Adaptive Transport Layer Protocols for Wireless Networks

Lian Wang

ChongQing University of Posts and Telecommunications

Abstract—**The future wireless networks should be a heterogeneous network with flexible and open architecture, capable of supporting various types of networks, terminals and applications. However how to integrate the protocols to meet the heterogeneous network environments becomes a significant challenge in the fourth generation wireless network. Adaptive protocols are proposed to solve heterogeneity problem in future wireless networks. The main objective of this paper is to explore the feasibility of adaptive transport layer protocols applied to the future wireless networks, based on the analysis of adaptive congestion mechanism compared with TCP congestion control, and the characteristics of the adaptive transport layer protocols proposed.**

***Index Terms*—adaptive transport layer protocols, adaptive congestion control, AIMD**

I. INTRODUCTION

TCP is an end-to-end, reliable, sender-centric transport layer protocol. It is initially designed for network with less link error without consideration more about wireless networks. TCP performance degradation is unavoidable in wireless networks due to its significant limitations. For example, TCP is unfit for real-time multimedia applications due to its burst transmission and inflexible retransmission rules. TCP cannot dynamically adjust its congestion control and rate control schemes to adapt heterogeneous wireless networks environments in the future wireless networks. For example satellite networks with wide coverage area, broadcast capability and immunity to the adverse geographic conditions, have a largely adverse impact on TCP performance, due to large propagation delay, high link errors, and link asymmetry etc. Moreover the third-generation (3G) wireless networks, with advantages of multi-megabit Internet access, omnipresent access, have significant TCP throughput degradation due to high packets losses, spurious TCP retransmission and link asymmetry etc. TCP seems more and more inflexible and inefficient in the wireless heterogeneous networks.

Thus adaptive transport layer protocol is proposed to cope with heterogeneity problem and maintain high performance in the future wireless networks [1]. Adaptive congestion control is vital function addressed in adaptive transport layer protocol based on the Additive-Increase Multiplicative-Decrease (AIMD) algorithm. For example the sender-centric adaptive transport layer protocol TCP-ATL is back compatible and TCP-friendly, proposed for reliable data transport [2]. And receiver-centric transport layer protocol Reception Control Protocol (ReCP) adopts an adaptive congestion control algorithm that dynamically monitors the wireless random loss rate and delay, and adjusts its congestion control adaptation parameters as offsets to the loss rate and delay components introduced by the wireless link [3]. ReCP integrates the advantages of receiver-centric mechanism and adaptive congestion control scheme.

The paper is organized as follows. Section 2 focuses on TCP congestion control mechanism, especially the Additive-Increase Multiplicative-Decrease(AIMD) algorithm as a crucial part of congestion control mechanism is introduced; Compared with the conventional TCP congestion control, section 3 highlights the adaptive congestion control mechanism; section 4 discusses the characteristics of Sender-centric and Receiver-centric adaptive transport layer protocols based on the analysis of multiple transport layer protocols. Finally conclusion presents in section 5.

II. TCP CONGESTION CONTROL MECHANISM

TCP congestion control adopts the increase-by-one decrease-to-half strategy. The Additive-Increase Multiplicative-Decrease (AIMD) algorithm is a part of TCP congestion control. The TCP uses a congestion window to control the sending rate of a connection in response to fluctuations of the network capacity. The size of the window is updated by the AIMD mechanism: it is increased by one if an acknowledgment received per RTT, or decrease to half upon each packet loss detected. The additive and multiplicative coefficients (α, β) are fixed, (1, 0.5). The window size ω at time t is adjusted as:

$$\omega(t+RTT) = \omega(t)+\alpha, \quad \text{if no loss in } [t, t+RTT] \tag{1}$$

$$\omega(t+\Delta t) = \beta\omega(t), \quad \text{if a loss occurs at time } t,$$

where α >0 and 0< β <1.

Yang and Lam first proposed the analytic equation to calculate the throughput of a TCP connection, as shown in Equation (2), where the parameters are RTT (round- trip time), T_0 (timeout value) and b (the number of packets acknowledged by each ACK). The equation has taken the fair access to the TCP flows into consideration [4][5][6], so β can safely selected from $[0.5,1)$, and under the TCP-friendly constrain, α (α >0) can be dynamically adjusted according to the current link conditions.

$$T_{\alpha,\beta}(p, RTT, T_0, b) = \frac{1}{RTT\sqrt{\frac{2b(1-\beta)p}{\alpha(1+\beta)}} + T_0(3\sqrt{\frac{(1-\beta^2)bp}{2\alpha}})p(1+32p^2)} \tag{2}$$

In TCP connection, (α, β) is fixed constant (1, 0.5), so plug the value of (α, β) into Equation (2), the throughput of the TCP connection is derived as:

$$T_{1,\frac{1}{2}}(p\ , RTT, T_0, b) = \frac{1}{RTT\sqrt{\frac{2bp}{3}} + T_0 \min(1,3\sqrt{\frac{3bp}{8}})p(1+32p^2)} \tag{3}$$

The wired TCP throughput can be expressed as $T_{1,\frac{1}{2}}(p_c, R_c, T_{0c}, b)$, where p_c is the packets loss probability due to congestion; R_c is the end-to-end RTT of the connection without any additional wireless link delay. $T_{1,\frac{1}{2}}(p_c, R_c, T_{0c}, b)$ is the upper bound for the throughput due to TCP-friendly consideration. TCP-friendly is important because it is desired that the realtime applications and traditional TCP applications have fair accesses to network resources. The fair accesses mean TCP flow and non-TCP flow can have the same bandwidth under identical conditions. An (α, β) combination is said to be TCP-friendly if it satisfied the following Equation:

$$\alpha = 4(1-\beta^2)/3 \tag{4}$$

In summary, TCP is unfit realtime application due to its inflexible retransmission rules and burst transmission, not mention the heterogeneous environments in the future wireless networks. So adaptive congestion control and adaptive rate control mechanisms are proposed to dynamically adjust of AIMD parameters to adapt the wireless link conditions and simultaneously maintain TCP-friendly.

III. ADAPTIVE CONGESTION CONTROL

The aim of adaptive congestion control method is to achieve $T_{1,\frac{1}{2}}(p_c, R_c, T_{0c}, b)$ by dynamical adjustment AIMD parameters regardless of the heterogonous physical environments and architectures. $T_{1,\frac{1}{2}}(p_c, R_c, T_{0c}, b)$ is the upon bound for the target throughput because of the TCP-friendly consideration, where p_c is the probability due to congestion; R_c is the end-to-end RTT of the connection without any additional wireless link delay.

he TCP sources achieve lower throughput in wireless etworks than in wired networks, because the increased packet loss probability and RTT. This situation can be expressed as the following inequation for $p > p_c$ and $R > R_c$.

$$T_{1,\frac{1}{2}}(p\ , R\ , T_0, b) \quad < \quad T_{1,\frac{1}{2}}(p_c, R_c, T_{0c}, b)$$

5)

where the loss probability p and the round-trip time R, both include two parts the packet loss and the round-trip time caused by congestion and wireless link.

$$p = 1-(1-p_w)(1-p_c) \tag{6}$$

$$R = R_c + 2d_w \tag{7}$$

where p_c is the packet loss probability due to congestion, and p_w is the packet loss probability due to wireless link errors. R_c is the end-to-end RTT of the connection without any additional wireless link delay. And d_w is the one-way wireless link delay.

To achieve high performance in different wireless environments in the future wireless networks, and to be TCP-friendly to the wired TCP sources sharing the same bottleneck; The adaptive congestion control mechanism can dynamically adjust its AIMD parameters (α, β) according to the current wireless environments. So the object of the adaptive congestion control can be expressed as the following equation to achieve high performance in the future wireless networks.

$$T_{\alpha,\beta}(p\ ,R\ ,T_0,b) \approx T_{1,\frac{1}{2}}(p_c,R_c,T_{0c},b) = \hat{T}$$
(8)

where $\hat{T}$ is the upper bound of the target throughput by the wireless TCP sources, $\hat{T} = T_{1,\frac{1}{2}}(p_c,R_c,T_{0c},b)$. If the p, R, T_0 and d_w are given, from the Equation (6) (7), p_c and R_c can be calculated as:

$$p_c = \frac{p - p_w}{1 - p_w} \quad (9)$$

$$R_c = R - 2d_w$$
(10)

The throughput of the TCP connection $T_{\alpha,\beta}(p\ ,R\ ,T_0,b)$ is expressed in Equation (2). So the throughput up bound $\hat{T}$ can be calculated with the given p_c, R_c, T_{0c} and b. where b is initialized as one, each packet is acknowledged by an ACK. And T_{oc} is set as 4 R_c according to the TCP-friendly consideration [5].

$$\hat{T} = \frac{1}{R\sqrt{\frac{2b(1-\beta)p}{\alpha(1+\beta)}} + T_0(3\sqrt{\frac{(1-\beta^2)bp}{2\alpha}})p(1+32p^2)} \quad (11)$$

As a result, the equation of the additive-increase parameter α can be achieved based on Equation (11).

$$\alpha = \frac{bp(1-\beta)}{2(1+\beta)}[\hat{T}(2R + 3T_o p(1+32p^2)(1+\beta))]^2 \quad (12)$$

Once p_w, d_w, p and R are given, $\hat{T}$ can be calculated by Equation (11); Then the additive-increase parameter α can be achieved from Equation (12) for $\forall\beta \in [0.5,1)$. (α, β) pair is determined by the Equation (2), and Equation (2) has taken fairness into consideration, so β can be safely selected from $[0.5,1)$.

The adaptive congestion control has high throughput by dynamically adjustment AIMD parameters, and maintains TCP-friendly under the upper bound throughput constrain.

IV. ADAPTIVE TRANSPORT LAYER PROTOCOLS FOR WIRELESS NETWORKS

TCP has exposed numerous problems in the wireless networks. Thus multiple transport layer protocols that each transport protocol is tailored to adapt the specific networks have been proposed. For example, the snoop protocol is proposed primarily for wireless LANs [7], WTCP and TCP-Westwood are proposed for 3G cellular networks, TCP-peach++ is to improve TCP performance in the satellite networks, and end-to-end wireless multimedia streaming TCP-friendly protocol (WMSTFP) is proposed for multimedia over wireless Internet [9]. Moreover, different mechanisms are required to implement handover while mobile users roam between the heterogeneous wireless architectures. Significant overheads are required for state maintenance, buffering, switching and synchronizing. As a result, different transport protocols in the future wireless networks are not practical due to processing, memory and power limitations of wireless terminals.

Therefore adaptive protocols are proposed to adapt to the varying wireless networks environments and to maintain high performance in the future wireless networks. Adaptive congestion control is vital function addressed in adaptive transport layer protocol based on the Additive-Increase Multiplicative-Decrease (AIMD) algorithm.

A. Adaptive Sender-centric Transport Layer Protocols

TCP-ATL

TCP-ATL is developed on TCP protocols and utilizes the adaptive congestion control schemes to address the heterogeneous environments in the future wireless networks. Furthermore, TCP-ATL adopts the selective acknowledgement (SACK) options to guarantee reliability, because SACK is very efficient to recover multiple packet losses in one TCP window, especially in high bandwidth-delay product networks. TCP-ATL can dynamically adjust AIMD parameters according to link conditions, i.e. $p_w, d_w, p,$ and *R*. To address the heterogeneous environments, the link condition parameters, for example the wireless link delay d_w and the probability of wireless packet loss p_w , can be obtained from the underlying adaptive MAC layer. The access delay d_w can be achieved directly from MAC layer because the access delay available in MAC layer is used to calculate the timeout values. The packet loss probability p_w can be obtained from the MAC layer information by the simple cross-layer communication.

So TCP-ATL can address the heterogeneous environments based on the link condition information from MAC layer. Once the link condition parameters and R are given, p_c and R_c can be calculated according to (9),(10). Hence, $\hat{T} = T_{1,\frac{1}{2}}(p_c, R_c, T_{oc}, b)$ can be calculated based on (2), then plug $\hat{T}$ into (12) to decide α for $\forall\beta \in [0.5,1)$. The multiplicative-decrease parameter β can be selected safely from $[0.5,1)$. The higher the β value is selected, the higher performance can be achieved in high bandwidth-delay products, because selected higher β value from $[0.5,1)$ can compensate the adverse effects from the high propagation delay on the performance.

At the beginning of the connection, the pair (α, β) is initialized as (1, 0.5). TCP-ATL is backward compatible, if no wireless link involves, $p_w = 0$ and $d_w = 0$, the normal TCP mechanism is adopted to maintain TCP-friendly. If packet losses occur due to congestion or link errors, the AIMD parameter will be dynamically adjusted to adapt the varying environments. Furthermore at each handoff, the AIMD parameters will be updated by ATL to address the varying wireless link environments. On the other hand, the adaptive congestion control algorithm of TCP-ATL can indirectly address the blackout situations due to mobility or fading by the environmental obstacles. p_w is changed according to the varying of the blackout situation. When the blackout occurs, the p_w increases. Once the signal is recovered, the p_w quickly decreases. The valued of α is dynamically adjusted according to the varying p_w to maintain the high throughput. So TCP-ATL can adapt to different link conditions to avoid throughput degradation.

The default TCP-ATL operation is based on the wireless/wired scenario corresponding to the sender and the receiver terminals. Hence the TCP-ATL operation is required to be modified according to the different source/receiver combinations to calculate p_c and R_c . In wireless source/wired receiver case, no modification is required because it is the default ATL operation scenario. In wired source/wireless receiver case, the wireless link parameters, i.e. p_w and d_w , are forwarded to the source within data ACK packets to avoid additional overhead. Similarly, handoff event is also informed to the source by the data ACK packet. In wired source/wired receiver case, no modification is needed because TCP-ATL is backward compatible when the absence of wireless link involves, i.e. $p_w = 0$ and $d_w = 0$. In wireless source/wireless receiver case, let p_w^s, d_w^s and p_w^r, d_w^r be the packet loss probability and the delay of the wireless link at the source and the receiver of the end-to-end path, respectively. The p_w^r, d_w^r can be forwarded to the source by the ACK packets. The original equations (6), (7) to calculate p and R should be modified as:

$$p = 1 - (1 - p_w^s)(1 - p_c)(1 - p_w^r) \qquad (13)$$

$$R = R_c + 2(d_w^s + d_w^r) \qquad (14)$$

According to (13) (14), the original equations to calculate p_c and R_c should be modified as:

$$p_c = 1 - \frac{1 - p}{(1 - p_w^s)(1 - p_w^r)}$$

(15)

$$R_c = R - 2(d_w^s + d_w^r)$$

(16)

Once p_c and R_c are achieved according to (15) (16), the original TCP-ATL can perform well in wireless source/wireless receiver scenario.

RaCP-ATL

RaCP-ATL is an end-to-end rate control protocol for multimedia traffic. It use adaptive rate control scheme to realize rate control. RaCP-ATL is developed on RTP/RTCP and UDP. The source does not retransmit the loss packet because of the timely requirements in multimedia applications. However, the receiver will send ACK to the source when each packet is received. If the RTP/RTCP is used, theses ACKs are the receiver reports that contain the details about reception quality. So RaCP-ATL can achieve p and *R* from the RTP receiver reports. Otherwise, if RTP/RTCP is not used, p and *R* can be obtained from the data ACKs. RaCP-ATL can address the heterogeneous environments relying on the information from ACKs or the receiver reports.

The goal of RaCP-ATL is to achieve high throughput by dynamical rate control schemes, and simultaneously maintains the TCP-friendliness. Dynamical rate control is based on AIMD scheme in response to the varying network environments. In case of packet loss, the sending rate (S) decrease multiplicatively, $S = S \cdot \beta$. In case of each ACK received per RTT, the sending rate increase additively, $S = S + \alpha$. Once p, R, p_w and d_w are known, p_c and R_c can be achieve by (9) (10). So the up bound throughput $\hat{T} = T_{1,\frac{1}{2}}(p_c, R_c, T_{oc}, b)$ can be calculated according to (2). Then plug $\hat{T}$ into (16) to obtain α for a given β for $\forall \beta \in [0.5,1)$. The value of β will affect the rate variation smoothness. Moreover, RaCP-ATL has incorporated an adaptive multimedia media encoding schemes. RaCP-ATL can provide the available bandwidth $S(t)$ to the adaptive encoder. So the encoder can dynamically adjust its encoding rate $R(t)$ in terms of TCP-friendly rule. By this way, congestion and abrupt quality variation can be minimized in multimedia delivery.

As mention in TCP-ATL, RaCP-ATL is required some modification in terms of p_c and R_c according to the different source/receiver combinations. The only difference is the source gets information about vertical handoff event, p_w and d_w from RTP receiver reports rather than ACKs.

B. Adaptive Receiver-centric Transport Layer Protocols

ReCP (Reception Control Protocol) is a receiver-centric transport protocol that is a TCP clone[7]. But the receiver in ReCP has transposed the main functionalities including congestion control, loss recovery and power management form the sender. The intelligence of transport protocol at the receiver is more neighboring to the wireless link. Hence the receiver can get the first-hand information and have quicker responses to the varying network environments to achieve high performance. More importantly, mobile hosts are increasingly being equipped with multiple interfaces to adapt to the heterogeneous wireless networks in the future wireless networks. ReCP can provide powerful and comprehensive transport layer solutions to the multi-homed hosts. Firstly, more effective congestion control approaches. There are two options to solve congestion control problem, scalable solution and adaptive congestion control mechanism. In scalable solution multiple congestion control protocols are needed to perform congestion control in an interface-specific fashion. However adaptive congestion control approach is a more effective and cost-effective solution. Hence ReCP can be a more powerful adaptive transport layer protocol, if it incorporates the advantages of receiver-centric schemes and the adaptive congestion control method; Secondly, seamless server migration capacity during handoffs; Thirdly effective bandwidth aggregation when receiving data through multiple interfaces.

R^2CP (Radial ReCP) is designed for multi-homed mobile hosts. R^2CP is a multi-state extension of ReCP at the receiver for the higher layer application[7]. A R^2CP connection has multiple independent ReCP senders with their corresponding ReCP receiver, and the R^2CP is responsible for coordinating receivers when a mobile hosts handoff from on interface to another during a live connection in heterogeneous networks. R^2CP can provide the following functionalities: seamless handoffs without relying on infrastructure support, server migration for achieving continuous service and bandwidth aggregation using multiple active interfaces.

V. CONCLUSION

At transport layer, the conventional TCP should be improved or substituted becomes unavoidable trend in the future wireless networks due to its inflexible retransmission mechanism and changeless congestion control scheme. Adaptive transport layer protocols are feasible schemes to cope with heterogeneity problem and maintain high performance in the future wireless networks. Sender-centric and receiver-centric adaptive transport layer protocols both adopt adaptive congestion control mechanism that can dynamically adjust AIMD parameters to offset the loss rate and delay components introduced by the heterogeneous wireless link. TCP-ATL is back compatible and TCP-friendly that satisfies a low complexity adaptive transport layer requirement for the mobile terminal with processing and power limitations. Furthermore ReCP can be improved to integrate the advantages of receiver-centric mechanism and adaptive congestion control scheme without increasing the complexities of transport protocol and the mobile terminal design. Furthermore the implementation of adaptive protocols is found on the reachable technologies.

ACKNOWLEDGMENT

First of all, I wish to express my profound gratitude to my supervisor Dr. Guray Acar for his guidance and support. His patient explanations and insightful suggestions help me overcome the difficult times. Finally, I should be grateful to my family members for their unfailing encourages and their profound love.

REFERENCES

[1] Ian Skyildiz and Yucel Altunbasak, “AdaptNet: An Adaptive Protocol Suite for the Next-Generation Wireless Internet,” IEEE Communication Magazine

[2] H. L. Z.Naor, “A Centralized Dynamic Access Probability Protocol for Next Generation Wireless Networks,” vol.2, 2001

[3] George Thomas, “A Rate-Adaptive Protocol For Wireless Data Channel,” IEEE, 2000

[4] Yang Richard, “General AIMD Congestion Control,” IEEE 2000

[5] Liang Zhao, Yuanwei Jing, “Improvement of AIMD Algorithm Considering Multimedia Transfer over Internet,” IEEE 2003

[6] Yih Huang, Genwen Xu, “Adaptive AIMD Rate Control for Smooth Multimedia Streaming,” IEEE 2003

[7] H.Y. Hsieh and R. Sivakumar, “A Receiver-centric Transport Protocol for Mobile Hosts with Heterogeneous Wireless Interfaces,” Proc. ACM MOBICOM, San Diego, CA, Sept. 2003.

Lian Wang received the B.S. degree in computer science from ChengDu University of Technology, Cheng Du, China, in 1999, and the M.S. degree in telecommunications and computer networks engineering from London South Bank University, UK, in 2004

She is a lecturer of ChongQing University of Posts and Telecommunications. Her research interests are in the areas of wireless communication technologies。