MPTCP Throughput Enhancement by Q-learning for Mobile Devices

Esmaeil Fakhimi
Gheshlagh Mohammad Beig
Faculty of Computer Science and Information Technology
University Of Eyvanekey
Eyvanekey, Iran
Fakhimi@eyc.ac.ir

Ali akbar Movassagh
Department of Medical Physics and Biomedical Engineering, School of Medicine
Tehran University of Medical Sciences
Tehran, Iran
a_movassagh@razi.tums.ac.ir

Parisa Daneshjoo
Department of Computer Engineering,
College of Engineering,
West Tehran Branch,
Islamic Azad University,
Tehran, Iran
Daneshjoo.p@wtiau.ac.ir

Ramin Karimi
Faculty of Information Technology,
Azad University Mallard,
Tehran, Iran
Rakarimi@iaumalard.ac.ir

Saeid Rezaei
Young Researchers and Elite Club,
West Tehran Branch,
Islamic Azad University,
Tehran, Iran
saeid.rezaei@wtiau.ac.ir

Yongrui Qin
School of Computing and Engineering
University of Huddersfield
Huddersfield, United Kingdom
yongrui.qin@gmail.com

***Abstract*—Mobile devices are able to leverage diverse heterogeneous network paths by Multi-Path Transmission Control Protocol (MPTCP); nevertheless, boosting MPTCP throughput in wireless networks is a real bear. Not only the best path(s) should be selected, but also the optimal congestion control mechanism should be chosen. We investigate the impact of different paths and congestion control for different signal quality states. Consequently, we present the novel MPTCP algorithm augmenting the end user throughput by understating the best policy in different situations by Q-learning. The Results reveal a tremendous effect of switching between the different interfaces and changing the congestion control mechanism on throughput and delay. By and large, the proposed framework achieves 10% more throughput compared to base MPTCP.**

Keywords— Multi-path TCP, Q-learning, Throughput, Congestion control

I. INTRODUCTION

The interest in mobile Internet data is dramatically boosting in wireless networks, mobile devices are armed with different network access interfaces, such as, Wi-Fi or Long-Term Evolution (LTE). Not long ago, MPTCP was introduced to enable end users to exploit multi-homed devices. In fact, mobile computing is one of the leading applications of MPTCP contributing to higher throughput and redundancy.

Despite the appreciated benefit of MPTCP to mobile users, MPTCP could put end users in a dicey situation since network availability and wireless properties fluctuate frequently. Undoubtedly, wireless uncertainty turns some different kinds of optimization on MPTCP as a challenging task. In particular, it could increase the delay and energy consumption when it is not tuned properly [16,24].

MPTCP optimization could fall into the category of network association and resource allocation in heterogeneous wireless networks.

The optimal network state can be achieved by employing either user-oriented or network-oriented approach. Nonetheless, in network-oriented approach network controllers should find the optimal user association and resource allocation which would be feasible in the age of software-defined networking (SDN). On the other side, in user-oriented networks, end users find out the optimal policy for joining different networks. Also, they learn how to compete and gain resource in the network. In this paper we consider a user-orientated approach where users are able to aggregate different access networks by applying MPTCP.

Three main motives are behind this paper. First of all, the ways we can achieve higher throughput by selecting the best network(s) in MPTCP context. Secondly, in terms of resource allocation, what kind of congestion control mechanism could help the end user to achieve higher data rate. Finally, in what ways a mobile phone could learn the optimized parameters based on different network scenarios. To the extent of our knowledge this is for the first time that network selection and congestion control resource allocation have been jointly considered. Therefore, this paper has three contributions as follow:

•We study the throughput enhancement problem in MPTCP context by state and reward uncertainty when the mobile node doesn't know the state transition probability distribution.

•We jointly find out the best policy toward network selection and the best congestion control mechanism for resource allocation on a mobile device by Q-learning.

•We investigate the effect of throughput optimization on MPTCP delay and the TCP and MPTCP congestion control mechanism performance on single sub flows.

The rest of the paper is organized as follows. Section II provides the related works. Section III describes the energy, and Q-learning models and the proposed algorithm. Section IV describes the simulation scenarios and discusses the results. Finally, we conclude in Section V.

II. RELATED WORK

The effect of different MPTCP congestion control algorithms and other parameters tuning have been investigated in large-scale simulation by [19]. Arzani et.al investigated the impact of path selection and packet scheduling on MPTCP throughput [2]; Nevertheless, they did not highlight the difference between congestion control mechanism and how to select the best initial path. In contrast, Shamani et.al [23] proposed selecting the best initial path by using the Bayesian game and they achieved more throughput in some conditions. In [7] authors show that MPTCP is not beneficial for short flows. Nonetheless, they did not consider finding the optimal decision for selecting different interfaces and congestion control mechanisms. Raiciu et.al [21] provides the first discussion of MPTCP in mobile networks by simulations and real experiments with WiFi and 3G. The authors show that MPTCP has a better throughput and energy efficiency; Nonetheless, in this paper we showed the effect of congestion control mechanism. The effect of throughput and size file on MPTCP energy consumption is studied in [15] and [24,25]. Authors trying to optimize energy consumption based on expected throughput. The benefit of MPTCP for WiFi visualization has been considered in [18]; However, the authors did not discuss the interference problem when one node downloads from the same channel but from different access points. By and large, none of the mentioned works have addressed the MPTCP in a game theoretic context when users are competing for shared resources. [17] Nagahvi et.al formulated the problem of Radio access technology (RAT) selection as an incomplete information game where players converge to Nash equilibrium by using Q-learning based algorithm. Finally, [22] employ Q-Learning to enhance the user energy efficiency by selecting the optimal path. However, none of these studies have looked at the possibility of using Q-learning for maximizing throughput of flows in divergent networks.

III. SYSTEM MODEL

The TCP extension which allows leveraging of multiple TCP sub flows for data transfer is MPTCP [9]. Two different types of congestion control mechanisms are introduced for MPTCP, coupled and uncoupled. An uncoupled congestion control which is the simplest form of control mechanism maintains a separate congestion window for each sub flow. Nevertheless, lack of resource pooling and agility in response time are the main features of this protocol. On the other hand, coupled congestion control endeavors to merge the congestion windows for different sub flows; Nonetheless, exploiting resource pooling and long response time are the main issues of coupled congestion control mechanisms. For instance, TCP Reno [1] and Cubic [12] are considered as uncoupled, while weighted vegas (wVgas) [28], linked increase adaptation (LIA) [20], opportunistic linked increase adaptation (OLIA) [13] and balanced linked adaptation (BALIA) [27] fall in the category of coupled congestion control. In this paper, we are looking to find out the best interface(s) and congestion control mechanism for different signal quality strength and different file size which improve the end user throughput. Thus, we employ Q-learning to learn the optimal decision for MPTCP mobile users.

The general model, where we have a number of states, actions, rewards, and transition probabilities are considered as a Markov Decision Process. (MDP). Nonetheless, defining probability distribution on the set of states for a mobile node is not possible because of the scenario complexity. Multi-agent reinforcement learning (MARL) could alleviate this problem by learning of the unknowns. The mobile device figures out the state and then based on the present state it selects an action, in consequence the environment makes a transition to the next state and the node will receive the reward. The agent consider the reward and the process is repeated.

IV. Q-LEARNING MODEL

Suppose the mobile device environment consists of finite states. Any action chosen from the action set introduces a new throughput data rate which is considered as rewards to the mobile user. Q-learning helps us find a policy which maximizes the throughput for the mobile node. A Q-learning game is defined as follows:

1)Agent *i* is a mobile device equipped with different interfaces.

2)The states are defined as a compound state of agent *i* based on signal strength for each interface and the transfer size file. The signal quality set for user *i* with interface *j* denoted by, $b_{ij} = \{b_{ij}^1, b_{ij}^2, \ldots, b_{ij}^k\}$, where b_{ij}^1 is the minimum and b_{ij}^k is the maximum signal quality state. Then the signal quality state for user *i* is defined as $B_i = \{b_{i1}b_{i2}, \ldots, b_{im}\}$. Let F_i denote the file size set of user *i* $F_i = \{f_{i1}, f_{i2}, \ldots, f_{im}\}$, where f_{i1} is the minimum and f_{im} is the maximum file size state. Thus, the compound state for user *i* is defined as $S_t = (Bi, F_i)$.

3) Actions are defined as a set of compound action of agent *i* based on available interfaces and congestion control mechanism. Without loss of generality, in this paper the set of available network interfaces are, N_i = {WiFi, LTE, MPTCP} and (MP)TCP congestion control mechanism set for user *i* is presented by C_i= {lia, Olia, wvegas, cubic, reno, westwood}. Therefore, the compound set of actions is $A_i = (N, C)$.

4)Reward is defined as the immediate throughput data rate return which is experienced due to the selection of any action from the compound action set A_i. The agent *i* is looking for a policy π_i^* which maximizes the throughput r=r (s, a) over time. The expected discounted reward of doing action *a* in states is defined as Q (s, a). It is worth noting that discounted reward should be considered since the return value shows the reward a mobile device incurs during one-time period in the future; Thus, we may discount it by a discount factor γ where ($0 \ll \gamma < 1$). The mobile device objective is maximizing the expected long-term reward which is defined as:

$$V^{\pi}(s) = \max_{\pi} E\{\sum_{(t=0)} \gamma^{t} r^{\pi}(s_t, a^{\pi}(s_t))\}. \tag{1}$$

The values of rewards and transition probabilities p ($s_{(t+1)} \mid s_t$) can only be learnt over time. Therefore, we should employ Q-learning to learn these values gradually. Let's define a policy π, which maps states to actions [29-33]. Q-learning would find the policy that maximizes the stochastic rewards. Consider $v(s_{t+1})$ as value at state s_{t+1}, then in each time slot we aim for an action which has the maximum reward. By using optimal action $a_t^*(s_t)$, the optimal value is equal to:

$$V^*(s_t)=R(s_t, a^*(s_t)) + \gamma V(s_{t+1}(s_t, a^*(s_t))) \tag{2}$$

Let us define a Q-value which is a discounted reward after selection of an action a for any policy π as

$$Q^*(s, a) = R(s, a) + \gamma \sum_{v \epsilon S} P_{s,v}(a) V^{\pi}(v) \tag{3}$$

Q-learning algorithm is used iteratively to find out the Q-value. The mobile device seeks to select all the actions with random probability $0 < \epsilon < 1$, then the mobile devices use a learning rate to update the Q-values θ by

$$Q(s,a) \leftarrow Q(s,a) + \theta[r + \gamma \max_a Q(v,a) - Q(s,a)] \tag{4}$$

algorithm 1:

Q-learning based algorithm for MPTCP throughput enhancement

Input: ε, γ, number of states

Begin

Let t=0, $Q_0(0,0)=0$

for time t do

if r<ε then

select action randomly from possible actions

else

select action a_t which has the maximum Q value

end

receive an immediate reward r_t

observe new state s_{t+1},update the Q table as given in (4)

end

End

V. Evaluation and results

We use NS3 in order to evaluate the proposed framework. The simulation parameters are shown in Table1. Four different states for Wi-Fi and LTE have been considered. For each state we have simulated all available actions. For instance, we test all congestion control mechanisms for LTE, Wi-Fi and MPTCP. Also, we considered different scenarios by adding competitors in each simulation. We consider 0, 2 and 4 competitors to understand the different congestion control mechanism reaction to different scenarios. Moreover, we consider two different file size to evaluate the MPTCP congestion control behavior. The agent moves between all states in his Wi-Fi and LTE range randomly and he gradually learns the throughput reward in each state. All other nodes try to download the 32MB file size from the same server and they have a static position. Simulation Parameters are listed in Table 1.

Table 1. Simulation Parameters

Parameters	Values
Wi-Fi states(dB)	(-85,-75),(-75,-65),(-55,-45),(-45,-35)
LTE states(dB)	(-124,-108),(-108,-92),(-92,-76),(-76,-60)
Download size	512KB,32MB
Point to point delay	1ms
Point to point bandwidth	40Mbps
MPTCP Congestion control	LIA , OLIA , Cubic , Wvegas
TCP Congestion control	Reno, Vegas
MPTCP buffer size	128KB
TCP buffer size	64KB

The number of competitors is equally distributed between the LTE base station and Wi-Fi access point but the Wi-Fi and LTE state for each node is randomly selected at the beginning of each simulation.

When the mobile agent is not in competition with other nodes, the highest throughput achieved by M-cubic for small and large file transfer will be based on Fig.1 and Fig.2. However, the lowest RTT belongs to Wvegas over Wi-Fi. It is worth noting that LIA has the lowest RTT among MPTCP congestion control. When the mobile agent is competing with 2 nodes, the highest throughput achievement by gain M-cubic and still Wvegas has the lowest RTT among all and LIA among MPTCP congestion control mechanism. Contrary to the previous simulations in 4 nodes scenario, M-Wvegas has the highest throughput and lowest RTT belongs to W-vegas [25-28].

The throughput values for various congestion control algorithms for large and small file transfer When the mobile agent is competing with 0 nodes are shown in Fig. 1 and Fig. 2.

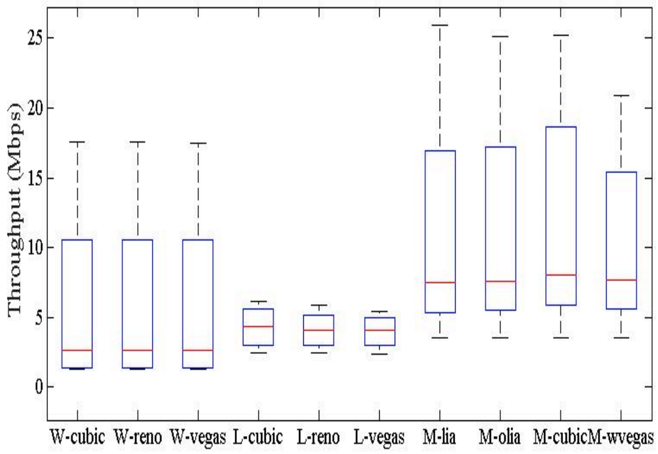

Fig. 1. 32 MB transfer in completion of 0 node.

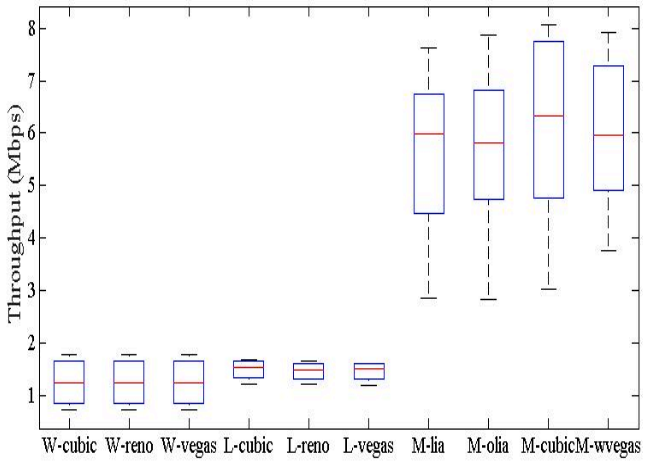

Fig. 2. 512 KB transfer in completion of 0 node.

The throughput values for various congestion control algorithms for large and small file transfer When the mobile agent is competing with 2 nodes are shown in Fig. 3 and Fig. 4.

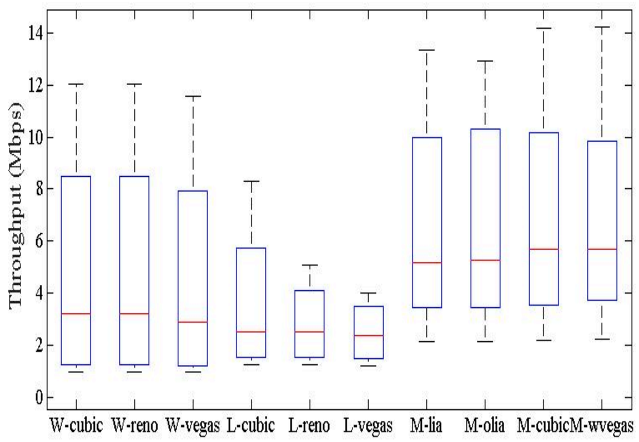

Fig. 3. 32 MB transfer in completion of 2 nodes.

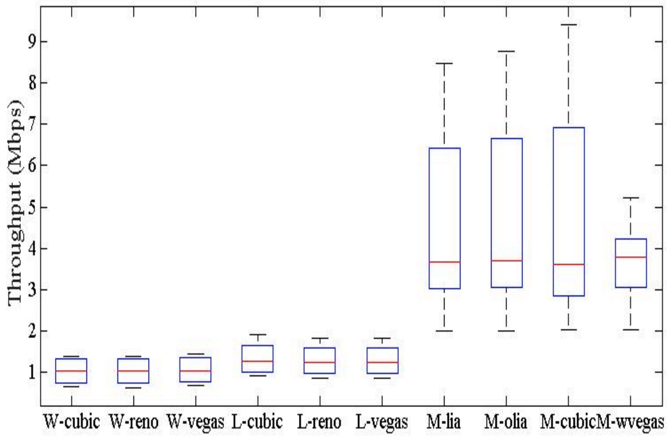

Fig. 4. 512 KB transfer in completion of 2 nodes.

The throughput values for various congestion control algorithms for large and small file transfer When the mobile agent is competing with 4 nodes are shown in Fig. 5 and Fig. 6.

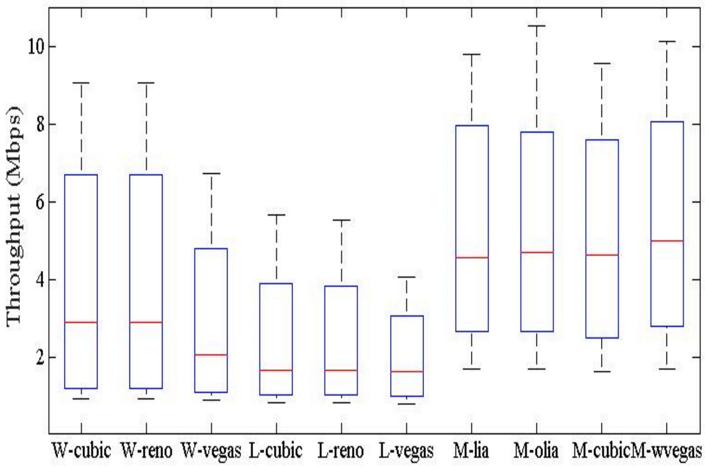

Fig. 5. 32 MB transfer in completion of 4 nodes.

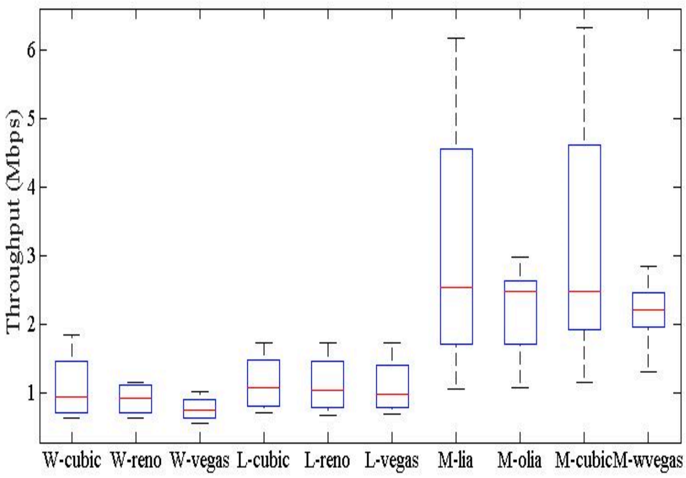

Fig. 6. 512 KB transfer in completion of 4 nodes.

The RTT values for various congestion control algorithms in Wifi link are shown in Fig. 8 and throughput values for Wi-Fi and LTE links are shown in Fig. 7 and Fig. 9 respectively.

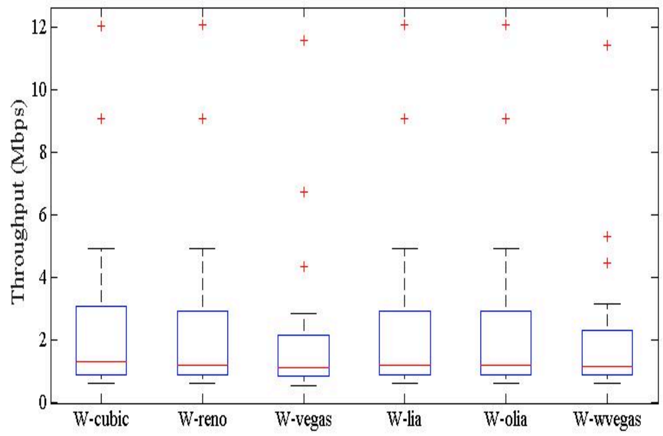

Fig. 7. Wifi Throuput comparison.

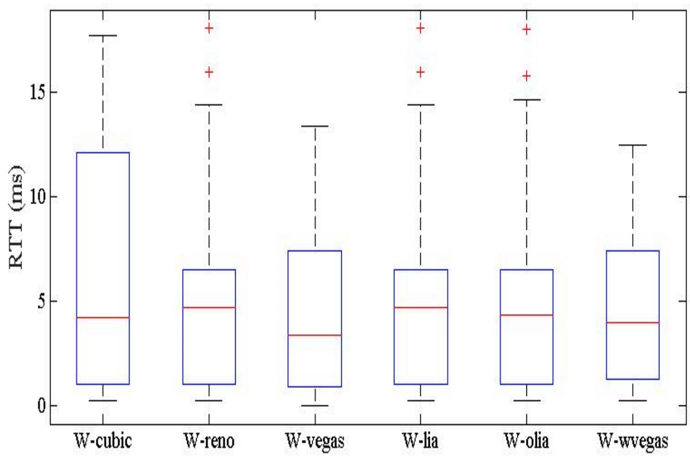

Fig. 8. Wifi RTT comparison.

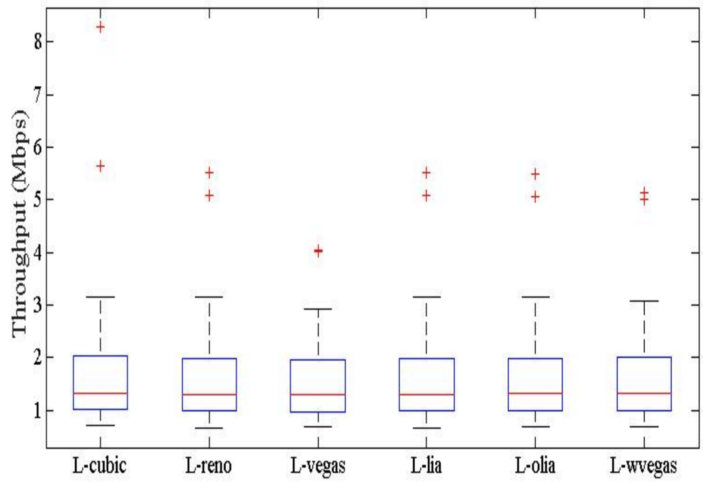

Fig. 9. LTE Throuput comparison.

Considering all scenarios there is no difference between TCP and MPTCP backup modes in terms of throughput; nevertheless, Vegas for both Wi-Fi and LTE gains the lowest RTT. By and large, the proposed mechanism enhances the MPTCP by 10%, which mainly is based on intelligent switching between LTE, WIFI, MPTCP and different congestion control mechanism.

VI. CONCLUSION

The optimal transmission strategy for the MPTCP mobile device to maximize throughput in heterogeneous wireless networks depends not only on the other users but also on the congestion control mechanisms employed by the mobile node. In this paper we introduced a Q-learning approach for determining the best strategy as it intelligently chooses the best action between selecting different interfaces and the optimal congestion control mechanism based on the previous experience. The proposed scheme boost the MPTCP throughput by 10% which is a significant achievement.

REFERENCES

1. Mark Allman, Vern Paxson, and Ethan Blanton. 2009. TCP congestion control. Technical Report.
2. Behnaz Arzani, Alexander Gurney, Shuotian Cheng, Roch Guerin, and Boon Thau Loo. 2014. Impact of path characteristics and scheduling policies on MPTCP performance. In Advanced Information Networking and Applications Workshops (WAINA), 2014 28th International Conference on. IEEE, 743–748.
3. Sbastien Barré, Christoph Paasch, and Olivier Bonaventure. 2011.Multipath TCP: from theory to practice. NETWORKING 2011(011), 444–457.
4. Mehdi Bennis and Dusit Niyato .2010. A Q-learning based approach to interference avoidance in self-organized femtocell networks. In GLOBECOM Workshops (GCWkshps), 2010 IEEE.IEEE, 706–710.

5. Vinh Bui and Weiping Zhu. 2008. A game theoretic framework for multipath optimal data transfer in multiuser overlay networks. In Communications, 2008. ICC'08. IEEE International Conference on. IEEE, 401–407.
6. Cisco Visual Networking Index Cisco. 2014. Global mobile data traffic forecast update, 2013–2018. white paper (2014).
7. Shuo Deng, Ravi Netravali, Anirudh Sivaraman, and Hari Balakrishnan. 2014. Wifi, lte, or both? Measuring multi-homed wireless internet performance. In Proceedings of the 2014 Conference on Internet Measurement Conference. ACM,181–194.
8. Pingping Dong, Jianxin Wang, Jiawei Huang, Haodong Wang, and Geyong Min. 2016. Performance Enhancement of MultipathTCP for Wireless Communications with Multiple Radio Interfaces. IEEE Transactionson Communications 64,8(2016), 3456–3466.
9. Alan Ford, Costin Raiciu, Mark Handley, and Olivier Bonaventure. 2013. TCP extensions for multipath operation with multiple addresses. Technical Report.
10. Fangwen Fuand Mihaelavander Schaar. 2009. Learning to compete for resources in wireless stochastic games. IEEE Transactions on Vehicular Technology 58,4 (2009),1904–1919.
11. Ana Galindo-Serrano and Lorenza Giupponi. 2010. Distributed Q-learning for aggregated interference control in cognitive radio networks. IEEE Transactions on Vehicular Technology 59,4(2010),1823–1834.
12. Sangtae Ha, Injong Rhee, and Lisong Xu. 2008. CUBIC: a new TCP-friendly high-speed TCP variant. ACM SIGOPS Operating Systems Review 42,5(2008), 64–74.
13. Ramin Khalili, Nicolas Gast, Miroslav Popovic, and Jean-Yves Le Boudec. 2013. MPTCP is not pareto-optimal: performance issues and a possible solution. IEEE/ACM Transactions on Networking 21,5(2013),1651–1665.
14. Husheng Li. 2009. Multi-agent Q-learning of channel selection in multi-user cognitive radio systems: a two by two case. In Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on. IEEE,1893–1898.
15. Yeon-sup Lim, Yung-Chih Chen, Erich MNahum, Don Towsley, and Richard J Gibbens. 2014. How green is MultipathTCP for mobile devices? In Proceedings of the 4th workshop on All things cellular: operations, applications, & challenges. ACM,3–8.
16. Yeon-sup Lim, Yung-Chih Chen, Erich MNahum, Don Towsley, and Kang-Won Lee. 2014. Cross-layer path management in multi-path transport protocol for mobile devices. In INFOCOM, 2014 Proceedings IEEE.IEEE,1815–1823.
17. Parnia Naghavi, Seyed Hamed Rastegar, Vahid Shah-Mansouri, and Hamed Kebriaei. 2016. Learning RAT selection game in 5G heterogeneous networks. IEEE Wireless Communications Letters 5,1(2016),52–55.
18. Kien Nguyen, Yusheng Ji, and ShigekiYamada. 2014. Across-layer approach for improving WiFi performance. In Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International.IEEE,458–463.
19. Christoph Paaschetal. 2014. Improving Multipath TCP. Ph.D. Dissertation. UCL.
20. Costin Raiciu, Mark Handley, and Damon Wischik. 2011. Coupled congestion control for multipath transport protocols. Technical Report.
21. Costin Raiciu, Dragos Niculescu, Marcelo Bagnulo, and Mark James Handley. 2011. Opportunistic mobility with multipath TCP. In Proceedings of the sixth international workshop on MobiArch.ACM,7–12.
22. Mohammad Javad Shamani, Saeid Rezaei, Aruna Seneviratne, and Guillaume Jourjon. 2017. A Q-Learning Framework for Multi-Path TCP Energy Saving in Mobile Devices. In Local Computer Networks, 2017 42th International Conference on. IEEE.
23. Mohammad Javad Shamani, Saeid Rezaei, Aruna Seneviratne, and Hamed Kebriaei. (accepted)2017. Multi-Path TCP Incomplete in formation Repeated Bayesian Game. In Vehicular Technology Conference, 2017 86th International Conference on. IEEE.
24. Mohammad Javad Shamani, Weiping Zhu, andSaeidRezaie.2016. On the Energy Inefficiency of MPTCP for Mobile Computing. In International Conference on Wired/Wireless Internet Communication.Springer,93–105.
25. Mohammad Javad Shamani, Weiping Zhu, Saeid Rezaie, and Vahid Naghshin. 2016. Signal aware multi-path TCP. In Wireless On-demand Network Systems and Services (WONS), 2016 12th Annual Conference on. IEEE,1–4.
26. M. Gheisari, G. Wang, and M. Z. A. Bhuiyan. A survey on deep learning in big data. In 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), volume 2, pages 173–180, July 2017.
27. M. Gheisari, G. Wang, M. Z. A. Bhuiyan, and Wei Zhang. Mapp: A modular arithmetic algorithm for privacy preserving in iot. In 15th IEEE International Symposium on Parallel and Distributed Processing with Applications (IEEE ISPA 2017) Guangzhou, China,, December 2017.
26. Hassan Sinky, Bechir Hamdaoui, and Mohsen Guizani. 2016. Proactive Multipath TCP for Seamless Handoff in Heterogeneous Wireless Access Networks. IEEE Transactions on Wireless Communications 15,7(2016),4754–4764.
27. Anwar Walid, Qiuyu Peng, Jaehyun Hwang, and SLow.2015. Balanced linked adaptation congestion control algorithm for MPTCP. IETF, Individual Submission, Internet Draft draft-walid-MPTCP-congestion-control-03 (2015).
28. Mingwei Xu, Yu Cao, and Enhuan Dong. 2015. Delay-based Congestion Control for MPTCP. IETF, work in progress, Internet-draft mptcp congestion control-01(2015).
29. M. Gheisari, A. A. Movassagh, Y. Qin, J. Yong, X. Tao, J. Zhang, and H. Shen. Nsssd: A new semantic hierarchical storage for sensor data. In 2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), pages 174–179, May 2016
30. Tao Peng, Qin Liu, Dacheng Meng, and Guojun Wang. Collaborative trajectory privacy preserving scheme in location-based services. Information Sciences, 387:165 – 179, 2017.
31. Kaveh Bashash Shahram Bashokian mehdi gheisari Manaf Sharifzadeh, saeid aragy. A comparison with two semantic sensor data storages in total data transmission. Networking and Internet Architecture, 2014
32. Mehdi Gheisari and Mehdi Esnaashari. Data storages in wireless sensor networks to deal with disaster management. Smart Technologies for Emergency Response and Disaster Management, page 196222, 2018.
33. Mehdi Gheisari. Design, Implementation, and Evaluation of SemHD: A New Semantic Hierarchical Sensor Data Storage. Indian J. Innovations Dev., Vol. 1, No. 3 (Mar 2012) ISSN 2277 – 5390