JIA ZHANG

joycezhangjia@gmail.com | joycezhangjia | +86 13121958631

Education

Tsinghua University China
Ph.D. Student with Department of Computer Science and Technology *2018 – 2023*
- Advisor: Prof. Mingwei Xu.

Tsinghua University China
B.Eng. with Department of Electronic Engineering, minor in Business Administration *2014 – 2018*

Work Experience

Research Intern | *Beijing Sankuai Online Technology Co., Ltd., China* Fall 2021 - Summer 2023
- Algorithm Strategy, Mentor: Mr. Xiaotian Li.

Research Intern | *Baidu Inc., China* Fall 2019 - Spring 2021
- BFE Team, Mentor: Mr. Sijie Yang.

Visiting Research Assistant | *University of Surrey, UK* Summer 2019
- 5G Innovation Centre, Advisor: Prof. Ning Wang.

Research Assistant | *Tsinghua University, China* Fall 2016 - Spring 2017
- Lab of New Generation Network Technology & Applications, Advisor: Prof. Yongfeng Huang

Research

Research Interest

- Routing Security
- Resource Public Key Infrastructure
- Transport Layer Performance Measurement and Improvement

Publications

- **(WWW'24) Jia Zhang**, Haitian Tong, Enhuan Dong, Xin Qian, Mingwei Xu, Xiaotian Li, Zili Meng. Cold Start or Hot Start? Robust Slow Start in Congestion Control with A Priori Knowledge for Mobile Web Services, TheWebConf 2024.
- **(ATC'23) Jia Zhang**, Yixuan Zhang, Enhuan Dong, Yan Zhang, Shaorui Ren, Zili Meng, Mingwei Xu, et al. Bridging the Gap between QoE and QoS in Congestion Control: A Large-scale Mobile Web Service Perspective. 2023 USENIX Annual Technical Conference.
- **(ToN'23) Jia Zhang**, Shaorui Ren, Enhuan Dong, Zili Meng, Yang Yuan, Xu Mingwei, Sijie Yang, Miao Zhang, Yang Yue. Reducing Mobile Web Latency through Adaptively Selecting Transport Protocol, Transactions on Networking (ToN).
- **(WWW'21) Jia Zhang**, Enhuan Dong, Zili Meng, Yuan Yang, Mingwei Xu, Sijie Yang, Miao Zhang Yang Yue. WiseTrans: Adaptive Transport Protocol Selection for Mobile Web Service, TheWebConf 2021.
- **(ToN'21)** Zili Meng, Yaning Guo, Yixin Shen, Jing Chen, Chao Zhou, Minhu Wang, **Jia Zhang**, Mingwei Xu, Chen Sun, Hongxin Hu, Practically Deploying Heavyweight Adaptive Bitrate Algorithms With Teacher-Student Learning, IEEE/ACM Transactions on Networking (ToN), 2021.
- **(SIGIR'17)** Fangzhao Wu, **Jia Zhang**, Zhigang Yuan, Sixing Wu, Yongfeng Huang, and Jun Yan. 2017. Sentence-level sentiment classification with weak supervision. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2017. 973–976.

Research Experience

- **Cold Start or Hot Start? Robust Slow Start in Congestion Control with A Priori Knowledge for Mobile Web Services (WWW'24)**

  Mobile web services value a quick loading of contents in the first page, which is quantified by the above-the-fold time of the first page (first AFT) and is likely to fall into the slow start phase in congestion control. However, the widely deployed slow start mechanism is "cold start", which manually hardcodes the parameters and is not suitable for the first AFT of heterogeneous mobile web services. We revisit the slow start mechanism and find that it could be optimized with a priori knowledge. However, blindly relying on a priori knowledge is not robust enough to handle the fluctuating mobile networks and unpredictable application traffic. In this paper, we propose WiseStart, a "hot-start-based" slow start mechanism. WiseStart utilizes the priori knowledge to set the initial parameters, continuously probes the new connection to handle the fluctuating network conditions, and carefully adapts to the application-limit scenarios. We implement WiseStart in a popular mobile web service online in production. Comprehensive experiments demonstrate that WiseStart reduces the First AFT by 25.43% and the average RCT at connection establishment by 16.15% compared to the default slow start mechanism and other state-of-the-art baselines.
- **Bridging the Gap between QoE and QoS in Congestion Control: A Large-scale Mobile Web Service Perspective (ATC'23)**

  To improve the user experience of mobile web services, various congestion control algorithms (CCAs) have been proposed, yet the performance of the application is still unsatisfactory. We argue that the suboptimal performance comes from the gap between what the application needs (i.e., Quality of Experience (QoE)) and what the current CCA is optimizing (i.e., Quality of Service (QoS)). However, optimizing QoE for CCAs is extremely challenging due to the convoluted relationship and mismatched timescale between QoE and QoS. To bridge the gap between QoE and QoS for CCAs, we propose Floo, a new QoE-oriented congestion control selection mechanism, as a shim layer between CCAs and applications to address the challenges above. Floo targets request completion time as QoE, and conveys the optimization goal of QoE to CCAs by always selecting the most appropriate CCA in the runtime. Floo further adopts reinforcement learning to capture the complexity in CCA selection and supports smooth CCA switching during transmission. We implement Floo in a popular mobile web service application online. Through extensive experiments in production environments and on various locally emulated network conditions, we demonstrate that Floo improves QoE by about 14.3% on average.
- **Reducing Mobile Web Latency Through Adaptively Selecting Transport Protocol (ToN'23)**

  To improve the performance of mobile web services, a new transport protocol, QUIC, has been recently proposed as a substitute for TCP. However, with pros and cons of QUIC, it is challenging to decide whether and when to use QUIC in large-scale real-world mobile web services. Complex temporal correlation of network conditions, high user heterogeneity in a nationwide deployment, implementation diversity of QUIC variants limited, and resources on mobile devices all affect the selection of transport protocols. In this paper, we present WiseTrans, an adaptive transport protocol selection mechanism, to switch transport protocols for mobile web services online and improve the completion time of web requests. WiseTrans introduces machine learning techniques to deal with temporal heterogeneity, makes decisions with historical information to handle spatial heterogeneity, adopts an online learning method to keep pace with implementation variation, and switches transport protocols at the request level to reach high performance with acceptable overhead. We implement WiseTrans on two platforms (Android and iOS) in a popular mobile web service application of Baidu. Comprehensive experiments demonstrate that WiseTrans can reduce request completion time by up to 25.8% on average compared to the usage of a single protocol.
- **WiseTrans: Adaptive Transport Protocol Selection for Mobile Web Service (WWW'21)**

  To improve the performance of mobile web service, a new transport protocol, QUIC, has been recently proposed. However, for large-scale real-world deployments, deciding whether and when to use QUIC in mobile web service is challenging. Complex temporal correlation of network conditions, high spatial heterogeneity of users in a nationwide deployment, and limited resources on mobile devices all affect the selection of transport protocols. In this paper, we present WiseTrans to adaptively switch transport protocols for mobile web service online and improve the completion time of web requests.

WiseTrans introduces machine learning techniques to deal with temporal heterogeneity, makes decisions with historical information to handle spatial heterogeneity, and switches transport protocols at the request level to reach both high performance and acceptable overhead. We implement WiseTrans on two platforms (Android and iOS) in a popular mobile web service application of Baidu. Comprehensive experiments demonstrate that WiseTrans can reduce request completion time by up to 26.5% on average compared to the usage of a single protocol.

- **Practically Deploying Heavyweight Adaptive Bitrate Algorithms With Teacher-Student Learning (ToN'21)**
  Major commercial client-side video players employ adaptive bitrate (ABR) algorithms to improve the user quality of experience (QoE). With the evolvement of ABR algorithms, increasingly complex methods such as neural networks have been adopted to pursue better performance. However, these complex methods are too heavyweight to be directly deployed in client devices with limited resources, such as mobile phones. Existing solutions suffer from a trade-off between algorithm performance and deployment overhead. To make the deployment of sophisticated ABR algorithms practical, we propose PiTree, a general, high-performance, and scalable framework that can faithfully convert sophisticated ABR algorithms into decision trees with teacher-student learning. In this way, network operators can train complex models offline and deploy converted lightweight decision trees online. We also present theoretical analysis on the conversion and provide two upper bounds of the prediction error during the conversion and the generalization loss after conversion. Evaluation on three representative ABR algorithms with both trace-driven emulation and real-world experiments demonstrates that PiTree could convert ABR algorithms into decision trees with <3% average performance degradation. Moreover, compared to original deployment solutions, PiTree could save considerable operating expenses for content providers.
- **Sentence-level Sentiment Classification with Weak Supervision (SIGIR'17)**
  Sentence-level sentiment classification is important to understand users' fine-grained opinions. Existing methods for sentence-level sentiment classification are mainly based on supervised learning. However, it is difficult to obtain sentiment labels of sentences since manual annotation is expensive and time-consuming. In this paper, we propose an approach for sentence-level sentiment classification without the need of sentence labels. More specifically, we propose a unified framework to incorporate two types of weak supervision, i.e., document-level and word-level sentiment labels, to learn the sentence-level sentiment classifier. In addition, the contextual information of sentences and words extracted from unlabeled sentences is incorporated into our approach to enhance the learning of sentiment classifier. Experiments on benchmark datasets show that our approach can effectively improve the performance of sentence-level sentiment classification.
- **Extend BBR congestion control algorithm into MPQUIC**
  In the context of future Internet development, emerging applications such as connected vehicles and eHealth require low-latency performance. Traditional protocols like TCP and congestion control mechanisms based on packet losses are inadequate for meeting these new demands. As a result, researchers have turned to the QUIC transmission protocol and the BBR congestion control algorithm. This study focuses on congestion control in MPQUIC, extending the BBR algorithm to improve page download speed and reduce network delay in MPQUIC environments with random loss. Experimental results demonstrate the effectiveness of BBR in handling random packet loss and its robustness across different network paths. The study also compares the performances of MPQUIC and MPTCP, highlighting the need for further research and providing guidance for deployment and improvement of QUIC and MPQUIC.

Teaching and Tutoring

Teaching Assistant Fall 2020
Computer Network Architecture, Department of Computer Science and Technology *Tsinghua University*

Undergraduate Tutor Fall 2018 - Fall 2020
Department of Electronic Engineering *Tsinghua University*