Correlation Analysis of Available Bandwidth Estimators for Mobile HCI

Doohyung Lee, Chihoon Lee, Jahwan Koo, and Jinwook Chung

School of Information and Communication Engineering, Sungkyunkwan University
Chunchun-dong 300, Jangan-gu, Suwon 440-746, South Korea
drlee07@skku.edu, chlee10@naver.com
jhkoo@songgang.skku.ac.kr, jwchung@songgang.skku.ac.kr

Abstract. Transport protocol in mobile devices of HCI (human-computer interaction) operates both in wired and wireless heterogeneous network. It has been developed from reactive congestion control schemes to proactive congestion control schemes suitable for both wired and wireless networks environment to overcome the poor wireless environment and restricted mobility. Nevertheless, recently proposed TCP's proactive congestion control schemes could not reflect network environment status in detail, there is also a limit to support new mobile services of HCI users. In this paper, we described issues and problems of rate adjustment metrics used for reducing packet loss and congestion loss in TCP-Jersey's available bandwidth estimators. Also, we presented the guideline of selecting more adequate metrics to improve wireless TCP performance of mobile HCI.

Keywords: Available Bandwidth Estimators, TCP schemes, wireless, Mobile HCI.

1 Introduction

Mobile Devices of HCI consist of five sensing parts (visual, hearing, smell, taste and tactual), middleware part, application part, operating system part and communication part. For example, these devices are mobile PCs, PDAs, cellular phones and all of hybrid devices which support for user mobility. While moving case of use to mobile devices have been presented HCI related issues such as diverse hybrid devices, seamless application roaming [1] . However the mobile wireless environment about several distinguished character of reliability and QoS have not been considered the features of existing focus of transport protocol layer[2]. This is important with aspects of reliability and QoS that accurate delivery to diverse HCI services cope with congestion and packet loss. In early days, in order to reliability and QoS the TCP congestion control have been developed difference between wired and wireless network. However, with increasing user mobility in the wired and wireless heterogeneous network have trended the enabling efficient congestion control of TCP schemes. We summarizes the features to control the congestion detection, notification, metrics of rate adjustment the

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2007, LNCS 4557, pp. 687–696, 2007.

existing presented schemes and outlines issues, problems centering around the TCP-Jersey recent proposed of proactive congestion control schemes on ABE of respectively metrics. Next this paper focus a selection for more suitable metrics which can reflect the bottleneck link state and its implementation.

The paper is organized as follows. Section 2 describes the existing researches related to the paper. In section 3, the major issues found in existing TCP proactive congestion control schemes. In section 4, the available bandwidth estimator is simulated to make a more accurate estimation for available bandwidth, and the correlation analysis between metrics in determining optimum congestion window is performed. A new metric to estimate available bandwidth is presented also in section 4, and section 5 concludes the paper and presents the future work in research for mobile HCI of TCP scheme.

2 Previous Works for TCP Schemes

Nowadays, mobile HCI devices for seamless application roaming consider easy access to information. for example, a museum or hospital's at indoor/outdoor for local or remote roaming is important to event or guiding services. Because occur packet loss or congestion with this services during roaming, it is well know that TCP's algorithm role has a prevented throughput degrade. Recent congestion control algorithms include the capability of distinguishing from the cause of loss. If the cause is link error, TCP transmission rate is maintained as it is, but if the cause is congestion at a router, the size of congestion window is adjusted through transmission rate control.The congestion control algorithms suited for wired and wireless heterogeneous network environments that have been proposed up to now can be divided into reactive congestion control schemes [4], [5], [6] and [7] and proactive congestion control schemes [8], [9], [10] and [11].

2.1 Reactive Schemes

This method is an algorithm preventing the TCP performance from degrading through early recovery of packet loss such as this of the traditional Reno [3]. they uses the acknowledgements of receiver for the segments sent by the TCP sender and generate timeout or congestion events, and by lowering the value of congestion window, the transmission rate is adjusted.

I-TCP [4], M-TCP [5] and WTCP [6] divide the connection between the sender and receiver into a wired transmission interval and wireless transmission interval. The previous TCP congestion control mechanism is used in the wired network interval between the sender and base station while a congestion control mechanism more suitable for the previous TCP or wireless environment is used in the wireless interval between the base station and receiver. SNOOP [7] adds a TCP-aware module in the link layer (layer 2) which enables the base station to look at the TCP header without having to disconnect the TCP connection. The advantages of such reactive schemes is that the TCP is not affected by the packet loss in the network layer. However, these schemes have a major disadvantage in that they violate end-to-end TCP semantics. Consequently, packets sent by TCP may not match with those arriving at the receiver. The base station must have a TCP layer and there is the burden of having to separately manage the two connections.

2.2 Proactive Schemes

The reactive approach to congestion control estimates the available bandwidth in advance based on the congestion experiences of TCP flow and sends the calculated value of TCP congestion window. This method obtains good goodput by avoiding congestion and lowering the retransmission rate. Based on such fundamental ideas, the congestion window is lowered before congestion arises in the sender. This accounts to the fact that if congestion arises and RTT becomes longer, the TCP transmission efficiency drops since TCP performance is inversely proportional to RTT.

In TCP-Vegas [8], the sender measures RTT (Round Trip Time) at the beginning of connection and sets it as BaseRTT. If a smaller RTT is measured during the packet transmission, the current BaseRTT is renewed to the smaller value. Also, the expected sending rate for the state in which all queues of middle nodes are empty can be obtained by dividing the current window size by BaseRTT, and the sender measures RTT and the number of bytes for packets acknowledged by the sender to obtain the actual sending rate.

TCP-Peach [9] and TCP-Westwood [10] measure the available bandwidth of the current connection and adjust the window size if a packet loss arises, regardless of its cause. TCP-Jersey [11] distinguishes congestion and transmission loss and adjusts the window size. The advantage of these schemes [9],[10],[11] are that the TCP window size is adjusted dynamically, fundamentally preventing packet loss due to congestion. On the other hand, especially in TCP-Vegas, the number of total packets saved in the buffer of middle nodes can be controlled, but the number of packets saved in each queue cannot be controlled.

3 Open Issues of Previous Solutions

The mobile characteristics must be considered sufficiently in the wired and wireless heterogeneous network environment. That is, the cause for packet loss must be clearly identified, and In the case of a congestion, a rate control scheme for more sophisticated control is needed while for a transmission error, the current congestion window (CW) is maintained while using the available bandwidth as much as possible, to display optimal performance.

The most important issue is the need for a rate control scheme which reflects accurate circumstances to bottlenecks. Matters to be considered are described in the following. First issue is Differentiation and detection of causes for packet loss. The second issue is notification of network status. The third issue is selection of metrics for transmission rate adjustment

The problems will be identified and possible measures discussed in the viewpoints of the above-mentioned issues based on the recent literature on TCP-Jersey.

3.1 Differentiation and Detection of Causes for Packet Loss

In previous wired networks (Tahoe, Reno, New-reno, SACK [12]), there was no need to differentiate the cause for packet loss. This is because most of the loss was due to congestion. However, with the increase of wireless network environments, it has become necessary to identify the cause for packet loss. The latest loss identification methods proposed determine the cause for packet loss by a threshold value for a bottleneck of a buffer state. In TCP-Jersey, the threshold value was set at 30 percent of the buffer based on test results. That is, loss under the threshold value at the sender is judged as a transmission error while the value exceeds the threshold as congestion. The problem here is that this causes the waste of buffer resources in router. Another problem is that TCP transmission states are changing frequently in a wired and wireless heterogeneous environment, and this cannot be coped with flexibly by holding a fixed threshold value such as the TCP-Jersey.

These problems can be resolved at a router by setting a dynamic threshold value or accurately identifying the buffer capacity and approximating the actual buffer value.

3.2 Notification of Network Status

As shown in the figure 1, the current method in transmitting the bottleneck situation is to use the ECN (Explicit Congestion Notification) scheme. But, TCP-Jersey uses Congestion Warning(CW) flag instead of ECN. CW operates similarly to ECN and uses the CE bit in the IP header and ECE and CWR in the TCP header. The difference is that ECN uses a maximum and minimum threshold value. If the maximum value is surpassed, the CW flag is marked unconditionally, and if under the minimum value, the flag is not marked. Finally, if the value is between the maximum and minimum values, it is marked by a probability value calculated. On the other hand, TCP-Jersey is simplified by using only one threshold value, and if the value is over the threshold value, the CW flag is marked and sent, but if under, it is not marked.

The problem here is that only 1 bit is used, placing a limit in expressing and sending the various states of congestion in the router. Accordingly, a mechanism which uses the standard ECN while making intelligent judgments at the sender is required.

3.3 Selection of Metrics for TCP's Congestion Control

The rate control is a function performed by the sender. It determines the sending rate based on contextual information of the network. This is a prerequisite for proactive congestion control, and selecting effective metrics to be used determines the success of accurate estimation of available bandwidth.

In the proactive congestion control method, the sender adjusts the congestion window value in precedence based on the feedback information(RTT, ACK) collected in the bottleneck interval. By performing this, the sender can intelligently

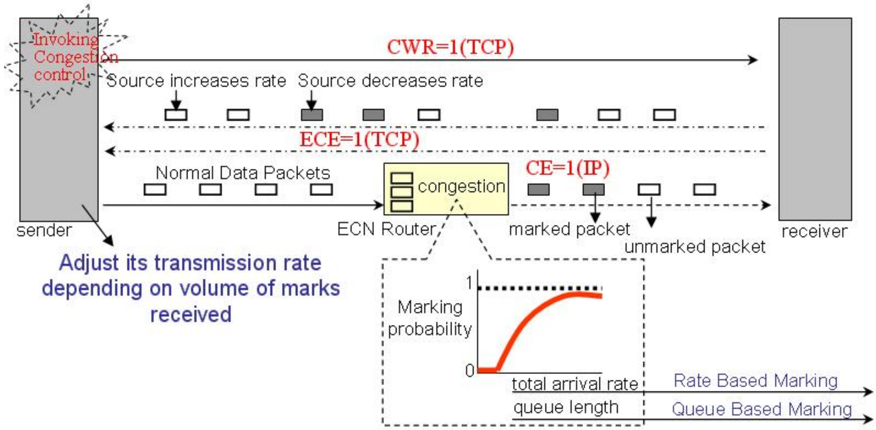


Fig. 1. The Operation of Mark Packets for ECN scheme

cope with the network situation or cause of packet loss. Also, undesirable network conditions(congestion and decrease in unnecessary congestion window) can be prevented in advance. In rate control methods proposed until now, ACK and RTT have been used as major metrics. These metrics have been brought about with the idea of bandwidth delay product [13], and hold a significant importance as they have been used for rate control in the end-to-end approach scheme.

The rate computation in TCP-Jersey was derived from Time Sliding Window(TSW) proposed by David D. Clark et al. [14], and is shown in the following Eq. (1). This equation is applied to the network router and is used to calculate the bandwidth allocated by each individual TCP flow.

$$R_n = \frac{T_w \times R_{n-1} + L_n}{(t_n - t_{n-1}) + T_w} \tag{1}$$

In the above equation, R_n is the calculated bandwidth for packet n when it arrives at time t_n. t_n is the arrival time of the *nth* packet, and t_{n-1} is the previous packet. L_n is the packet size and T_w is a time constant depending on the network situation. R_{n-1} is the previously calculated rate value. TCP-Jersey proposes the following equation in which TW is substituted to RTT. R_n (rate of available bandwidth) in Eq. (2) is multiplied by RTT, the estimation bandwidth and delay, and divided by segment data to obtain the $Ownd$ (Optimum Congestion Window) value in Eq. (3). This was proposed to perform efficient TCP rate control for the estimated available bandwidth.

$$R_n = \frac{RTT \times R_{n-1} + L_n}{(t_n - t_{n-1}) + RTT} \tag{2}$$

TCP-Jersey derives the $Ownd$ to set $Cwnd$ (Congestion Window) and *ssthresh* as Eq (3) with the rate computed in Eq. (2). A contradiction can be seen here as RTT was used to obtain R_n as well as to calculate $Ownd$. That is, the value of RTT is suitable variable to determine $Ownd$ but not for determining the rate. This calls the need for new separate metrics, and this will be verified in section 4 through tests.

$$Ownd = \frac{RTT \times R_n}{seg\_size} \tag{3}$$

The ACK value mostly affects the most in rate control. However, as the transmission path to the receiver and path returning back the ACK value may differ, the delay may differ depending on the buffer queuing value of the router. Also, if an ACK drop of compression occurs, the ACK rate loses its meaning of an adequate sample value. In order to settle this, timestamp should be used or an additional mechanism reflecting only valid ACK values is necessary.

In addition, TCP-Jersey does not perform rate control in the *slowstart* phase. Accordingly, if rate control is performed in the *slowstart* phase, *ssthresh* will generally be formed in a higher point, meaning that the available bandwidth could be reached more quickly. Also, if RTO occurs in the *slowstart* phase, *ssthresh* will not decrease by half as it would in TCP-Reno[3] or TCP-Jersey [11].

4 Verification of Transmission Rate Calculation in TCP-Jersey

In previous section, we have drawn the problems from three aspects of Available Bandwidth Estimator (ABE) in Eq. (2) and Optimum Congestion Window (Ownd) in Eq. (3) of proposed in TCP-Jersey among the proactive congestion control schemes and presented measures to cope with the problems. This section provides the verification through simulation. In particular, the correlation between dependent variables for the rate calculation in Eq. (2) used to compute available bandwidth, and $Ownd$ in Eq. (3) used to compute optimal congestion window have been analyzed to verify the effectiveness of each dependent variable.In addition, verification will be focused on how the R_n and $Ownd$ values change when each independent variable is changed and the correlation with the independent variable using a statistical method.

4.1 Gathering Test Data and Verification Method

Fixed TCP segment data of 600 bytes was generated 500 times/sec for the test data in a topology structure environment shown in Fig. 2 in which there is one input and one output. Fig. 3 shows the value of R_n generated in the test environment. This value has approximated to 1.5 Mbps, similar to that in a TCP-Jersey testing environment. With this, it has been verified to be appropriate data for the test.

The input values RTT, L_n, $t_n - t_{n-1}$ and $RTT * R_{n-1}$ used in Eq. (2) to compute the rate and $RTT * R_n$ in Eq. (3) to compute the optimal congestion window were used to find the correlation with output values R_n and $Ownd$ for the verification. As for the correlation analysis, Pearson's correlation method was used.

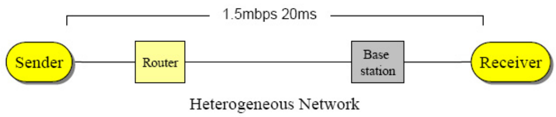


Fig. 2. Experiment environment

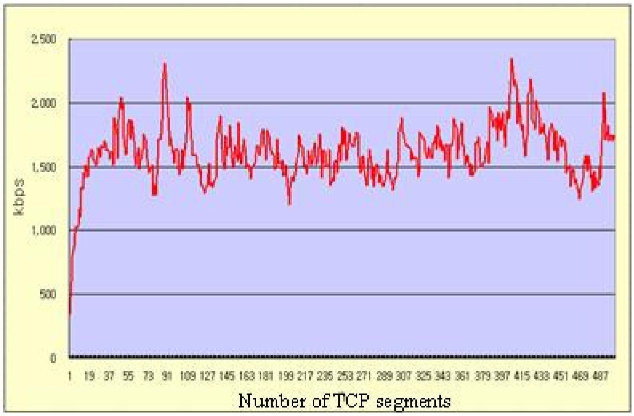


Fig. 3. Raw data gathered from the experiment environment

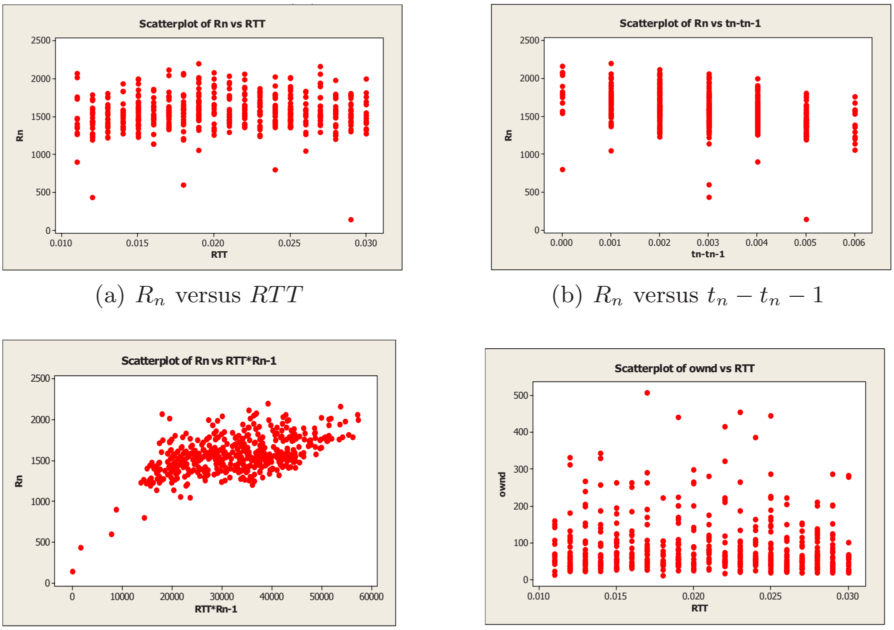
The significance level was set at P-value 0.05 in this test, That is, if P-value is lower than 0.05 in the 95 percent confidence interval, there is high correlativity, and if higher, then there is low correlativity. + correlativity means positive correlativity and - means negative correlativity. If the correlativity factor comes close to an absolute value of 1, then the correlativity is high, and if it reaches near 0, there is low correlativity. The scatter-plot shows the relation between two quantitative variables visibly.

4.2 Verification of Transmission Rate Adjustment Value Calculation

The correlation between R_n, each metric in Eq. (2) and the available bandwidth estimator for TCP-Jersey will be verified here. Results from analyzing the correlation between R_n and RTT show that P-value is 0.831. Since P-value is higher than 0.05, it is judged to have low correlativity.

Fig. 4(a) shows the scatter-plot between RTT and R_n, and when taking a close look, it can be seen that RTT and R_n do not change in proportion or inverse proportion. Accordingly, this shows that R_n does not change when RTT changes.

The second analysis was performed on the correlativity between R_n and $t_n - t_{n-1}$. P-value turned out to be 0.000, statistically meaning they are significant and are correlative. The correlativity factor was -0.424, showing there was a negative correlation. Fig. 4(b) shows the scatter-plot of R_n and $t_n - t_{n-1}$, and it



(a) R_n versus RTT

(b) R_n versus $t_n - t_n - 1$

(c) R_n versus $RTT * R_{n-1}$

(d) $Ownd$ versus RTT

Fig. 4. Correlation analysis of available bandwidth estimator in TCP-Jersey

can be seen that R_n is inversely proportional to $t_n - t_{n-1}$. The figure too shows the negative correlation.

The third analysis was performed on the correlativity between R_n and $RTT * R_{n-1}$. Just by looking at the scatter-plot in Fig. 4(c), it can bee seen that there is considerable correlation (correlation factor 0.863). This means that there is correlativity with $RTT * R_{n-1}$, but this is not from RTT but by R_{n-1}.

In particular, the scatter-plot in Fig. 4(d) shows a positive correlativity in which R_n makes close to linear changes when $RTT*R_{n-1}$ changes. That is, there is quite a close relation, but the relational factor is based on the previous value of R_n, which is R_{n-1}, rather than RTT. Accordingly, RTT has no correlativity with R_n. It can be seen in the above test that RTT does not have close correlativity in computing the rate value. This means that the rate in estimating available bandwidth is determined by some other variable. This produces counterevidence that RTT should be substituted with a more suitable value.

4.3 Verification of Optimum Congestion Window

The $Ownd$ value is computed by using the available bandwidth estimator R_n, RTT and seg_size. Results from analyzing the correlativity with RTT show that there is a negative correlation (correlation factor -0.024, P-value = 0.004) as can be seen in Fig. 4(d). Last of all, analyzing the correlativity between $Ownd$ and R_n show that a positive correlation (correlation factor 0.206, P-value=0.000) exists. $Ownd$ increases in proportion to R_n. Through the above test, it can be seen that $Ownd$ has correlation with RTT and R_n.

4.4 Discussion and Solution

Combining the test results, we can see that RTT is an effective value in producing $Ownd$ in Eq. (3) based on Bandwidth Delay Product[13], but not appropriate in obtaining R_n in Eq. (2). Consequently, a new metric in producing the rate is needed. This should be substituted as a constant which considers properly the bottleneck link situation. Accordingly, if the wired/wireless border router can send the buffer value correctly to the sender, the above-mentioned problem can be solved. However, since the only realistic method is the 1 bit ECN, sufficient information cannot be sent.

After all, the solution to this problem is to make an extension for additional information fields by modifying TCP or stochastically marking and sending congestion information to ECN so that the sender can use the statistical information in estimating the buffer value. The advantage of the former is that even though TCP modification is necessary, network congestion state may be transmitted more accurately. And the advantage of the latter is that TCP does not have to be modified but then it will be more difficult to send accurate information. If congestion status of bottlenecklink can be identified accurately through new metrics, two effects are anticipated of proactive TCP congestion control methods for mobile HCI services of seamless application. The first is that if the cause is not due to congestion(CW=0,CE=0), that is, random error, the current rate is maintained and threshold value increased by a reasonable level, enhancing transmission performance. The second is that if it is due to congestion, $Cwnd$ is not decreased inconsiderately by half and the buffer threshold value of the router is utilized as mush as possible. This removes resource waste and brings performance improvement in TCP transmission.

5 Conclusions and Future Work Directions

We have described the problems and issues by identifying the cause of packet loss, inspecting the congestion information transmission of existing TCP schemas and testing the validity of metrics in computing the rate of the proactive congestion control method proposed in TCP-Jersey, and presented its solutions. In particular, RTT used in TCP-Jersey was found to be useful in computing congestion window but not proper in computing transmission rate through a correlation analysis. Accordingly, we have presented another metric to be used in palace of RTT in computing the rate. This makes it possible to send a more accurate buffer capacity from the wired and wireless border router to the sender in order to produce an accurate rate. Through this, proactive congestion control in a wired and wireless heterogeneous environment can be performed by generating an optimum TCP transmission rate for congestion, and if loss is not due to congestion, the current rate may be maintained as it is. Future research in the two measures mentioned above should be done in depth to create an optimum rate computation model and mechanism for the mobile network environment.

References

1. Bardram, J., Nielsen, T.K.C.: Supporting Local Mobility in Healthcare by Application Roaming Among Heterogeneous Devices. In: Chittaro, L. (ed.) Mobile HCI 2003. LNCS, vol. 2795, pp. 161–176. Springer, Heidelberg (2003) 2
2. Sohraby, K.C.W., Daneshmand, M B.L., Hu, Y.: A survey of transport protocols for wireless sensor networks. Network, May-June 2006, vol. 20(3), pp. 34–40. IEEE, New York (2006)
3. Stevens, W.: TCP slow start, congestion avoidance, fast retransmit and fast recovery algorithms. IETF, RFC 2001 (1997)
4. Bakre, A., Badrinath, B.R.: I-TCP: indirect TCP for mobile hosts. In: Proceedings of the 15th International Conference on Distributed Computing Systems, June 1995, pp. 136–143 (1995)
5. Brown, K., Singh, S.: M-TCP: TCP for mobile cellular networks. ACM Computer Communications Review 27, 19–43 (1997)
6. Sinha, P., et al.: WTCP: A reliable transport protocol for wireless wide-area networks. In: Proceedings of ACM Mobicom, pp. 301–316 (1999)
7. Caceres, R., Iftode, L.: Improving the performance of reliable transport protocols in combilecomputing environments. IEEE Journal on Selected Areas in Communications 13(5), 850–857 (1995)
8. Brakmo, L.S., Peterson, L.L.: TCP Vegas: end to end congestion avoidance on a global Internet. IEEE Journal on Selected Areas in Communications 13(8), 1465–1480 (1995)
9. Akylidiz, I.F., Morabito, G., Palazzo, S.: TCP-Peach: a new congestion control scheme for satellite IP networks. IEEE/ACM Transactions on Networking 9(3), 307–321 (2001)
10. Casett, C., et al.: TCP Westwood: bandwidth estimation for enhanced transport over wireless links. In: Proceedings of ACM Mobicom, pp. 287–297 (2001)
11. Xu, K., Tian, Y., Ansari, N.: TCP-Jersey for wireless IP communications. IEEE Journal on Selected Areas in Communications 22(4), 747–756 (2004)
12. Mathis, M. et al.: TCP Selective ACKnowledgment Options, RFC 2018 (October 1996)
13. Chen, K., Xue, Y., Shah, S.H., Nahrstedt, K.: Understanding Bandwidth-Delay Product in Mobile Ad Hoc Networks. Elsevier Computer Communications (ComCom) Journal. Special Issue on Protocol Engineering for Wired and Wireless Networks (2004)
14. Clark, D.D., Fang, W.: Explicit allocation of best effort packet delivery service. IEEE/ACM Trans. Networking 6, 362–373 (1998)