Joongheon Kim's Selected Publications

Associate Professor
Korea University – School of Electrical Engineering, Seoul, Republic of Korea
https://joongheon.github.io/
joongheon@korea.ac.kr

August 22, 2021

BOOKS AND BOOK CHAPTERS

Book Chapters

[IntechOpen'20] Chapter 6. Dynamic Decision-Making for Stabilized Deep Learning Software Platforms,
Soohyun Park, Dohyun Kim, Joongheon Kim,
Advances and Applications in Deep Learning, IntechOpen, September 2020. (Editor: Marco A. Aceves-Fernandez)

[Wiley'17] Chapter 9. Device-to-Device Communications,
Andreas F. Molisch, Mingyue Ji, Joongheon Kim, Daoud Burghal, Arash Saber Tehrani,
Towards 5G: Applications, Requirements and Candidate Technologies, Wiley, January 2017. (Editors: Rath Vannithamby and Shilpa Talwar)

[CRC'16] Chapter 19. Millimeter-Wave (mmWave) Medium Access Control: A Survey,
Joongheon Kim,
Opportunities in 5G Networks: A Research and Development Perspective, CRC Press Taylor & Francis Group, April 2016. (Editor: Fei Hu)

[CRC'16] Chapter 17. Millimeter-Wave (mmWave) Radio Propagation Characteristics,
Joongheon Kim,
Opportunities in 5G Networks: A Research and Development Perspective, CRC Press Taylor & Francis Group, April 2016. (Editor: Fei Hu)

[CRC'12] Chapter 22. Weighted Localized Clustering: A Coverage-Aware Reader Collision Arbitration Protocol in RFID Networks,
Joongheon Kim, Eunkyo Kim, Wonjun Lee†, Dongshin Kim, Jihoon Choi, Jaewon Jung, Christian K. Shin,
Handbook on Mobile and Ubiquitous Computing: Status and Perspective, CRC Press Taylor & Francis Group, October 2012. (Editors: Laurence T. Yang, Evi Syukur, and Seng W. Loke)

[McGraw'08] Chapter 2.5.4.1. Coverage-Time Optimized Dynamic Clustering for Two-Tiered WM2Nets,
Joongheon Kim, Wonjun Lee†, Eunkyo Kim, Timothy K. Shih,
Wireless Mesh Networking, McGraw-Hill, August 2008. (Editor: George Aggelou)

MAGAZINES AND JOURNALS

Magazines and Guest Editorials

[PIEEE'21.05] Communication-Efficient and Distributed Learning Over Wireless Networks: Principles and Applications,
Jihong Park, Sumudu Samarakoon, Anis Elgabli, Joongheon Kim, Mehdi Bennis, Seong-Lyun Kim, Mérouane Debbah,
Proceedings of the IEEE, 109(5):796–819, May 2021.
(Google Scholar Citations: 26+) *Machine learning (ML) is a promising enabler for the fifth-generation (5G) communication systems and beyond. By imbuing intelligence into the network edge, edge nodes can proactively carry out decision-making and, thereby, react to local environmental changes and disturbances while experiencing zero communication latency. To achieve this goal, it is essential to cater for high ML inference accuracy at scale under the time-varying channel and network dynamics, by continuously exchanging fresh data and ML model updates in a distributed way. Taming this new kind of data traffic boils down to improving the communication efficiency of distributed learning by optimizing communication payload types, transmission techniques, and scheduling, as well as ML architectures, algorithms, and data processing methods. To this end, this article aims to provide a holistic overview of relevant communication and ML principles and, thereby, present communication-efficient and distributed learning frameworks with selected use cases.*

[CM'19.03] New Challenges of Wireless Power Transfer and Secured Billing for Internet of Electric Vehicles,
Laihyuk Park, Seohyeon Jeong, Demeke Shumeye Lakew, Joongheon Kim, Sungrae Cho
IEEE Communications Magazine, 57(3):118–124, March 2019.
(Google Scholar Citations: 10+) *Smart grid modernization and integration of Internet of electric vehicles (IoEVs) have attracted much attention due to their processing capability and convenience. In addition, the developments of magnetic resonance or inductive-based wireless power transfer (WPT) technologies can improve the efficiency of power transmission and user convenience. In particular, the use of renewable energy resources (RERs) is very attractive for low-carbon clean energy, and thus we can imagine integration services that utilize RERs and IoEVs. Accordingly, energy trading markets among IoEVs and power stations will appear in future smart grids. Thus, this article proposes an auction mechanism for energy trading markets of IoEV and RER power stations. Moreover, we can imagine a convenient system where IoEVs can automatically be charged from the WPT charging pad when they are parked in an IoEV parking lot. However, to realize such a convenient system, there are several problems that need to be resolved. Suppose that many IoEVs are parked in the parking lot; then due to power system constraints, it is impossible to simultaneously charge all IoEVs at maximum power. Therefore, an efficient load scheduling algorithm that allows several electric vehicle service equipments (EVSEs) to charge many IoEVs is required. Furthermore, when IoEVs are automatically charged in the anticipated convenient system, it is particularly challenging to design a suitable and secure billing system. In this article, we introduce the advantages of a convenient system that integrates IoEVs and a smart grid. We also discuss the challenges and solutions for future charging and billing systems.*

[VTM'17.03] The Useful Impact of Carrier Aggregation: A Measurement Study in South Korea for Commercial LTE-Advanced Networks,
Sangmin Lee, Seungheon Hyeon, Joongheon Kim, Heejun Roh, Wonjun Lee,
IEEE Vehicular Technology Magazine, 12(1):55–62, March 2017.
(Google Scholar Citations: 16+) *Carrier aggregation (CA) is one of the main features of the long-term evolution (LTE)-Advanced network that was introduced in Third-Generation Partnership Project (3GPP) Release 10 (Rel-10). CA was applied to commercial cellular networks in South Korea in the middle of 2013; however, the performance of CA in commercial networks has not yet been well studied. This article describes how CA technology is applied to a commercial network and how it performs. An extensive field drive test was conducted to compare CA technology performance with that of non-CA technology by using commercial evolved node B (eNB) and user equipment (UE) in a dense urban area and a suburban area. Downlink (DL) CA with two component carriers (CCs) of 30-MHz aggregated bandwidth (BW) was used in the network with one CC of 20-MHz BW at Band 7 and the other with 10-MHz BW at Band 5. The measurement results verified that the maximum DL data rate of CA reached 203 Mb/s, close to the theoretical peak bit rate of 225 Mb/s, and the average DL data rate was 76 Mb/s during the suburban drive test. As a comparison, the maximum DL data rate of single-carrier Band 7 was 141 Mb/s, and the average DL data rate was 51 Mb/s in the same area.*

IEEE Communications Society (Top-Tiers: JSAC/TMC/TON/TWC)

[TMC.accept] Supremo: Cloud-Assisted Low-Latency Super-Resolution in Mobile Devices,
Juheon Yi, Seongwon Kim, Joongheon Kim, Sunghyun Choi,
IEEE Transactions on Mobile Computing, v(n):ppp–ppp, Month Year.
(Google Scholar Citations: 6+) *We present Supremo, a cloud-assisted system for low-latency image Super-Resolution (SR) in mobile devices. As SR is extremely compute-intensive, we first further optimize state-of-the-art DNN to reduce the inference latency. Furthermore, we design a mobile-cloud cooperative execution pipeline composed of specialized data compression algorithms to minimize end-to-end latency with minimal image quality degradation. Finally, we extend Supremo to video applications by formulating a dynamic optimal control algorithm to design Supremo-Opt, which aims to maximize the impact of SR while satisfying latency and resource constraints under practical network conditions. Supremo upscales 360p image to 1080p in 122 ms, which is 43.68 x faster than on-device GPU execution. Compared to cloud offloading-based solutions, Supremo reduces wireless network bandwidth consumption and end-to-end latency by 15.23 x and 4.85 x compared to baseline approach of sending and receiving whole images, and achieves 2.39 dB higher PSNR compared to using conventional JPEG to achieve similar data size compression. Furthermore, Supremo-Opt guarantees robust performance in practical scenarios.*

[TMC'21.06] A Personalized Preference Learning Framework for Caching in Mobile Networks,
Adeel Malik, Joongheon Kim, Kwang Soon Kim, Won-Yong Shin,
IEEE Transactions on Mobile Computing, 20(6):2124–2139, June 2021.
(Google Scholar Citations: 8+) *This paper comprehensively studies a content-centric mobile network based on a preference learning framework, where each mobile user is equipped with a finite-size cache. We consider a practical scenario where each user requests a content file according to its own preferences, which is motivated by the existence of heterogeneity in file preferences among different users. Under our model, we consider a single-hop-based device-to-device (D2D) content delivery protocol and characterize the average hit ratio for the following two file preference cases: the personalized file preferences and the common file preferences. By assuming that the model parameters such as user activity levels, user file preferences, and file popularity are unknown and thus need to be inferred, we present a collaborative filtering (CF)-based approach to learn these parameters. Then, we reformulate the hit ratio maximization problems into a submodular function maximization and propose two computationally efficient algorithms including a greedy approach to efficiently solve the cache allocation problems. We analyze the computational complexity of each algorithm. Moreover, we analyze the corresponding level of the approximation that our greedy algorithm can achieve compared to the optimal solution. Using a real-world dataset, we demonstrate that the proposed framework employing the personalized file preferences brings substantial gains over its counterpart for various system parameters.*

[TWC'21.04] Probabilistic Caching and Dynamic Delivery Policies for Categorized Contents and Consecutive User Demands,
Minseok Choi, Andreas F. Molisch, Dong-Jun Han, Dongjae Kim, Joongheon Kim, Jaekyun Moon,
IEEE Transactions on Wireless Communications, 20(4):2685–2699, April 2021.
(Google Scholar Citations: 4+) *Wireless caching networks have been extensively researched as a promising technique for supporting the massive data traffic of multimedia services. Many of the existing studies on real-data traffic have shown that users of a multimedia service consecutively request multiple contents and this sequence is strongly dependent on the related list of the first content and/or the top referrer in the category. This paper thus introduces the notion of "temporary preference", characterizing the behavior of users who are highly likely to request the next content from a certain target category (i.e., related content list). Based on this observation, this paper proposes both probabilistic caching and dynamic delivery policies for categorized contents and consecutive user demands. The proposed caching scheme maximizes the minimum of the cache hit rates for all users. In the delivery phase, a dynamic helper association policy for receiving multiple contents in a row is designed to reduce the delivery latency. By comparing with the content placement optimized for one-shot requests, numerical results verify the effects of categorized contents and consecutive user demands on the proposed caching and delivery policies.*

[TWC'20.12] Joint Distributed Link Scheduling and Power Allocation for Content Delivery in Wireless Caching Networks,
Minseok Choi, Andreas F. Molisch, Joongheon Kim,
IEEE Transactions on Wireless Communications, 19(11):7810–7824, December 2020.
(Google Scholar Citations: 9+) *In wireless caching networks, the design of the content delivery method must consider random user requests, caching states, network topology, and interference management. In this article, we establish a general framework for content delivery in wireless caching networks without stringent assumptions that restrict the network structure and interference model. Based on the framework, we propose a dynamic and distributed link scheduling and power allocation scheme for content delivery that is assisted by belief-propagation (BP) algorithms. The proposed scheme achieves three critical purposes of wireless caching networks: 1) limiting the delay of user request satisfactions, 2) maintaining the power efficiency of caching nodes, and 3) managing interference among users. In addition, we address the intrinsic problem of the BP algorithm in our network model, proposing a matching algorithm for one-to-one link scheduling. Simulation results show that the proposed scheme provides almost the same delay performance as the optimal scheme found through an exhaustive search at the expense of a little additional power consumption and does not require a clustering method and orthogonal resources in a large-scale D2D network.*

[TWC'19.12] Markov Decision Policies for Dynamic Video Delivery in Wireless Caching Networks,
Minseok Choi, Albert No, Mingyue Ji, Joongheon Kim,
IEEE Transactions on Wireless Communications, 18(12):5705–5718, December 2019.
(Google Scholar Citations: 18+) *This paper proposes a video delivery strategy for dynamic streaming services which maximizes time-average streaming quality under a playback delay constraint in wireless caching networks. The network where popular videos encoded by scalable video coding are already stored in randomly distributed caching nodes is considered under adaptive video streaming concepts, and distance-based interference management is investigated in this paper. In this network model, a streaming user makes delay-constrained decisions depending on stochastic network states: 1) caching node for video delivery, 2) video quality, and 3) the quantity of video chunks to receive. Since wireless link activation for video delivery may introduce delays, different timescales for updating caching node association, video quality adaptation, and chunk amounts*

are considered. After associating with a caching node for video delivery, the streaming user chooses combinations of quality and chunk amounts in the small timescale. The dynamic decision making process for video quality and chunk amounts at each slot is modeled using Markov decision process, and the caching node decision is made based on the framework of Lyapunov optimization. Our intensive simulations verify that the proposed video delivery algorithm works reliably and also can control the tradeoff between video quality and playback latency.

[TWC'19.10] Dynamic Power Allocation and User Scheduling for Power-Efficient and Delay-Constrained Multiple Access Networks,
Minseok Choi, Joongheon Kim, Jaekyun Moon,
IEEE Transactions on Wireless Communications, 18(10):4846–4858, October 2019.
(Google Scholar Citations: 14+) *In this paper, we propose a joint dynamic power control and user pairing algorithm for power-efficient and delay-constrained hybrid multiple access systems. In a hybrid multiple access system, user pairing determines whether the transmitter serves as a certain user by orthogonal multiple access (OMA) or non-orthogonal multiple access (NOMA). The proposed optimization framework minimizes the long-term time-average transmit power expenditure while reducing the queuing delay and guaranteeing the minimum time-average data rates. The proposed technique observes both channel and queue state information and adjusts queue backlogs to avoid an excessive queueing delay by appropriate user pairing and power allocation. Furthermore, the flexible use of resources is captured in the proposed algorithm by employing NOMA. The data-intensive simulation results show that the proposed scheme for power allocation and user scheduling achieves a balance among multiple performance goals, i.e., power efficiency, queueing delay, and data rate.*

[TMC'19.07] Seamless Dynamic Adaptive Streaming in LTE/Wi-Fi Integrated Network under Smartphone Resource Constraints,
Jonghoe Koo, Juheon Yi, Joongheon Kim, Mohammad Ashraful Hoque, Sunghyun Choi,
IEEE Transactions on Mobile Computing, 18(7):1647–1660, July 2019.
(Google Scholar Citations: 17+) *Exploiting both LTE and Wi-Fi links simultaneously enhances the performance of video streaming services in a smartphone. However, it is challenging to achieve seamless and high quality video while saving battery energy and LTE data usage to prolong the usage time of a smartphone. In this paper, we propose REQUEST, a video chunk request policy for Dynamic Adaptive Streaming over HTTP (DASH) in a smartphone, which can utilize both LTE and Wi-Fi. REQUEST enables seamless DASH video streaming with near optimal video quality under given budgets of battery energy and LTE data usage. Through extensive simulation and measurement in a real environment, we demonstrate that REQUEST significantly outperforms other existing schemes in terms of average video bitrate, rebuffering, and resource waste.*

[JSAC'18.11] SGCO: Stabilized Green Crosshaul Orchestration for Dense IoT Offloading Services,
Nhu-Ngoc Dao, Duc-Nghia Vu, Woongsoo Na, Joongheon Kim, Sungrae Cho,
IEEE Journal on Selected Areas in Communications, 36(11):2538–2548, November 2018.
(S.I. on Emerging Technologies in Tactile Internet and Backhaul/Fronthaul Networks)
(Google Scholar Citations: 19+) *The next-generation mobile network anticipates integrated heterogeneous fronthaul and backhaul technologies referred to as a unified crosshaul architecture. The crosshaul enables a flexible and cost-efficient infrastructure for handling mobile data tsunami from dense Internet of things (IoT). However, stabilization, energy efficiency, and latency have not been jointly considered in the optimization of crosshaul performance. To overcome these issues, we propose an orchestration scheme referred to as the stabilized green crosshaul orchestration (SGCO). SGCO utilizes a Lyapunov-theory-based drift-plus-penalty policy to determine the optimal amount of offloaded data that should be processed either at the eastbound or westbound computing platforms to minimize energy consumption. To achieve system stability, the cache buffer is considered as the main constraint in developing the optimization process. Moreover, the amount of offloaded data transmitted via crosshaul links is selected by adopting the binary min-knapsack problem. Accordingly, a lightweight heuristic algorithm is proposed. As the cache buffer is stabilized and the computations are controlled, the SGCO ensures adjustable computing latency threshold for various IoT services. The performance analysis shows that the proposed SGCO scheme exposes effective energy consumption compared to other existing schemes while maintaining system stability considering latency.*

[JSAC'18.06] Wireless Video Caching and Dynamic Streaming under Differentiated Quality Requirements,
Minseok Choi, Joongheon Kim, Jaekyun Moon,
IEEE Journal on Selected Areas in Communications, 36(6):1245–1257, June 2018.
(S.I. on Caching for Communication Systems and Networks)
(Google Scholar Citations: 41+) *This paper considers one-hop device-to-device-assisted wireless caching networks that cache video files of varying quality levels, with the assumption that the base station can control the video quality but cache-enabled devices cannot. Two problems arise in such a caching network: file placement problem and node association problem. This paper suggests a method to cache videos of different qualities, and thus of varying file sizes, by maximizing the sum of video quality measures that users can enjoy. There exists an interesting tradeoff between video quality and video diversity, i.e., the ability to provision diverse video files. By caching high-quality files, the cache-enabled devices can provide high-quality video, but cannot cache a variety of files. Conversely, when the device caches various files, it cannot provide a good quality for file-requesting users. In addition, when multiple devices cache the same file but their qualities are different, advanced node association is required for file delivery. This paper proposes a node association algorithm that maximizes time-averaged video quality for multiple users under a playback delay constraint. In this algorithm, we also consider request collision, the situation where several users request files from the same device at the same time, and we propose two ways to cope with the collision: scheduling of one user and non-orthogonal multiple access. Simulation results verify that the proposed caching method and the node association algorithm work reliably.*

[TON'16.08] Quality-Aware Streaming and Scheduling for Device-to-Device Video Delivery,
Joongheon Kim, Giuseppe Caire, Andreas F. Molisch,
IEEE/ACM Transactions on Networking, 24(4):2319–2331, August 2016.
(Best Reading Papers in Device-to-Device Communications by IEEE Communications Society)
(Google Scholar Citations: 116+) *On-demand video streaming is becoming a killer application for wireless networks. Recent information-theoretic results have shown that a combination of caching on the users' devices and device-to-device (D2D) communications yields throughput scalability for very dense networks, which represent critical bottlenecks for conventional*

cellular and wireless local area network (WLAN) technologies. In this paper, we consider the implementation of such caching D2D systems where each device pre-caches a subset of video files from a library, and users requesting a file that is not already in their library obtain it from neighboring devices through D2D communication. We develop centralized and distributed algorithms for the delivery phase, encompassing a link scheduling and a streaming component. The centralized scheduling is based on the max-weighted independent set (MWIS) principle and uses message-passing to determine max-weight independent sets. The distributed scheduling is based on a variant of the FlashLinQ link scheduling algorithm, enhanced by introducing video-streaming specific weights. In both cases, the streaming component is based on a quality-aware stochastic optimization approach, reminiscent of current Dynamic Adaptive Streaming over HTTP (DASH) technology, for which users sequentially request video "chunks" by choosing adaptively their quality level. The streaming and the scheduling components are coupled by the length of the users' request queues. Through extensive system simulation, the proposed approaches are shown to provide sizeable gains with respect to baseline schemes formed by the concatenation of off-the-shelf FlashLinQ with proportional fair link scheduling and DASH at the application layer.

IEEE Communications Society (Others)

[JCN'21.04] Stabilized Adaptive Sampling Control for Reliable Real-Time Learning-based Surveillance Systems,
Dohyun Kim, Soohyun Park, Joongheon Kim, Jae Young Bang, Soyi Jung,
IEEE/KICS Journal of Communications and Networks, 23(2):129–137, April 2021.
In modern security systems such as CCTV-based surveillance applications, real-time deep-learning based computer vision algorithms are actively utilized for always-on automated execution. The real-time computer vision system for surveillance applications is highly computation-intensive and exhausts computation resources when it performed on the device with a limited amount of resources. Based on the nature of Internet-of-Things networks, the device is connected to main computing platforms with offloading techniques. In addition, the real-time computer vision system such as the CCTV system with image recognition functionality performs better when arrival images are sampled at a higher rate because it minimizes missing video frame feeds. However, performing it at overwhelmingly high rates exposes the system to the risk of a queue overflow that hampers the reliability of the system. In order to deal with this issue, this paper proposes a novel queue-aware dynamic sampling rate adaptation algorithm that optimizes the sampling rates to maximize the computer vision performance(i.e., recognition ratio) while avoiding queue overflow under the concept of Lyapunov optimization framework. Through extensive system simulations, the proposed approaches are shown to provide remarkable gains.

[JCN'21.04] Dynamic Video Delivery using Deep Reinforcement Learning for Device-to-Device Underlaid Cache-Enabled Internet-of-Vehicle Networks,
Minseok Choi, MyungJae Shin, Joongheon Kim,
IEEE/KICS Journal of Communications and Networks, 23(2):117–128, April 2021.
This paper addresses an Internet-of-vehicle network that utilizes a device-to-device (D2D) underlaid cellular system, where distributed caching at each vehicle is available and the video streaming service is provided via D2D links. Given the spectrum reuse policy, three decisions having different timescales in such a D2D underlaid cache-enabled vehicular network were investigated: 1) The decision on the cache-enabled vehicles for providing contents, 2) power allocation for D2D users, and 3) power allocation for cellular vehicles. Since wireless link activation for video delivery could introduce delays, node association is determined in a larger time scale compared to power allocations. We jointly optimize these delivery decisions by maximizing the average video quality under the constraints on the playback delays of streaming users and the data rate guarantees for cellular vehicles. Depending on the channel and queue states of users, the decision on the cache-enabled vehicle for video delivery is adaptively made based on the frame-based Lyapunov optimization theory by comparing the expected costs of vehicles. For each cache-enabled vehicle, the expected cost is obtained from the stochastic shortest path problem that is solved by deep reinforcement learning without the knowledge of global channel state information. Specifically, the deep deterministic policy gradient (DDPG) algorithm is adopted for dealing with the very large state space, i.e., time-varying channel states. Simulation results verify that the proposed video delivery algorithm achieves all the given goals, i.e., average video quality, smooth playback, and reliable data rates for cellular vehicles.

[IOTJ'20.10] Multiagent DDPG-Based Deep Learning for Smart Ocean Federated Learning IoT Networks,
Dohyun Kwon, Joohyung Jeon, Soohyun Park, Joongheon Kim, Sungrae Cho,
IEEE Internet of Things Journal, 7(10):9895–9903, October 2020.
(S.I. on Internet of Things for Smart Ocean)
(Google Scholar Citations: 19+) *This article proposes a novel multiagent deep reinforcement learning-based algorithm which can realize federated learning (FL) computation with Internet-of-Underwater-Things (IoUT) devices in the ocean environment. According to the fact that underwater networks are relatively not easy to set up reliable links by huge fading compared to wireless free-space air medium, gathering all training data for conducting centralized deep learning training is not easy. Therefore, FL-based distributed deep learning can be a suitable solution for this application. In this IoUT network (IoUT-Net) scenario, the FL system needs to construct a global learning model by aggregating the local model parameters that are obtained from individual IoUT devices. In order to reliably deliver the parameters from IoUT devices to a centralized FL machine, base station like devices are needed. Therefore, a joint cell association and resource allocation (JCARA) method is required and it is designed inspired by multiagent deep deterministic policy gradient (MADDPG) to deal with distributed situations and unexpected time-varying states. The performance evaluation results show that our proposed MADDPG-based algorithm achieves 80% and 41% performance improvements than the standard actor-critic and DDPG, respectively, in terms of the downlink throughput.*

[JCN'20.08] Self-Adaptive Power Control with Deep Reinforcement Learning for Millimeter-Wave Internet-of-Vehicles Video Caching,
Dohyun Kwon, Joongheon Kim, David Mohaisen, Wonjun Lee,
IEEE/KICS Journal of Communications and Networks, 22(4):326–337, August 2020.
Video delivery and caching over the millimeter-wave (mmWave) spectrum is a promising technology for high data rate and efficient frequency utilization in many applications, including distributed vehicular networks. However, due to the short handoff duration, calibrating both optimal power allocation of each base station toward its associated vehicles and cache allocation are challenging for their computational complexity. Heretofore, most video delivery applications were based on on-line or off-line algorithms, and they were limited to compute and optimize high dimensional objectives within low-delay in large scale vehicular networks. On the other hand, deep reinforcement learning is shown for learning such scale of a problem with an optimized policy learning phase. In this paper, we propose deep deterministic policy gradient-based power control of mmWave base station (mBS) and proactive cache allocation toward mBSs in distributed mmWave Internet-of-vehicle (IoV) networks. Simulation results validate the performance of the proposed caching scheme in terms of quality of the provisioned video and playback stall in various scales of IoV networks.

[JCN'20.02] Numerical Approximation of Millimeter-Wave Frequency Sharing between Cellular Systems and Fixed Service Systems,
Sungmin Han, Ji-Woong Choi, Joongheon Kim,
IEEE/KICS Journal of Communications and Networks, 22(1):37–45, February 2020.
This paper presents numerical analysis and simulation results in order to study the impact of interference between fixed

service (FS) systems and 5G cellular networks at 28 GHz, 38 GHz, and 60 GHz millimeter-wave (mmWave) bands. For this study, two different scenarios were considered, i.e., the aggregation of interference from small cells into an FS receiver from (i) base stations (BSs) to their associated user equipment (UE) (downlink); (ii) from UEs to their associated BSs (uplink). The simulation results determined how much interference rejection is required to protect the operation of the FS. This study is essential when using mmWave technologies in cellular networks, to determine whether the newly deployed cellular systems can co-exist with pre-deployed FS systems or not. This paper presents closed-form numerical approximation results with Taylor series approximation along with intensive simulation results. Finally, this paper confirmed that the numerical approximation results were precise, i.e., that here were only marginal differences to the intensive simulation results.

[IOTJ'19.10] Two-Stage IoT Device Scheduling with Dynamic Programming for Energy Internet Systems,
Laihyuk Park, Chunghyun Lee, Joongheon Kim, Aziz Mohaisen, Sungrae Cho,
IEEE Internet of Things Journal, 6(5):8782–8791, October 2019.
With the rapid evolution of electric systems, there has been a significant demand for energy Internet (EI) systems that allow sustainable and environmentally friendly energy management. Several research efforts regarding EI systems have been aimed at providing reliable, efficient, and cost-effective techniques. In this paper, we propose a novel algorithm and system for real-time electricity pricing and scheduling. Our algorithm consists of a two-stage operation. The first stage performs real-time pricing to determine the maximum electricity consumption while the second stage performs Internet of Things (IoT) device scheduling. In the second stage, the optimization framework for scheduling is modeled as a 0-1 Knapsack problem; therefore, the solutions to the optimization problem are computed using a dynamic programming framework. Through intensive simulations with well-defined parameters, it is verified that the proposed scheme provides several features, especially reductions in electricity bills with the appropriate parameter settings.

[IOTJ'18.12] Internet of Things for Smart Manufacturing System: Trust Issues in Resource Allocation,
Seohyeon Jeong, Woongsoo Na, Joongheon Kim, Sungrae Cho,
IEEE Internet of Things Journal, 5(6):4418–4427, December 2018.
(S.I. on Industrial Internet-of-Things for Smart and Sensing Systems: Issues, Trends, and Applications)
(Google Scholar Citations: 44+) *In industrial Internet of Things (IIoT) applications for smart manufacturing system, efficient allocation of the carrier and computing resources is crucial. However, existing resource assignment schemes in smart manufacturing system cannot provide timely provision of resources to the inherently dynamic and bursty user demands. To reflect real-time supply and demand for smart manufacturing resources, several research results on auction-style resource assignments have been introduced; however, security, privacy, and trust computing related issues are not actively discussed in the results. The resources should be assigned to devices according to the system policy, which depends on the information provided by IIoT devices. If there are any resource demanding devices, they can report manipulated malicious information for their own interest to obtain more resources. That is, the smart manufacturing system may be vulnerable due to selfish smart manufacturing devices' behaviors. This reduces the efficiency of the entire system and moreover ceases the plant-wide process. While many research contributions related to the trust computing aim at detecting malicious nodes, this paper presents a novel view of trust computing by showing why devices inside the smart manufacturing system have to act honestly. In this paper, a Vickrey–Clarke–Groves auction-based hierarchical trust computing algorithm is proposed for: 1) computing carrier resources required for wireless communication between IIoT devices and gateways and 2) distributing CPU resources for processing data at central processing controller. Last, simulation results demonstrate that the utilities of each participant are maximized when the IIoT devices and gateways are trustful.*

[IOTJ'18.02] Energy-Efficient Mobile Charging for Wireless Power Transfer in Internet of Things Networks,
Woongsoo Na, Junho Park, Cheol Lee, Kyoungjun Park, Joongheon Kim, Sungrae Cho,
IEEE Internet of Things Journal, 5(1):79–92, February 2018.
(Google Scholar Citations: 59+) *The Internet of Things (IoT) is expected to play an important role in the construction of next generation mobile communication services, and is currently used in various services. However, the power-hungry battery significantly limits the lifetime of IoT devices. Among the various lifetime extension techniques, this paper discusses mobile charging, which enables wireless power transfer based on radio frequency with mobile chargers (MCs). MCs function as traveling target IoT networks that provide energy to battery-operated IoT devices. However, MCs with an energy-constrained battery result in limitation of travel-time. This paper formulates a problem to minimize energy consumption for charging IoT devices by determining the path of motion of an MC and efficient charging points, and proves that the problem is NP-hard. An efficient algorithm, named best charging efficiency (BCE), is proposed to solve the problem and the upper bound of the BCE algorithm is guaranteed using the duality of linear programming. In addition, an improved BCE algorithm called branching second best efficiency algorithm with additional searching techniques is introduced. Finally, this paper analyzes the difference in performance among the proposed algorithms, optimal solutions, and the existing algorithm and concludes that the performance of the proposed algorithm is near optimal, within 1% of difference ratio in terms of charging efficiency and delay.*

[IOTJ'17.10] Feasibility Study of 60 GHz Millimeter-Wave Technologies for Hyperconnected Fog Computing Applications,
Joongheon Kim, Wonjun Lee,
IEEE Internet of Things Journal, 4(5):1165–1173, October 2017.
(S.I. on Fog Computing in the Internet of Things)
This paper conducts the feasibility study for discussing about the interference impacts in advanced fog computing (FC) networks with 60 GHz millimeter-wave wireless technology. With the concept of FC, cloud computing services can be utilized with local computing devices (FC devices) those are located nearby user terminals [edge devices (EDs)] for low-latency communications. Since EDs can be densely deployed in urban areas, wireless hyperconnection should be supportable in FC networks. For the system, 60 GHz millimeter-wave technology can be used as one of wireless networking methods because it can provide: 1) high rates for visual information delivery and 2) high directionality for spatial reuse in hyperconnected FC networks. Therefore, this performance simulation study under the consideration of interference impacts determines whether utilizing 60 GHz millimeter-wave wireless technology for hyperconnected FC networks is feasible or not. Under the consideration of various interference scenarios, the interference impacts are calculated. As presented in simulation-based performance evaluation results, we verify

that 1.5 Gb/s high rate can be supportable even though more than 1000 EDs are deployed in the 100 m-by-100 m size small-scale networks (upper bound). Therefore, we confirm that hyperconnection can be realized with 60 GHz millimeter-wave technology for FC networks.

[JCN'14.10] Fast Millimeter-Wave Beam Training with Receive Beamforming,
Joongheon Kim, Andreas F. Molisch,
IEEE/KICS Journal of Communications and Networks, 16(5):512–522, October 2014.
(Google Scholar Citations: 83+) *This paper proposes fast millimeter-wave (mm-wave) beam training protocols with receive beamforming. Both IEEE standards and the academic literature have generally considered beam training protocols involving exhaustive search over all possible beam directions for both the beamforming initiator and responded However, this operation requires a long time (and thus overhead) when the beamwidth is quite narrow such as for mm-wave beams (1° in the worst case). To alleviate this problem, we propose two types of adaptive beam training protocols for fixed and adaptive modulation, respectively, which take into account the unique propagation characteristics of millimeter waves. For fixed modulation, the proposed protocol allows for interactive beam training, stopping the search when a local maximum of the power angular spectrum is found that is sufficient to support the chosen modulation/coding scheme. We furthermore suggest approaches to prioritize certain directions determined from the propagation geometry, long-term statistics, etc. For adaptive modulation, the proposed protocol uses iterative multi-level beam training concepts for fast link configuration that provide an exhaustive search with significantly lower complexity. Our simulation results verify that the proposed protocol performs better than traditional exhaustive search in terms of the link configuration speed for mobile wireless service applications.*

[CL'14.09] Joint Coding and Stochastic Data Transmission for Uplink Cloud Radio Access Networks,
Song-Nam Hong, Joongheon Kim,
IEEE Communications Letters, 18(9):1619–1622, September 2014.
(Google Scholar Citations: 10+) *We study the uplink cloud radio access networks using quantize and forward as an underlying relaying scheme. In all previous works, each radio unit (RU) is assumed to choose a compression rate equal to an instantaneous backhaul capacity. In this letter, we consider a more general network model for which each RU is equipped with a queue whose arrival rate is equal to a compression rate and departure rate is equal to a current backhaul capacity. For each time slot, it can choose a higher compression rate than the given backhaul capacity as long as the queue is stable. Using the principle of stochastic network optimization, we present a distributed algorithm to find the compression rate of each RU such that a time-averaged sum-rate is maximized subject to the queue stability. Under the optimal compression policy, some of RUs are active with a higher compression rate and the others are silent, for a given time slot, which is completely different from the conventional approach that all RUs are always active.*

[CL'14.07] A Low-Complexity Algorithm for Neighbor Discovery in Wireless Networks,
Song-Nam Hong, Joongheon Kim,
IEEE Communications Letters, 18(7):1119–1122, July 2014.
We study a neighbor discovery problem in wireless networks for which each node wishes to identify the so-called neighboring nodes within a single-hop communication. This problem can be optimally addressed using maximum a posteriori (MAP) estimation, but its implementation is notoriously difficult in practice. In this letter, we present a low-complexity algorithm consisting of two stages: 1) we solve such problem using LASSO estimator that is a convex relaxation of MAP estimator to encourage a sparse solution; and 2) we find a desired binary vector (e.g., indicator of neighbor nodes) by taking a "hard-decision" with threshold, carefully chosen by exploiting fading statistics. Finally, we provide some numerical results to confirm that the proposed algorithm performs quite well.

[CL'14.03] Fast and Low-Power Link Setup for IEEE 802.15.3c Multi-Gigabit/s Wireless Sensor Networks,
Joongheon Kim, Aziz Mohaisen, Jong-Kook Kim,
IEEE Communications Letters, 18(3):455–458, March 2014.
(Google Scholar Citations: 11+) *In this paper, we propose a polynomial-time framework for fast and low-power link setup between two 60 GHz IEEE 802.15.3c wireless sensor devices. Due to the narrow beamwidth of 60 GHz radio beams, the beam training procedure needs to evaluate a lot of search spaces. Thus, the running time to establish an IEEE 802.15.3c radio link is relatively longer than the running time of other schemes. In addition, evaluating a large search space requires a lot of training packets transmission, which increases the communication power consumption. Therefore, our proposed framework reduces the number of training packet transmission (i.e., the number of search space evaluation) using multi-resolution beam training for fast and low-power link setup.*

[CL'07.01] Optimized Transmission Power Control of Interrogators for Collision Arbitration in UHF RFID Systems,
Joongheon Kim, Wonjun Lee, Eunkyo Kim, Dongshin Kim, Kyoungwon Suh,
IEEE Communications Letters, 11(1):22–24, January 2007.
(Google Scholar Citations: 31+) *The emergence of UHF RFID as one of the dominant technology trends has posed numerous unique challenges to researchers. This letter presents a novel, theoretically-grounded collision arbitration protocol, called TPC-CA, which optimally controls transmission power of RFID interrogators and thereby reducing redundant interrogator collisions.*

IEEE Vehicular Technology Society

[TVT'21.08] Infrastructure-Assisted On-Driving Experience Sharing for Millimeter-Wave Connected Vehicles,
Soyi Jung, Joongheon Kim, Marco Levorato, Carlos Cordeiro, Jae-Hyun Kim,
IEEE Transactions on Vehicular Technology, 70(8):7307–7321, August 2021.
(S.S. on Connected Vehicles)
This paper proposes on-driving experience sharing algorithms at junctions in infrastructure-assisted vehicles-to-everything networks. For the purpose, a millimeter-wave (mmWave) technology is used because it provides multi-Gbps data rates which is helpful for handling users' short stay times at junctions and spatial reuse due to high beam directionality which is helpful for interference-avoidance among densely deployed vehicles at junctions. To realize on-driving experience sharing, the proposed algorithms focus on joint resource allocation and scheduling for 3GPP-compliant multiple unicast vehicle-to-vehicle (V2V) communications where the vehicles are group leaders (GLs) in 3GPP Mode 4(d). The resource allocation stands for the roadside unit (RSU) allocation to scheduled V2V GL links where RSU is essentially required for overcoming blockage by establishing two-hop relaying. Because vehicles stay for short times at junctions, this paper designs two algorithms without or with delay considerations. Without delay considerations, the joint optimization of RSU allocation and scheduling was originally formulated as mixed 0-1 non-convex optimization. However our proposed algorithm reformulates the problem into mixed 0-1 convex optimization, which is computationally easier to solve. With delay considerations, our proposed algorithm dynamically controls video contents frame rates for time-average on-driving video sharing quality maximization subject to delay constraints, inspired by Lyapunov optimization. Extensive simulation results demonstrate that our algorithms can significantly outperform in a variety of scenarios. Furthermore, we conduct the cost analysis for the proposed algorithms in terms of capital expenditure (CAPEX) and operating expenditure (OPEX).

[TVT'21.06] Orchestrated Scheduling and Multi-Agent Deep Reinforcement Learning for Cloud-Assisted Multi-UAV Charging Systems,
Soyi Jung, Won Joon Yun, MyungJae Shin, Joongheon Kim, Jae-Hyun Kim,
IEEE Transactions on Vehicular Technology, 70(6):5362–5377, June 2021.
(S.S. on Vehicular Networks in the era of 6G: End-Edge-Cloud Orchestrated Intelligence)
This paper proposes a cloud-assisted joint charging scheduling and energy management framework for unmanned aerial vehicle (UAV) networks. For charging the UAVs those are extremely power hungry, charging towers are considered for plug-and-play charging during run-time operations. The charging towers should be cost-effective, thus it is equipped with photovoltaic power generation and energy storage systems functionalities. Furthermore, the towers should be cooperative for more cost-effectiveness by intelligent energy sharing. Based on the needs and setting, this paper proposes 1) charging scheduling between UAVs and towers and 2) cooperative energy managements among towers. For charging scheduling, the UAVs and towers should be scheduled for maximizing charging energy amounts and the scheduled pairs should determine charging energy allocation amounts. Here, two decisions are correlated, i.e. , it is a non-convex problem. We re-formulate the non-convex to convex for guaranteeing optimal solutions. Lastly, the cooperative energy sharing among towers is designed and implemented with multi-agent deep reinforcement learning and then intelligent energy sharing can be realized. We can observe that the two methods are related and it should be managed, coordinated, and harmonized by a centralized orchestration manager under the consideration of fairness, energy-efficiency, and cost-effectiveness. Our data-intensive performance evaluation verifies that our proposed framework achieves desired performance.

[TVT'19.10] Blind Signal Classification for Non-Orthogonal Multiple Access in Vehicular Networks,
Minseok Choi, Daejung Yoon, Joongheon Kim,
IEEE Transactions on Vehicular Technology, 68(10):9722–9734, October 2019.
In this paper, blind signal classification and detection in a non-orthogonal multiple access (NOMA) system are explored. Since a NOMA scheme superposes the multiple-user signals within nonorthogonal resources, classical modulation classification methods used in orthogonal multiple access (OMA) systems are not sufficient to process the superposed NOMA signal. NOMA receivers require information about the multiple access schemes such as modulation order and need interference cancellation of the co-scheduled user's signal; therefore, a NOMA system causes more high-layer signaling overheads than OMA during packet scheduling. Blind detection algorithms used for multiplexing information are considered to be possible solutions; however, they pose various challenges and could cause performance loss while performing blind modulation classification in the order of OMA/NOMA classification, co-scheduled user's modulation classification, and classification of the signal, due to the necessity for successive interference cancellation. To improve the performance of blind detection, we propose a NOMA transmission scheme that applies phase rotation to data or pilot symbols depending on the NOMA multiplexing format, as an aid to the blind detection. The proposed classification algorithm can implicitly provide essential information on NOMA multiplexing without the need for any extra high layer signaling or resources. The performance improvement is verified through simulation studies, and it is found that the proposed algorithm provides a gain of more than 1 dB compared to the existing blind signal classification methods and shows almost equivalent performance as the genie information scheme.

[TVT'19.05] Auction-Based Charging Scheduling With Deep Learning Framework for Multi-Drone Networks,
MyungJae Shin, Joongheon Kim, Marco Levorato,
IEEE Transactions on Vehicular Technology, 68(5):4235–4248, May 2019.
(S.S. on Machine Learning-based Internet of Vehicle: Theory, Methodology, and Applications)
(Google Scholar Citations: 48+) *State-of-the-art drone technologies have severe flight time limitations due to weight constraints, which inevitably lead to a relatively small amount of available energy. Therefore, frequent battery replacement or recharging is necessary in applications such as delivery, exploration, or support to the wireless infrastructure. Mobile charging stations (i.e., mobile stations with charging equipment) for outdoor ad-hoc battery charging is one of the feasible solutions to address this issue. However, the ability of these platforms to charge the drones is limited in terms of the number and charging time. This paper designs an auction-based mechanism to control the charging schedule in multi-drone setting. In this paper, charging time slots are auctioned, and their assignment is determined by a bidding process. The main challenge in developing this framework is the lack of prior knowledge on the distribution of the number of drones participating in the auction. Based on optimal second-price-auction, the proposed formulation, then, relies on deep learning algorithms to learn*

such distribution online. Numerical results from extensive simulations show that the proposed deep-learning-based approach provides effective battery charging control in multi-drone scenarios.

[TVT'18.04] Adaptive Detector Selection for Queue-Stable Word Error Rate Minimization in Connected Vehicle Receiver Design,
Minseok Choi, Joongheon Kim, Jaekyun Moon,
IEEE Transactions on Vehicular Technology, 67(4):3635–3639, April 2018.
High-speed and low-latency communications are one of the major requirements for connected vehicle environments. Safety messages in vehicular networks must arrive on time for each vehicle, and the optimization of driving paths also strongly depends on low-latency traffic information. To comply with the demands, a base station take important roles to provide a short service delay for the vehicles, even at the expense of performance. The relationship between delay time and decoding performance at receiver sides in physical layers can be explained by the performance/complexity tradeoff as follows. To achieve better performance in terms of the minimization of word error rates in decoding, the receiver usually requires higher complexity, which results in larger delays. We analyze this tradeoff based on the receiver queue model. This paper proposes a novel algorithm to adaptively select one of receiver candidates under the constraint of queue stability so as not to degrade low-latency communications. Based on the configurable architecture for multi-input multioutput (MIMO) detectors, this paper shows that the proposed algorithm works well on adaptive MIMO detector selection.

[TVT'16.12] Performance of Video Streaming in Infrastructure-to-Vehicle Telematic Platforms With 60-GHz Radiation and IEEE 802.11ad Baseband,
Joongheon Kim, Seok-Chul Kwon, Giwan Choi,
IEEE Transactions on Vehicular Technology, 65(12):10111–10115, December 2016.
(Google Scholar Citations: 22+) *This paper proposes feasible and satisfactory system design parameters to mitigate the impact of interference on real-time high-definition video streaming in infrastructure-to-vehicle (I2V) telematic platforms utilizing 60-GHz radiation and the corresponding IEEE 802.11ad baseband. The analysis captures the impact of interference on the 60-GHz I2V vehicle (IV), which is caused by multiple interference sources through the 60-GHz wireless transmissions from nearby I2V base stations (IBSs) to their associated IVs. The impact of the interference on the quality of main 1080p at 30-frame/s (30 1080p frames/s) and 1080p at 60-frame/s (60 1080p frames/s) streaming is analyzed and estimated for various simulation settings.*

IEEE Industrial Electronics Society

[TII'20.05] Cooperative Management for PV/ESS-Enabled Electric-Vehicle Charging Stations: A Multiagent Deep Reinforcement Learning Approach,
MyungJae Shin, Dae-Hyun Choi, Joongheon Kim,
IEEE Transactions on Industrial Informatics, 16(5):3493–3503, May 2020.
(S.S. on Advanced Informatics for Energy Storage Systems in Electrified Vehicles and Smart Grids)
(Google Scholar Citations: 28+) *This article proposes a novel multiagent deep reinforcement learning method for the energy management of distributed electric vehicle charging stations with a solar photovoltaic system and energy storage system. In the literature, the conventional method is to calculate the optimal electric vehicle charging schedule in a centralized manner. However, in general, the centralized approach is not realistic under certain environments where the system operators for multiple electric vehicle charging stations handle dynamically varying data, such as the status of the energy storage system and electric vehicle-related information. Therefore, this article proposes a method that can compute the scheduling solutions of multiple electric vehicle charging stations in a distributed manner while handling run-time time-varying dynamic data. As shown in the data-intensive performance evaluation, it can be observed that the proposed method achieves a desirable performance in terms of reducing the operation costs of electric vehicle charging stations.*

[TIE'19.02] Joint Geometric Unsupervised Learning and Truthful Auction for Local Energy Market,
Laihyuk Park, Seohyeon Jeong, Joongheon Kim, Sungrae Cho,
IEEE Transactions on Industrial Electronics, 66(2):1499–1508, February 2019.
(S.S. on Methods and Systems for a Smart Energy City)
(Google Scholar Citations: 29+) *Development of smart grid technologies has created a promising atmosphere for smart cities and energy trading markets. Especially, traditional electricity consumers evolve into prosumers who produce as well as consume electricity in modern power electric systems. In this evolution, the electric power industry has tried to introduce the notion of local energy markets for prosumers. In the local energy market, prosumers purchase electricity from distributed energy generators or the other prosumers with surplus electricity via a local power exchange center. For this purpose, this paper proposes joint geometric clustering and truthful auction schemes in the local energy markets. The proposed clustering scheme is designed for distribution fairness of the distributed energy generator for serving prosumers, where the scheme is inspired by expectation and maximization based unsupervised learning. Moreover, this paper proposes an auction mechanism for truthful electricity trading in a local energy market. In order to guarantee truthful electricity trading, the proposed auction mechanism is constructed based on the Vickrey-Clarke-Groves auction, which was proven to guarantee truthful operations. The Hungarian method is also considered in addition to the auction. The simulation results for the auction verify that the utilities of local market energy entities are maximized when the prosumers are truthful.*

[TII'17.12] Residential Demand Response for Renewable Energy Resources in Smart Grid Systems,
Laihyuk Park, Yongwoon Jang, Sungrae Cho, Joongheon Kim,
IEEE Transactions on Industrial Informatics, 13(6):3165–3173, December 2017.
(S.S. on Smart Grid and Renewable Energy Resources: Information and Communication Technologies With Industry Perspective)
(Google Scholar Citations: 72+) *With the current state of development in demand response (DR) programs in smart grid systems, there have been great demands for automated energy scheduling for residential customers. Recently, energy scheduling in smart grids have focused on the minimization of electricity bills, the reduction of the peak demand, and the maximization of user convenience. Thus, a user convenience model is proposed under the consideration of user waiting times, which is a nonconvex problem. Therefore, the nonconvex is reformulated as convex to guarantee optimal solutions. Moreover, mathematical formulations for DR optimization are derived based on the reformulated convex problem. In addition, two types of pricing policies for electricity bills are designed in the mathematical formulations, i.e., real-time pricing policy and progressive policy. With real-time pricing policy, convexity is guaranteed whereas progressive policy cannot. Then, heuristic algorithms are finally designed for obtaining approximated optimal solutions in progressive policy.*

[TII'15.12] Energy-Efficient Dynamic Packet Downloading for Medical IoT Platforms,
Joongheon Kim,
IEEE Transactions on Industrial Informatics, 11(6):1653–1659, December 2015.
(S.S. on Energy Efficient Technology in Sensor Networks)
(Google Scholar Citations: 50+) *This paper proposes a polynomial-time algorithm for energy-efficient dynamic packet downloading from medical cloud storage to medical Internet-of-Things (IoT) devices. The medical cloud can distribute its own medical data to medical IoT devices via access points. Therefore, network disconnection can happen between the medical cloud and medical IoT devices when power/energy management in each access point is not efficient. This situation is especially harmful in in-hospital network architectures, because the architecture usually has strict requirements in terms of reliability. Therefore, this paper proposes a dynamic energy-efficient algorithm, which computes the amount of power allocation in each access point based on the buffer backlog size and channel states under the consideration of buffer stability. With the proposed adaptive algorithm, each access point calibrates its own parameters for more adaptive power/energy management. The performance of the proposed algorithm is evaluated in terms of network lifetime, and it is observed that the proposed algorithm achieves the desired performance.*

IEEE Systems Council

[ISJ.accept] Securing Heterogeneous IoT with Intelligent DDoS Attack Behavior Learning,
Nhu-Ngoc Dao, Trung V. Phan, Umar Sa'ad, Joongheon Kim, Thomas Bauschert, Dinh-Thuan Do, Sungrae Cho,
IEEE Systems Journal, (Early Access).
(Google Scholar Citations: 20+) *The rapid increase of diverse Internet of Things (IoT) services and devices has raised numerous challenges in terms of connectivity, interoperability, and security. The heterogeneity of the networks, devices, and services introduces serious vulnerabilities to security, especially distributed denial-of-service (DDoS) attacks, which exploit massive IoT devices to exhaust both network and victim resources. As such, this article proposes FOGshield, which is a localized DDoS prevention framework leveraging the federated computing power of the fog computing-based access networks to deploy multiple smart endpoint defenders at the border of relevant attack-source/destination networks. Cooperation among the smart endpoint defenders is supervised by a central orchestrator. The central orchestrator localizes each smart endpoint defender by feeding appropriate training parameters into its self-organizing map component, based on the attacking behavior. Performance of the FOGshield framework is verified using three typical IoT traffic scenarios. Numerical results reveal that the FOGshield outperforms existing solutions.*

[ISJ.accept] LiteZKP: Lightening Zero-Knowledge Proof-based Blockchains for IoT and Edge Platforms,
EunSeong Boo, Joongheon Kim, JeongGil Ko,
IEEE Systems Journal, (Early Access).
This article presents LiteZKP a framework for supporting multiple anonymous payments using a smart contract-based zero-knowledge proof (ZKP) protocol on resource-limited devices. Specifically, to address challenges related to minimizing the computational overhead and offer a fully anonymous system, LiteZKP includes novel schemes such a new merkle tree mechanism to reduce the burden of ZKP operations, and integrates smart contract-based ZKP with an off-chain payment channel to reduce the amount of ZKP operations when performing continuous data exchange. We present evaluation results of LiteZKP from both PC-scale and resource-limited client devices and our results suggest that LiteZKP reduces the latency and energy consumption by more than 55% on Internet of Things (IoT)/mobile edge computing platforms, while requiring only 8% of block processing fee compared to a naive ZKP-based scheme.

[ISJ'21.09] Intelligent Active Queue Management for Stabilized QoS Guarantees in 5G Mobile Networks,
Soyi Jung, Joongheon Kim, Jae-Hyun Kim,
IEEE Systems Journal, 15(3):ppp–ppp, September 2021.
The 5G new radio standard defines various functions for forwarding treatments under differentiated quality of services requirements. In particular, a gNB base station helps combat overbuffering due to massive packet flows from associated mobile devices. Several active queue management (AQM) schemes have been proposed previously, with controlled delay (CoDel) proving to be an efficient method that can properly handle excessive buffering, that is called bufferbloat. This article proposes a novel and enhanced AQM scheme based on CoDel for gNB that guarantees low queuing delays, minimizes packet drop rate, and furthermore stabilizes the built-in queue. According to data-intensive performance evaluation results, the proposed scheme verifies that the queuing delay for served packets are decreased and supports ultrareliable low latency communication applications. Especially, the proposed algorithm achieves the packet drop rates and queuing delays tradeoff as [O(1/V),O(V)].

[ISJ'21.03] Multiscale LSTM-Based Deep Learning for Very-Short-Term Photovoltaic Power Generation Forecasting in Smart City Energy Management,
Dohyun Kim, Dohyun Kwon, Laihyuk Park, Joongheon Kim, Sungrae Cho,
IEEE Systems Journal, 15(1):346–354, March 2021.
Photovoltaic power generation forecasting (PVGF) is an attractive research topic for efficient energy management in smart city. In addition, the long short-term memory recurrent neural network (LSTM/RNN) has been actively utilized for predicting various time series tasks in recent years due to its outstanding ability to learn the feature of sequential time-series data. Although the existing forecasting models were obtained from learning the sequential PVGF data, it is observed that irregular factors made adverse effects on the forecasting results of very-short-term PVGF tasks, thus, the entire forecasting performance was deteriorated. In this regard, multiscale LSTM-based deep learning which is capable for forecasting very-short-term PVGF is proposed for efficient management. The model concatenates on two different scaled LSTM modules to overcome the deterioration that is originated from the irregular factors. Lastly, experimental results present the proposed framework can assist to forecast the tendency of PVGF amount steadily.

[ISJ'20.03] Towards Characterizing Blockchain-based Cryptocurrencies for Highly-Accurate Predictions,
Muhammad Saad, Jinchun Choi, DaeHun Nyang, Joongheon Kim, Aziz Mohaisen,
IEEE Systems Journal, 14(1):321–332, March 2020.
(IEEE Systems Journal Best Paper Award, Top 7 among 793 accepted papers in 2019: 0.88%)
(Google Scholar Citations: 69+) *Recently, the Blockchain-based cryptocurrency market witnessed enormous growth. Bitcoin, the leading cryptocurrency, reached all-time highs many times over the year leading to speculations to explain the trend in its growth. In this article, we study Bitcoin and Ethereum and explore features in their network that explain their price hikes. We gather data and analyze user and network activity that highly impact the price of these cryptocurrencies. We monitor the change in the activities over time and relate them to economic theories. We identify key network features that help us to determine the demand and supply dynamics in a cryptocurrency. Finally, we use machine learning methods to construct models that predict Bitcoin price. Based on our experimental results using two large datasets for validation, we confirm that our approach provides an accuracy of up to 99% for Bitcoin and Ethereum price prediction in both instances.*

Representative Journals in Other Societies

[T-CAD'19.09] TEI-ULP: Exploiting Body Biasing to Improve the TEI-Aware Ultra-Low Power Methods,
Woojoo Lee, Taewook Kang, Jae-Jin Lee, Kyuseung Han, Joongheon Kim, Massoud Pedram,
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 38(9):1758–1770, September 2019.
Temperature effect inversion (TEI) phenomenon in ultralow power (ULP) very large scale integration circuits has been identified as an important effect by both academia and industry. Although a number of ULP methods that attempt to exploit the TEI phenomenon have been proposed, the small size of the design exploration space when applying these methods to ULP circuits hinders them from achieving their full potential. This is mainly due to the limited granularity of the supply voltage level control. Starting with an intuition that the body biasing (BB) technique is a key to overcome this limitation, this paper exploits the BB technique along with the TEI-aware voltage scaling (TEI-VS) method and TEI-aware frequency scaling (TEI-FS) method, so as to substantially increase the design spaces of these methods. Techniques for optimally combining the BB technique with TEI-VS and TEI-FS are introduced. Simulation results with the latest commercial CMOS process technologies for ULP designs demonstrate the effectiveness of the proposed methodology.

[T-SMC'15.11] Stochastic Decision Making for Adaptive Crowdsourcing in Medical Big-Data Platforms,
Joongheon Kim, Wonjun Lee,
IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45(11):1471–1476, November 2015.
(Google Scholar Citations: 25+) *This paper proposes two novel algorithms for adaptive crowdsourcing in 60-GHz medical imaging big-data platforms, namely, a max-weight scheduling algorithm for medical cloud platforms and a stochastic decision-making algorithm for distributed power-and-latency-aware dynamic buffer management in medical devices. In the first algorithm, medical cloud platforms perform a joint queue-backlog and rate-aware scheduling decisions for matching deployed access points (APs) and medical users where APs are eventually connected to medical clouds. In the second algorithm, each scheduled medical device computes the amounts of power allocation to upload its own medical data to medical big-data clouds with stochastic decision making considering joint energy-efficiency and buffer stability optimization. Through extensive simulations, the proposed algorithms are shown to achieve the desired results.*

[TBC'13.09] Joint Scalable Coding and Routing for 60 GHz Real-Time Live HD Video Streaming Applications,
Joongheon Kim, Yafei Tian, Stefan Mangold, Andreas F. Molisch,
IEEE Transactions on Broadcasting, 59(3):500–512, September 2013.
(Google Scholar Citations: 51+) *Transmission of high-definition (HD) video is a promising application for 60 GHz wireless links, since very high transmission rates (up to several Gbit/s) are possible. In particular we consider a sports stadium broadcasting system where signals from multiple cameras are transmitted to a central location. Due to the high pathloss of 60 GHz radiation over the large distances encountered in this scenario, the use of relays might be required. The current paper analyzes the joint selection of the routes (relays) and the compression rates from the various sources for maximization of the overall video quality. We consider three different scenarios: (i) each source transmits only to one relay and the relay can receive only one data stream, and (ii) each source can transmit only to a single relay, but relays can aggregate streams from different sources and forward to the destination, and (iii) the source can split its data stream into parallel streams, which can be transmitted via different relays to the destination. For each scenario, we derive the mathematical formulations of the optimization problem and re-formulate them as convex mixed-integer programming, which can guarantee optimal solutions. Extensive simulations demonstrate that high-quality transmission is possible for at least ten cameras over distances of 300 m. Furthermore, optimization of the video quality gives results that can significantly outperform algorithms that maximize data rates.*

[TCE'07.11] Movement-Aware Vertical Handoff of WLAN and Mobile WiMAX for Seamless Ubiquitous Access,
Wonjun Lee, Eunkyo Kim, Joongheon Kim, Inkyu Lee, Choonhwa Lee,
IEEE Transactions on Consumer Electronics, 53(4):1268–1275, November 2007.
(Google Scholar Citations: 111+) *This paper addresses a movement-aware vertical (MAV) handover algorithm between WLAN and Mobile WiMAX for seamless ubiquitous access. An MAV handover algorithm is proposed in this paper to exploit movement pattern for avoiding unnecessary handovers in the integrated WLAN and Mobile WiMAX networks. If a mobile station (MS)'s velocity is high and its movement pattern is irregular, unnecessary handovers likely occur more frequently. Therefore, the MS velocity and moving pattern are important factors for the handover decision procedure. To avoid unnecessary handovers, the MAV handover algorithm adjusts the dwell time adaptively and predicts the residual time in the cell of target base station (BS). Consequently, the adaptive dwell timer of MAV handover algorithm allows an MS a better connection as long as possible. Our simulation results show that the reduction of unnecessary handovers by leads to significant throughput improvements.*

[TCE'07.05] Coverage-Time Optimized Dynamic Clustering of Networked Sensors for Pervasive Home Networking,
Joongheon Kim, Wonjun Lee, Eunkyo Kim, Doe-Wan Kim, Hyeokman Kim,
IEEE Transactions on Consumer Electronics, 53(2):433–441, May 2007.
This paper proposes a novel energy-efficient coverage-time optimized dynamic clustering scheme for two- tiered wireless sensor networks (WSNs) used in an outdoor monitoring application of home networking systems. The coverage-time is defined as the time until one of cluster heads (CHs) runs out of energy in clustering-based WSNs, thereby resulting in an incomplete coverage. DC-CTO scheme regulates cluster radii for balanced energy consumption among CHs to maximize coverage-time. There are several advantages of using the DC-CTO scheme. The first advantage is balanced energy consumption among CHs under NLP- based computations. The second advantage is minimized energy consumption in each CH to extend coverage-time. The last one is the guarantee of perfect coverage. If there is no area which is uncontrolled by a CH in a networked sensing field, we call it a perfect coverage. If there is an incomplete coverage, the event in the area cannot be identified. Therefore perfect coverage is important to robust monitoring systems. The novelty of DC-CTO scheme is demonstrated by various simulation-based performance evaluations.

IEEE Access

[Access.accept] Spatio-Temporal Split Learning for Privacy-Preserving Medical Platforms: Case Studies with COVID-19 CT, X-Ray, and Cholesterol Data,
Yoo Jeong Ha, Minjae Yoo, Gusang Lee, Soyi Jung, Sae Won Choi, Joongheon Kim, Seehwan Yoo,
IEEE Access, 9:ppp–ppp, September 2021.
Machine learning requires a large volume of sample data, especially when it is used in high-accuracy medical applications. However, patient records are one of the most sensitive private information that is not usually shared among institutes. This paper presents spatio-temporal split learning, a distributed deep neural network framework, which is a turning point in allowing collaboration among privacy-sensitive organizations. Our spatio-temporal split learning presents how distributed machine learning can be efficiently conducted with minimal privacy concerns. The proposed split learning consists of a number of clients and a centralized server. Each client has only has one hidden layer, which acts as the privacy-preserving layer, and the centralized server comprises the other hidden layers and the output layer. Since the centralized server does not need to access the training data and trains the deep neural network with parameters received from the privacy-preserving layer, privacy of original data is guaranteed. We have coined the term, spatio-temporal split learning, as multiple clients are spatially distributed to cover diverse datasets from different participants, and we can temporally split the learning process, detaching the privacy preserving layer from the rest of the learning process to minimize privacy breaches. This paper shows how we can analyze the medical data whilst ensuring privacy using our proposed multi-site spatio-temporal split learning algorithm on Coronavirus Disease-19 (COVID-19) chest Computed Tomography (CT) scans, MUsculoskeletal RAdiographs (MURA) X-ray images, and cholesterol levels.

[Access'21.06] Joint Mobile Charging and Coverage-Time Extension for Unmanned Aerial Vehicles,
Soohyun Park, Minseok Choi, Won-Yong Shin, Joongheon Kim,
IEEE Access, 9:94053–94063, June 2021.
In modern networks, the use of drones as mobile base stations (MBSs) has been discussed for coverage flexibility. However, the realization of drone-based networks raises several issues. One of critical issues is drones are extremely power-hungry. To overcome this, we need to characterize a new type of drones, so-called charging drones, which can deliver energy to MBS drones. Motivated by the fact that the charging drones also need to be charged, we deploy ground-mounted charging towers for delivering energy to the charging drones. We introduce a new energy-efficiency maximization problem, which is partitioned into two independently separable tasks. More specifically, as our first optimization task, two-stage charging matching is proposed due to the inherent nature of our network model, where the first matching aims to schedule between charging towers and charging drones while the second matching solves the scheduling between charging drones and MBS drones. We analyze how to convert the formulation containing non-convex terms to another one only with convex terms. As our second optimization task, each MBS drone conducts energy-aware time-average transmit power allocation minimization subject to stability via Lyapunov optimization. Our solutions enable the MBS drones to extend their lifetimes; in turn, network coverage-time can be extended.

[Access'20.06] Blind Signal Classification Analysis and Impact on User Pairing and Power Allocation in Nonorthogonal Multiple Access,
Minseok Choi, Joongheon Kim,
IEEE Access, 8:100916–100929, June 2020.
This paper studies blind signal classification (SC) in the nonorthogonal multiple access (NOMA) system, which especially determines whether or not the received NOMA signal requires successive interference cancellation (SIC) without a priori signal information. In this paper, two types of blind SC errors are analyzed: 1) the signal that has to cancel the superposed interference component is classified as one that does not require SIC and 2) the signal that has no need of performing SIC is classified as one that requires SIC. Here, the interesting observation is that the classification error of the first type decreases with the SNR, but that of the second type increases with the SNR. In this regard, this paper proposes the joint user pairing and power allocation policy for the NOMA system based on the above observation that the blind SC performance of the NOMA user who does not perform SIC decreases with the SNR. The joint optimization problem maximizes the sum-rate gain of NOMA over orthogonal multiple access (OMA) with constraints on the maximum classification error probability and the minimum data rate. Since the problem is nonconvex, we propose the numerical algorithm that iteratively finds the appropriate user scheduling and power allocation. Simulation results show that the proposed scheme outperforms existing user scheduling methods.

[Access'18.05] Soft Memory Box: A Virtual Shared Memory Framework for Fast Deep Neural Network Training in Distributed High Performance Computing,
Shinyoung Ahn, Joongheon Kim, Eunji Lim, Sungwon Kang,
IEEE Access, 6:26493–26504, May 2018.
(Google Scholar Citations: 10+) *Deep learning is one of the major promising machine learning methodologies. Deep learning is widely used in various application domains, e.g., image recognition, voice recognition, and natural language processing. In order to improve learning accuracy, deep neural networks have evolved by: 1) increasing the number of layers and 2) increasing the number of parameters in massive models. This implies that distributed deep learning platforms need to evolve to: 1) deal with huge/complex deep neural networks and 2) process with high-performance computing resources for massive training data. This paper proposes a new virtual shared memory framework, called Soft Memory Box (SMB), which enables sharing the memory of remote node among distributed processes in the nodes so as to improve communication performance via parameter sharing. According to data-intensive performance evaluation results, the communication time of deep learning using the proposed SMB is 2.1 times faster than that using the massage passing interface (MPI). In addition, the communication time of the SMB-based asynchronous parameter update becomes 2-7 times faster than that using the MPI depending on deep learning models and the number of deep learning workers.*

[Access'17.09] A Software-based Monitoring Framework for Time-Space Partitioned Avionics Systems,
Changmin Shin, Chaedeok Lim, Joongheon Kim, Heejun Roh, Wonjun Lee,
IEEE Access, 5:19132–19143, September 2017.

Recently, avionics systems have evolved into a time and space partitioning (TSP)-based integrated modular avionics (IMA) structure for integration into a single system from a variety of existing independently configured federated systems. The TSP-based IMA architecture is suitable for solving size, weight, and power problems in avionics systems. Partitioning real-time operating systems (RTOSs) to support TSP-based IMA have been researched, and the international aviation industry has established the ARINC 653 standard for a partitioning RTOS. The ARINC 653 standard has defined the health monitoring (HM) function for debugging. However, the HM of the ARINC 653 standard does not support monitoring and debugging functions, such as snapshot, cycle, and, redundancy monitor, which makes the system development hard. To this end, the purpose of this paper is to introduce a monitoring framework that supports high reliability and stability for RTOS and application software based on TSP structure used in avionics systems. The proposed monitoring framework is designed for Qplus-AIR, an RTOS based on the TSP structure that conforms to the ARINC 653 for aircraft systems. It is also applicable to other RTOSs based on TSP structure that does not conform to ARINC 653. It supports monitoring functions, such as snapshot, trigger, and cycle as well as various debugging functions. It also supports debugging and monitoring operations under the redundancy of avionics systems, and minimizes the intrusive effect, which is a disadvantage of the software-based debugging approach. These functionalities enable avionics system developers to monitor and measure the performance of TSP structure-based RTOS and application software in flight control system for unmanned aerial vehicles. Our evaluation results show that the proposed monitoring framework is suitable for monitoring and debugging of RTOS and application software based on TSP structure.

[Access'17.08] Energy-Efficient Stabilized Automatic Control for Multicore Baseband in Millimeter-Wave Systems,
Joongheon Kim, Jae-Jin Lee, Jong-Kook Kim, Woojoo Lee,
IEEE Access, 5:16584–16591, August 2017.
The fifth generation (5G) cellular network is upon us. Academia and Industry have intensively collaborated together to bring the power of 5G cellular networks to the masses, and now the 5G millimeterwave (mmWave) platforms come into being in the market. One of the most popular 5GmmWave platforms mounts the massive mmWave phased antenna arrays in order to transfer a huge number of bits in a second (e.g., more than ten gigabits-per-second) to the baseband in the platform. While exploiting chip multicore processors (CMPs) may be the best solution to process such huge data in the mmWave baseband platform, power dissipate by the CMPs should become critical. Starting from an intuition that utilizing all processors in every single time introduces inefficient energy consumption, this paper proposes an energy aware queue-stable control (EQC) algorithm to control the activation/deactivation of individual processors and antenna arrays for pursuing time average energy consumption minimization subject to the stability of queues in the 5G-mm Wave baseband. Results from intensive simulations based on realistic experimental setups demonstrate the efficacy of the proposed EQC that achieves significant energy savings while queue stability is maintained.

[Access'17.06] Adaptive Resource Balancing for Serviceability Maximization in Fog Radio Access Networks,
Nhu-Ngoc Dao, Junwook Lee, Duc-Nghia Vu, Jeongyeup Paek, Joongheon Kim, Sungrae Cho, Ki-Sook Chung, Changsup Keum,
IEEE Access, 5:14548–14559, June 2017.
(Google Scholar Citations: 32+) *Serviceability is the ability of a network to serve user equipments (UEs) within desired requirements (e.g., throughput, delay, and packet loss). High serviceability is considered as one of the key foundational criteria toward a successful fog radio access infrastructure satisfying the Internet of Things paradigm in the 5G era. In this paper, we propose an adaptive resource balancing (ARB) scheme for serviceability maximization in fog radio access networks wherein the resource block (RB) utilization among remote radio heads (RRHs) are balanced using the backpressure algorithm with respect to a time-varying network topology issued by potential RRH mobilities. The optimal UE selection for service migration from a high-RB-utilization RRH to its neighboring low-RB-utilization RRHs is determined by the Hungarian method to minimize RB occupation after moving the service. Analytical results reveal that the proposed ARB scheme provides substantial gains compared with the standalone capacity-aware, max-rate, and cache-aware UE association approaches in terms of serviceability, availability, and throughput.*

[Access'16.12] Numerical Simulation Study for Frequency Sharing between Micro-Cellular Systems and Fixed Service Systems in Millimeter-Wave Bands,
Joongheon Kim, Liang Xian, Ali S. Sadri,
IEEE Access, 4:9847–9859, December 2016.
(Google Scholar Citations: 15+) *This paper presents numerical simulation results to study the impact of the co-existence between a fixed service (FS) system and 5G small cell networks at 28-, 38-, and 60-GHz millimeter-wave (mmWave) frequency bands. For this paper, two scenarios are considered: aggregation of interference from small cells into an FS receiver from base stations (BSs) to their associated user equipment (UE) (downlink) and the aggregation of cellular interference at the FS receiver from UEs to their associated BSs (downlink). Moreover, mmWave-specific propagation characteristics and attenuation factors are considered for a more precise simulation study. The simulation results determine how much interference rejection is required to protect the operation of FS. In addition, currently available mmWave modular antenna array (MAA) architectures are introduced. Based on the information, additional mmWave frequency sharing study is performed using the realistic MAA radiation patterns. Last, we compare and analyze the performance differences between ITU standard models and MAA solutions.*

CONFERENCES PROCEEDINGS

Top-Tiers

[ICDCS'20] Understanding the Potential Risks of Sharing Elevation Information on Fitness Applications,
Ülkü Meteriz, Necip Fazil Yildiran, Joongheon Kim, David Mohaisen,
IEEE International Conference on Distributed Computing Systems (IEEE ICDCS),
Virtual, November/December 2020.
(Acceptance Rate: 17.98% (105/584))
The extensive use of smartphones and wearable devices has facilitated many useful applications. For example, with Global Positioning System (GPS)-equipped smart and wearable devices, many applications can gather, process, and share rich metadata, such as geolocation, trajectories, elevation, and time. For example, fitness applications, such as Runkeeper and Strava, utilize information for activity tracking, and have recently witnessed a boom in popularity. Those fitness tracker applications have their own web platforms, and allow users to share activities on such platforms, or even with other social network platforms. To preserve privacy of users while allowing sharing, several of those platforms may allow users to disclose partial information, such as the elevation profile for an activity, which supposedly would not leak the location of the users. In this work, and as a cautionary tale, we create a proof of concept where we examine the extent to which elevation profiles can be used to predict the location of users. To tackle this problem, we devise three plausible threat settings under which the city or borough of the targets can be predicted. Those threat settings define the amount of information available to the adversary to launch the prediction attacks. Establishing that simple features of elevation profiles, e.g., spectral features, are insufficient, we devise both natural language processing (NLP)-inspired text-like representation and computer vision-inspired image-like representation of elevation profiles, and we convert the problem at hand into text and image classification problem. We use both traditional machine learning-and deep learning-based techniques, and achieve a prediction success rate ranging from 59.59% to 95.83%. The findings are alarming, and highlight that sharing elevation information may have significant location privacy risks.

[IJCAI'19] Randomized Adversarial Imitation Learning for Autonomous Driving,
MyungJae Shin, Joongheon Kim,
International Joint Conference on Artificial Intelligence (IJCAI),
Macao, August 2019.
(Acceptance Rate: 17.89% (850/4752))
With the evolution of various advanced driver assistance system (ADAS) platforms, the design of autonomous driving system is becoming more complex and safety-critical. The autonomous driving system simultaneously activates multiple ADAS functions; and thus it is essential to coordinate various ADAS functions. This paper proposes a randomized adversarial imitation learning (RAIL) method that imitates the coordination of autonomous vehicle equipped with advanced sensors. The RAIL policies are trained through derivative-free optimization for the decision maker that coordinates the proper ADAS functions, e.g., smart cruise control and lane keeping system. Especially, the proposed method is also able to deal with the LIDAR data and makes decisions in complex multi-lane highways and multi-agent environments.

[ICDCS'18] ShmCaffe: A Distributed Deep Learning Platform with Shared Memory Buffer for HPC Architecture,
Shinyoung Ahn, Joongheon Kim, Eunji Lim, Wan Choi, Aziz Mohaisen, Sungwon Kang,
IEEE International Conference on Distributed Computing Systems (IEEE ICDCS),
Vienna, Austria, July 2018.
(Acceptance Rate: 20.63% (78/378))
(Google Scholar Citations: 14+) *One of the reasons behind the tremendous success of deep learning theory and applications in the recent days is advances in distributed and parallel high performance computing (HPC). This paper proposes a new distributed deep learning platform, named ShmCaffe, which utilizes remote shared memory for communication overhead reduction in massive deep neural network training parameter sharing. ShmCaffe is designed based on Soft Memory Box (SMB), a virtual shared memory framework. In the SMB framework, the remote shared memory is used as a shared buffer for asynchronous massive parameter sharing among many distributed deep learning processes. Moreover, a hybrid method that combines asynchronous and synchronous parameter sharing methods is also discussed in this paper for improving scalability. As a result, ShmCaffe is 10.1 times faster than Caffe and 2.8 times faster than Caffe-MPI for deep neural network training when Inception_v1 is trained with 16 GPUs. We verify the convergence of the Inception_v1 model training using ShmCaffe-A and ShmCaffe-H by varying the number of workers. Furthermore, we evaluate scalability of ShmCaffe by analyzing the computation and communication times per one iteration of deep learning training in four convolutional neural network (CNN) models.*

[MM'17] REQUEST: Seamless Dynamic Adaptive Streaming over HTTP for Multi-Homed Smartphone under Resource Constraints,
Jonghoe Koo, Juheon Yi, Joongheon Kim, Mohammad Ashraful Hoque, Sunghyun Choi,
ACM International Conference on Multimedia (ACM MM),
Mountain View, California, USA, October 2017.
(Acceptance Rate: 27.63% (189/684))
(Google Scholar Citations: 23+) *Exploiting both LTE and Wi-Fi links simultaneously enhances the performance of video streaming services in a smartphone. However, it is challenging to achieve seamless and high quality video while saving battery energy and LTE data usage to prolong the usage time of a smartphone. In this paper, we propose REQUEST, a video chunk request policy for Dynamic Adaptive Streaming over HTTP (DASH) in a smartphone, which can utilize both LTE and Wi-Fi. REQUEST enables seamless DASH video streaming with near optimal video quality under given budgets of battery energy and LTE data usage. Through extensive simulation and measurement in a real environment, we demonstrate that REQUEST significantly outperforms other existing schemes in terms of average video bitrate, rebuffering, and resource waste.*

[MobiSys'10] Energy-Efficient Rate-Adaptive GPS-based Positioning for Smartphones,
Jeongyeup Paek, Joongheon Kim, Ramesh Govindan,
ACM International Conference on Mobile Systems, Applications, and Services (ACM MobiSys),
San Francisco, California, USA, June 2010.
(Acceptance Rate: 19.84% (25/126))
(Google Scholar Citations: 600+) *Many emerging smartphone applications require position information to provide location-based or context-aware services. In these applications, GPS is often preferred over its alternatives such as GSM/WiFi based positioning systems because it is known to be more accurate. However, GPS is extremely power hungry. Hence a common approach is to periodically duty-cycle GPS. However, GPS duty-cycling trades-off positioning accuracy for lower energy. A key requirement for such applications, then, is a positioning system that provides accurate position information while spending minimal energy. In this paper, we present RAPS, rate-adaptive positioning system for smartphone applications. It is based on the observation that GPS is generally less accurate in urban areas, so it suffices to turn on GPS only as often as necessary to achieve this accuracy. RAPS uses a collection of techniques to cleverly determine when to turn on GPS. It uses the location-time history of the user to estimate user velocity and adaptively turn on GPS only if the estimated uncertainty in position exceeds the accuracy threshold. It also efficiently estimates user movement using a duty-cycled accelerometer, and utilizes Bluetooth communication to reduce position uncertainty among neighboring devices. Finally, it employs celltower-RSS blacklisting to detect GPS unavailability (e.g., indoors) and avoid turning on GPS in these cases. We evaluate RAPS through real-world experiments using a prototype implementation on a modern smartphone and show that it can increase phone lifetimes by more than a factor of 3.8 over an approach where GPS is always on.*