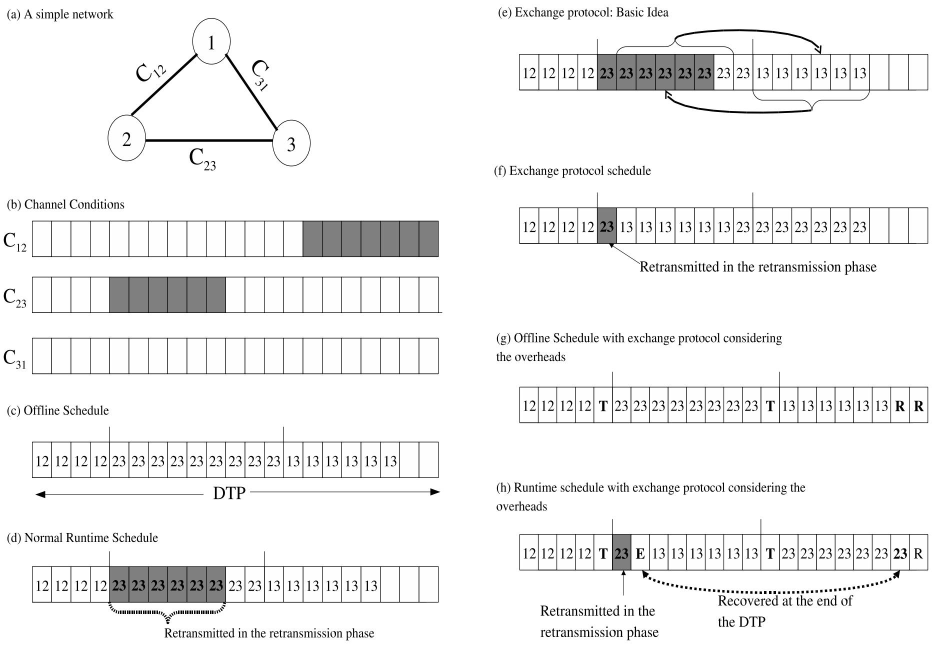


Fig. 2. Illustrative Example

5.2 Protocol details

The basic idea of the exchange protocol is to avoid transmissions on a channel in the bad state by passing control to a different transmitter-receiver pair whose channel is in good state. In order to preserve the schedulability guarantee, the exchange protocol incurs some control overhead.

When a exchange initiator wants to exchange its slots with an exchanged sender, a slot called the exchange slot (slot 7 indicated as E in Fig. 2(h)) is used in which a two way handshake is performed. The exchange initiator sends an exchange request (N_{req}) along with the maximum slots it want to exchange which is typically till the end of its data transmission phase and the exchanged sender replies with an ACK that denotes the actual number of slots it has available for exchange ($N_{available}$). In the example, $N_{exchg} = N_{available} = 6$.

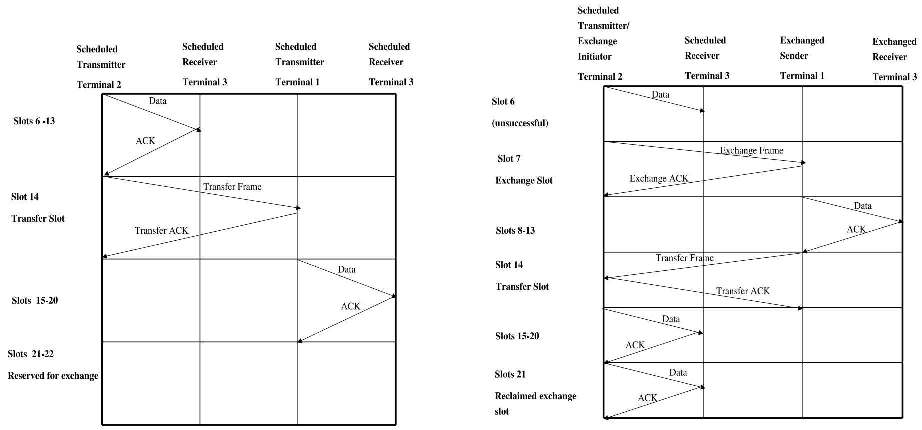


Fig. 3. (a) Timing diagram for the offline schedule (b) Timing diagram with exchange

Since for every exchange initiated, an exchange slot is being consumed, the number of exchanges that can be performed is limited to N_{exchg} in every super-frame. The scheduler broadcasts the N_{exchg} value to all nodes during the STP. To compensate for the exchange slot (to maintain the schedulability guarantee) which are being used by the exchange initiator from the scheduled slots that it has been allocated for transmission, N_{exchg} number of slots are reserved at the end of the super-frame (indicated by R in Fig. 2g). From this pool of reserved slots, every exchange initiator exclusively gets a slot for every exchange it has performed. To enable these functions, an exchange counter, N_{ctr}, is maintained that denotes the number of exchanges that has been performed in the super-frame until the current time. This exchange counter is passed on between the transmitting nodes by means of the transfer slot (indicated by T in Fig. 2g) occurring at the end of every message transmission. Therefore at the beginning of the transmission, each node knows how many more exchanges can be performed. Each time an exchange is performed, the exchange counter is decremented by the exchange initiator and the value of the exchange counter is passed onto the exchanged sender in the exchange slot. In this way, the exchanged sender knows how many more exchanges it can perform during the exchange period. After its exchange period, it passes on the value of the exchange counter to the next transmitting station in the transfer slot. If the exchange counter becomes 0, no more exchanges are performed. If any of the transfer or exchange slots are completed the exchange counter is reset to zero and the transmissions proceed as per the offline schedule.

Let N_{ctr} denote the current value of the exchange counter. When a node uses up an exchange slot for performing exchange, it decrements the exchange counter to $N_{ctr} - 1$ and $N_{exchg} - N_{ctr}$th slot is used by this exchange initiator from the reserved slots. In the above example, assume that $N_{exchg} = 2$. Therefore node 2 has the exchange counter of 2 before performing the exchange. During exchange, it decrements the exchange counter to 1 and uses the (2-1) = 1st slot from among the slots reserved for exchange(slot 21 in the example) since it is the first node performing the exchange. Note that when an exchange initiator performs an exchange, it is limited to its message boundary and it does not spill over into other transmissions.

Hence, by using the transfer slots and the exchange slots, the exchange counter is maintained in a distributed manner. This enables limiting the number of exchanges in every super-frame and thus enables controlling the number of actual slots available for data transmission. In addition, it allows for reclaiming the slots used up for exchange in a exclusive manner; thus preserving the actual number of slots allotted to each node for performing data transmission. Thus the protocol preserves the schedulability guarantee given for messages at the time of admission and effectively uses the channel resources.

The timing diagrams shown in the Fig. 3 explain the exact transmissions that take place for the above example during the working of the Exchange protocol.

6 Simulation studies

We simulated a single hop wireless network with 10 nodes over a 1Mbps channel with periodic messages of size 1050 bits and aperiodic of size 450 bits. Each slot has a time duration equal to the transmission time of 150 bits. Approximately 10% of slots in every super-frame is allocated for re-transmission. We simulated the different channel conditions using the Gilbert-Elliot model for different values of the model parameters. In our simulation studies we compared the performance of the above proposed protocols. The performance metric for all our simulation studies is the loss rate defined as the ratio of number of deadline violated to the number of messages admitted. P_{bb} represents the probability that the channel remains in a bad state given that it is in a bad state. P_{gg} represents the probability that the channel remains in the good state, given that the channel is in a good condition. N_e denotes the number of exchanges that can be performed in a given super-frame. M_s is the number of slots required for the complete transmission of a message. Total number of messages per super-frame is given by N_m.

6.1 Results and discussions

Effect of bad state probability (P_{bb}) : Figure 4(a) compares the loss rates incurred by the above three protocols by varying P_{bb}. The other parameters are assumed as follows: $P_{gg} = 0.9$, $S_l = 1$, $N_e = 11$, $M_s = 7$, $N_m = 10$. The graph has two distinct regions of interest corresponding to $P_{bb} < 0.8$ (small burst region) and $P_{bb} \geq 0.8$ (large burst region). In the small burst region, with low values of P_{bb} the channel quickly switches to the good state and the benefits of the exchange protocol are not very significant. In fact the overhead due to the exchange scheme overshadows the benefits of the protocol. On the other hand, in the large burst region (shown enlarged in Fig. 4(b))which depicts the typical industrial environment, the exchange protocol performs better than the basic

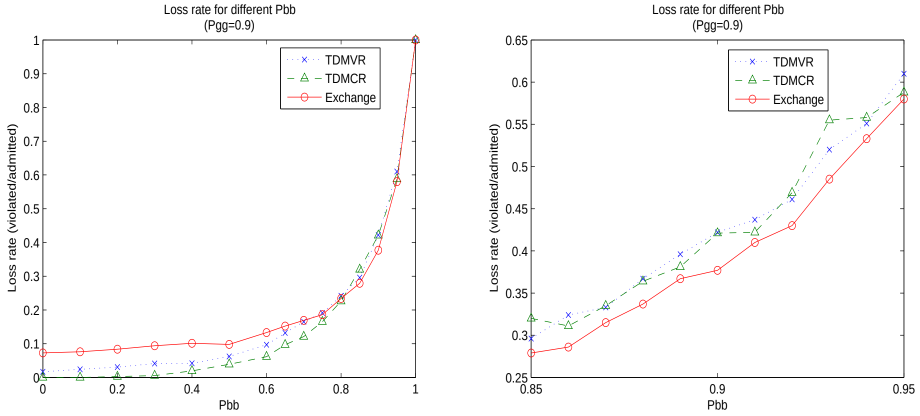


Fig. 4. (a) Effect of P_{bb} (b) Effect of the high P_{bb}

schemes due to the fact that the exchange protocol *exchanges* the slots of a bad channel at the beginning of the bad burst with a good channel which is not noisy. At $P_{bb} = 0.9$, the exchange protocol gives an improvement of 10.6% over TDMCR and 10.5 % over TDMCR. Interestingly, towards the end of the large burst region where $P_{bb} \geq 0.96$, exchange protocols behave similar to the basic protocols due to the fact that the channel experiences significantly long bursts that deferred transmissions also encounter the erroneous channel condition.

Effect of good state probability ($\boldsymbol{P_{gg}}$) : Fig. 5(a) compares the loss rates incurred by the three protocols by varying P_{gg}. The other parameters are assumed as follows: $P_{bb} = 0.9$, $S_l = 1$, $N_e = 11$, $M_s = 7$, $N_m = 10$. The graph has two distinct regions of interest corresponding to $P_{gg} < 0.8$ and $P_{gg} \geq 0.8$. At low values of P_{gg} the channel quickly switches to the bad state and hence experiences frequent bad state bursts whose size is depicted by the P_{bb} value. This results in an exchange being performed from a bad channel to another channel that also moves into bad state frequently; hence the benefits of the exchange protocol are not very significant. At high P_{gg} (shown enlarged in Fig. 5(b)), which is the typical scenario in an industrial environment, the exchange protocol performs better than the basic schemes because the channels are in good state for a longer time and when the channel is erroneous,the exchange protocol *exchanges* its slots with a good channel. At $P_{gg} = 0.91$, the exchange protocol gives an improvement of 14% over the basic schemes.Therefore, at very large values of P_{gg} the exchange protocol performs better than the others and at $P_{gg} = 1$, all the schemes show similar results.

Effect of number of exchange slots ($\boldsymbol{N_e}$) : We study the effect of the N_e by varying the message sizes and number of messages per super-frame. We have chosen $P_{gg} = P_{bb} = 0.9$ for these simulations.

- Effect of message size (M_s): Fig. 6(a) shows the effect of the N_e for different M_s values keeping N_m fixed at 10 . With the increasing N_e, the loss rate for all message sizes decrease due to the benefits of the exchange protocol.

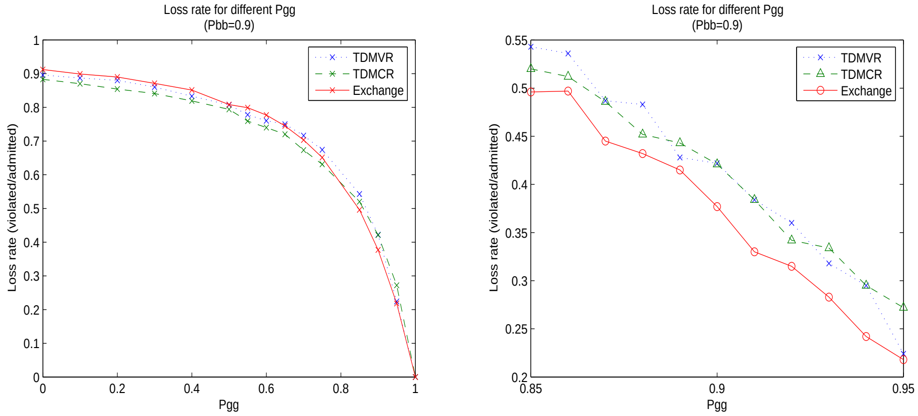


Fig. 5. (a) Effect of P_{gg} (b) Effect of high P_{gg}

However, after a point, N_e becomes more than the maximum number of exchanges that need to be performed and hence the loss rate saturates beyond that point. The saturation point depends on the message size, number of messages and channel parameters. For large message sizes, the saturation point is higher (12 for message of size 10 while it is 8 for message of size 4) since more exchanges can be performed.

– Effect of number of messages (N_m): Fig. 6(b) shows the effect of the N_e for different N_m values keeping M_s fixed at 7 . With increasing N_e, the loss rate for all N_m values decrease due to the benefits of the exchange protocol. However, after a point, N_e becomes more than the maximum number of exchanges that need to be performed and hence the loss rate saturates beyond that point. As in the previous case, the saturation point is higher for large number of messages (2 for N_m= 2 while it is 12 for N_m= 10) since the number of exchanges that can be performed is more when the number of messages increase.

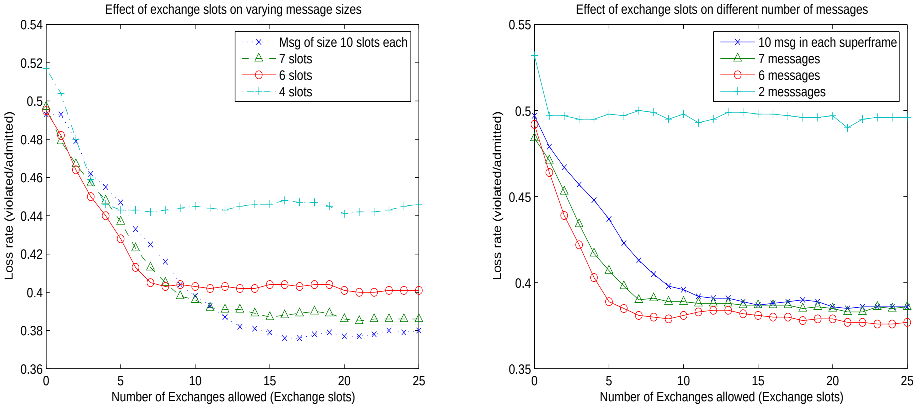


Fig. 6. (a) Effect of N_e for different M_s (b) Effect of N_e for different N_m

7 Conclusions

In this paper, we proposed a novel MAC protocol for real-time message scheduling which adapts to the channel conditions by exploiting spatial and temporal channel diversity characteristics of the wireless medium. Our simulation results show that the proposed exchange protocol provides better loss rate as compared to the generic protocols. In our future work, we would like to make the protocol distributed and extend it to multi-hop networks. We also plan to improve the protocol through channel estimation techniques.

References

1. A. Willig, M. Kubisch, C. Hoene, and A. Wolisz, "Measurements of a Wireless Link in an Industrial Environment using an IEEE 802.11-Compliant Physical Layer," *IEEE Trans. on Industrial Electronics*, vol. 43, no. 6, pp. 1265-1282, Dec. 2002.
2. Q. Wang, X. Liu, W. Chen, W. He and M. Caccamo,"Building Robust Wireless LAN for Industrial Control with DSSS-CDMA Cellphone Network Paradigm," *in Proc. of the IEEE Real-Time Systems Symposium*, Miami, Dec. 2005.
3. D. Eckhard and P. Steenkiste, "Measurement and analysis of the error characteristics of an in-building wireless network," *in Proc. Of ACM SIGCOMM 96 Conference*, pp. 243-254, California, Aug. 1996.
4. D. Duchamp and N. Reynolds, "Measured performance of wireless LAN," *in Proc. of 17th Conf. on Local Computer Networks*,Minneapolis, 1992.
5. D. Miorandi and S. Vitturi, "Analysis of Master-Slave Protocols for Real-Time Industrial Communications over IEEE802.11 WLANs," *in Proc. of INDIN04*, Berlin, June 2004.
6. A. Willig and A. Kpke, "The Adaptive-Intervals MAC protocol for a wireless PROFIBUS," *In Proc. of 2002 IEEE Intl. Symp. on Industrial Electronics*, L'Aquila, Italy, July 2002.
7. P. Morel, A. Croisier, and J.-D. Decotignie, "Requirements for wireless extensions of a FIP fieldbus," *in Proc. IEEE International Conference on Emerging Technologies and Factory Automation*, pp. 116-122, Nov. 1996.
8. P. Bhagwat, P. Bhattacharya, A. Krishna, and S. Tripathi, "Enhancing throughput over wireless LANs using Channel State Dependent Packet Scheduling," *Proc. of the IEEE INFOCOM*, vol. 3, pp. 1133-1140, March 1996.
9. A. Willig, "A MAC protocol and a scheduling approach as elements of a lower layers architecture in wireless industrial LANs," *2nd IEEE Int. Workshop on Factory Communication Systems*, Barcelona, 1997.
10. K. Sairam, N. Gunasekaran, S.R. Redd, "Bluetooth in Wireless Communication," *IEEE Communications*, Vol. 40, No. 6, June 2002.
11. A. Willig, "Antenna Redundancy for Increasing Transmission Reliability in Wireless Industrial LANs," *in Proc. 9th IEEE International Conference on Emerging Technologies and Factory Automation*, pp. 7-14, 2003.
12. A. Willig, "Exploiting redundancy concepts to increase transmission reliability in wireless industrial LANS," *IEEE Trans. on Industrial Electronics*, vol. 50 no. 11, Nov. 2003.
13. E.O. Elliot, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech.* J 42(9): 1977-1997, Sept. 1963.
14. E.N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech.* J. vol. 39, no. 8: 1253-1265, Sept. 1960