



Division for Information Technology
Department of Computer Science

Stefan Alfredsson

TCP in Wireless Networks: Challenges, Optimizations and Evaluations

Karlstad University Studies

2005:13

Stefan Alfredsson

TCP in Wireless Networks: Challenges, Optimizations and Evaluations

Stefan Alfredsson. *TCP in Wireless Networks: Challenges, Optimizations and Evaluations.*

Licentiate thesis

Karlstad University Studies 2005:13

ISSN 1403-8099

ISBN 91-85335-53-3

© The author

Distribution:

Karlstad University

Division for Information Technology

Department of Computer Science

SE-651 88 KARLSTAD

SWEDEN

+46 54-700 10 00

www.kau.se

Printed at Universitetstryckeriet, Karlstad 2005

TCP in Wireless Networks: Challenges, Optimizations and Evaluations

STEFAN ALFREDSSON

Department of Computer Science, Karlstad University

Abstract

This thesis presents research on transport layer behavior in wireless networks. As the Internet is expanding its reach to include mobile devices, it has become apparent that some of the original design assumptions for the dominant transport protocol, TCP, are approaching their limits. A key feature of TCP is the congestion control algorithm, constructed with the assumption that packet loss is normally very low, and that packet loss therefore is a sign of network congestion. This holds true for wired networks, but for mobile wireless networks non-congestion related packet loss may appear. The varying signal power inherent with mobility and handover between base-stations are two example causes of such packet loss. This thesis provides an overview of the challenges for TCP in wireless networks together with a compilation of a number of suggested TCP optimizations for these environments. A TCP modification called TCP-L is proposed. It allows an application to increase its performance, in environments where residual bit errors normally give a degraded throughput, by making a reliability tradeoff. The performance of TCP-L is experimentally evaluated with an implementation in the Linux kernel. The transport layer performance in a 4G scenario is also experimentally investigated, focusing on the impact of the link layer design and its parameterization. Further, for emulation-based protocol evaluations, controlled packet loss and bit error generation is shown to be an important aspect.

Keywords: Wireless networks, TCP/IP, TCP-L, Network emulation

Acknowledgments

I would first and foremost like to thank my supervisor Anna Brunström. Her support and encouragement during my research studies and article writings has been invaluable. I also thank my other co-authors, namely Annika Wennström, Mikael Sternad and Johan Garcia (in order of appearance). My co-supervisor Tony Ottoson deserves a special thanks, as does the other colleagues in the “Wireless IP” project and the Distributed Systems and Communications Research Group, for support and interesting discussions. Anna Widenius put a lot of effort into the construction and development of WIPEMU, which is greatly appreciated. My colleagues at the Computer Science department are well worth an acknowledgement. Thank you for providing such a nice and friendly working environment. Financial support has been provided (in various constellations) by the SSF, Vinnova, CMIT, and the PCC++ research school. I acknowledge this with gratitude.

Last but not least, a big “thank you” to my family and friends for their love, support and patience, and for enduring my sometimes strange working hours.

Karlstad, May 2005

Stefan Alfredsson

List of Appended Papers

This thesis is based on the work presented in the following five papers. References to the papers will be made using the Roman numbers associated with the papers.

- Paper I Annika Wennström, Stefan Alfredsson, and Anna Brunstrom. *TCP over Wireless Networks*. Karlstad University Studies 2004:21, Karlstad University, Sweden, May 2004.
- Paper II Stefan Alfredsson and Anna Brunstrom. Bit Error Transparent Multimedia Transport. Chapter 10 in *Perspectives on Multimedia - Communication, Media and Information Technology*, Wiley 2004.
- Paper III Stefan Alfredsson and Anna Brunstrom. TCP-L: Allowing Bit Errors in Wireless TCP. In *Proceedings of the IST Mobile & Wireless Communications Summit*, Aveiro, Portugal, June 2003.
- Paper IV Stefan Alfredsson, Anna Brunstrom, and Mikael Sternad. Emulation and Validation of a 4G System Proposal. To appear in *Proceedings of RadioVetenskap och Kommunikation*, Linköping, Sweden, June 2005.
- Paper V Johan Garcia, Stefan Alfredsson, and Anna Brunstrom. The Impact of Loss Generation on Emulation-based Protocol Evaluation. *Under submission*.

Some of the papers have been subjected to minor editorial changes.

Contents

Abstract	i
Acknowledgments	iii
List of Appended Papers	v
Introductory Summary	1
1 Introduction	3
2 Data Transmission Background	5
3 Overview of Research Area	8
4 Research Questions	9
5 Research Methodology	10
6 Main Contributions	10
7 Paper Summary	11
7.1 Paper I - TCP over Wireless Networks	11
7.2 Paper II - Bit Error Transparent Multimedia Transport	11
7.3 Paper III - TCP-L: Allowing Bit Errors in Wireless TCP	12
7.4 Paper IV - Emulation and Validation of a 4G System Proposal	12
7.5 Paper V - The Impact of Loss Generation on Emulation-based Protocol Evaluation	12
8 Future Work	13
Paper I: TCP over Wireless Networks	17
1 Introduction	19
2 Wireless Networks	21
2.1 Wireless LANs	21
2.2 Wireless WANs	22

2.3	Conclusion	29
3	TCP	30
3.1	Slow Start and Congestion Avoidance	30
3.2	Fast Retransmit and Fast Recovery	32
3.3	TCP Options	32
3.4	Other Mechanisms	33
3.5	TCP Variants	34
4	Problems with TCP in Wireless Networks	35
5	Proposed Optimizations	36
5.1	Link Layer	36
5.2	Split Connection	38
5.3	Explicit Notification	39
5.4	End-to-end	41
5.5	Summary	44
6	Concluding Remarks	45
Paper II: Bit Error Transparent Multimedia Transport		51
1	Introduction	53
2	Transport Protocol Background	54
2.1	TCP	55
2.2	TCP over Wireless Links	56
2.3	TCP-L	57
3	Performance Evaluation	58
4	Experiment Results	59
4.1	Randomly Distributed Errors	59
4.2	Impact of Burstiness	60
4.3	Experiment Conclusions	61
5	Further Readings	62
6	Concluding Remarks	64
Paper III: TCP-L: Allowing Bit Errors in Wireless TCP		69
1	Introduction	71
2	Header Decoding/Recovery	73
2.1	TCP Header Overview	74
2.2	Phases of Recovery	76
2.3	Recovery Algorithm	78
3	Experiments	78
3.1	Experiment Setup	79
3.2	Experiment Results	80

4	Conclusions	82
Paper IV: Emulation and Validation of a 4G System Proposal		87
1	Introduction	89
2	WIPEMU	90
3	Transport Protocols	92
4	Experiment Overview	94
5	Results	97
6	Conclusions	102
7	Acknowledgements	102
Paper V: The Impact of Loss Generation on Emulation-based Protocol Evaluation		105
1	Introduction	107
2	Network Emulation Background	109
3	Controlled Packet Loss	110
	3.1 Positional Dependencies	111
	3.2 Repeatability	114
4	Controlled Bit-errors	117
	4.1 Time-driven Bit-errors	118
	4.2 Data-driven Bit-errors	120
5	Implementation Details	122
6	Conclusions	123

Introductory Summary



1 Introduction

During the last decades, computing has become more and more pervasive. Personal computers are commonly found in homes, cars are controlled by microchips, and mobile phones have more computing power than an average PC a decade ago.

Closely related to computing is communication, for what use is a computation if its result cannot be communicated? Either to the end user, via screen and keyboard, or to another computing device. By enabling communication between devices, their power is combined, and their efficacy is multiplied many times more.

Communication in its simplest form consists of transmitting information from one place to another, via a medium. In the case of computers, communication is often performed via metallic mediums, such as copper on a circuit board, coaxial cable, or twisted pair wire, popular in the pervasive Ethernet technology. In the case of metallic medium, electrons are used to build up the signals. Glass fiber is also commonly used for communication, but here photons (light) are used instead of electrons. All these mediums require some form of mechanical connection between the communication end-points.

Starting with Marconi and Tesla some 100 years back, we also have the ability to communicate without using wires, as electro-magnetic waves do not require any medium to be present. This allows for a great amount of communication freedom, as phones and computers can communicate without having a mechanical connection.

During the last decade, wireless communication equipment has become readily available for use with personal computers. Wireless local area networking is commonly done with the IEEE 802.11 standard, or enhancements thereof [28]. Wider area coverage is made possible by utilizing the existing mobile phone infrastructure for data transmission in the form of for example GPRS [3, 11]. The integration between packet data and voice communications is further evolved with the deployment of 3G/UMTS [27]. This system is specifically designed to carry packet data, video and voice communications with much higher capacity than previous wide area coverage networks.

Parallel to this development is the expansion of the Internet, which has seen an explosive growth since the introduction of the World Wide Web in the early 1990s. It seems like a natural evolution that the Internet should be accessible wirelessly, just like the mobile phone is superseding the fixed phone. The number of people with only a mobile phone subscription is increasing, and it can

be expected that this trend will continue as the call rates become cheaper and cheaper. Likewise more and more Internet users can be expected to use a wireless instead of a wired connection.

The core of the Internet is the best effort IP packet network, which only forwards packets from one destination to another. On top of this are protocols that provide a data transport service, where the major reliable transport protocol used today is TCP [20]. Although TCP was designed over two decades ago, it has survived the evolution in networking. For example, the bandwidth capacity has increased many orders of magnitudes, while the size of the network has also grown immensely. Thanks to the extensibility of the protocol (see for example [12]) and the key idea that intelligence should be pushed to the network edge (the “end-to-end argument” [22]), TCP still performs very well with wired network technologies. Even as the Internet is used over wireless networks, it works, but the limitations in its design is starting to show [6].

The pervasiveness of the Internet in combination with the increased use of wireless technologies makes TCP over wireless an important research topic, and it is thus a background motivation for this thesis. The reported limitations inspired a further analysis of the workings of TCP in wireless networks. What types of networks are used, how does TCP behave in these, what optimizations have been proposed? The result of this analysis is reported in Paper I. Earlier research on a partially reliable transport protocol [4], the UDP-Lite [13] protocol, and the characteristics of wireless networks, inspired the investigation of the concept of allowing bit errors in TCP. Papers II and III report on the findings of this research. To evaluate the protocols, experimental network test-beds were prepared. Within the test-beds, investigations were also performed on the interactions between the lower layers and the transport layer. Paper IV presents research on the impact of the link layer design and associated parameter settings on the transport layer. Another related aspect is considered in Paper V where the impact of controlled loss for network emulation on protocol evaluation is studied.

The remainder of this introductory summary is organized as follows. Section 2 contains a general background on data transmission, focused on the physical layer. This is followed by an overview of the research area in section 3. Section 4 outlines the research questions for the thesis. The research methodology is discussed in Section 5. This is followed by the main contributions of the thesis in Section 6. Section 7 provides a summary of the included papers. Finally, thoughts on future work are presented in Section 8.

2 Data Transmission Background

Data transmission on the transport layer is extensively discussed in the included papers. Data transmission at the lower layers are not, however. Therefore, this section provides a short introduction to modulation, the basic technology for data transmission. This gives a better understanding for the workings of the lower layers, and explains differences between wired and wireless communication. This section is aimed at readers not already familiar with these concepts.

To facilitate data transmission in a medium on the physical layer, a carrier signal is often used (Figure 1). This carrier is modulated, or modified, in certain ways according to the data being transmitted. The most frequent modulation techniques are amplitude, phase and frequency modulation. For example, common broadcast radio transmissions used amplitude modulation for a long time, known as AM radio. This modulation scheme uses a fixed carrier frequency, and the amplitude of the signal indicates the information transmitted. For binary transmission, the "full" amplitude may represent a 1, and half of that amplitude may represent a 0. This is illustrated in Figure 3, where the carrier has been modulated with the signal in Figure 2.

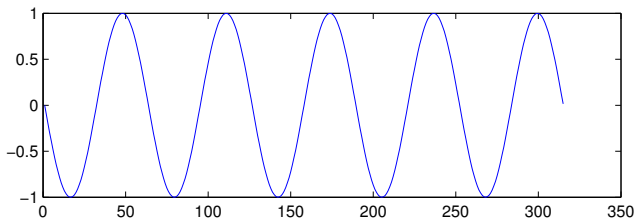


Figure 1: An example carrier signal

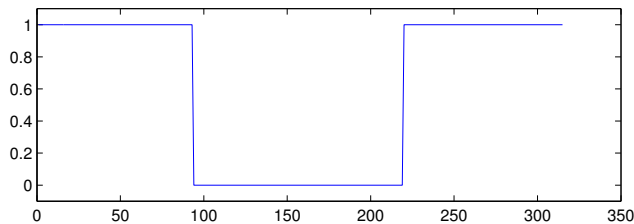


Figure 2: A data signal

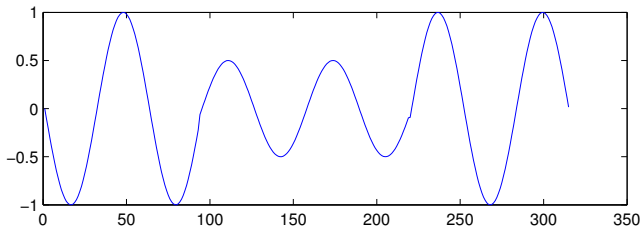


Figure 3: Amplitude modulation

When phase modulation is used, the phase of the carrier signal is shifted. With our binary example, a 1 may be represented with a phase shift of zero degrees, while a 0 may be represented by a 180 degree phase shift, as seen in Figure 4.

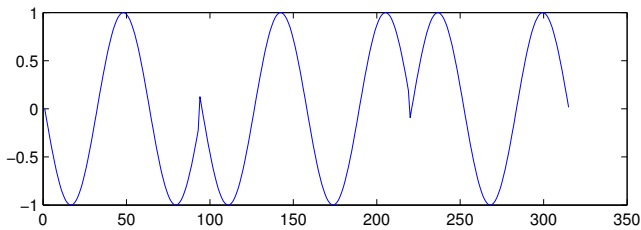


Figure 4: Phase shift modulation

The third way of modulation is frequency modulation (which is widely used in FM-radio broadcasting). Here, the frequency of the carrier signal is varied according to the data that is to be transmitted. For example, a 1 may be represented by a frequency of 800 Hz, while a 0 may be represented by a frequency of 400 Hz, as illustrated by Figure 5.

These examples assume transmission of only 0's and 1's, by using two distinct modulation levels. By adding more levels, for example 25%, 50%, 75% and 100% of the maximum amplitude, more bits can be represented per time unit (i.e. per transmitted *symbol*). This can be increased even further, from 4 to 8, 16, or 32 levels and so on. The limit for the number of modulation levels (or modulation order) is dependent on how clear the signal is compared to the noise¹, named the signal to noise ratio. For example, 25% (of the maximum amplitude) may represent 00 and 50% may represent 01. The sender transmits

¹The noise may also include interference from the sender itself, or other sources.

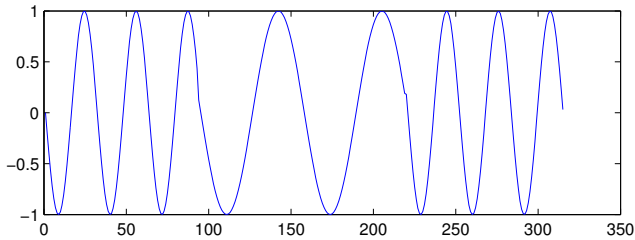


Figure 5: Frequency modulation

at 25%, but on the path to the receiver the signal may have become amplified to 38%. This symbol may then be erroneously demodulated as 01 by the receiver. The better the signal, the more levels can be used, and thus a higher channel capacity is obtained.

To further improve capacity, different modulation techniques can be combined. For example, amplitude and phase modulation is commonly combined. By using two modulation levels for both amplitude and phase, four different combinations (symbols) are created, while still being as robust to noise as each of the modulations themselves. The combination of amplitude and phase modulation is commonly referred to as QAM (Quadrature Amplitude Modulation).

The cause for amplification and attenuation is due to the properties of radio waves. Considering an omni-directional or semi-directional antenna, the radio waves that reach the receiver may arrive directly from the transmitter. The communications is said to be in *line of sight*. Often there are additional components in the received signal, caused by waves that have bounced off buildings, off the ground, or off other objects in the path. These extra rays then contribute to a fuzziness in the reception, as parts of the signals are delayed and all these signals are combined at the point of reception. In urban terrain, line of sight may be rare because of buildings and other objects in the path from the transmitter to the receiver. This complicates the detection further, as all received signals are reflections from other objects. This is characteristic of radio transmission that is non-directional. For example directional laser or microwave point-to-point wireless transmission, does not experience such problems.

Comparing wireless to wired transmission, the wire is a guided and often shielded medium. This means that most of the transmitted energy reaches the receiver with a low amount of interference. The shield also works the other way - transmissions in a wire does not normally² cause interference. This means a

²Of course, there is always a magnetic field around a wire with flowing current. Also, the wire

much greater freedom in choosing frequencies to use, compared to the heavily regulated radio spectrum for wireless transmission. This then allows for a much higher bandwidth, with a lower amount of errors, in wired transmission compared to wireless transmission.

3 Overview of Research Area

As discussed in the previous sections, moving from wired to mobile wireless communication provides additional problems to handle. In general, these can be attributed to higher bit error rates, disconnect/handover events, and limited and varying bandwidth. The proposed optimizations for these problems can be coarsely classified into four groups.

Link layer optimizations adhere to the philosophy that the network should provide the best possible service to the transport layer, by hiding the heterogeneity. For example “snoop” [2] performs caching and local retransmissions over the wireless link to avoid triggering congestion avoidance at the sender.

Split connection approaches instead use two connections. The mobile has one connection to the base station over the wireless link. The base station then has a connection to the server, over a wired network. TCP can then be used between the server and base station, and a more suitable protocol can be used over the wireless link, as is for instance done in WAP [18, 29].

Explicit notification type of optimizations reason that the communication is best served by having information available at the end-point(s). This way, the sender can for example be able to distinguish congestion from data loss in the wireless network. In [10], a scheme is proposed where “ICMP-DEFER” messages can be sent by a base station to the sender. The sender can then avoid shrinking the congestion window.

End-to-end optimizations are implemented in one or both of the communicating end-points. A common argument is that the end-point is the only place that can completely and correctly implement a functionality in question [23]. An example of this is TCP Westwood [5], that modifies the bandwidth estimation algorithm in the sender, which has been shown to improve the throughput in both wired as well as wireless links.

Another area related to our research is the concept of *partial reliability*. The TCP protocol provides full reliability, while the UDP [19] protocol provides no reliability. The *partially reliable transport protocol*, PRTP [4], is an example of

may act as an antenna. For our discussion these issues are not considered.

a TCP modification that allows the application to choose its required reliability level. PRTP operates on a per-packet level. A similar idea, but operating on the *bit-level*, is used by TCP-L. Here, the needed reliability (or rather integrity) of received packets can be set by the application. This is inspired by the UDP-Lite [14] protocol, which uses a resembling concept but for UDP.

The research area also expands into experimental network emulation. Emulation and simulation has for a long time been important tools to evaluate computer networks [1, 8, 9, 16, 17]. Our research continues this tradition, with the emulation of a 4G downlink scenario and emulation of controlled packet loss and bit errors.

To summarize, our research area is focused on experimental evaluations of transport protocols, investigating link-layer and end-to-end optimizations for wireless networking.

4 Research Questions

The overall research question for this thesis is

- *How can performance of the wireless Internet be improved from the transport layer point of view?*

This is a quite broad question, and it is therefore further divided to put it into context. A natural first step is to investigate the current state of the art, leading to the question

- *To which existing wireless networks do problems and proposed optimizations of TCP over wireless networks apply?*

We find that there are many suggestions from the research community, but none that is accepted as a good general solution. All proposed optimizations also conserve the semantics of full reliability used in TCP.

Inspired by the work in e.g. UDP-Lite [13] and PRTP [4], a question is raised about the reliability in relation to the performance of TCP. This question is formulated as

- *What performance can be expected from a more flexible transport service where reliability may be traded for performance?*

For this investigation, we also need to look into the interactions between the transport layer and the lower layers. The link layer for example can provide a fully reliable service at the expense of increased delay, drop erroneous packets, or forward them although erroneous. This raises the question of

- *How do the lower layers and the transport layer interact?*

5 Research Methodology

To answer the questions in the previous section it is good to follow a methodology. There are many methods of performing research. Problems can for example be analyzed analytically, where a system is modeled with mathematics and solved with equations. Another method is experimental research, where a model is subjected to specific inputs, and the outputs are studied. This is especially useful in complex environments where an analytical model is too abstract.

The research methodology used for evaluations in this thesis can be categorized as experimental network emulation. The conceived ideas have been implemented in network test beds. Experiments have then been performed in these test beds. Measured output (the metrics) from these experiments can then be compared with the output from experiments with a base-line (standard) configuration. From such comparisons conclusions can be drawn as to what effect a certain modification or implementation has. It is often impossible, or impractical at best, to test all combinations of parameter settings and metrics, as these grow exponentially for each new parameter. The scope of the evaluated parameters has therefore been limited by necessity, but is believed to capture relevant aspects of the investigated protocols and systems.

6 Main Contributions

The contributions of this thesis consist mainly of three parts.

First, a number of TCP optimizations have been summarized, together with a problem description of TCP in wireless networks. The optimizations have been related to a number of relevant wireless data transmission technologies. This contribution is presented in Paper I.

Second, we have developed a TCP modification that allows the application to choose a more flexible transport service. The application can, by accepting errors in its data stream, achieve a higher performance (compared to using regular TCP) in cases where residual bit errors are the cause of packet loss in the receiver. The

actual performance improvement is dependent upon the amount of errors in the network. For example, Paper II presents a scenario where TCP-L achieves the double throughput of TCP at a residual BER of 10^{-5} . The intended use case for TCP-L is in a wireless receiver that forwards damaged IP packets, running an application that is capable of handling errors in the data stream. The modification only needs to be implemented in the receiver, and does not need support from the network or the sender. It can be argued whether such a modification can be compared to other TCP optimizations, or even TCP itself, as the protocols differ in the semantics on transfer reliability. Nevertheless, it allows a TCP-L receiver to communicate with a TCP sender. Further, a study of header recovery in TCP-L is presented in Paper III.

Third, the thesis investigates the interactions between the lower layers and the transport layer. A network emulator, WIPEMU, has been implemented and configured according to the parameters of a 4G system proposal [26]. Experiments with TCP, TCP-L and TCP Westwood+ [15] have been performed and are reported in Paper IV. The results indicate that for best effort traffic, a highly persistent link layer with adaptive modulation should be used to obtain the maximum throughput. On a similar subject, the thesis further contributes knowledge on the impact of loss generation on emulation-based protocol evaluation. By being able to precisely place packet loss and bit errors, protocol characteristics can be studied that are otherwise lost in the averaging process over randomly distributed losses and errors. This is presented in Paper V.

7 Paper Summary

7.1 Paper I - TCP over Wireless Networks

The first paper provides an overview and background of common wireless technologies used today, with emphasis on data transmission. It further outlines the TCP protocol, with a special focus on the mechanisms that are problematic in wireless networks. A number of TCP optimizations and modifications have been proposed in the research community. The paper summarizes many of these proposals in relation to the investigated wireless technologies.

7.2 Paper II - Bit Error Transparent Multimedia Transport

The second paper provides insight into the requirements of multimedia applications, and how they may have different requirements both on network transport reliability and data integrity than traditional applications such as file transfer and

E-mail. The paper introduces the idea of a TCP modification, called TCP-L, that allows bit errors in the data stream. The paper also presents experiments comparing the throughput of TCP-L to TCP with various amounts of residual bit errors.

7.3 Paper III - TCP-L: Allowing Bit Errors in Wireless TCP

The third paper further expands on the ideas in the previous paper. The main contribution is a concept for *header recovery* in TCP-L. A damaged TCP packet may have errors in the payload, in the header, or in both. Errors in the payload part need to be handled by the application, while errors in the header need to be handled by the transport layer. The paper analyzes the header in detail and discusses how it can be recovered to a safe state. The paper also presents some experimental results on the performance of TCP-L for the case where header recovery is always successful.

7.4 Paper IV - Emulation and Validation of a 4G System Proposal

The fourth paper presents an emulator for a proposed 4G downlink specification from the “Wireless IP” project [24]. The emulator is used to examine the impact of modulation and retransmissions in the link layer on the transport layer. This evaluation is done in the context of three TCP variants: TCP, TCP-L and TCP-Westwood+. The paper further supports the performance improvements of TCP-L in presence of residual bit errors. It concludes that the link layer in question should use highly persistent link layer retransmissions in combination with adaptive modulation.

7.5 Paper V - The Impact of Loss Generation on Emulation-based Protocol Evaluation

A major part of the thesis work involves experimental network studies. The fifth paper presents background work on network emulation. The key point of the paper is that by being able to control packet loss and bit errors in detail, aspects of network protocols can be studied that are otherwise lost in averaging, assuming a random loss process. The implementation of bit error generation described in this paper has also been used for the experiments reported in Papers II and III.

8 Future Work

There exists a number of continuations of this work. The WIPEMU emulator can be extended to support multiple users together with channel scheduling, frame soft-combining and forward error coding techniques. This allows for investigation of different resource scheduling algorithms and criteria, such as [7]. It may also highlight effects not directly visible in the currently investigated one-user scenario, such as how the total system capacity is affected by different number of link layer retransmissions and modulation switching levels. Also, until now only bulk data transfer has been studied. Short-lived flows, or long-lived flows with sporadic transmissions, would be good to investigate as these are also common traffic patterns.

The current research focus has been on TCP. Instead of modifying and optimizing this protocol, new protocols such as DCCP [21] and SCTP [25] are emerging. Are applications perhaps better served by moving on from TCP and instead use new transport protocols?

With an increased interaction between layers, the concept of cross-layer information is an interesting research area. For example, the transport layer could have access to information normally only available to the link or physical layer. "Soft information" from the signal decoding in the physical layer could be useful when the transport layer or application layer is decoding data whose correctness is uncertain. How should this information be conveyed in a good way?

References

- [1] S. Bajaj, L. Breslau, D. Estrin, K. Fall, S. Floyd, P. Haldar, M. Handley, A. Helmy, J. Heidemann, P. Huang, S. Kumar, S. McCanne, R. Rejaie, P. Sharma, K. Varadhan, Y. Xu, H. Yu, and D. Zappala. Improving simulation for network research. Technical Report 99-702, University of Southern California, March 1999.
- [2] H. Balakrishnan, S. Seshan, E. Amir, and R. H. Katz. Improving TCP/IP performance over wireless networks. *In proc. 1st ACM Int'l Conf. on Mobile Computing and Networking (Mobicom)*, November 1995.
- [3] C. Bettstetter, H.-J. Vögel, and J. Eberspächer. GSM phase 2+ general packet radio service GPRS: Architecture, protocols, and air interface. *IEEE Communication Surveys*, 2(3), 1999.

-
- [4] A. Brunstrom. Research project proposal: Analysis and implementation of a partially reliable transport protocol for multimedia applications. *Project Proposal, University of Karlstad*, February 1998.
 - [5] C. Casetti, M. Gerla, S. Mascolo, M. Y. Sanadidi, and R. Wang. TCP Westwood: Bandwidth estimation for enhanced transport over wireless links. In *Proceedings of ACM Mobicom*, pages 287–297, Rome, Italy, July 2001.
 - [6] H. Elaarag. Improving TCP performance over mobile networks. *ACM Computing Surveys*, 34(3):357–374, August 2002.
 - [7] N. C. Ericsson. *Revenue Maximization in Resource Allocation*. PhD thesis, Signals and Systems, Uppsala University, 2004.
 - [8] K. Fall. Network emulation in the VINT/NS simulator. *Proceedings of the fourth IEEE Symposium on Computers and Communications*, July 1999.
 - [9] K. Fall and S. Floyd. Simulation-based comparisons of Tahoe, Reno and SACK TCP. *Computer Communication Review*, 26(3):5–21, July 1996.
 - [10] S. Goel and D. Sanghi. Improving TCP performance over wireless links. *Proceedings of TENCON 98 - IEEE Region Ten Conference on Global Connectivity in Energy, Computer Communication and Control*, 1998.
 - [11] H. Granbohm and J. Wiklund. GPRS - general packet radio service. *Ericsson Review*, (2):82–88, 1999.
 - [12] V. Jacobson, R. Braden, and D. Borman. RFC 1323: TCP Extensions for High Performance, May 1992.
 - [13] L.-Å. Larzon, M. Degermark, and S. Pink. UDP Lite for real-time multimedia applications. *Proceedings of the IEEE International Conference of Communications (ICC) - Quality of Service Mini Conference*, June 1999.
 - [14] L.-A. Larzon, M. Degermark, S. Pink, L.-E. Jonsson, and G. Fairhurst. RFC 3828: The lightweight user datagram protocol (udp-lite), July 2004.
 - [15] S. Mascolo, L. A. Grieco, R. Ferorelli, P. Camarda, and G. Piscitelli. Performance evaluation of Westwood+ TCP congestion control. *Performance Evaluation*, 55(1-2):93–111, January 2004.
 - [16] B. Melander and M. Björkman. Trace-driven network path emulation. Technical report 2002-037, Department of Information Technology, Uppsala University, Sweden, 2002.

-
- [17] B. Noble, M. Satyanarayanan, G. Nguyen, and R. Katz. Trace-based mobile network emulation. In *Proceedings of ACM SIGCOMM '97*, September 1997.
 - [18] S. Pehrson. WAP - the catalyst of the mobile Internet. *Ericsson Review*, (1):14 – 19, 2000.
 - [19] J. Postel. RFC 768: User Datagram Protocol, Aug. 1980.
 - [20] J. Postel. RFC 793: Transmission control protocol, Sept. 1981.
 - [21] E. K. Sally Floyd, Mark Handley. Problem statement for DCCP. *draft-ietf-dccp-problem-00.txt*, *Work in progress*, October 2002.
 - [22] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end arguments in system design. *ACM Transactions on Computer Systems*, 2(4):277–288, Nov. 1984.
 - [23] J. H. Saltzer, D. P. Reed, and D. D. Clark. End-to-end arguments in system design. *ACM Transactions on Computer Systems*, 2(4):277–288, Nov. 1984.
 - [24] M. Sternad. The wireless IP project, June 2002.
 - [25] R. Stewart. RFC 2960: Stream Control Transmission Protocol, October 2000.
 - [26] A. Svensson, A. Ahlén, A. Brunstrom, T. Ottosson, and M. Sternad. An OFDM based system proposal for 4G downlinks. In *Proceedings of Multi-Carrier Spread Spectrum Workshop*, Oberpfaffenhofen, Germany, 2003.
 - [27] M. Taferner and E. Bonek. *Wireless Internet Access over GSM and UMTS*. Springer-Verlag, 2002.
 - [28] The Working Group for WLAN Standards. IEEE 802.11 wireless local area networks (webpage). <http://grouper.ieee.org/groups/802/11/>, checked 2005-04-12.
 - [29] WAP Forum Ltd. Wireless Application Protocol - architecture specification. <http://www.wapforum.org>, April 1998.

TCP in Wireless Networks: Challenges, Optimizations and Evaluations

This thesis presents research on transport layer behavior in wireless networks. As the Internet is expanding its reach to include mobile devices, it has become apparent that some of the original design assumptions for the dominant transport protocol, TCP, are approaching their limits. A key feature of TCP is the congestion control algorithm, constructed with the assumption that packet loss is normally very low, and that packet loss therefore is a sign of network congestion. This holds true for wired networks, but for mobile wireless networks non-congestion related packet loss may appear. The varying signal power inherent with mobility and handover between base-stations are two example causes of such packet loss. This thesis provides an overview of the challenges for TCP in wireless networks together with a compilation of a number of suggested TCP optimizations for these environments. A TCP modification called TCP-L is proposed. It allows an application to increase its performance, in environments where residual bit errors normally give a degraded throughput, by making a reliability tradeoff. The performance of TCP-L is experimentally evaluated with an implementation in the Linux kernel. The transport layer performance in a 4G scenario is also experimentally investigated, focusing on the impact of the link layer design and its parameterization. Further, for emulation-based protocol evaluations, controlled packet loss and bit error generation is shown to be an important aspect.