

DONGEUN LEE

Computer Science & Information Systems, College of Science and Engineering, Texas A&M University - Commerce, Commerce, TX

E-mail: Dongeun.Lee@tamuc.edu

EDUCATION

- **Seoul National University (SNU)**, Seoul, Korea
Ph. D. in Electrical Engineering and Computer Science February 2014
Advisor: Prof. Heonshik Shin
Thesis: *Analysis for Scalable Coding of Quality-Adjustable Sensor Data.*

- **Seoul National University (SNU)**, Seoul, Korea
B.S. in Computer Science and Engineering February 2006

RESEARCH INTERESTS

- **Big Spatio-Temporal Data Processing/Analytics**
 - Data Reduction using Source Coding
 - Machine Learning for Streaming Data Analysis
 - Nonlinear Dimensionality Reduction
 - Efficient Algorithms for Resource Limited Devices

APPOINTMENTS

- **Assistant Professor, Texas A&M University - Commerce (TAMUC), Commerce, TX** 2016—Present
- **Faculty Research Affiliate, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA** 2016–2018
 - Scientific Data Management Group, Computational Research Division
- **Computer Systems Engineer, Postdoctoral Research Affiliate, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA** 2015–2016
 - Researched big streaming data analysis problems.
 - Participated in two projects: “Open Framework for High-Performance Streaming Analytics” and “Behavior Analysis on Residential Electricity Usage Data.”
- **Postdoctoral Research Associate, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Korea** 2014–2016
 - Researched big spatio-temporal data archiving and sensing problems.
 - Studied machine learning algorithms and led the project “Failure Prediction and Diagnosis Algorithm for RMS.”
 - Advised students in research.
 - Published three international conference papers (UAI 2016, SDM 2015, IEEE BigData 2014).
- **Research Intern, Hokkaido University, Sapporo, Japan** Fall 2009, Spring 2010
 - Research theme: “Energy-aware and Fault-tolerant Peat Fire Monitoring with Wireless Sensor Network” under guidance of Prof. Toshihisa Honma.
 - Deployed and operated 20 sensor motes at experimental farms in Hokkaido university.

-TinyOS programming for routing of monitored environmental sensor data and sending alarm in case of fire.

-Published a domestic conference paper and subsequently an international conference paper (AINA 2012).

- **Instructor, BIT Computer Academy, Seoul, Korea** **2008–2010**
-Lectured three one-month courses: “UNIX/LINUX System Programming,” “Client/Server Programming,” and “MS Visual Studio 2008.”
- **Software Engineer, SimsLine Inc., Seoul, Korea** **2003–2004**
-Developed/maintained text-to-speech applications using L&H API in MS Windows platform.
-Developed/maintained device-sync applications that can synchronize music with MP3 players using WDM (Windows driver model) in MS Windows platform.
- **Software Engineer, Hyun-seung Inc., Seoul, Korea** **2002–2003**
-Maintained a graphical desktop sharing system similar to VNC written in MFC/Win32 API.
- **Software Engineer, Betaland Inc., Seoul, Korea** **2001**
-Developed an educational purpose application in MS Windows platform that could record Windows desktop and annotate on the recorded desktop.

TEACHING

-Big Data Computing and Analytics (CSCI 573), TAMUC, 2017—Present.

-Fundamentals of Programming C/C++ (CSCI 515), TAMUC, 2016—Present.

-Programming Fundamentals II (COSC 1337), TAMUC, 2016—Present.

AWARDS

- **DataCom 2015 Best Paper Award**
-Taehoon Kim, **Dongeun Lee**, Jaesik Choi, Anna Spurlock, Alex Sim, Annika Todd, and Kesheng Wu, “Extracting Baseline Electricity Usage Using Gradient Tree Boosting,” *1st Intl. Conf. Big Data Intelligence and Computing (DataCom)*, pp. 734-741, Dec. 2015.
- **IEEE ISCC 2010 Best Paper Award - Student Category**
-**Dongeun Lee**, Jonghun Lee, Yonghee Lee, Heejung Lee, and Heonshik Shin, “Low-Complexity Aggregation of Collected Images with Correlated Fields of View in Wireless Video Sensor Networks,” *IEEE 15th Symp. Computers and Communications (ISCC)*, pp. 765-771, Jun. 2010.
- **Efficient and Stable Video Transmission over Wireless Video Sensor Network, Korea Student Aid Foundation (KOSAF) Graduate Research Grant, Korea, Sept. 2008—Aug. 2009.**
-Reviewed multiview video coding (MVC) for an application in video sensor network.
-The grant served as a basis for research of the ISCC 2010 paper.

PUBLICATIONS

- **Thesis**
 1. **Dongeun Lee**, [*Analysis for Scalable Coding of Quality-Adjustable Sensor Data*](#), Ph. D. Thesis, Department of Electrical Engineering and Computer Science, Seoul National University, 2014.

• Book Chapter

2. **Dongeun Lee**, “Big Sensor Data Acquisition and Archiving with Compression,” in *Big Data and Visual Analytics*, Sang C. Suh and Thomas Anthony, Eds. Springer International Publishing, pp. 115–143, 2017.

• Journals

3. Chanyoung Park, Yoonsoo Jo, **Dongeun Lee**, and Kyungtae Kang, “Change Your Cluster to Cold: Gradually Applicable and Serviceable Cold Storage Design,” *IEEE Access*, accepted for publication.
4. Jiwoong Won, Oseok Kwon, Junhee Ryu, **Dongeun Lee**, and Kyungtae Kang, “iFetcher: User-Level Prefetching Framework with File-System Event Monitoring for Linux,” *IEEE Access*, vol. 6, pp. 46213–46226, 2018.
5. Junhee Ryu, **Dongeun Lee**, Kang G. Shin, and Kyungtae Kang, “ClusterFetch: A Lightweight Prefetcher for Intensive Disk Reads,” *IEEE Transactions on Computers*, vol. 67, no. 2, pp. 284–290, Feb. 2018 (corresponding author).
6. Jaemyoun Lee, Haegwon Jeong, Won-Joo Lee, Hyo-Joong Suh, **Dongeun Lee**, and Kyungtae Kang, “Advanced Primary-Backup Platform with Container-Based Automatic Deployment for Fault-Tolerant Systems,” *Wireless Personal Communications*, vol. 98, no. 4, pp. 3177–3194, Feb. 2018.
7. Taehoon Kim, Jaesik Choi, **Dongeun Lee**, Alex Sim, Anna Spurlock, Annika Todd, and Kesheng Wu, “Predicting Baseline for Analysis of Electricity Pricing,” *International Journal of Big Data Intelligence*, vol. 5, nos. 1/2, pp. 3–20, 2018.
8. Junhee Ryu, **Dongeun Lee**, Changhee Han, Heonshik Shin, and Kyungtae Kang, “File-System-Level Storage Tiering for Faster Application Launches with No Mapping Overhead,” *IEEE Access*, vol. 4, pp. 3688–3696, 2016 (corresponding author).
9. **Dongeun Lee**, Jaesik Choi, and Heonshik Shin, “[A Scalable and Flexible Repository for Big Sensor Data](#),” *IEEE Sensors Journal*, vol. 15, no. 12, pp. 7284–7294, Dec. 2015.
10. **Dongeun Lee**, Junhee Ryu, and Heonshik Shin, “[Scalable Management of Storage for Massive Quality-Adjustable Sensor Data](#),” *Computing*, vol. 97, no. 8, pp. 769–793, Aug. 2015.
11. **Dongeun Lee**, Jaesik Choi, and Heonshik Shin, “Low-Complexity Compressive Sensing with Down-sampling,” *IEICE Electronics Express*, vol. 11, no. 3, pp. 20130947, Feb. 2014.
12. Heejung Lee, Yonghee Lee, Jonghun Lee, **Dongeun Lee**, and Heonshik Shin, “Design of a Mobile Video Streaming System using Adaptive Spatial Resolution Control,” *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1682–1689, Aug. 2009.
13. Donggeon Noh, **Dongeun Lee**, and Heonshik Shin, “QoS-Aware Geographic Routing for Solar-Powered Wireless Sensor Networks,” *IEICE Transactions on Communications*, vol. 90, no. 12, pp. 3373–3382, Dec. 2007.

• Refereed Conferences

14. J. Kade Gibson, **Dongeun Lee**, Jaesik Choi, and Alex Sim, “Dynamic Online Performance Optimization in Streaming Data Compression,” *IEEE 6th Intl. Conf. Big Data (IEEE BigData)*, pp. 534–541, Dec. 2018 (equal contribution).
15. Kesheng Wu, **Dongeun Lee**, Alex Sim, and Jaesik Choi, “Statistical Data Reduction for Streaming Data,” *2nd New York Scientific Data Summit (NYSDS)*, pp. 1–6, Aug. 2017.
16. **Dongeun Lee**, Alex Sim, Jaesik Choi, and Kesheng Wu, “Improving Statistical Similarity Based Data Reduction for Non-Stationary Data,” *29th Intl. Conf. Scientific and Statistical Database Management (SSDBM)*, pp. 37:1–37:6, Jun. 2017.
17. **Dongeun Lee**, Alex Sim, Jaesik Choi, and Kesheng Wu, “Novel Data Reduction Based on Statistical Similarity,” *28th Intl. Conf. Scientific and Statistical Database Management (SSDBM)*, pp. 21:1–21:12, Jul. 2016.
18. **Dongeun Lee**, Rafael Lima, and Jaesik Choi, “[Improving Imprecise Compressive Sensing Models](#),” *32nd Conf. Uncertainty in Artificial Intelligence (UAI)*, pp. 397–406, Jun. 2016.

19. Taehoon Kim, **Dongeun Lee**, Jaesik Choi, Anna Spurlock, Alex Sim, Annika Todd, and Kesheng Wu, "Extracting Baseline Electricity Usage Using Gradient Tree Boosting," *1st Intl. Conf. Big Data Intelligence and Computing (DataCom)*, pp. 734–741, Dec. 2015.
20. Junhee Ryu, Haksu Jeong, **Dongeun Lee**, Heonshik Shin, and Kyungtae Kang, "ClusterFetch: A Lightweight Prefetcher that Responds to Intensive Disk Read Patterns," *IEEE 12th Intl. Conf. Embedded Software and Systems (ICSS)*, pp. 1051–1056, Aug. 2015 (corresponding author).
21. **Dongeun Lee** and Jaesik Choi, "[Learning Compressive Sensing Models for Big Spatio-Temporal Data](#)," *SIAM 15th Intl. Conf. Data Mining (SDM)*, pp. 667–675, Apr./May 2015.
22. **Dongeun Lee** and Jaesik Choi, "[Low Complexity Sensing for Big Spatio-Temporal Data](#)," *IEEE 2nd Intl. Conf. Big Data (IEEE BigData)*, pp. 323–328, Oct. 2014.
23. Ikkjune Yoon, Dong Kun Noh, **Dongeun Lee**, Rony Teguh, Toshihisa Honma, and Heonshik Shin, "Reliable Wildfire Monitoring with Sparsely Deployed Wireless Sensor Networks," *IEEE 26th Intl. Conf. Advanced Information Networking and Applications (AINA)*, pp. 460–466, Mar. 2012.
24. **Dongeun Lee**, Heonshik Shin, and Eunjeong Park, "Modeling Recovery Strategies in Service-Oriented Architecture using a Markov Decision Process," *IEEE 13th Intl. Symp. High-Assurance Systems Engineering (HASE)*, pp. 285–290, Nov. 2011.
25. **Dongeun Lee**, Jonghun Lee, Yonghee Lee, Heejung Lee, and Heonshik Shin, "Low-Complexity Aggregation of Collected Images with Correlated Fields of View in Wireless Video Sensor Networks," *IEEE 15th Symp. Computers and Communications (ISCC)*, pp. 765–771, Jun. 2010.
26. Heejung Lee, Yonghee Lee, **Dongeun Lee**, Jonghun Lee, and Heonshik Shin, "Implementing Rate Allocation and Control for Real-Time H.264/SVC Encoding," *IEEE 28th Intl. Conf. Consumer Electronics (ICCE)*, pp. 269–270, Jan. 2010.
27. **Dongeun Lee**, Yonghee Lee, Heejung Lee, Jonghun Lee, and Heonshik Shin, "Determining Efficient Bit Stream Extraction Paths in H.264/AVC Scalable Video Coding," *42nd Asilomar Conf. Signals, Systems, and Computers (Asilomar)*, pp. 2233–2237, Oct. 2008.
28. Heejung Lee, **Dongeun Lee**, Yonghee Lee, and Heonshik Shin, "Luminance Scalable Coding using H.264/AVC SVC Extensions for Mobile Video Applications," *IEEE Intl. Conf. Multimedia and Expo (ICME)*, pp. 1025–1028, Jun. 2008.
29. Donggeon Noh, **Dongeun Lee**, and Heonshik Shin, "Mission-Oriented Selective Routing for Wireless Sensor Network," *2nd Intl. Conf. Communications and Networking in China (CHINACOM)*, pp. 809–813, Aug. 2007.
30. Hyuntaek Kwon, Donggeon Noh, Junu Kim, Joonho Lee, **Dongeun Lee**, and Heonshik Shin, "Low-Latency Routing for Energy-Harvesting Sensor Networks," *4th Intl. Conf. Ubiquitous Intelligence and Computing (UIC)*, pp. 422–433, Jul. 2007.
31. Donggeon Noh, Junu Kim, Joonho Lee, **Dongeun Lee**, Hyuntaek Kwon, and Heonshik Shin, "Priority-Based Routing for Solar-Powered Wireless Sensor Networks," *2nd Intl. Symp. Wireless Pervasive Computing (ISWPC)*, pp. 53–58, Feb. 2007.

- **Abstracts**

32. **Dongeun Lee**, Alex Sim, Jaesik Choi, and Kesheng Wu, "Expanding Statistical Similarity Based Data Reduction to Capture Diverse Patterns," *27th Data Compression Conf. (DCC)*, p. 445, Apr. 2017.
33. Haksu Jeong, Junhee Ryu, **Dongeun Lee**, Jaemyoun Lee, Heonshik Shin, and Kyungtae Kang, "ClusterFetch: A Lightweight Prefetcher for General Workloads," *ACM/SPEC 6th Intl. Conf. Performance Engineering (ICPE)*, pp. 99–100, Jan./Feb. 2015.
34. Changhee Han, Junhee Ryu, **Dongeun Lee**, Jaemyoun Lee, Kyungtae Kang, and Heonshik Shin, "File-System-Level Flash Caching for Improving Application Launch Time on Logical Hybrid Disks," *IEEE 33rd Intl. Performance Computing and Communications Conf. (IPCCC)*, pp. 1–2, Dec. 2014.

- **Manuscript in Preparation**

–Junhee Ryu, **Dongeun Lee**, Kyungtae Kang, and Heonshik Shin, "Paralfetch: A Fast Application Launch for Consumer Devices."

-Duanshun Li, Jing Liu, Noseong Park, **Dongeun Lee**, Giridhar Ramachandran, Ali Seyedmazloom, Kookjin Lee, Chen Feng, Vadim Sokolov, and Rajesh Ganesan, "Solving Large-Scale 0-1 Knapsack Problems and Its Application to Point Cloud Resampling."

RESEARCH PROJECTS

- **Open Framework for High-Performance Streaming Analytics, LBNL** 2015–2016
 - Studied an effective streaming data analysis framework for high-velocity data from smart grid.
 - Implemented streaming data compression schemes with the Locally Exchangeable Measures technique developed for data reduction and pattern discovery.
 - Released an open software for streaming data compression (IDEALEM) and published an international conference paper (SSDBM 2016).
- **Behavior Analysis on Residential Electricity Usage Data, LBNL** 2015
 - Authored two manuscripts for publication.
 - Published an international conference paper (DataCom 2015) and prepared a manuscript.
- **Failure Prediction and Diagnosis Algorithm for RMS (Remote Monitoring System), UNIST** 2015
 - Led the project with Doosan Heavy Industry.
 - Applied various machine learning algorithms to power plant sensor data for detecting failure.
- **Breakpoint-Based Prefetching Techniques to Improve the Responsiveness of Mobile Applications, Hanyang University** 2014–2016
 - Studied diverse prefetching/caching techniques on memory hierarchy that can improve application launch times.
 - Published three international conference papers (ICISS 2015, ICPE 2015, IPCCC 2014) and prepared two manuscripts.
- **Development of Next Generation File System for Urban Computing, SNU** 2009–2012
 - Devised new data management scheme for huge amount of data generated in urban areas.
 - Analyzed conventional distributed file system codes for an adaptation to comply with characteristics of urban sensing data.
 - Focused on data quality adjustment for efficient management of data storage.
 - Derived the idea for the Ph. D. dissertation.
 - Published two journal papers (IEEE Sensors Journal, Computing).
- **An Adaptive Service Composition Technique for Reliable Service-Oriented Architecture, SNU** 2009–2011
 - Surveyed numerous literature regarding service-oriented computing with emphasis on reliability and fault-tolerance.
 - Published an international conference paper (HASE 2011).
- **A Study on Scalable Video Server for Heterogeneous Network Environment, SNU** 2007–2008
 - Studied H.264/AVC SVC (scalable video coding) extension standard and analyzed its reference software JSVM (joint scalable video model).
 - Conducted profiling of JSVM and evaluated coding efficiencies on diverse parameter combinations.
 - Surveyed numerous literature regarding image and video coding issues.
 - Published an international conference paper (Asilomar 2008).

• Development of Ubiquitous Storage Dust, SNU

2006

- Researched data synchronization problems between a central storage server and many storage dusts with lower capacity and bandwidth (Intel's XScale PXA270 - ARMv5TE ISA compliant).
- Developed data synchronization and conflict resolution module using cross compiler that could merge modified replicas across storage dusts and also support disconnected operation and asynchronous update.

PATENT

-Junhee Ryu, **Dongeun Lee**, and Kwangjin Ko, "Prefetching Method for Flash Memory Device and Recording Medium in Which Method is Recorded," Korean Patent No. 1020140061018, Korea, May 2014.

SOFTWARE PACKAGE RELEASE AND DEMO

- Implementation of Dynamic Extensible Adaptive Locally Exchangeable Measures (IDEALEM), *ACM/IEEE 29th Intl. Conf. High Performance Computing, Networking, Storage, and Analysis (SC)*, [Demo](#), Nov. 2016.
- Implementation of Dynamic Extensible Adaptive Locally Exchangeable Measures (IDEALEM), [LBNL S/W Disclosure No. 2016-045](#), under the modified BSD license, USA, Feb. 2016.

INVITED TALKS

- Challenges in Statistical Similarity Based Data Reduction*, Department of Computer Science, Texas A&M University-Commerce (TAMUC), Nov. 2018.
- Novel Data Reduction Based on Statistical Similarity*, Computational Research Division, Lawrence Berkeley National Laboratory (LBNL), May 2016.
- Big Sensor Data Acquisition and Archiving*, Computational Research Division, Lawrence Berkeley National Laboratory (LBNL), Nov. 2015.
- Efficient Big Data Signal Acquisition by Compressive Sensing and Random Sampling*, Department of Electrical and Computer Engineering, Ulsan National Institute of Science and Technology (UNIST), Mar. 2015.
- An Introduction to Compressive Sensing and Big Data Applications*, Department of Computer Science and Engineering, Hanyang University, Nov. 2014.
- An Introduction to Compressive Sensing and Big Data Applications*, Department of Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Aug. 2014.
- An Introduction to Compressive Sensing*, School of Electronic Engineering, Soongsil University, Jan. 2014.

PROFESSIONAL ACTIVITIES

- Guest Editor, *Transactions on Emerging Telecommunications Technologies - SI on Real-Time Internet of Things (IoT) and Cyber-Physical Systems (CPS)*, 2018.
- TPC, IEEE International Conference on Communications (IEEE ICC) - Cloud Communications and Networks Track, Selected Areas in Communications Symposium, 2017.
- TPC, IEEE Global Communications Conference (GLOBECOM) - Big Data Track, Selected Areas in Communications Symposium, 2017—Present.
- TPC, International Workshop on Systems and Network Telemetry and Analysis (SNTA) in conjunction with ACM International Symposium on High-Performance Parallel and Distributed Computing (ACM HPDC), 2019.
- Reviewer, AMS Mathematical Reviews, 2016—Present.
- Reviewer, IEEE Internet of Things Journal, 2018.
- Reviewer, IEEE Systems Journal, 2015–2016.
- Reviewer, Conference on Uncertainty in Artificial Intelligence (UAI), 2016.
- Reviewer, AAAI Conference on Artificial Intelligence (AAAI), 2015.
- Reviewer, Conference on Neural Information Processing Systems (NIPS), 2015.

-Reviewer, IEEE Journal on Selected Areas in Communications, 2014.