Optimisation of IrLAP IrDA Protocol Including Processor Speed

P. Chatzimisios and A. Boucouvalas
Multimedia Communications Research Group,
School of Design, Engineering and Computing,
Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK
{pchatzimisios, tboucouv}@bournemouth.ac.uk
tel. +44-1202-595435
fax: +44-1202-595314

topic: IrDA communication protocol
keywords: IrDA, infrared, processor speed, IrLAP, wireless protocols

ABSTRACT

This paper presents a throughput performance analysis of the IrDA data link layer access protocol (IrLAP) for a 16 Mbit/s wireless infrared link. The optimisation of the link by *adapting frame and window size simultaneously* to optimum values is proposed. The optimum values take into account delays due to preparation, processing and acknowledgement of the packets during transmission and reception. The results of the analysis include calculated maximum throughput and is validated with exact numerical methods. Simple equations for the optimum values for window and frame size are derived which include the effect of processor speed. Moreover, time delays involved in the IrLAP operation are examined.

1. Introduction

The widespread reliance on networking in business and the meteoric growth of the Internet and online services are strong proves of the benefits of shared data and shared resources. Portable devices have seen a rapid growth in recent years, which has led to an increasing demand for wireless data connectivity. These devices range from mobile phones and personal digital assistants (PDAs) to portable computers, digital cameras and printers. The ideal goal of wireless communication is that users can access real-time information anywhere without the need to be wired.

The purpose of development of the IrDA standards is to address indoor, high speed, short range infrared links requiring low power provided with low cost. IrDA offers the advantage of being easy to implement and simple to use, in addition to the high data rates achieved. The IrDA standard defines links offering half-duplex, point-to-point links of data rates ranging from 115.2 Kbit/s [1] to 16 Mbit/s with high-speed extensions [2]. Table 1 shows the various IrDA standards through the years and the data rate for each specification. Also, [3] presents a performance analysis of IrDA operating at data rate of 40 and 100 Mbit/s. Such high data rates maybe considered for future generations of IrDA links.

The IrDA protocol stack specifies three mandatory layers: the physical layer (IrPHY), the data link layer (IrLAP) and the Link Management Protocol (IrLMP) layer. A detailed description of the operation of each component and the optional extensions of IrDA are given in [4].

IrLAP specifies the parameters that infrared devices must negotiate before establishing a link for data transfer. These parameters are data rate, maximum turn-around time, minimum turn-around time, maximum window size and maximum frame size. The negotiation process takes place at 9600 bit/s and the devices choose the best set of parameters that are mutually supported, in order to make data transfers as efficient as possible.

Specification	Data rate	Year
SIR (Serial Infrared)	115.2 Kbit/s	1994
FIR (Fast Infrared)	576 Kbit/s	1995
FIR (Fast Infrared)	1.152 Mbit/s	1995
FIR (Fast Infrared)	4 Mbit/s	1995
VFIR (Very Fast Infrared)	4 Mbit/s	1999
VFIR (Very Fast Infrared)	16 Mbit/s	1999

Table 1 *IrDA specifications and data rates*

2. IrLAP Protocol description

IrLAP, a derivative of the HDLC protocol [5], is presented in detail in [6]. IrLAP uses primary and secondary station roles. Role assignment takes place during the link negotiation and only one station is assigned the primary role and all other stations are assigned the secondary role. If more than two stations are involved in the data exchange, all traffic passes through the primary station.

In the current analysis the saturation case is considered. The transmitter always has information ready for transmission. A window of N frames will be transmitted before the link reverses direction. Window size N represents the maximum number of frames that the transmitter can send without an acknowledgement. The maximum value for N is 127 for the data rates of 4Mbit/s and 16 Mbit/s wireless links.

Setting the Poll/Final (P/F) bit of a transmitted frame reverses link direction. The transmitter is setting the Poll bit (P-bit) to request a response from the receiver. Similarly, the receiver is setting the Final bit (F-bit) to respond to a previous poll command.

Following the transmission of the last frame, the transmitter needs a specific time duration t_{ta} to handle with hardware Tx/Rx turnaround latency and waits for a response from the receiver. If the last frame that contains the P-bit is not correctly received, the receiver doesn't respond. The transmitter waits for a period t_{Fout} and sends a S-frame (Supervisory frame) forcing the receiver to acknowledge the received frames.

The receiver sends a response indicating the next expected data frame if all frames are received without errors. If at least one frame is received in error, then the receiver responds pointing out the first erroneous frame. Correct frame transmissions following an erroneous frame transmission in the same window are considered out of sequence and have to be retransmitted.

A station is allowed to hold the transmission control for a maximum turn-around time equal to T_{max}. This parameter is agreed in the negotiation process between two stations. For data rates less than 115.2 Kbit/s maximum turn-around time must be 500ms. Alternatively, for higher data rates T_{max} may have smaller values. Maximum turn-around time, combined with frame size and link data rate may limit the window size applied.

It is considered here a system architecture represented by the timing diagram of figure 1, in order to derive the IrLAP throughput including the effect of processing time. An example window transmission of 7 frames is considered. Since the distance between wireless devices is up to 1m, it is assumed that propagation delay is very small and considered to be negligible.

The transmitter prepares the first frame *f1* consuming time p_1 and transmits the frame. The receiver is consuming time p_2 to process the received frame and during the same time the transmitter prepares the next frame *f2*. After sending the last frame *f7*, the transmitter waits for an acknowledgement packet. The receiver processes the last frame and the station's transmitter circuit requires time t_{ta} to revert to a transmitting mode. Following this turn-around time, the receiver sends an acknowledgement packet (ACK) and it takes p_3 time for the transmitter to process it. The transmitter had plenty of time to revert to listening mode and subsequently can continue with a new window transmission of frames.

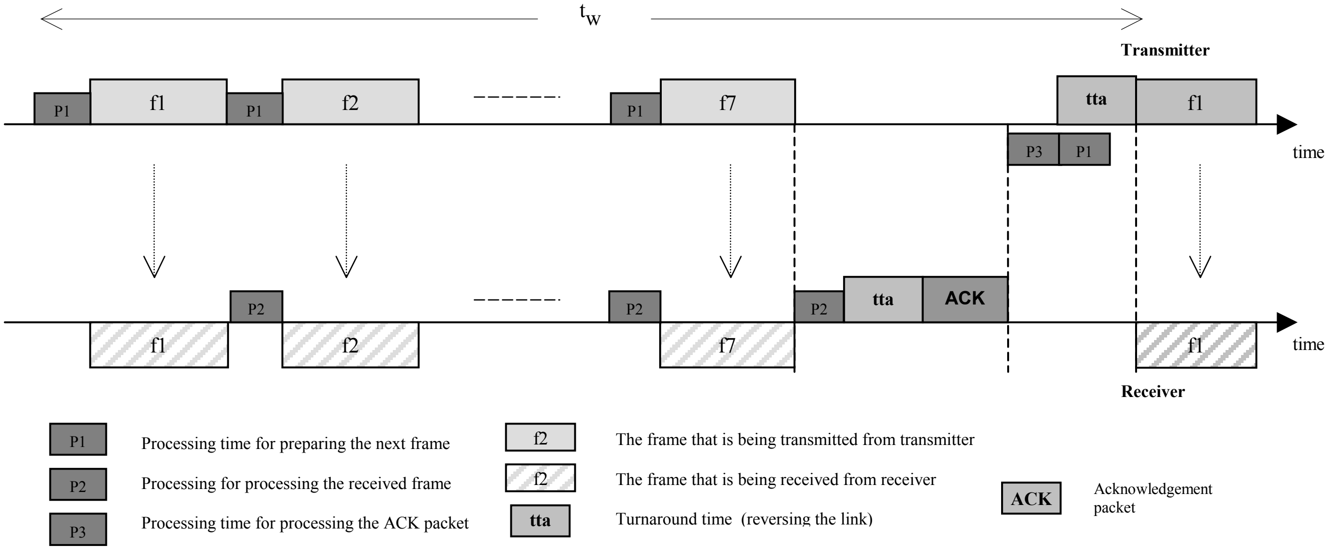


Figure 1 *Timing diagram of data frame transmission for window size of 7*

3. IrLAP Mathematical Modelling

In a previous analysis [7], a mathematical model has been developed using the concept of "Window Transmission Time" (WTT) so the link throughput can be calculated. WTT represents the time needed for a complete window frame transmission and for acknowledgements and delays concerning this transmission. WTT includes the time needed for I-frame (Information frame) transmissions, acknowledgement for the received frames, for reversing the direction of the link and time consumed in possible timer time-out delays.

According to [8], $p_1 = p_3 = 4\text{x}10^3/v$, where v is the processor speed in MHz. Upon preparation and reception of a frame there is time consumed on calculating the 32-bit CRC (p_1,p_2 respectively) in order to check the frame for errors. As a result, it can be assumed that $p_1 = p_2$ and finally the processing times are assumed $p_1 = p_2 = p_3 = 4\text{x}10^3/v$. The parameters used in the current analysis are shown in Table 2.

Parameter	Description	Unit
C	Link data rate	bits /sec
p_b	Link bit error rate	-
p	Frame error probability	-
l	I-frame message data length	bits
l'	S-frame length / I-frame overhead	bits
t_I	Transmission time of an I-frame	sec
t_S	Transmission time of an S-frame	sec
t_{ta}	Minimum turn-around time	sec
T_{max}	Maximum turn-around time	sec
t_{ack}	Acknowledgement time	sec
t_{Fout}	F-timer Time-out period	sec
D_f	Frame throughput	frames/sec
D_b	Data throughput	bits/sec
p_1	Preparation time of an I-frame	sec
p_2	Processing time of an received I-frame	sec
p_3	Processing time of an S-frame	sec
v	Processor Speed	MHz

Table 2 *Parameters used in modelling IrLAP throughput*

The values for t_S, t_I, and p are given by:

$$t_S = \frac{l'}{C}\ , \quad t_I = \frac{l+l'}{C}\ , \quad p = 1-(1-p_b)^{l+l'} \tag{1}$$

Also the data throughput D_b of the link is equal to :

$$D_b = l \cdot D_f \tag{2}$$

and the acknowledgement time t_{ack} is:

$$t_{ack} = 2t_{ta} + t_S \tag{3}$$

If the last frame in the window transmission is correctly received, irrespective of how many frames in this particular window were received in error, the window transmission time is given by:

$$t_w = Nt_I + t_{ack} + (N-1)p_1 + p_2 + p_3 \tag{4}$$

If the last frame is not correctly received, the P-bit is lost and a delay of $t_{Fout} + t_s$ is added in order for the receiver to finally respond. In this event, WTT is:

$$t_w = Nt_I + p(t_{Fout} + t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3 \tag{5}$$

In [9] link throughput was derived without processing times and it is given by:

$$D_b = l\frac{1-p}{p}\frac{\left(1-(1-p)^N\right)}{t_w} \tag{6}$$

Combining (3), (5) and (6), the data throughput D_b , including the effect of processing times, is given by:

$$\boxed{D_b = l\frac{1-p}{p}\frac{(1-(1-p)^N)}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3}} \tag{7}$$

In [6] the time that IrLAP spends on various tasks has been calculated. Considering frame processing times, time utilized on acknowledgements T_{tack} and on transmitting frame overheads $T_{l'}$, are given by:

$$T_{tack} = \frac{t_{ack}}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3} \tag{8}$$

$$T_{l'} = \frac{Nl'/C}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3} \tag{9}$$

Also the time consumed on the P-bit loss T_{Fout} is:

$$T_{Fout} = \frac{p(t_{Fout}+t_s)}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3} \tag{10}$$

Time utilized on retransmission of frames in error T_{error} is given by:

$$T_{error} = \frac{pNl/C}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3} \tag{11}$$

Finally, the delay that processing times introduce, T_{proc} , is:

$$T_{proc} = \frac{(N-1)p_1 + p_2 + p_3}{Nt_I + p(t_{Fout}+t_s) + t_{ack} + (N-1)p_1 + p_2 + p_3} \tag{12}$$

4. Optimum Values Analysis

In a previous work [10], optimum values for either frame or window size have been derived. If window and frame size link parameters can be simultaneously adjusted, throughput performance can be maximized for any processor speed. To derive simultaneously optimum *N* and *l* values we take $\frac{\partial D_b}{\partial N} = \frac{\partial D_b}{\partial l} = 0$.

To begin with, equation (7) is differentiated versus *N* and set to zero.

For low values of *p,* it can be assumed that: $(1-p)^N \approx 1 - Np + \frac{N(N-1)}{2}p^2$

After some calculations, we obtain:

$$N_{opt} = \sqrt{\frac{2(t_{ack}-p_1)}{pt_I + pp_1}} \tag{13}$$

Considering that for $l >> l'$, $p \approx lp_b$ and $t_I \approx \frac{l}{C}$, finally, optimum window size as a function of *l* is given simply by:

$$\boxed{N_{opt} = \sqrt{\frac{2(t_{ack}-p_1)\cdot C}{l\cdot p_b(l+p_1\cdot C)}}} \tag{14}$$

The optimum N value derived can be substituted to throughput equation (7). Throughput D_b becomes a function of frame size l for optimum N values.

The derivative $\frac{\partial D_b}{\partial l}$ can now be taken and set equal to zero to derive optimum l values. N_{opt} given by (13) should be used as the assumption $l >> l'$ is no longer valid as optimum l values may be significantly small. The valid approximation $p = 1-(1-p_b)^{l+l'} \approx (l+l')p_b$ can be assumed and the first derivative versus l is set equal to zero. After some approximations, optimum frame size is found as:

$$\boxed{l_{opt} = \frac{-p_1 \cdot C + \sqrt{p_1{}^2 \cdot C^2 + \frac{4(l' + p_1 \cdot C)}{p_b}}}{2}} \qquad (15)$$

By substituting (15) to (14), we obtain the optimum window size N_{opt}.

5. Validation Analysis

Figure 2, compares the throughput values obtained from (7) using (14) and (15) and those from exact numerical methods with v=500MHz and v=50MHz. As shown in the figure, a perfect match is observed. As a result, all the approximations used to derive (14) and (15) are validated.

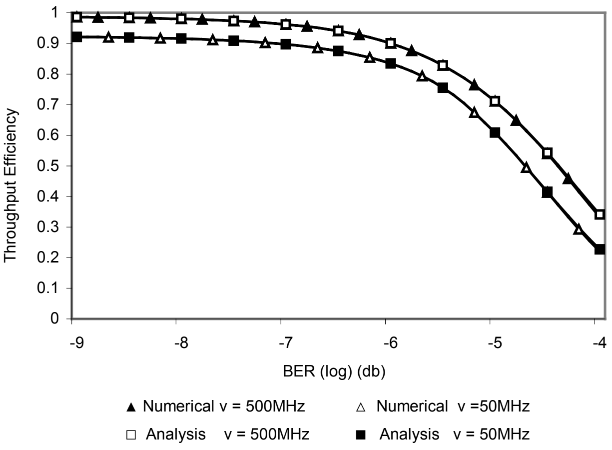


Figure 2 *Throughput Efficiency vs BER for C=16Mbit/s*

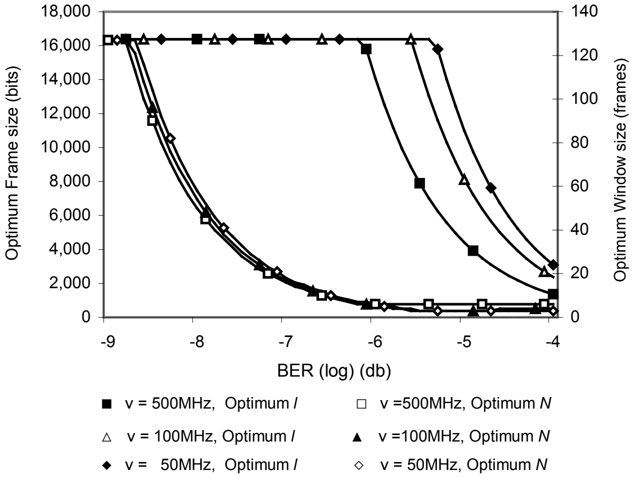


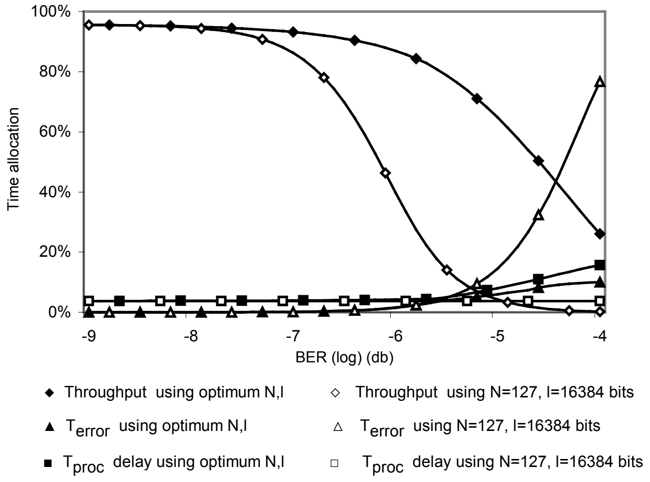
Figure 3 *Optimum frame and window size vs BER for C=16Mbit/s*

6. Analysis Results

Figure 3 shows simultaneously optimum values for frame and window size that maximize throughput against BER for three different processor speeds. For any processor speed, the optimum window size decreases. Particularly, when BER is high (BER<10^{-6}) the window size drops to values under 10 frames in order to cope with the increase of transmission errors. For low BER values, l_{opt} values are above the allowed from IrLAP specification limit of 16K. Additionally, for BER values greater than 10^{-6} , optimum frame size decreases. The decrease level seems to depend on the processor speed. For instance, if v=500MHz , l_{opt} values are descending faster than the other two lower processor speeds (100 and 50 MHz).

Figure 4 plots the time portion consumed for various tasks during the IrLAP operation for v=100MHz. The results from equations (7), (11) and (12) for simultaneously optimum N, l values are compared against those with N=127, l=16384 bits. If optimum values are used, the throughput performance is improved and delays are minimized compared to the non-optimum situation. For N=127, l=16384 bits, the delay T_{proc} that processing times introduce, is almost constant as from (12) it is understandable that the factor $(N-1)p_1 + p_2 + p_3$ is constant and only frame error probability p changes. In addition to this, for BER greater than 10^{-6} the retransmission of frames in error is high. On the contrary, when optimum values are employed the probability of transmitting frames in error is greatly decreased.

Figure 5 shows the throughput efficiency against processor speed for BER equal to 10^{-8} and 10^{-6}. For each BER, two curves are been plotted. The first curve is produced by implementing optimum frame size and window size values simultaneously and the second one with N=127, l=16384. As shown in the figure, there is a significant increase on throughput using optimum values given by (14) and (15), comparing to the non-optimum case for high BER values. Especially as many errors take place (BER=10^{-6}) , the utilization of the link is under 40% . For good quality links (BER=10^{-8}) , optimum values don't improve throughput performance. The situation is explained by considering that, for low BER values, there is no need to adapt optimum values, as the performance of the communication is satisfactory. Moreover, it is observed that throughput increases as the processor speed increases for any BER value.



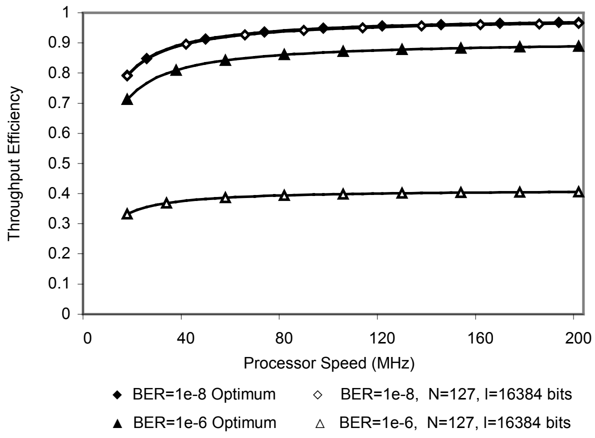


Figure 4 *Time allocation vs BER for processor speed 100 MHz and C=16Mbit/s*

Figure 5 *Throughput Efficiency vs processor speed for C=16Mbit/s*

7. Conclusions

In this work, we derive simple formulas for optimisation of the IrLAP throughput performance. We show that when optimum values of frame and window size are simultaneously employed, the throughput of the link is maximised. The optimum values are especially beneficial for high performance IrLAP especially when BER is high. Additionally, when the processor speed is increased , throughput is improved. This improvement is achieved because high processor speed means that the time consumed on preparing frames for transmission or on checking frames upon reception is reduced. We conclude that maximum benefit is obtained when the processor clock frequency is 5-10 times higher that the data rate.

8. References

[1] IrDA: Serial Infrared Link Access Protocol (IrLAP) – Version 1.1, Infrared Data Association, (1996).
[2] IrDA: Serial Infrared Physical Layer Specification for 16Mb/s (VFIR) –Infrared Data Association, (1999)
[3] A. C. Boucouvalas, V.Vitsas "100 Mb/s IrDA Protocol Performance Evaluation", Proceedings of IASTED International Conference on Wireless Optical Communications (WOC 2001) June 27-29, 2001 Banff, Canada, pp.49-57.
[4] Williams S., "IrDA: Past, Present and Future", IEEE Personal Communications, Vol. 7, No.1, pp 11-19, Feb. 2000.
[5] Bux W., Kummerle K. & Truong H.L., "Balanced HDLC procedures: A performance Analysis", IEEE Trans. Comm, 1980, vol.28, pp.1889-1898.
[6] V. Vitsas and A.C. Boucouvalas "Optimisation of IrDA IrLAP Link Access protocol", Accepted for publication in IEEE Journal on Selected Areas in Communications.
[7] V. Vitsas, A. C. Boucouvalas "IrDA IrLAP Protocol Throughput Performance", Proceedings of IWNA'2000, New Jersey, USA, 30th Nov-1st Dec. 2000.
[8] T. Ozugur, M. Naghshineh and P. Kermani, Comparison of ARQ and SREJ Modes of HDLC over half-duplex and full-duplex links and the effect of window size and processor speed in utilization, Proc. of IEEE PIMRC'98, pp. 708-712, Boston, Sept.1998.
[9]V.Vitsas and A.C. Boucouvalas "Simultaneous Optimization of Window and frame Size for IrDA IrLAP Links" IEE Electronic Letters, 2nd August 2001, Vol.37, No.16, pp1042-1043.
[10] P. Chatzimisios and A. C. Boucouvalas "IrLAP IrDA Protocol Throughput Dependence on Processor Speed", Accepted for publication on International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP 2002).