JOONHO BAE

E2-1 Room 4217
Korea Advanced Institute of Science and Technology
Republic of Korea

+82-10-8288-8792
lifefor1@kaist.ac.kr
joonhobae.github.io

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST), South Korea 2017-Current

- M.S. in Industrial & Systems Engineering (expected February 2019), GPA 3.89/4.0
  Thesis: "Multi-Output Log Gaussian Processes for Change Point Detection"

Seoul National University (SNU), South Korea 2009-2016

- B.S. in Statistics and Financial Economics, Cumulative GPA 3.50/4.0

RESEARCH INTERESTS

System Identification

- Development of data-driven, probabilistic models for complex dynamic systems
- Bayesian nonparametric methods (e.g., sparse & multi-output Gaussian Process regression)
- Sparse & low-rank tensor reconstruction (e.g., recommender systems for mobile applications)

System Monitoring & Prognostics

- Metamodeling of dynamic systems based on multi-sensors (e.g., degradation process modeling)
- Real-time system monitoring and anomaly/change-point detection

System Control & Optimization

- Stochastic optimization and model predictive control (e.g., optimal operations for Energy Storage System)
- Operations management and management science

RESEARCH IN PROGRESS (Extended Abstracts on Page 4-7)

1. **J. Bae** (with J. Park, and J. Choi). "Uncovering Dynamic Preferences and Recommending Game Applications" (under review at *Management Science)*
2. **J. Bae** (with J. Park). "Frequency-based Change Detection via Multi-Output Log Gaussian Cox Processes" (to be submitted to *KDD 2019* in February 2019)
3. **J. Bae** (with S. Lee, H. Sim, and J. Park). "Optimal Management of Energy Storage Systems for Wind Turbines"
4. **J. Bae** (with J. Park, S. Soh, and J. Choi). "Mobile Contents and Consumer Engagement: Modeling Contents Usage via Stochastic Processes"

RESEARCH EXPERIENCE

***Research Assistant*, KAIST** (Advisor: Prof. Jinkyoo Park, Sponsor: **LG CNS**) 2018-current

- Prediction of State of Health (SOH) for Energy Storage Systems (ESS)
  - Developed a data-driven, nonparametric approach to estimate SOH
  - Implemented a hierarchical/multi-task strategy exploiting correlations between multi-sensors
- Suggestion of a new criterion of operations management for ESS
  - Optimized the operations of variables (e.g., C-rate, DoD) to maximize cost-efficiency for battery lifecycle
- Real-time system monitoring for ESS
  - Examined real-time prognosis models for monitoring fault of an ESS battery

***Research Assistant*, KAIST** (Advisor: Prof. Jeonghye Choi, Sponsor: **Kantar TNS**) 2017-current

- Prediction of future usage for mobile applications from log data
  - Developed a nonparametric time series model to estimate a user-specific usage for mobile applications
  - Recommended personalized lists of mobile applications based on the estimated future usage
- Development of a targeting strategy
  - Proposed a new segmenting and targeting strategy to maximize the hitting ratio
  - Conducted statistical analysis to estimate the suggested strategy

***Research Assistant*, Seoul National University** (Advisor: Prof. Sinsup Cho, **Time Series Lab.**) 2012-2013

- Revised the Korean Educational Statistics Software (KESS) to enable statistical analysis on Excel
- Predicted Altman Z-score by extracting features based on companies' financial figures

CONFERENCE PRESENTATIONS & INVITED TALKS

1. "The Recommender System for Mobile Applications", *2018 Global Marketing Conference*, Tokyo, Japan, July 26-29, 2018
2. "Frequency-based Anomaly Detection via Multi-Output Log Gaussian Cox Processes", *2018 Stochastic Processes and their Applications*, Gothenburg, Sweden, June 11-15, 2018.
3. "An Application of Doubly Stochastic Poisson Process for Detecting Abnormalities", *INFORMS Annual Meeting 2018*, Phoenix, AZ, U.S., November 4-7, 2018
4. "The Recommender System for Mobile Applications", Department of Business Graduate Seminar, Yonsei University, South Korea, May 29, 2018

AWARDS & HONORS

- National Science & Technology Scholarship, KOSAF, 2009-2015
- Social Venture Idea ($3,000), Ministry of Employment & Labor, South Korea, 2014
- Enactus National Competition, Enactus, 2014
- Hope Advertisement, Seoul Metropolitan Government, 2014
- SCH Social Venture Idea ($1,000), Soonchunhyang University, 2013
- SK Social Enterprise ($10,000), SK Happiness Foundation, 2013
- Army Commendation Medal, 8th U.S. ARMY, 2012

WORK EXPERIENCE

***Marketing Manager*, SK TELECOM** 2015-2017

- Managed more than 40 stores in Seoul
- Developed optimizing tools for distribution of the cell phones to maximize the profits

***Co-Founder & Team Leader*, CHAM SON GIL Cooperative** 2013-2015

- Established a healing center for blind masseurs and developed a unique B2B service and products targeting 20-30s, who were not the main customers of the original market
- Expanded branches to nationwide and opened six stores, ensuring $3,000 monthly income for each masseur on average

***Sergeant*, HHC, 8th U.S. ARMY** 2010-2012

- Served in the U.S. Army as a Korean Augmentation to the United States Army

TEACHING EXPERIENCE

***Teaching Assistant*, KAIST** 2018-current

- Data-Driven Decision Making and Control (Fall 2018)
- Engineering Statistics I (Spring 2018)

TECHNICAL STRENGTHS

- Programming Languages: C, C++, Java, Python, MATLAB
- Statistical Languages: R, SAS, Stata, SPSS, SQL, Excel VBA
- Machine Learning Modules: GPy, GPyOpt, gpflow, tensorflow, keras, PyMC3, GPML, gpstuff

SELECTED COURSEWORK

Operations Research/Management Science

- Stochastic Modeling I — Dr. Kyoung-Kuk Kim
- Stochastic Modeling II — Dr. Kyoung-Kuk Kim
- Convex Optimization — Dr. Woo-Chang Kim
- Game Theory with Engineering Applications — Dr. Jinkyoo Park

Machine Learning/Statistical Learning

- Applications of AI/Data Mining Technology — Dr. Il-Chul Moon
- Statistical Learning Theory — Dr. Changdong Yoo
- Deep Learning for Computer Vision — Dr. Junmo Kim
- Mathematical Foundation of Reinforcement Learning — Dr. Song Chong
- Bayesian Estimation and its Application — Dr. Joohwan Chun

Graduate & Advanced Undergraduate Level from Seoul National University

- Time Series Analysis
- Mathematical Statistics I / II
- Data Mining Methods
- Experimental Design & Survey Practice
- Discrete Data Analysis
- Statistical Computing

REFERENCES

Jinkyoo Park (M.S. Advisor)
Assistant Professor
Department of Industrial & Systems Engineering
KAIST, Daejeon, South Korea
Phone: +82-42-350-3133
E-mail: jinkyoo.park@kaist.ac.kr

Jeonghye Choi (Research Co-Advisor)
Associate Professor
Department of Marketing
Yonsei University, Seoul, South Korea
Phone: +82-2-2123-6575
E-mail: jeonghye@yonsei.ac.kr

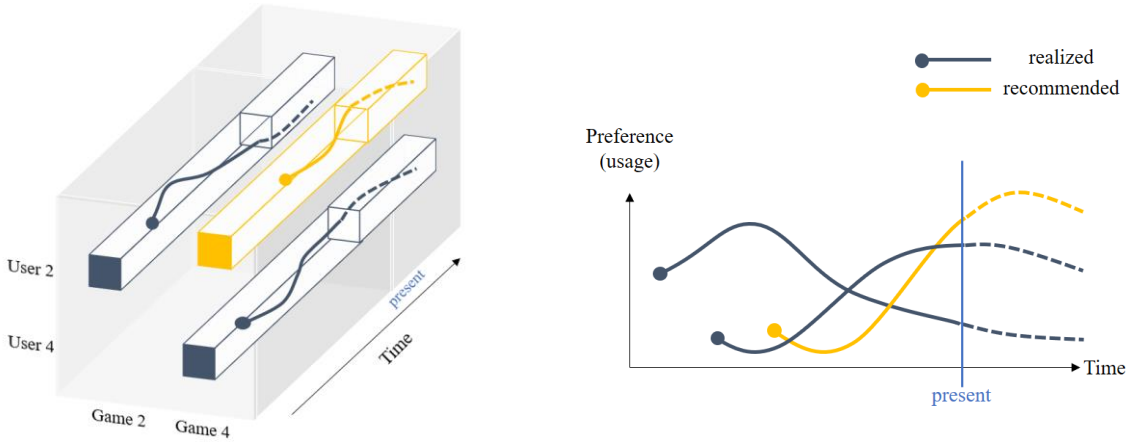
Kyoung-Kuk Kim (Thesis Committee Member)
Associate Professor
Department of Industrial & Systems Engineering
Department of Mathematical Sciences (Affiliate)
KAIST, Daejeon, South Korea
Phone: +82-42-350-3128
E-mail: catenoid@kaist.ac.kr

Seung Bum Soh (Co-Researcher)
Assistant Professor
Department of Operations, Decisions and Information
Yonsei University, Seoul, South Korea
Phone: +82-2-2123-6562
E-mail: sbsoh@yonsei.ac.kr

[1] Uncovering Dynamic Preferences and Recommending Game Applications

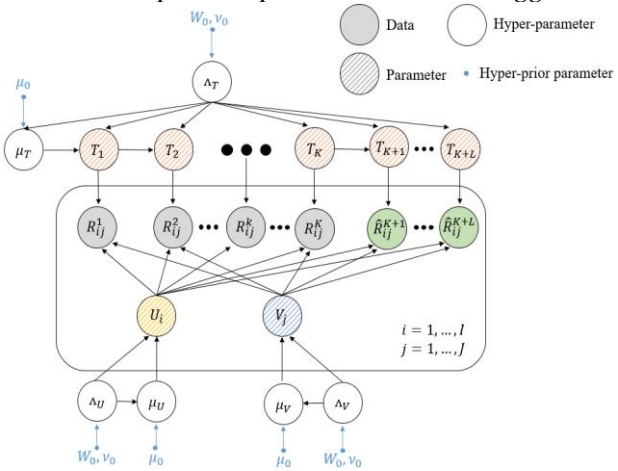
I developed a new recommender system for mobile applications which predicts and uses a future usage pattern. I first defined time-varying preferences to capture the dynamics of user preferences based on the daily usage time extracted from mobile application log data over 2,000 panels and suggested a data-driven customer targeting strategy for managerial purposes.

Scheme 1. Modeling and predicting dynamic preferences developed based on the amount of usage time

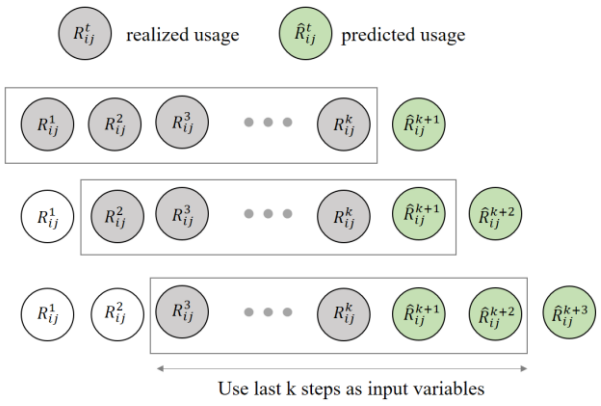


I combined two discrete Bayesian models: 1. Bayesian probabilistic tensor factorization (BPTF) to interpolate and predict for unrealized user-item pairs; and 2. Gaussian process nonlinear autoregressive (GPNAR) – an adaptive nonparametric model – to predict future preferences for "realized" user-item pairs. This research will be submitted to *Management Science* (1st author)[1] in September 2018.

Model 1. Graphical representation of the suggested model (BPTF-GPNAR)



BPTF: Fully Bayesian tensor factorization



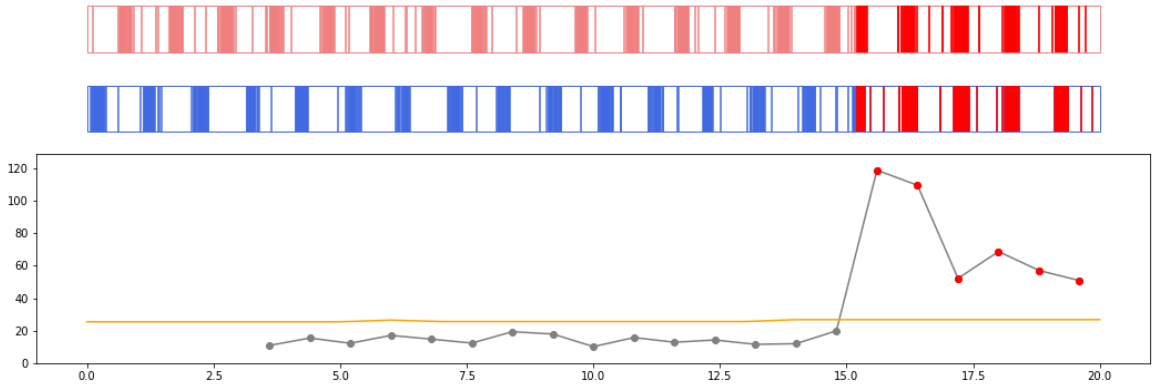
GPNAR: Gaussian process nonlinear autoregressive

[1] **J. Bae** (with J. Park, and J. Choi). "Uncovering Dynamic Preferences and Recommending Game Applications" (under review at *Management Science*)

[2] Frequency-based Change Detection via Multi-Output Log-Gaussian Cox Processes

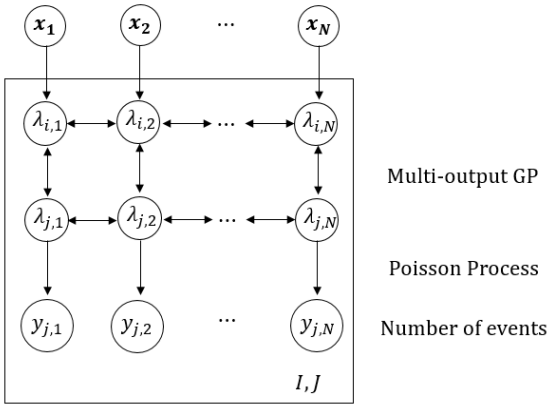
I developed a new change-point detection approach based on the multi-output Log-Gaussian Cox process (MOLGCP), which generalizes Poisson processes to embed correlations among distributed sensors. Adaptive and scalable decision-making strategies were suggested by introducing a receding window for sequential updates. Extreme value theory (EVT) and streaming peaks-over-threshold (SPOT) algorithm were implemented in determining the threshold for detecting change-points.

Scheme 2. Detecting a change-point (out-of-phase → in-phase) using the model (MOLGCP)

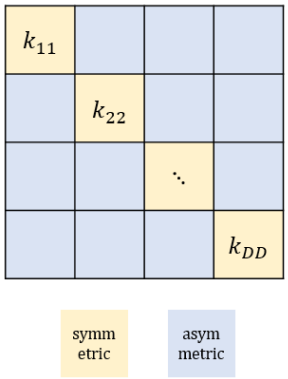


I employed a cross-spectral mixture kernel, which approximates any stationary kernel function with arbitrary precision and also captures interrelated dynamics between channels. The kernel function constructs asymmetric sub-matrices for off-diagonal parts, which enables to capture the temporal and environmental context of a complex/distributed system. In specific, both negative and positive correlations and the phase differences between channels can be explicitly learned by the model. This research will be submitted to *KDD 2019* by February (1st author)[2].

Model 2. Graphical representation of the model and covariance matrix via cross-spectral mixture kernels



[Graphical model for MOLGCP]



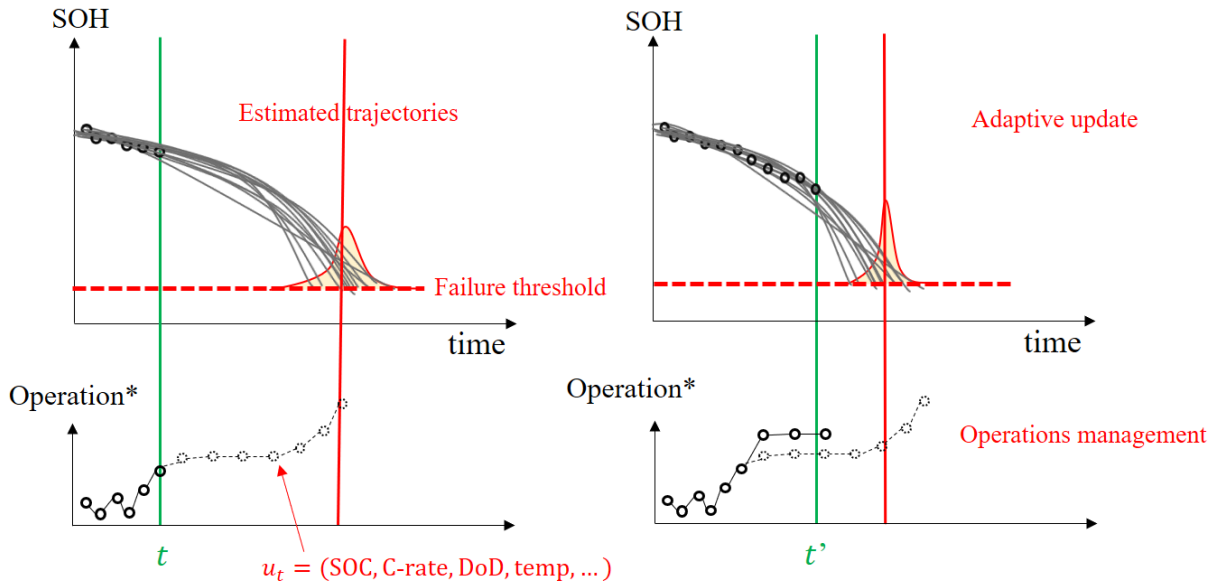
[Covariance matrix for MOLGCP]

[2] **J. Bae** (with J. Park). "Frequency-based Change Detection via Multi-Output Log Gaussian Cox Processes" (to be submitted to *KDD 2019* in February 2019)

[3] Optimal Management of Energy Storage Systems for Wind Turbines

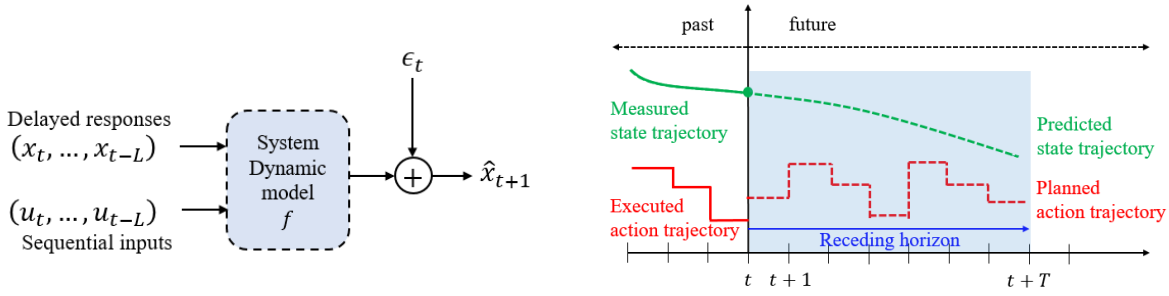
I implemented an adaptive, nonparametric model – Gaussian process nonlinear autoregressive with exogenous variables (GPNARX) – to predict State of Health (SOH) for the Energy Storage Systems (ESS) of wind farms. I used multi-dimensional streaming input variables such as C-rate, State of Charge, Depth of Discharge, and temperature along with the history of SOH, all stemmed from the multi-sensors attached to ESS. I also suggested electric fatigue damage, a new feature extraction that measures cumulative damage to a failure (defined as an event in which SOH drops to 70%). I am extending the model into a hierarchical/multi-output frame to capture the interrelationship between the paralleled batteries.

Scheme 3. Adaptive prediction and operations management for ESS



I will conduct further research to optimize operations for ESS in order to maximize the lifespan of the batteries. In particular, model predictive control (MPC) is used to select the first action and to sequentially update the model by considering future trajectories. This research is still a work in progress.[3]

Model 3. Graphical representation of the suggested model (GPNARX) and model predictive control (MPC)

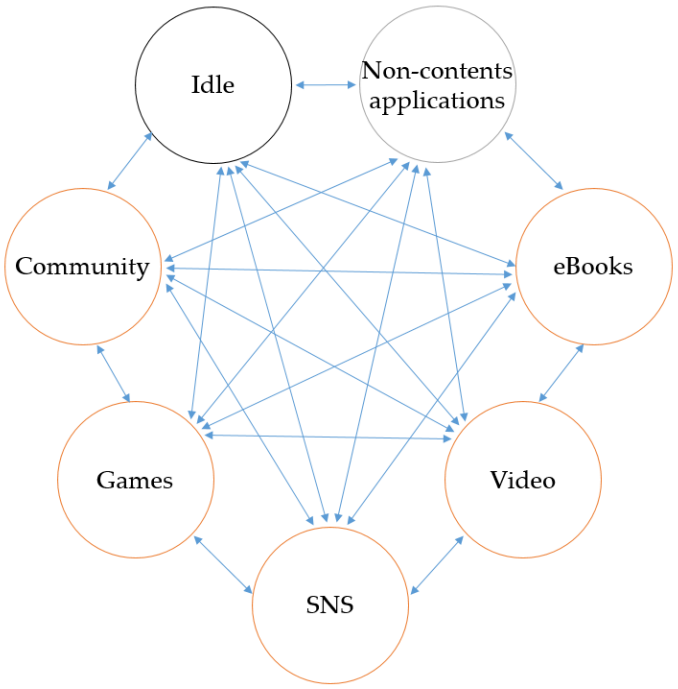


[3] **J. Bae** (with S. Lee, H. Sim, and J. Park). "Optimal Management of Energy Storage Systems for Wind Turbines" (work in progress)

[4] Mobile Contents and Consumer Engagement: Modeling Contents Usage via Stochastic Processes

I am developing a time-heterogeneous continuous-time Markov chain (CTMC) to analyze the usage patterns of mobile applications. In particular, the focal area includes the content industry classified into five sub-categories: community, games, SNS, video, and eBooks. To analyze the frequencies and the transitions between and within the categories simultaneously, I formulated the problem as a time-evolving CTMC model. To consider a user's contextual circumstances, I performed multivariate analysis using network conditions (LTE, 3G, and Wi-Fi) and user-specific demographics along with contextual information such as battery-level, display size, and telecom fee to derive managerial conclusions. This research is still a work in progress.[4]

Scheme 4. Graphical representation of the suggested continuous-time Markov chain model



$a_{ij}(t)=f$ (network(3G, LTE, Wi-Fi), battery-level, display size, telecom fee, demographics, …)

[4] **J. Bae** (with J. Park, S. Soh, J. Choi). "Mobile Contents and Consumer Engagement: Modeling Contents Usage via Stochastic Processes" (work in progress)