

Optimising IrDA throughput by including processing time with physical layer consideration

Pi Huang, P. Chatzimisios and A.C. Boucouvalas

Multimedia Communications Research Centre, School of Design, Engineering and Computing,
Bournemouth University, Fern Barrow, Poole, BH12 5BB, U.K.

(phuang@bournemouth.ac.uk, pchatzimisios@ieee.org, tboucouv@bournemouth.ac.uk)

In this work, we derive new equations linking the IrDA data link layer (IrLAP) throughput to physical layer by providing Bit Error Rate (BER) in relation to link distance and various physical layer parameters. Furthermore, we optimise the link by simultaneously adapting frame and window size. The optimum equations include the effect of processor speed. The mathematical analysis is validated by comparison with exact numerical methods. Analytical results presented in the paper explore the effect of frame processing time on improving the performance of IrDA links and identify the improvement in throughput performance by utilizing the derived optimum values.

1. Introduction

In recent years, portable devices have seen a rapid growth and consequently there is an increasing demand for wireless data connectivity. Recent advances in wireless technology have made it possible to equip wireless connection capabilities with many portable devices that allow network communication. Wireless links aim to replace cables between devices such as laptop computers, personal digital assistants (PDA's), video cameras, mobile phones and printers.

In 1993 an industry-based group of companies founded the Infrared Data Association (IrDA). Its members have been responsible for creating a simple, low-cost, low-power, high speed indoor optical wireless communication standard. The developed IrDA platform supports point-to-point links utilizing the infrared spectrum. The IrDA communication ports have been installed in over 600 million electronic devices with a growth rate that still exceeds 20% annually [1]. The 'point and shoot' nature of the IrDA user model requires line-of-sight link alignment, and as a result short data transfer time is important. As the trend of using larger files combined with the need for faster file transfer time continues to become more important, there will be a great demand for much faster short range wireless communication. The IrDA specifications define half-duplex, short range links of data rates ranging from 2.4 Kbit/s [2] to 16 Mbit/s with high-speed extensions [3]. Currently, work continues both in IrDA and the research community for future evolution to 100 Mbit/s and beyond [4]. The IrDA protocol stack specifies three mandatory layers; the physical layer (IrPHY), the Link Access Protocol (IrLAP) and the Link Management Protocol (IrLMP). The present work is mainly focused on IrLAP [5] performance by also considering the interaction with IrPHY.

In this paper, we investigate the IrDA physical layer and identify different physical parameters contributing to the Signal to Noise Ratio (SNR). Subsequently the relation between SNR and Bit Error Rate (BER) for different modulation schemes is derived. Then, an inclusive analysis is carried out on IrLAP. We extend previous work [6] and [7] by developing a mathematical model that takes into account the frame processing time. New equations are derived linking IrLAP throughput to physical layer by providing BER in relation to link distance and various physical layer parameters. Furthermore, we propose the optimization of the link by simultaneously adapting frame and window size. Finally, we identify the potential improvement in throughput performance achieved by utilizing the derived optimum values.

Our paper is outlined as follows: Section 2 introduces IrDA physical layer and derives the relation of SNR and BER. In section 3, we develop a comprehensive mathematical model and derive new equations for the link layer throughput performance. Section 4 validates our mathematical analysis and provides several numerical results, which examine the system performance and demonstrate the improvement on throughput by adopting the optimised link parameters. Finally, conclusions are presented in section 5.

2. IrDA Physical Layer and Bit Error Rate

The IrDA Serial Infrared Physical Layer (IrPHY) defines standards for half-duplex links at several data rates up to 16 Mbit/s. Different modulation schemes are used for different data rates. IrPHY ver. 1.0 Serial Infrared (SIR) [2] defines hardware specifications for 2.4 Kbit/s to 115.2 Kbit/s data rates using conventional serial Universal Asynchronous Receiver Transmitter (UART) chips. It employs on-off keying (OOK) with return-to-zero (RZ) pulses having a duty cycle of 0.1875. IrPHY ver. 1.1 Fast Infrared (FIR) [8] introduced 0.576 Mbit/s and 1.152 Mbit/s data rates utilizing OOK with RZ pulses having a duty cycle of 0.25, and the 4 Mbit/s data rate employing four-pulse-position modulation (4PPM). IrPHY ver. 1.4 Very Fast Infrared (VFIR) [3] added the 16 Mbit/s data rate by using the dedicated HHH encoding, a code rate of $2/3$, $(d, k) = (1, 13)$ Run Length Limited (RLL) modulation.

According to [9], by assuming the ambient light as the dominant noise source, the peak electrical line-of-sight signal to noise ratio (SNR) is given in (1):

$$SNR = \frac{RAI_0^2}{2d^4qE_B B} \quad (1)$$

where R is the detector responsivity (A/W), A is the detector area (m^2), I_0 is the transmitter axial radiant intensity (W/Sr), d is the link distance, E_B is the background ambient irradiance (W/m^2), q the electron charge (1.6×10^{-19} C) and B is the receiver bandwidth (Hz), the minimum required B for SIR, 0.576 and 1.152 Mbit/s links, 4PPM and HHH are $\frac{16}{3}C$, $4C$, $2C$ and $\frac{3}{2}C$ respectively, C represents the link data rate.

For an on-off-keying modulation scheme, considering a threshold voltage v_T , an error will occur when the receiver output voltage $v < v_T$ for a '1' transmitted, or $v > v_T$ for a '0' transmitted. The average BER or probability of bit error p_b , assuming an equal probability of a '1' and '0' bit being in error is given in [10]:

$$p_b = \frac{1}{2} Pr[v < v_T | 1] + \frac{1}{2} Pr[v > v_T | 0] = \frac{1}{2} Q\left(\frac{v_1 - v_T}{\sigma_1}\right) + \frac{1}{2} Q\left(\frac{v_T}{\sigma_0}\right) \quad (2)$$

where $Q(x)$ is the Gaussian tail integral function:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-y^2/2} dy \quad (3)$$

The values σ_1 and σ_0 are the noise variance for a signal '0' and '1'. We consider the noise power is dominated by the ambient light induced shot noise, this gives:

$$\sigma_1 = \sigma_0 = \sqrt{N_0 B} \quad (4)$$

where N_0 is the double-sided noise power spectral density (A^2/Hz).

We assume the threshold v_T is chosen at the half of the received signal level ($v_T = v_1/2$). Applying this to equation (2) gives:

$$p_b = Q\left(\frac{v_1}{2\sqrt{N_0 B}}\right) \quad (5)$$

Since $\frac{v_1}{\sqrt{N_0 B}} = \sqrt{SNR}$, a relation between BER p_b and electrical peak SNR is obtained:

$$p_b = Q(\sqrt{SNR} / 2) \quad (6)$$

PPM encoding is achieved by subdividing data symbol duration into a set of equal time slices called "chips". Each chip position within a data symbol represents one of the possible bit combinations. For 4PPM modulation, each data symbol has four chips and represents 2 data bits (code rate of 1/2). Since each chip has the same BER as the OOK symbol (6), the BER for 4PPM can be derived as:

$$p_b = 1 - (1 - Q(\sqrt{SNR} / 2))^2 \quad (7)$$

In a similar fashion, the BER for the 2/3 code rate HHH modulation is given by:

$$p_b = 1 - (1 - Q(\sqrt{SNR} / 2))^{3/2} \quad (8)$$

Cyclic Redundancy Check (CRC) coding is applied on each IrLAP frame for error checking. Erroneous frames will be detected and recovered by the IrLAP protocol.

In Fig.1, the average BER is plotted against link distance in the range of 0.2-2m for four different data rates. According to [3], the following physical parameters are used: $A=0.16\text{cm}^2$, $R=44 \mu\text{A/mW}$, $I_0=109\text{mW/Sr}$, $E_B=490 \mu\text{W/cm}^2$. Note that the results are shown in the logarithmic scale. The average bit error rate, as expected, increases with the link distance. IrLAP standard [3] specifies the optimum working condition at $\text{BER} < 10^{-9}$. For 115.2 Kbit/s and 0.576 Mbit/s, even at a distance of 2m, the BER is still lower than 10^{-9} . However, the high speed links, 4 Mbit/s and 16 Mbit/s, are more susceptible over the long link distance. In order to have a low BER environment, the link distance should not be over 1 m and 1.5 m respectively for the 16 Mbit/s and 4 Mbit/s links.

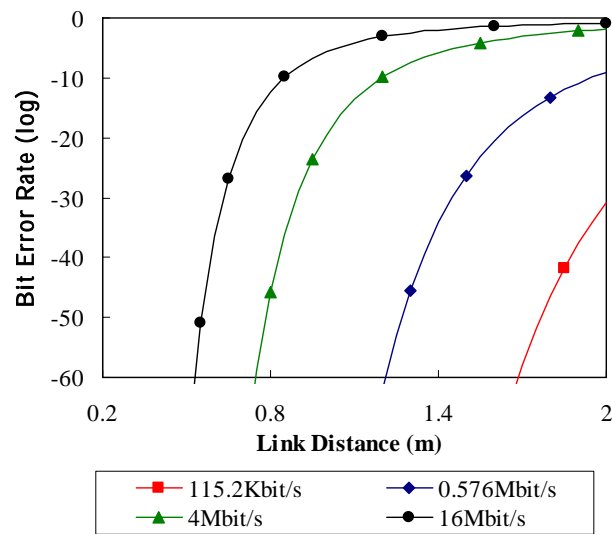


Fig. 1. The average Bit Error Rate for the link distance in the range of 0.2-2 m

3. IrLAP modelling and optimisation

I. IrLAP properties

The IrDA Link Access Protocol (IrLAP) implements a Go-Back-N Automatic Repeat Request (ARQ) scheme in order to transmit data within a window of N Information (I-) frames and each frame with size of l_{LAP} . IrLAP manages the transmission by assigning primary and secondary stations. The primary station initiates transfers to the secondary station and manages the link. When the primary completes the transmission of N (window size) I-frames, it then sets the Poll (P) bit in the last I-frame to signal link turnaround and requests an acknowledgement (a Supervision (S-) frame) from the secondary. Once P bit is set, the secondary can start sending data. The secondary changes P bit to 0 to turnaround the link when it finishes transmission. Referring to the IrDA specifications [3] and [5], the window and frame size range from 1-127 and 128-16384 bits, respectively.

The considered performance model examines data transfer from primary to secondary stations. We assume the saturation case, where the transmitter (primary) always has information ready for transmission to the receiver (secondary) and a window of N frames is transmitted before reversing link direction. Our model utilizes the concept of “Window Transmission Time” (WTT), which is defined as the average time needed for a complete window transmission. WTT represents the time needed from the beginning of the window’s first frame transmission to the beginning of the next window’s first frame transmission. WTT incorporates time needed for a complete window frame transmission of I-frames, acknowledgements for the received frames, packet processing time, delays for reversing the link direction and time consumed in possible timer time-out delays.

Research in [6] pointed out that WTT should include the time delays in order to prepare each frame for transmission (p_1), to process the received frame by the receiver (p_2) and to process the acknowledgement (p_3).

According to [6], $p_1 = p_3 \approx 4 \cdot 10^3 / w$, where w is the processor speed in MHz. The processing times p_1 and p_2 are mainly consumed in calculating the 32-bit CRC (upon preparation and reception of a frame respectively) for the purpose of checking the frame for errors. We assume transmitter and receiver have the same processor ($p_1 = p_2$), thus, $p_1 = p_2 = p_3 = 4 \cdot 10^3 / w$. WTT should also include the link propagation delay (t_p). Although propagation time has minor effect on the throughput with current IrDA data rates [7], it would become more important for the future high data rates. Thus, we also include t_p in our model, $t_p = c/d$, where c is the light speed (3×10^8 m/s) and d is the link distance (m).

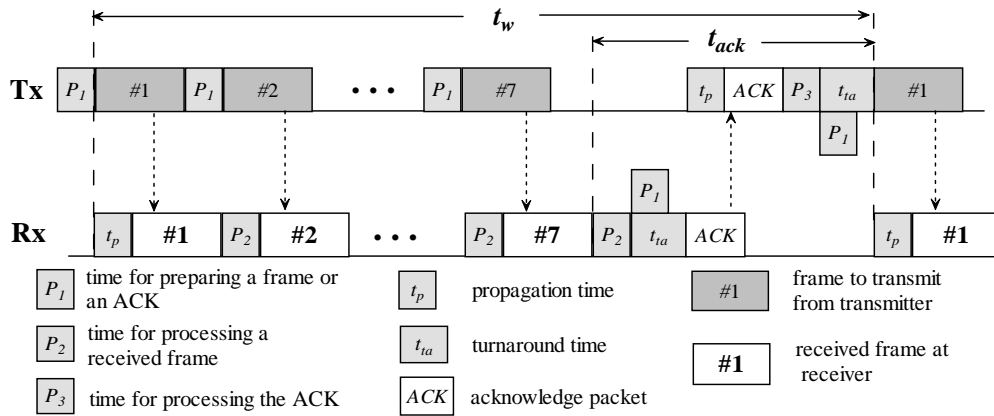


Fig. 2. IrLAP timing diagram for window size of 7

The timing diagram presented in Fig.2, considers both processing time and propagation time for a window transmission of 7 frames. More specifically, the transmitter prepares the first frame (#1) consuming time equals to p_1 . The frame is transmitted and arrives at the receiver after t_p due to propagation time. The receiver needs time p_2 in order to process the received frame and the transmitter prepares #2. After sending the last frame #7, the transmitter waits for the acknowledgement packet. The receiver processes #7 and the receiver circuit requires turnaround time t_{ra} to recover after the reception and to revert to transmitting mode. After t_{ra} , the receiver sends an acknowledgement (ACK) packet and it takes time p_3 to process it. Subsequently, the sender continues with a new window transmission of frames.

II. IrLAP throughput analysis

According to [5], the S-frame transmission time t_s , I-frame transmission time t_l , frame error probability p and F-timer t_{Fout} are given by:

$$t_s = l'/C, \quad t_l = l_{LAP} + l'/C, \quad p = 1 - (1 - p_b)^{l_{LAP} + l'}, \quad t_{Fout} = t_l + 2t_{ra} \quad (9)$$

where l_{LAP} and l' are the I-frame payload and overhead size respectively and p_b is the bit error rate. In [10], the throughput of an IrDA wireless link was derived without considering either the influence of processing time or propagation delay:

$$D_b = l_{LAP} \frac{1-p}{p} \frac{(1-(1-p)^N)}{t_w} \quad (10)$$

By including frame processing time and propagation delay (Fig.2), the window transmission time becomes:

$$t_w = Nt_I + p(t_{Fout} + t_s + 2t_p) + t_{ack} + (N-1)p_1 + t_p \quad (11)$$

where t_{ack} is the acknowledgement time of a frame and includes frame processing times:

$$t_{ack} = 2t_{ta} + t_s + p_2 + p_3 + t_p = 2t_{ta} + t_s + 2p_1 + t_p \quad (12)$$

Thus, the link throughput D_b taking into account the effect of processing and propagation times is given by:

$$D_b = l_{LAP} \frac{1-p}{p} \frac{(1-(1-p)^N)}{Nt_I + p(t_{Fout} + t_s + 2t_p) + t_{ack} + (N-1)p_1 + t_p} \quad (13)$$

III. Optimum window and frame size

Increasing frame size will reduce the relative frame overhead and reduce the frequency of link turnaround following transmission of a window, thus tending to increase throughput. However, increasing frame size also increases the frame error rate, thus increasing the average number of frame retransmission in a window and decreasing throughput. Similarly, increasing the window size reduces the frequency of link turnaround, tending to increase throughput. However, increasing the window size also increases the error probability, thus increasing the average retransmission window and reducing throughput. Therefore, there is an optimum frame size and window size for a particular data rate, BER and parameter set that will maximize throughput.

In order to achieve maximum throughput performance, optimum window size N values for fixed frame size l_{LAP} will be derived next. Equation (13) is a function of both N and l_{LAP} . The optimum window and frame size can be derived by letting the first order derivative of throughput equal zero in the following equation:

$\frac{\partial D_b}{\partial N} = \frac{\partial D_b}{\partial l_{LAP}} = 0$. To begin with, we first calculate $\frac{\partial D_b}{\partial N} = 0$. For small p , we assume that:

$$(1-p)^N \approx 1 - Np + \frac{N(N-1)}{2} p^2 \quad (14)$$

We approximate: $t_{ack} + 2t_p - p_1 + p(t_{Fout} + t_s + 2t_p) \approx t_{ack} + 2t_p - p_1$, $p \ll 2$, $t_I \approx \frac{l_{LAP}}{C}$ and $p \approx l_{LAP} p_b$.

After some algebra, the optimum window size N_{opt} for a fixed frame size is given by:

$$N_{opt} = \text{round} \left(\sqrt{\frac{2(t_{ack} + t_p - p_1)C}{l_{LAP} p_b (l_{LAP} + p_1 C)}} - \frac{(t_{ack} + t_p - p_1)C}{l_{LAP} + p_1 C} \right) \quad (15)$$

where *round* means round the result to the nearest integer. The optimum window size is substituted to (13).

Thus, we have a throughput equation of frame size l_{LAP} for N_{opt} . $\frac{\partial D_b}{\partial l_{LAP}} = 0$ is then taken to derive the optimum frame size. After some approximations and calculations, optimum frame size l_{opt} is given in (16):

$$l_{opt} = \text{round} \left(\sqrt{(l' + CP_1)^2 + \frac{l' + CP_1}{P_b}} - (l' + CP_1) \right) \quad (16)$$

By substituting (16) to (15), the optimum window size N_{opt} independent of l_{LAP} is obtained.

4. Performance results

A few assumptions are made in deriving the optimum window and frame size. In order to examine the accuracy of (15) and (16), verification is carried out in this section. The comparison of the exact optimum results and the analytical optimum results are given next. The exact optimum results are obtained using a numerical algorithm which locates maximum throughput (using (13)) by numerically cycling the integer values of N or l_{LAP} in the range 1-127 and 512bit-16Kbit respectively for different BER's.

Fig. 3, 4 and 5 confirm the accuracy of the considered assumptions in our mathematical analysis. The figures provide performance results (throughput efficiency (TPE), simultaneously optimum frame and window size) versus BER values for two different data rates, 4 Mbit/s and 16 Mbit/s. The values of the other parameters are given as follows: $t_m = 10^{-4}$ s, $d = 1$ m and $w = 300$ MHz. Unless otherwise specified, these parameter values are used throughout this paper.

As a very good match between values given by (15) and (16) and optimum values derived by using numerical methods is observed, approximations made to derive (15) and (16) are validated. Slight differences are observed because optimum N values given by the mathematical analysis and (15) are real values and have to be rounded as N can take only integer values. These differences result in negligible difference in TPE (Fig.3) for both 4 and 16 Mbit/s data rates. We also observe that throughput highly depends on BER and in fact degrades for high BER values.

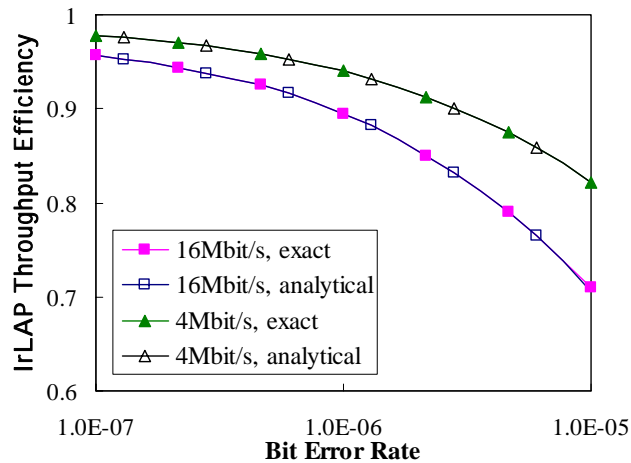


Fig. 3. Comparison of TPE by using exact algorithm and analytical formula for 4 and 16 Mbit/s links

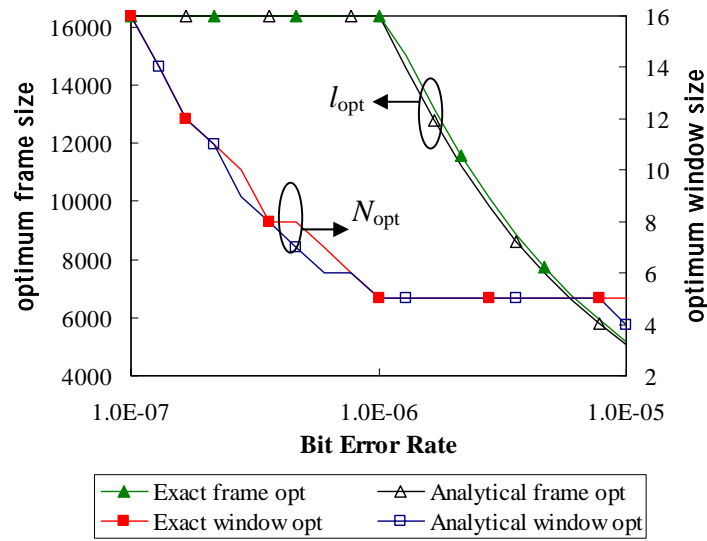


Fig. 4. The corresponding optimum frame and window size for 16 Mbit/s links

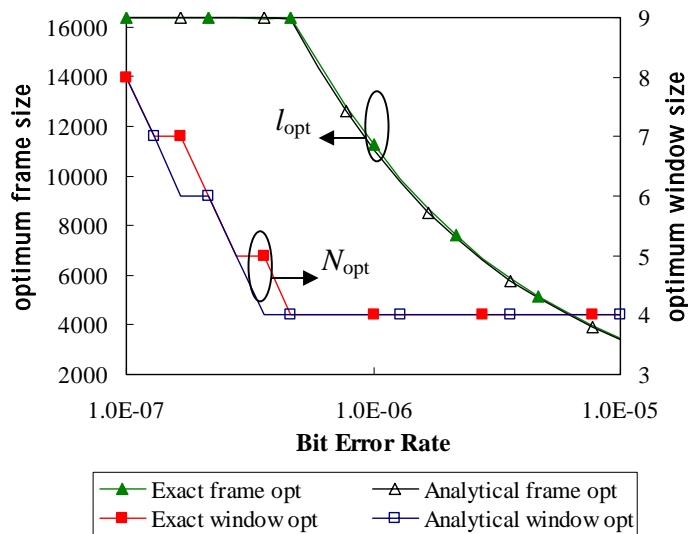


Fig. 5. The corresponding optimum frame and window size for 4 Mbit/s links

The simultaneously optimum frame and window size values versus BER are shown in Fig. 4 and 5, for 4 Mbit/s and 16 Mbit/s, respectively. As expected, the optimum frame size decreases with increasing BER. Note that at very low BER, the optimum frame size values should be greater than the maximum value of 16 Kbits allowed by the IrLAP specifications. Optimum window size also decreases with BER. At very low BER, similarly the optimum window size values should be also greater than the maximum value of 127 allowed by the IrLAP specifications.

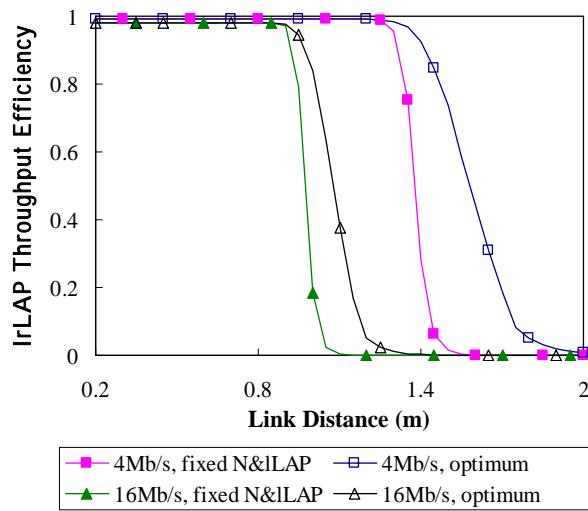
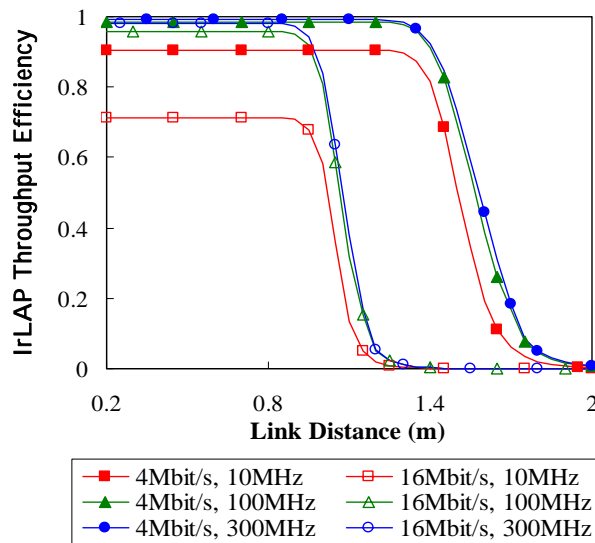
Fig. 6. IrLAP TPE comparison by using optimum and fixed N & l_{LAP} size

Fig. 7. IrLAP TPE of 4Mbit/s and 16Mbit/s links for three different processor speeds

The TPE's of fixed and optimum N & l_{LAP} are compared in Fig.6 for the data rates of 4 Mbit/s and 16 Mbit/s. The fixed N and l_{LAP} are set at 127 and 16384 respectively. As shown in Fig.6, for both data rates, there is a clear 'cut-off' point (~ 1.5 m for 4 Mbit/s, ~ 1 m for 16 Mbit/s) after which the TPE decreases significantly with increasing link distance. This is due to the corresponding increment of BER (Fig.1). It can be seen that the use of optimized N and l_{LAP} yields a more gradual 'cut-off' in throughput with link distance, while the use of fixed N and l_{LAP} produces a sharp decrease in throughput. In practice this would result in a link not suddenly failing beyond a critical distance but more gradually reducing performance.

Fig. 7 investigates the dependency of TPE on the processor speed and the link distance by utilizing the optimum N & l_{LAP} . The figure studies 4 and 16 Mbit/s data rates and reports three different processor speeds. The results show that processor speed highly affects the performance of an IrDA link. In particular, by employing high processor speeds, performance is improved significantly since the time in packet preparation and processing is minimized. Typically a processor 10 times the data rate should be used in order to avoid significant reduction on performance due to processor speed. However, TPE considerably drops when the link distance increases (thus BER increases) regardless the employed processor speed as this depends on the link protocol.

5. Conclusions

In this article, new equations are derived linking the IrDA data link layer throughput to physical layer by providing bit error rate in relation to link distance and various physical layer parameters. An inclusive analysis of IrLAP protocol has been carried out by incorporating the processing speed and propagation delay. Furthermore, the optimization of the link throughput due to bit errors is proposed by simultaneously adapting IrLAP frame and window size. The derived mathematical analysis is validated by comparison with the exact numerical method. Analytical results explore the effect of processing speed on the performance of IrDA and demonstrate the potential throughput improvement when utilizing the derived optimum values.

References

1. C. D. Knutson with J. M. Brown, "IrDA Principles and Protocols: The IrDA Library", volume 1, MCL Press, 2004.
2. IrDA: "Serial Infrared Physical Layer Specification (IrPHY) – version 1.0", (IrDA, 1994).
3. IrDA: "Serial Infrared Physical Layer Specification (IrPHY) – version 1.4", (IrDA, 2001).
4. A. C. Boucouvalas and V. Vitsas, "100 Mbit/s IrDA Optical Wireless Throughput Performance Analysis", IEE Proceedings Optoelectronics, vol. 150, no. 5, pp. 465-472, Oct. 2003.
5. IrDA, Serial Infrared Link Access Protocol (IrLAP), Version 1.1, Infrared Data Association, 1996.
6. P. Chatzimisios and A. C. Boucouvalas "IrLAP IrDA Protocol Throughput Dependence on Processor Speed", CSNDSP 2002, Staffordshire, UK, pp. 272-275, July 2002.
7. Pi Huang and A. C. Boucouvalas, "A Link Layer analytical model for high speed full-duplex Free Space Optical links", International Optical Wireless Communications Symposium, Sept 2003, Warwick, UK.
8. IrDA: "Serial Infrared Physical Layer Specification (IrPHY) – version 1.1", (IrDA, 1995).
9. Kahn J.M. and Barry J.R., "Wireless infrared communications", Proc. of the IEEE, vol. 85, pp. 265-298, Feb. 1997.
10. P. Barker, "Analytical and Simulation Performance Modelling of Indoor Infrared Wireless Data Communications Protocols", PhD Thesis, Bournemouth University, 2003.