MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
IEEE COMMUNICATIONS SOCIETY
http://mmc.committees.comsoc.org/

MMTC Communications – Review

IEEE COMMUNICATIONS SOCIETY

Vol. 14, No. 5, October 2023

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the October 2023 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises three reviews that cover multiple facets of multimedia communication research including privacy-reserving and resilient mobile robotic systems, low-latency video streaming, and secure and efficiently searchable IoT communication. These reviews are briefly introduced below.

The first paper, published in IEEE Wireless Communications edited by Dr. Shengjie Xu, present a decentralized, distributed deep learning framework for mobile robotics application to overcome unreliable wireless connections by aggregating trained parameters asynchronously in a peer-to-peer manner instead of using a centralized approach.

The second paper, edited by Dr. Luca De Cicco, was published in IEEE Transactions on Multimedia. This paper investigates a delivery architecture for low-latency video streaming.

The third paper, edited by Dr. Qichao Xu, was published in IEEE Internet of Things Journal. The authors proposed an IoT communication data management model based on blockchain.

All the authors, reviewers, editors, and others who contribute to the release of this issue deserve appreciation with thanks.

IEEE ComSoc MMTC Communications – Review Directors

Yao Liu
Rutgers University, USA
Email: yao.liu@rutgers.edu

Wenming Cao
Shenzhen University, China
Email: wmcao@szu.edu.cn

Dongfeng (Phoenix) Fang
California Polytechnic State University, USA
Email: dofang@calpoly.edu

Ye Liu
Macau University of Science and Technology, Macau, China Email: liuye@must.edu.mo

Decentralized P2P Federated Learning for Privacy-Preserving and Resilient Mobile Robotic Systems

A short review for "Decentralized P2P Federated Learning for Privacy-Preserving and Resilient Mobile Robotic Systems"

Edited by Shengjie Xu

X. Zhou et al., "Decentralized P2P Federated Learning for Privacy-Preserving and Resilient Mobile Robotic Systems," in IEEE Wireless Communications, vol. 30, no. 2, pp. 82-89, April 2023, doi: 10.1109/MWC.004.2200381.

In this article, the authors present a decentralized, distributed deep learning framework for mobile robotics applications called P2P-PPAFL. It uses virtual coordination servers, each corresponding to a group of smart devices. These servers only exchange parameters and update models by communicating with nearby clients. Multiple virtual clusters manage decentralized federated learning among groups of mobile robots. To overcome unreliable wireless connections, trained parameters are aggregated asynchronously in a peer-to-peer manner instead of using a centralized approach.

The paper first presents the overview of federated learning in mobile robotics systems. Distributed machine learning faces security threats in data collection, model training, parameter transmission, and model aggregation. Distributed machines may be attacked to tamper with training data, interfere with the learning process, or invade privacy. Mobile applications are especially vulnerable to malware attacks on deep learning models. Eavesdroppers can compromise data offloading in mobile robotics through cognitive eavesdropping.
Data poisoning attacks can interfere with distributed model training by injecting malicious samples. Autonomous robots may face adversarial environments where attackers influence their decisions through physical or software attacks. Compromised sensors in robots like self-driving vehicles are a major concern.
Federated learning (FL) provides a distributed framework for privacy-preserving machine learning by training models across many participants.

FL systems may or may not use a central coordinator. P2P networks can ensure security by enabling direct communication without a coordinator. Privacy protection is crucial but challenging in FL. Attackers can obtain parameters to reconstruct or reason about private data. Proposed privacy schemes fall into two categories: encryption based on secure multi-party computation and differential privacy. These aim to make FL more secure and reliable against attacks to destroy data privacy.

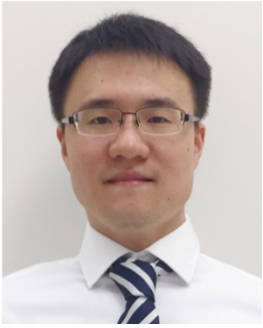
The paper then presents the system model. First, A distributed mobile robotics system, like unmanned aerial vehicles in cities or multi-agent manufacturing, uses different robots with varying capabilities that must collaborate despite resource constraints and unreliable connections. To enable efficient decentralized learning in this environment, a peer-to-peer federated learning framework trains an accurate global model across robots in temporary clusters without a central server. Second, the proposed peer-to-peer federated learning framework incorporates asynchronous model aggregation based on reputation-aware coordination and secure local training using secret sharing communication and stochastic gradient descent with autoencoders and Gaussian noise. This enables privacy-preserving, resilient decentralized learning across mobile robots by dynamically clustering them for encrypted peer-to-peer communication and secure localized model updates.

The paper then presents the implementation of mobile robotic applications. First, to enable resilient federated learning among mobile robots, models are aggregated asynchronously in a peer-to-peer manner instead of relying on a central server, with robots dynamically forming temporary clusters for each training epoch after evaluating neighbors based on reputation from factors like recent performance and parameter

diversity. This asynchronous, decentralized aggregation prevents wasted idle resources waiting for synchronization while allowing models to be robustly trained through diverse peer contributions in unreliable environments. Second, to enable secure communication, secret sharing encrypts and distributes model parameters as shards across peers so that the originals can only be recovered when enough shards are collected, using a scheme where secrets are split into n shards and k are needed to recover them. This protects privacy during model aggregation by letting peers cooperatively recover averaged parameters encrypted via secure multi-party computation of shard subtotals, ensuring federated learning without exposing raw model data. Third, to enable secure localized training, stochastic gradient descent randomly samples data to simplify and safeguard each iteration's updates, adding Gaussian noise during gradient calculations and using autoencoders to anonymize the updated weight parameters after backpropagation. This perturbation and reconstruction preserve privacy in each robot's individual training while the randomness and encryption of stochastic methods protects the parameters during decentralized learning.

A case study of underground tunnel construction is used to demonstrate the proposed approach for mobile robotics applications involving different sensors and devices functioning in harsh environments. First, three communication scenarios are considered between devices and network infrastructure in underground tunnels: 1) High-speed direct connections within a section; 2) Relayed low-throughput connections within a tunnel; 3) External high-latency links between tunnels. To evaluate these, 9 devices across 3 network partitions simulated Google Compute Engine bandwidths and latencies for intra-section, inter-section, and inter-tunnel communications. The non-IID EMNIST handwritten digit dataset was distributed across devices in the partitions to assess federated learning efficiency under the 3 scenarios. Second, synchronous and asynchronous P2P federated learning were evaluated on distributed EMNIST data, with asynchronous converging 42% faster in accuracy and 37% faster overall. Compared to vanilla federated learning, the proposed PPAFL framework achieved better efficiency in low-, medium- and high-latency networks while requiring less frequent aggregation. Experiments demonstrate PPAFL enables efficient, privacy-preserving decentralized learning for mobile robotics, with asynchronous aggregation and infrequent coordination improving performance in unreliable conditions.

The paper concludes the study and states that mobile robotics systems with heterogeneous devices struggle with unreliable, insecure communications in high-latency networks, therefore peer-to-peer federated learning is proposed to enable flexible, privacy-preserving decentralized learning for robots. The introduced reputation and secret sharing mechanisms facilitate asynchronous, encrypted model aggregation and training in dynamically formed clusters of robots for scalable and reliable learning in challenging environments.



Shengjie Xu [SM'14-M'19] received a Ph.D. degree in Computer Engineering from the University of Nebraska-Lincoln and an M.S. degree in Telecommunications from the University of Pittsburgh. Before that, he held a B.E. degree in Computer Science and Information Security. He is an assistant professor in the Management Information Systems Department at the Fowler College of Business at San Diego State University. His research interests are cybersecurity, trustworthy AI and robust machine learning, secure edge computing, and critical infrastructure protection. He serves as a Technical Editor for IEEE Wireless Communications. He is the recipient of the IET Journals Premium Award for Best Paper. He holds multiple professional certifications in cybersecurity and computer networking.

A Delivery Architecture for Low-Latency Video Streaming

A short review for "HxL3: Optimized Delivery Architecture for HTTP Low-Latency Live Streaming"

Edited by Luca De Cicco

F. Tashtarian, A. Bentaleb A. Erfanian, H. Hellwagner, C. Timmerer, R. Zimmermann, "HxL3: Optimized Delivery Architecture for HTTP Low-Latency Live Streaming," IEEE Transactions on Multimedia, vol. 25, pp. 2585-2600, 2023

Today, the widespread use of video applications has led to a mature technological ecosystem which enables real-time communication and the distribution of prerecorded or live videos over the Internet using the HTTP Adaptive Streaming (HAS) paradigm. In this context, standardization has been a key element which has driven the rapid growth of video-based services. As a matter of fact, two independent standards are dominating the video distribution landscape, the MPEG-Dynamic Adaptive Streaming over HTTP (MPEG-DASH) [1] and the HTTP Live Streaming (HLS) [2] proposed by Apple.

Despite the maturity of those technologies, distributing videos at scale is still a challenging research issue which essentially entails finding a balance between service costs and user's perceived quality to provide a sustainable service. Video delivery systems require satisfying the ever-increasing user's expectations fundamentally in terms of visual quality and playback continuity in a changing scenario where contents become more and more resource hungry such as in the case of new immersive formats (360° videos and 6DoF videos).

The popularity of live video streaming services is increasing due to two trends: (i) the increasing number of users shifting from classical broadcast services to Web TV services and (ii) the widespread use of services such as f.i., YouTube, Twitch, Facebook, allowing users to produce and stream high-quality live videos directly from their devices.

Besides guaranteeing a seamless experience characterized by playback without stalls and high visual quality, live video streaming services should also strive to improve interactivity and thus attract users [3]. To this purpose, it is important to decrease to a few seconds the *glass-to-glass latency*, i.e., the time elapsed between source video capture and player video playback. Such issues, that are specific to *low latency live* (L3) streaming, can be tackled considering the MPEG's Common Media Application Format (CMAF) standard [4] and the HTTP/1.1 Chunked Transfer Encoding (CTE) [5].

In this paper, authors propose *HTTP/x-based Low-Latency Live* (HxL3), an end-to-end HAS video distribution architecture for L3 streaming. In a nutshell, HxL3 is a protocol-agnostic architecture designed to work with several codecs, streaming formats, and transport protocols. To provide an L3 service to users, authors propose to employ an architecture which makes use of Virtual Reverse Proxies (VRP) at the origin (near the video producer) and at the edge (near the user). At the video contribution side, the live source is captured and uploaded to an ABR encoder which publishes the encoded segments to the VRP origin which serves as the origin of the CDN that is responsible for distributing the video to the user. At the video distribution side, the VRP at the edge is designed to allow caching, prefetching, and transcoding segments that are pulled from the CDN. The user's HAS player then requests video segments from the VRP edge.

The main contribution of the paper resides in the design of the edge VRP. This entity allows aggregating the player requests based on the user's requested live channel and delivers the requested video segments to the viewers at the same time to ensure fair latency and synchronized playback. The edge VRP also supports transcoding a higher quality bitrate segment to the requested quality to avoid downloading such segment from the origin in case of cache misses. Another possible action, in the case of cache miss, is to serve a given request performance. The results suggest that the to a lower quality that is available in the edge VRP cache. Clearly, all those choices entail different trade-offs in terms of achieving good visual quality, avoiding rebuffering, and obtaining a low glass-to-glass

delay. To this purpose, authors propose a Binary Integer Linear Programming (BILP) problem whose cost function is a combination of several terms which consider: 1) the computational and latency cost in case a segment has to be transcoded at the edge VRP, 2) the bandwidth consumed for receiving segments from the origin VRP, 3) a penalty which occurs when serving a segment at a lower bitrate instead of the requested one. Next, the BILP problem is solved by a greedy-based heuristic algorithm.

Authors implemented the HxL3 architecture employing for the client the popular dash.js player, FFMpeg to encode videos, whereas the origin and edge VRP were implemented in python. An extensive experimental evaluation has been conducted to demonstrate the performance of the proposed HxL3 architecture in several scenarios. In particular, authors implemented both a solution which employs UDP (HxL3-UDP) at the transport layer and another one using the TCP. Different segment sizes ranging from 0.5 to 5 seconds were considered and experiments were run over the public Internet. Results show that HxL3-UDP can obtain average latencies below 5 seconds when the segment duration is 0.5 seconds.

References:

[1] SO/IEC, 23009-1:2014, "Information Technology–Dynamic Adaptive Streaming Over HTTP (DASH)—Part 1: Media Presentation Description and Segment Formats", Nov. 2017, [online] Available: https://www.iso.org/standard/65274.html.

[2] HTTP Live Streaming, Nov. 2018, [online] Available: https://developer.apple.com/streaming/

[3] L. Federico, and A. Mathur, "Under the hood: Broadcasting live video to millions," Facebook Code 2015. [Online]. Available: https://engineering.fb.com/networking-traffic/under-the-hoodbroadcasting-live-video-to-millions/

[4] ISO/IEC, "23000-19:2020 information technology - multimedia application format (MPEG-A) - Part 19: Common media application format (CMAF) for segmented media."

[5] R. Fielding, J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, June 2014.

Luca De Cicco, Ph.D, is an Associate Professor at the Polytechnic University of Bari, Italy. He received M.S. degree in Computer Science Engineering and Ph.D degree both from the Polytechnic University of Bari in 2003 and 2008, respectively. His research interests include massive video distribution over the Internet, Adaptive Bitrate algorithm design, congestion control for real-time flows (WebRTC), QoE-fair resource allocation. He is an Associate Editor of the Internet Technology Letters (Wiley) and General Chair of the ACM Multimedia Systems 2024 conference (ACM MMSys). He has covered several visiting researcher positions: at University of New Mexico (Albuquerque, NM, USA) and at L2S (Laboratoire des signaux et systèmes) École Supérieure d'Électricité (Supelec), Gif-sur-Yvette (Paris), France; at the Laboratory of Information, Networking and Communication Sciences (LINCS).

Secure and Efficiently Searchable IoT Communication Data Management Model: Using Blockchain as a New Tool

A short review for "Secure and Efficiently Searchable IoT Communication Data Management Model: Using Blockchain as a New Tool"

Edited by Qichao Xu

H. Zhang, X. Zhang, Z. Guo, H. Wang, D. Cui and Q. Wen, "Secure and Efficiently Searchable IoT Communication Data Management Model: Using Blockchain as a New Tool," in IEEE Internet of Things Journal, vol. 10, no. 14, pp. 11985-11999, 15 July15, 2023, doi: 10.1109/JIOT.2021.3121482.

With the rapid development of the Internet of Things (IoT) technology, devices under the IoT realize a finer division of labor and more frequent collaboration, which makes the IoT technology more deeply applied in smart life. Different devices under the same network fulfill the whole task assigned to them by sharing data or switching commands. As a result, the behavior of a single IoT device may be part of a joint behavior resulting from the overall task, making it difficult for IoT users to locate and fix faults when certain failures occur. Currently, the most direct and feasible way to solve the above problems is to achieve a reliable data management system to record the communication information between IoT devices to facilitate rapid fault localization and recovery. Meanwhile, in view of the large number of IoT devices and different device models, the data management system should meet the six major requirements such as large storable capacity, high scalability, high reliability, confidentiality, efficient search and economy [1]. Since cloud servers have sufficient data storage capacity and powerful computing capability, the mainstream solution is to meet these requirements by uploading encrypted communication data files to the public cloud [2].

However, the method of dependent on a single central cloud server to store communication data faces the risk of a single point of failure [3]. Moreover, cloud servers and user devices do not fully trust each other, and in case of some data incidents, the responsibility for the incidents cannot be quickly determined. On one hand, the cloud server may lose the outsourced data [4]. On the other hand, IoT devices may not upload all data to the cloud. Of course, employing multiple cloud servers as data management centers can solve the above problems, but it will add more operational costs.

Therefore, the author's main contribution is to propose an IoT communication data management model (ICDM-BC) based on blockchain, in which the cloud provides off-chain storage for IoT data and the blockchain stores communication logs. To realize the efficient search of time sensitive IoT communication data, two layers of indexes are designed for the first time. The first layer is the "Late First Tree", which can efficiently locate the most recently generated blocks, and can also be built on other long-tailed databases. The second layer is the kd-tree, encrypted using the scalar-product preserving encryption (ASPE) algorithm, on which the encrypted half space range query (EhQ) approach can be used to efficiently find encrypted IoT communication logs.

The blockchain network contains three types of entities, IoT devices, the cloud server and users, with some idle devices playing the role of miners. Multiple devices will communicate with each other to switch commands or share data. Devices upload encrypted data with added signatures to the cloud server, and the uploaded data includes communication logs, index construction of communication logs, and signatures. Upon receiving a new broadcast message, the miner generates a new block, which is added to the main chain when half of the devices agree with its legitimacy. In addition, a part of the devices is responsible for handling the user's search request and they share with the user the key used to generate the trapdoor. They help the user to perform two queries [5]: The first one finds the matching block ID based on the timestamp or interval sent by the user. The second one finds the

matching communication log based on the trapdoor sent by the user. Once the cloud server receives the user's search request (i.e., the unique ID that identifies the encrypted communication data file), it sends the corresponding encrypted file and the associated signature to the user.

To realize high search efficiency, this paper proposes a two-layer index for the two search processes. The first layer is a Late First Tree (LF-tree) for locating blocks, which can quickly locate the blocks generated later. The second level index is an encrypted kd-tree for searching the communication logs. kd-tree is binary search tree in high-dimensional space, where each tree node of a kd-tree represents a hyper-rectangle, where the root node represents the entire space, and the leaf nodes represent the "smallest" hyper-rectangles that do not need to be partitioned. Each non-leaf node has two children, which represent two partitioned subspaces.

Extensive simulations have evaluated the time and space efficiency of the proposed scheme. Simulation experiments show an approximately linear increase in the construction time of the LF-tree as the number of blocks increases. The time cost of constructing the kd-tree in explicit form increases linearly with the increase of the number of communication logs, but when the number of communication logs is fixed, different spatial dimensions have almost no effect on the construction time. Meanwhile, the localization efficiencies of LF-tree and balanced binary tree (BB-tree) are compared, and the simulation shows that the block localization efficiency of LF-tree is better than that of BB-tree under the same conditions.

In summary, this paper proposes an IoT data management model based on blockchain technology to solve the single-point-of-failure problem. First, a secure range query scheme for encrypted communication logs stored in a blockchain database is proposed. A two-layer index is designed to improve the search efficiency. The first-layer index achieves higher efficiency in locating blocks that are closer in time, and the second-layer index supports efficient security range queries based on the EhQ method. Second, a hash algorithm like the Merkle tree root is utilized to compute the hash value of the root node of the second index, whose corresponding signature can provide verifiability for the search results. Finally, experiments on synthetic datasets demonstrate the time and space efficiency of the scheme.

References:

[1] K. V. Sowmya, V. Teju, and T. P. Kumar, "An extensive survey on IoT protocols and applications," in Proc. Int. Conf. Intell. Smart Comput. Data Anal., 2021, pp. 131–138.

[2] M. Gabel and J. Mechler, "Secure database outsourcing to the cloud: Side-channels, counter-measures and trusted execution," in Proc. IEEE Int. Symp. Comput. Based Med. Syst., 2017, pp. 799–804.

[3] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, "Blockchain for IoT security and privacy: The case study of a smart home," in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops), 2017, pp. 618–623.

[4] L. Rao, H. Zhang, and T. Tu, "Dynamic outsourced auditing services for cloud storage based on batch-leaves-authenticated merkle hash tree," IEEE Trans. Services Comput., vol. 13, no. 3, pp. 451–463, May/Jun. 2020.

[5] Z. Sun, Y. Wang, Z. Cai, T. Lin, X. Tong, and N. Jiang, "A twostage privacy protection mechanism based on blockchain in mobile crowdsourcing," Int. J. Intell. Syst., vol. 36, no. 5, pp. 2058–2080, 2021.

Qichao Xu, Ph.D, is an Assistant Professor in the School of Mechatronic Engineering and Automation, Shanghai University. He received the Ph.D. degree from Shanghai University, Shanghai, China, in 2019.

His research interests include Internet of Things, autonomous driving vehicles, and trust management. He has published more than 50 papers in prestigious journals such as IEEE TIFS, IEEE TMM, IEEE TITIS, IEEE TII, IEEE TVT, IEEE TBD and IEEE IoTs, in prestigious conferences such as IEEE ICC, IEEE INFOCOM.

Multimedia Communications Technical Committee Officers

MMTC examines systems, applications, services and techniques in which two or more media are used in the same session. These media include, but are not restricted to, voice, video, image, music, data, and executable code. The scope of the committee includes conversational, presentational, and transactional applications and the underlying networking systems to support them.