NeTS: JUNO2: Collaborative Research:

STEAM: Secure and Trustworthy Framework for Integrated Energy and Mobility in Smart Connected Communities

(2018-2021)

US-Japan
Collaboration

Kickoff Meeting

NICT Innovation Center, Tokyo

October 26, 2018

Our US-Japan Team

Sajal K. Das
(Lead PI)

MISSOURI S&T

Shameek
Bhattacharjee

Abhishek Dubey

VANDERBILT

Hayato Yamana

Waseda University

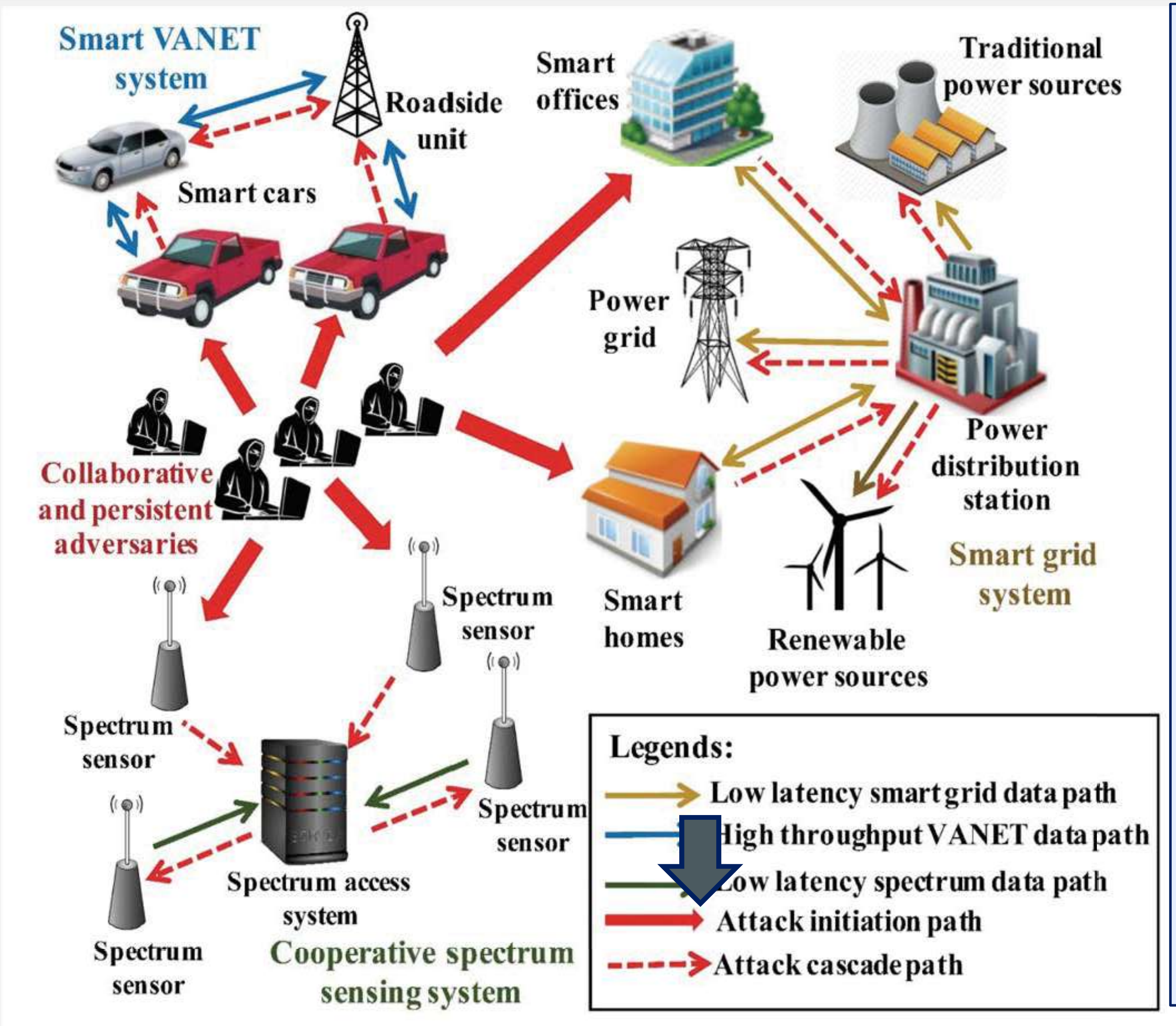
Hirozumi Yamaguchi

大阪大学
OSAKA UNIVERSITY

Keiichi Yasumoto

Nara Institute of Science and Technology
NAIST

Motivation: A Smart City Scenario



Smart Mobility & Energy:

- Interdependent, critical services (e.g., utilities, transportation)
- Infrastructures: **VANET and Smart Grid**

Building and PEV:

- A common link between smart energy systems and mobility.
- A smart car or plug-in vehicle (PEV) is often owned by individuals residing in buildings connected to the smart grid.

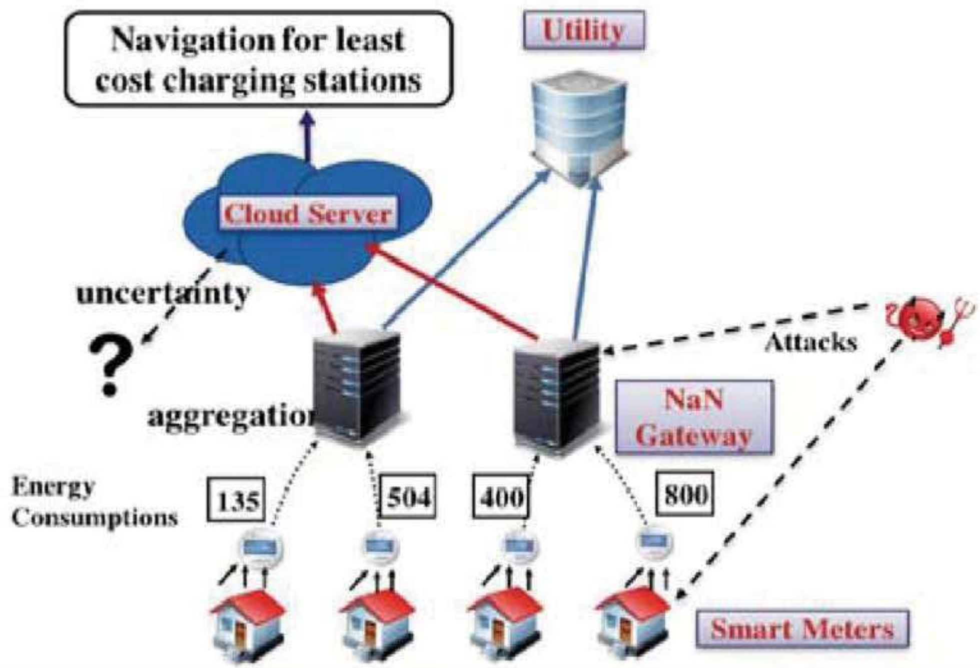
A integrated treatment of smart energy and smart mobility is important.

Integrated Treatment of Energy and Mobility

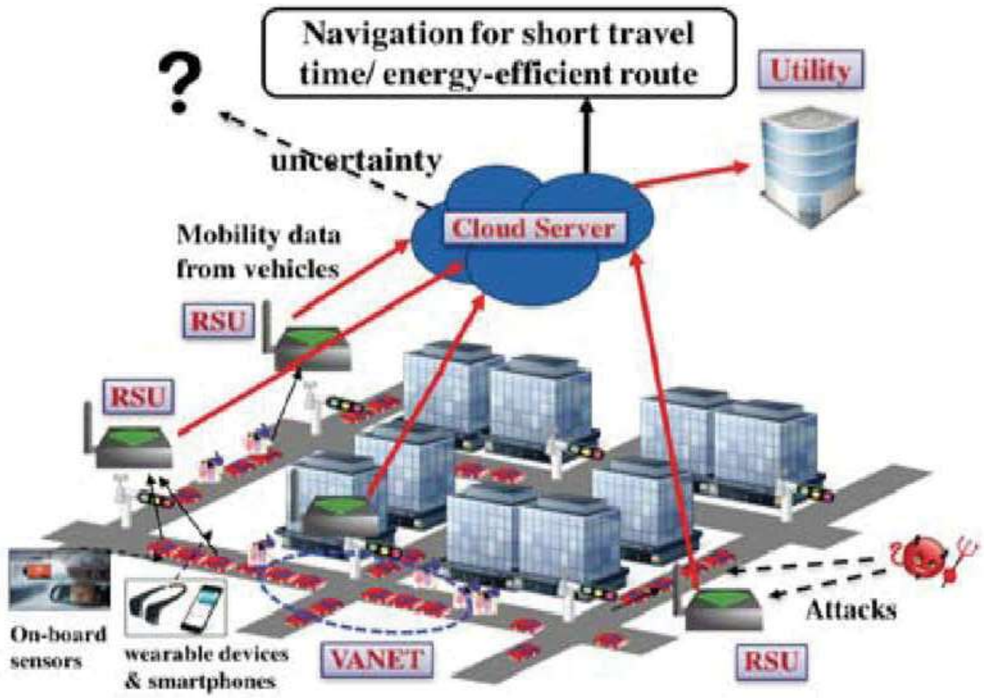
Similarity: Hierarchical structure, spatiotemporal relationships

Difference: Static vs. dynamic, delay-tolerant vs. less delay-tolerant

Smart Energy (AMI)



Smart Mobility (VANET)



STEAM Project

Hypothesis:

- Trust in Smart and Connected Community (SCC) Applications requires efficient mechanisms to handle conflicting goals of identifying anomalous operations while preserving privacy and integrity of the system at scale.

- State of the art literature in this area being nascent, we believe careful co-design and calibration of encryption and robust anomaly detection scheme will lead to significant new advances.

Goal: STEAM aims to develop integrated frameworks, models and algorithms to address security and trustworthiness challenges in smart mobility and energy, considering various threat models.

STEAM Project: Five Thrusts

Thrust 1: Secure and Trustworthy Decision Making under Uncertainty

Thrust 2: Privacy-preserving Computations using FHE (Fully Homomorphic Encryption)

Thrust 3: Characterizing Security, Privacy and Resource Trade-offs

Thrust 4: Developing Secure and Trustworthy Middleware Architecture

Thrust 5: Validation with Real Datasets for Smart Mobility and Smart Energy Applications

Thrust 1: Secure and Trustworthy Decision Making under Uncertainty

Lead: *Prof. Sajal K. Das; Dr. Shameek Bhattacharjee*
(Missouri Univ. of Science and Technology)

Challenges and Goals:

- How to ensure secure and trustworthy decisions across integrated domains of smart energy and smart mobility networks using data that could be manipulated, omitted, or delayed by adversarial behaviors, faults, etc.,?
- How to ensure the secure solution is FHE compatible by design and feasible under resource constraints and fluctuations.

➢ **Proposed Research: Design novel lightweight anomaly detection algorithm, and stochastic trust and robust decision models to minimize wrong decisions under uncertainty and risks.**

Tasks:

1.1 Anomaly Detection

1.2 Trust Models

1.3 Dependable Decision Models

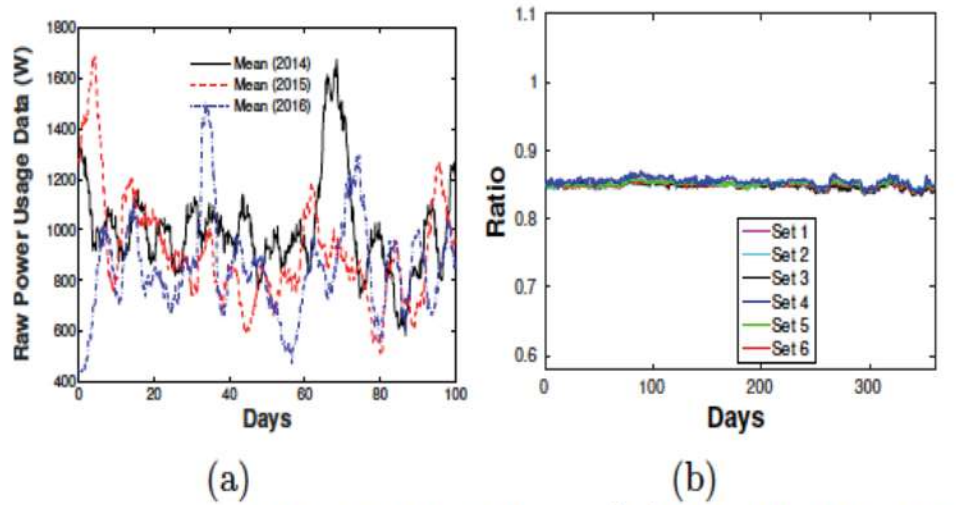
Task 1.1 Anomaly Detection

State of the Art:

- Existing techniques → Classification, State Monitors, Statistical, Clustering and Nearest Neighborhood, or Information Theory.
- Most require complex math operations → not suitable for privacy preserving computations, such as FHE and SMC.
- Need additional monitoring hardware/resources; assumptions on predictable error residuals in data is unrealistic.

Preliminary Results:

- Our study on three-year Texas data reveal that existing approaches for anomaly detection are not effective.
- We showed that Pythagorean mean based metrics (e.g. ratio of Harmonic mean to Arithmetic mean) leads to Stable Invariant, which is lightweight and suitable over FHE based cipher text.



(a) Unstable Mean (b) Stable Metric

Bhattacharjee, Das, et al. ACM AsiaCCS 2018

Task 1.2 Trust Models

State of the Art:

- ❖ Existing trust models include Subjective Logic, Josang's Belief, Information-theoretic Divergence, Dempster-Shafer Belief, and Bayesian Expectation.
- ❖ They are inadequate for SCC applications due to following reasons:
  - Unifying effect of both quantity and quality of (probability inferences) information sources into trust models is challenging.
  - Most of the existing trust models do not provide null invariant measures.
  - Most trust models build evidence from a human/sensor enabled feedback mechanism which themselves could be attacked.
  - Interdependence among SCC applications complicates trust evolution.

❖ **Preliminary Efforts:** Inadequacies and attack models pointed out in [Bhattacharjee, Das, et al. IEEE CNS 2017].

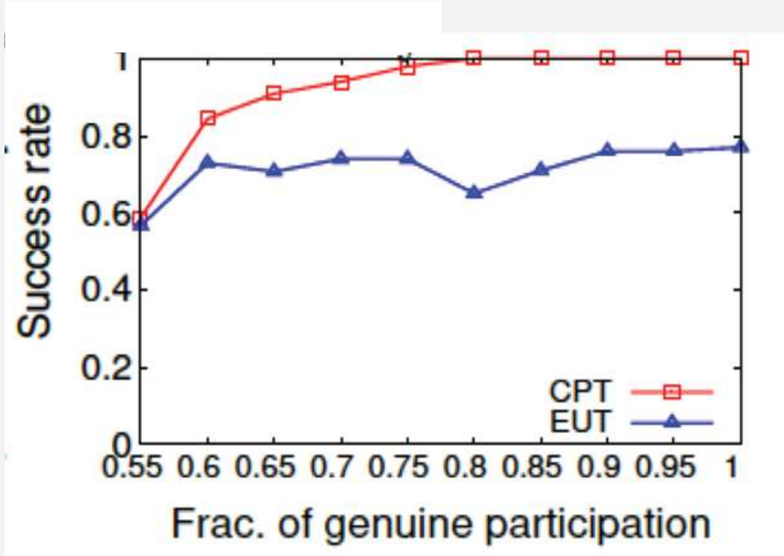
Task 1.3 Dependable Decision Models

State of the Art:

- Dependable decisions combine trustworthiness of components, with prior or contextual probabilities of outcomes (uncertainty) and associated payoffs (risks) for those outcomes.
- Expected Utility Theory and Decision Tree are commonly used techniques.
- Recent studies in behavioral decision theory and economics have shown that these approaches do not reflect or model well realistic behaviors under risk and uncertainty, often yielding sub-optimal results.

Preliminary Results:

- We exploited belief manipulation attacks using built knowledge against SCC, leading to higher errors under expected utility theory (EUT).
- Cumulative prospect theory (CPT) improves the success rate of publishing accurate decisions in vehicular crowdsensing.



Thrust 1: Proposed Approach

- Develop novel lightweight metrics to provide more stable invariants in unpredictable SCC systems by abstraction of correlation structures, robust statistics and subsampling methods.
- Establish trustworthiness of information or component with a local and short term scope termed as Topical Trust, Reputation Trust, Kernel Methods, Trust Management and update modules.
- Design simple utility value function, probability weighing function, and stochastic choice function to handle behavioral decisions under risk in transactive and smart energy, V2V and V2I settings.
- Case Studies: Irish Dataset (ISSDA), Texas Dataset (PeCan Street), City of Nashville (Vanderbilt Univ.), Waze App Dataset (Boston Municipality)

Thrust 2: Privacy-preserving Computations using FHE

Lead: *Prof. Hayato Yamana* (Waseda Univ.); *Shameek Bhattacharjee*

Challenge and Goal:

- Given the anomaly detection requires contextual information (e.g., smart meter data, energy pricing bids, smart vehicle trajectory), how to ensure computations and analytics that preserve privacy of end customers?

➢ Efficiently adopt Fully Homomorphic Encryption (FHE) with anomaly detection and trustworthy decision making at community scale, yet to be realized.

Tasks:

2.1 FHE Calculations with Table Search

2.2 Handling Range Search

2.3 Applying FHE to Secure Decisions

Thrust 2: Current Status

State of the Art:

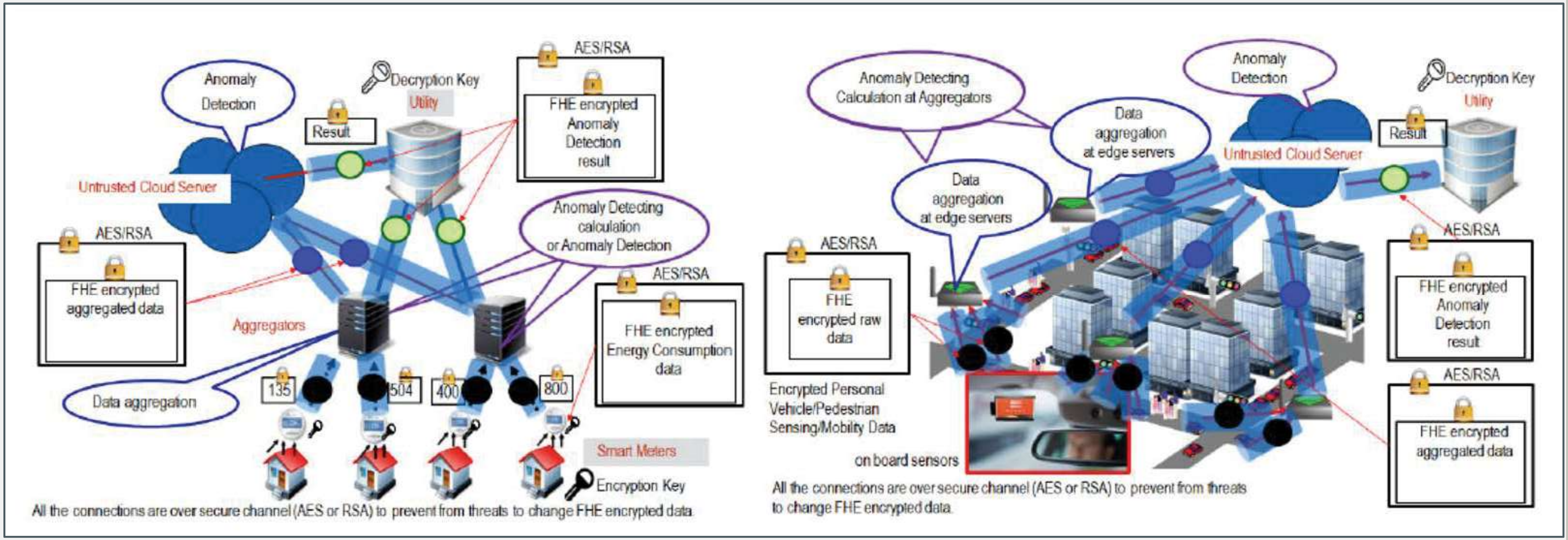
- Among all encryption schemes, only FHE enables additions and multiplications over encrypted data. By adopting FHE for calculation in Thrust 1, the individual raw data are not exposed to others.
- Drawbacks of existing FHE that prevent their adoption in SCC:
  - Too high complexity (space and time) → too slow calculations.
  - Lack of support for division and other rational number arithmetic operations (e.g., range search, logarithm).

Preliminary Work:

- Our efficient and secure frequent pattern mining scheme has smaller time and space complexity compared to state-of-the-art schemes: 430 times speed up and 94.7% reduction of memory usage.

Thrust 2: Proposed Approach

- Encrypt raw data at each communicating link and transfer to an aggregator over secure channel (RSA or AES), to prevent attackers from changing the FHE encrypted data. Only the utility provider can see the aggregated results.
- Implement anomaly detection in Thrust 1 by table search mechanism which does not require division operation.



Thrust 3: Security, Privacy and Resource Trade-offs

Lead: *Dr. Abhishek Dubey* (Vanderbilt University)

Challenge and Goal:

- Given that FHE is computationally slow and resource intensive, how to efficiently ensure security and privacy for resource-constrained SCC applications deploying IoTs and edge computing?

➢ Define metrics for analyzing trade-offs between resource utilization and timeliness (responsiveness) vs. security, privacy and trustworthiness in SCC decision support systems.

Tasks:

3.1 FHE Parameter Configuration

3.2 Anomaly Detection Threshold Configuration

3.3 Critical Sensors Identification

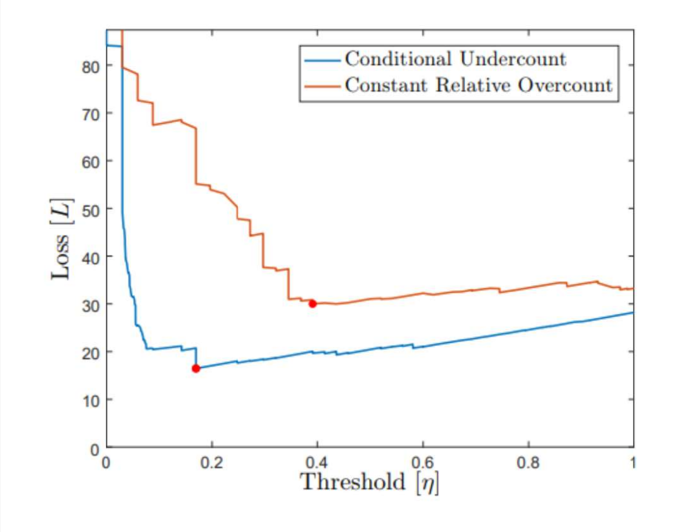
Thrust 3: Current Status

State of the Art:

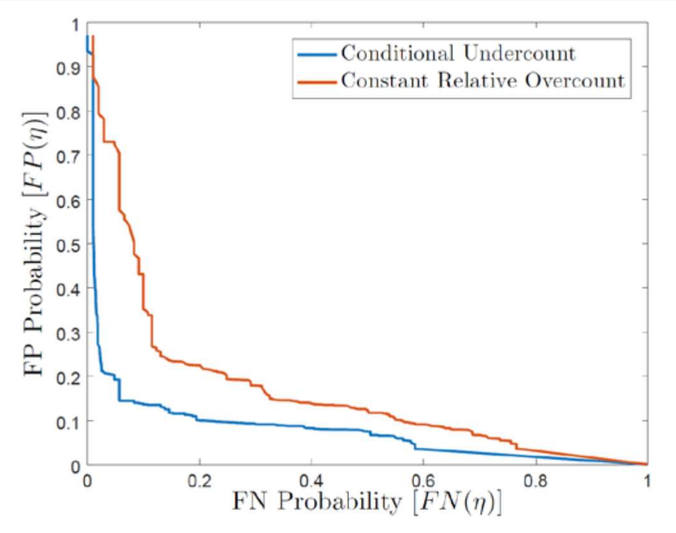
- SCC applications are typically data-driven, where sensors collect data at unprecedented scale and granularity. Since data privacy and integrity are two major concerns with such data, addressing their trade-off is crucial.
- Performance impact of state of the art methods on SCC depends on the choice of detection thresholds, mainly due to trade-offs between false-positive and false-negative rates.

Preliminary Work:

- Attacker- defender model for power systems: A strategic attacker tries to maximize the damage by identifying worst-case scenario (subset of attacking substations causing maximum damage) [PHM 2017, Resilience Week 2017, ISGT 2018]
- Application anomaly detection threshold configuration: We showed how to configure thresholds based on the impact on total travel time in a transportation scenario. [Scope 2017]



Application Impact vs Threshold

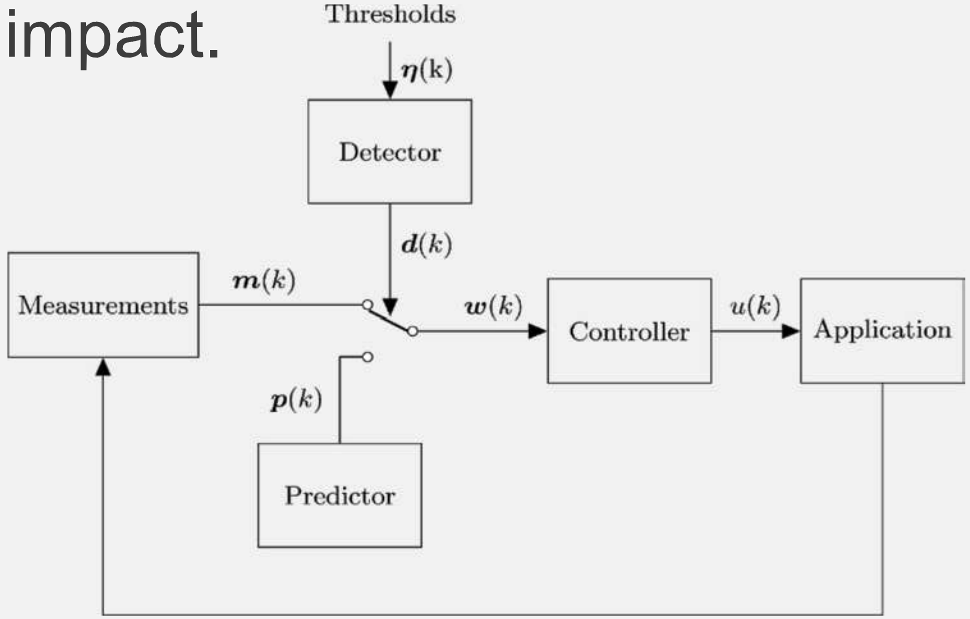


ROC curve – Precision vs Recall

Thrust 3: Proposed Approach

- Develop mathematical models for evaluating trade-offs between different configurations for anomaly detectors, each resulting in different detection performance and security overhead.
- To improve security of privacy-protected SCC application in presence of strategic adversary, design a bad-data detection system anticipating potential malicious attacks, while considering application-specific performance impact.

$$\boldsymbol{\eta}^* \in \underset{\boldsymbol{\eta}}{\operatorname{argmin}}\, L(\boldsymbol{\eta})$$



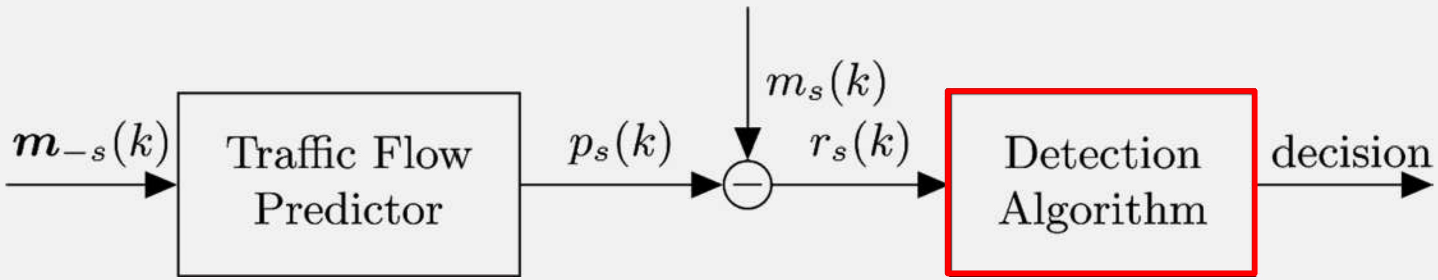
Example: Application Aware Anomaly Detection Thresholds

- Given normalized residuals $z_s(k) = \frac{m_s(k) - p_s(k)}{\sigma_s(k)}$, the upper and lower CUSUM statistics are

$$U_s(k) = \max(0, U_s(k-1) + z_s(k) - b_s) \quad L_s(k) = \min(0, L_s(k-1) + z_s(k) + b_s)$$

- For detector with **threshold η(k)**, an **alarm is raised** if

$$U_s(k) > \eta_s(k) \text{ or } L_s(k) < -\eta_s(k)$$



- Derive cost of false alarms (false positives and negatives), total performance loss, and optimal detection threshold.

Thrust 4: Secure and Trustworthy Middleware Architecture

Lead: *<u>Prof. Keiichi Yasumoto</u>* (NAIST); *Dr. Hirozumi Yamaguchi* (Osaka)

Challenges and Goal:

- How to build multi-domain architecture for smart mobility and smart energy?
- How and where to implement computations related to privacy, security, trust?
- What are computational/resource challenges for scalability?

➢ Propose a novel middleware framework that distributes security features across tasks and incorporates privacy, trustworthiness, resource constraints, and distributed decision support.

Tasks:

4.1 Distributed Aggregation

4.2 Secure Anonymization

4.3 Decision Making under Trade-offs

Thrust 4: Current Status

State of the Art:

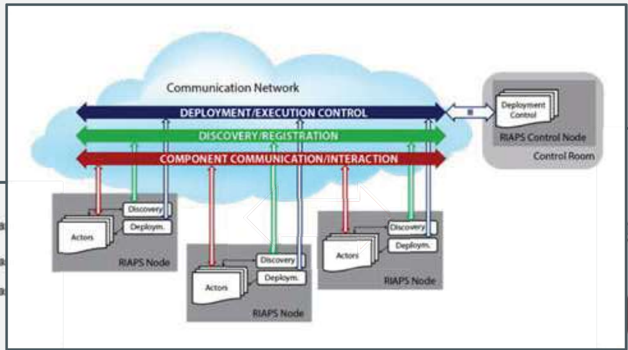
- Delay-awareness is essential to provide up-to-date mobility and energy information in a smart city for navigation services (e.g., movement and charging of PEVs).
- Edge/fog computing-based distributed architecture can minimize delay and wastage of bandwidth/resources for cloud servers.
- No attempt has been made on the edge-based solution for trustworthy services in SCC.

Preliminary Work:

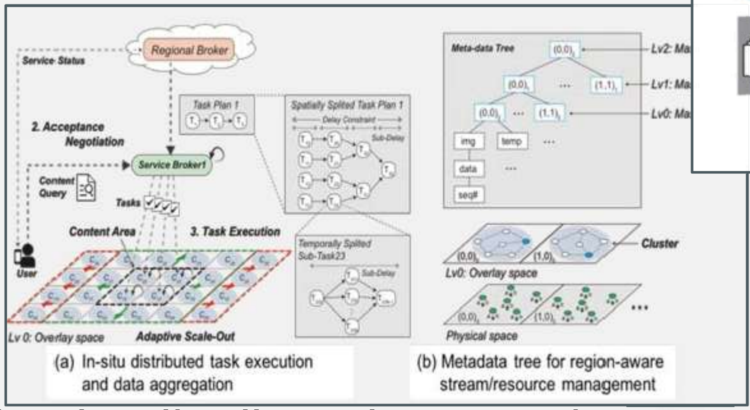
- IFoT framework: aims at mechanisms for unified treatment of IoT devices, distributed and delay-aware processing, intelligent content creation [ACM/IEEE SEC'18].
- EdgeCEP framework: realizes highly-adaptive complex event processing by dynamically forming mesh-like network and distributing task specifications written in an event-based language over the devices [IEEE DCOSS'17].
- RIAPS (Resilient Information Architecture Platform for Decentralized Smart Systems)
  - A component-based decentralized software platform for realizing smart grid infrastructure (https://riaps.isis.vanderbilt.edu/)

Thrust 4: Proposed Approach

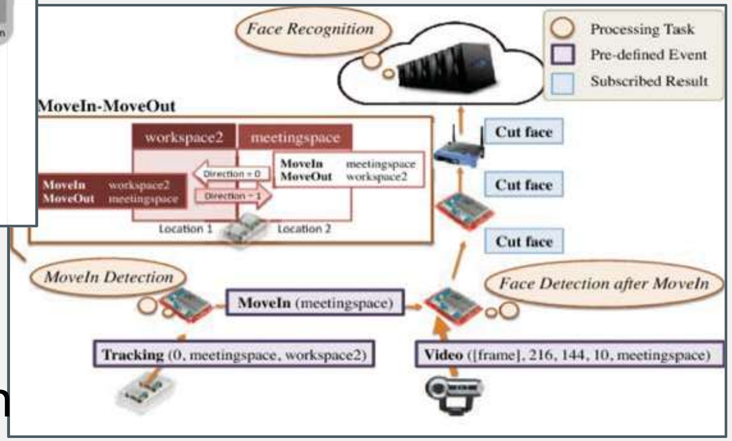
- We design and implement a novel edge computing middleware platform to aggregate user data in a secure and timely manner.
- On top of IFoT, EdgeCEP and RIAPS, we design and implement fully-distributed, lightweight anomaly detection, FHE computation & secure decision making



RIAPS for distributed smart grid design/ implementation



In-situ distributed processing with IFoT



Distributed CEP processing with EdgeCEP

FHE computation

Anomaly Detection

Decision Making

Thrust 5: Validation with Exemplar SCC Applications

Lead: *Dr. Hirozumi Yamaguchi* (Osaka); *Abhishek Dubey* (Vanderbilt)

Challenge and Goal:

- How to evaluate proposed models, solutions and resource provisioning architecture for trustworthy computations and control?

➢ Validate the proposed models and approaches using smart mobility and smart energy distribution/consumption scenarios with real-world datasets.

Tasks:

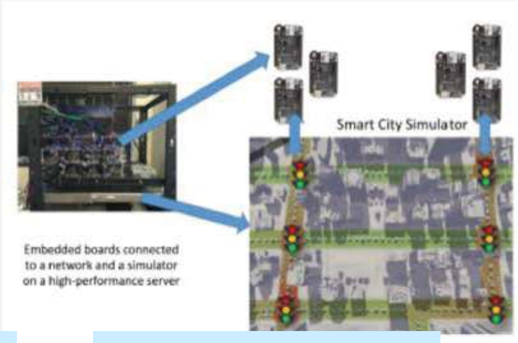
5.1 Smart Transportation Application

5.2 Smart Energy Application

Thrust 5: Testbeds and Datasets

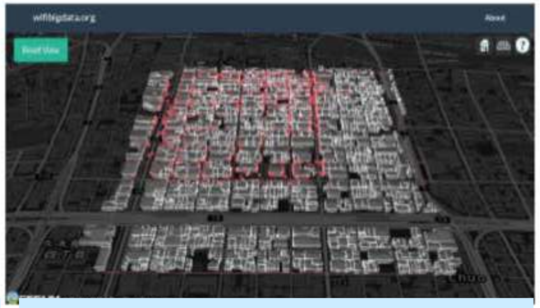
Thrust areas will be validated with real and simulated testbeds, data

PEV and charging facility



HW in the loop simulations

RSUs in Osaka city



3D Visualization WiFi access points in Osaka

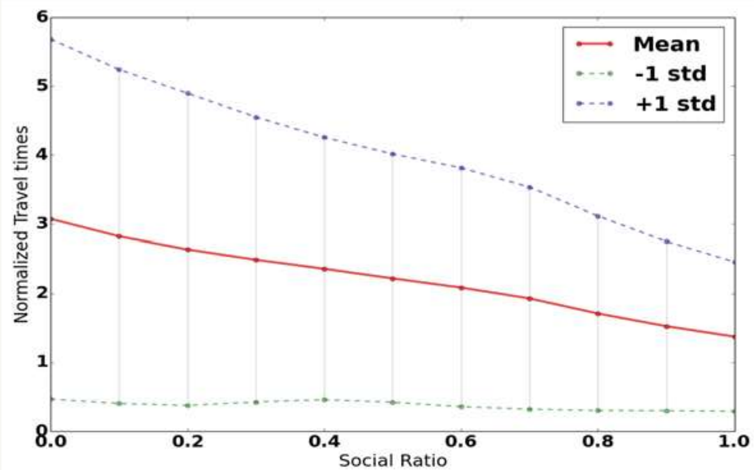
Available Testbeds and Datasets:

- **Missouri S&T:** Micro-grid and Smart Village. Access to huge smart meter datasets at thousands of households from Austin, Texas and Dublin, Ireland for over 3 years with hourly granularity.
- **Vanderbilt Univ:** A facility for hardware in the loop simulations on a 64-node beaglebone cluster integrated with SUMO, Opentrip planner, Matsim, GridLab-d and Opal-RT. Datasets include traffic speed, traffic incidents, weather conditions, and bus schedules in Nashville for last 3 years.
- **Waseda Univ:** High performance compute clusters for running large scale FHE calculations.
- **NAIST:** A small PEV testbed with three PEVs and two parking lots with charging facility.
- **Osaka Univ:** Access to VICS data via RSUs deployed in Osaka city as well as simulation testbed for real traffic in Sapporo, using the taxi probing data.

Thrust 5: Proposed Approach

Smart Transportation Application

- Integrated simulation framework to study effectiveness of social algorithms for multi-modal routing
  - Suggest travel routes considering both spatial and modal variations (bus, bike, car, walk, etc.)
  - Evaluate shared effect of riders on the overall efficacy of mobility services
- Extend simulator by integrating anomaly detection and trustworthy decision making algorithms to prevent anomalies/leakage of private location information.



Travel time and variance decreases as social ratio increases

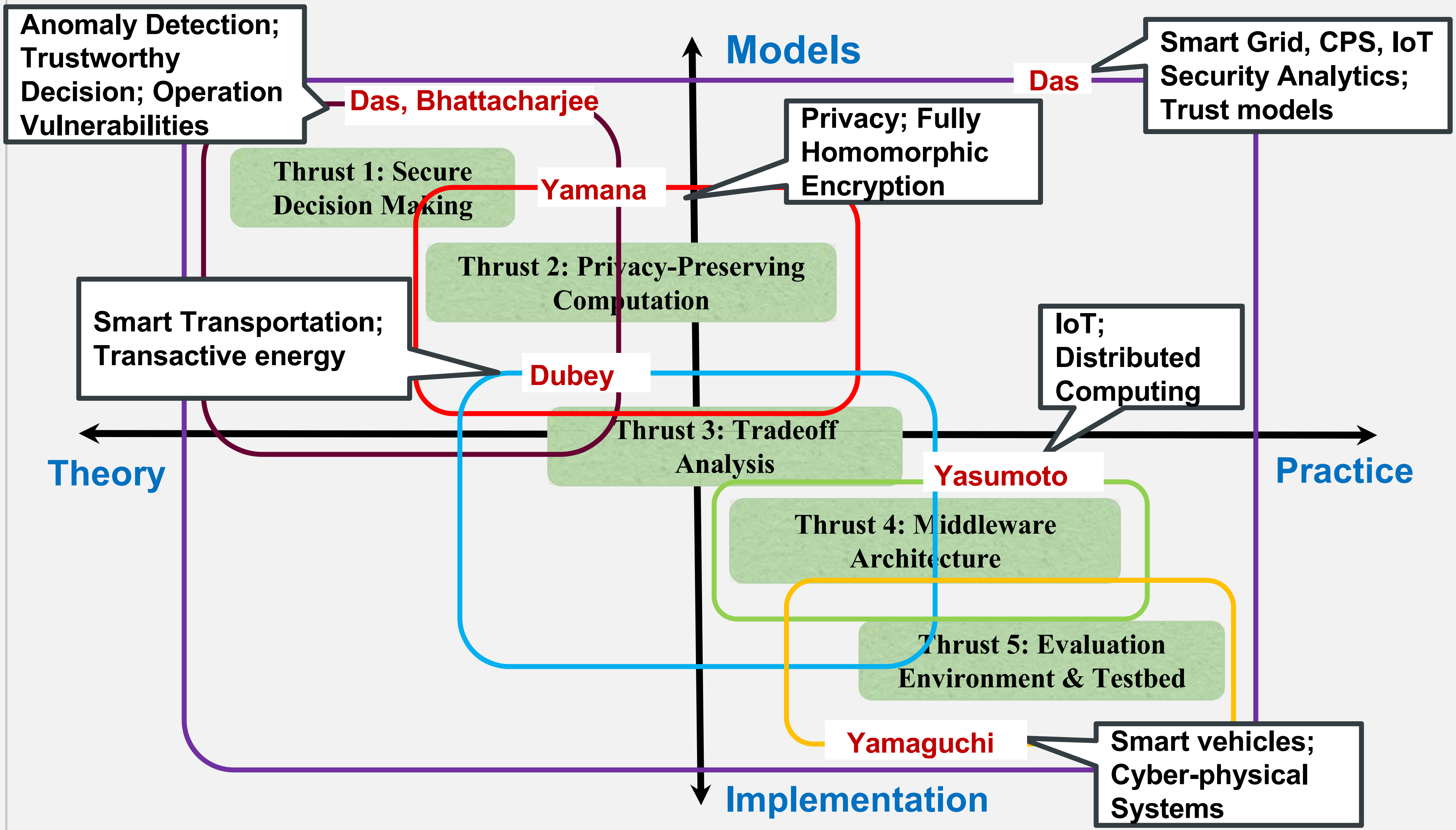
Smart Energy Application

- Distributed transactive energy work flow to enable energy trading in a secure and verifiable manner yet protecting privacy leakage about user's future energy consumption.
- Susceptible to attacks that compromise the integrity of smart meters

→ incorporate the transactive energy scenario in charging stations for PEVs as consumers

- Implementation will be tested both in simulation and on the PEV testbed

Roles and Responsibilities of PIs



Timeline and Project Schedule

M: Missouri &T; **W:** Waseda Univ.; **V:** Vanderbilt Univ.; **N:** NAIST; **O:** Osaka Univ.

Tasks	**Sep 2018**	**2019**	**2020**	**Aug 2021**
1.1 Anomaly Detection (M) 1.2 Trust Models (M) 1.3 Dependable Decision Models (M)				
2.1 Calculations with Table Search (W) 2.2 Handling Range Search (W) 2.3 Applying FHE to Secure Decisions (W)				
3.1 Privacy Threshold Configuration (V) 3.2 Dynamic Threshold Configuration (V) 3.3 Critical Sensors Identification (V)				
4.1 Distributed Aggregation (N) 4.2 Secure Anonymization (N) 4.3 Decision Making under Trade-offs (O)				
5.1 Smart Mobility Application (O) 5.2 Smart Energy Application (V)				

STEAM Project: Broader Impacts

Scientific Impact:

- Secure and trustworthy foundations for cyber-physical systems (CPS), IoT and smart connected communities (SCC), in particular smart mobility and smart energy applications
- Proposed lightweight solutions have potential for technology transfer

Education and Student Training:

- Unique platform for interdisciplinary education and experiential learning for undergraduate and graduate students, inspiring them to pursue higher studies and careers in STEAM fields
- PIs will integrate research findings into courses they teach (CPS, WSN, data and applications security)

New Course Development:

- Jointly offer a course on *Securing Smart and Connected Communities* via distance learning

Workshop Organization:

- Das and Yamana established the Big Data and IoT Security in Smart Computing (BITS) workshop in conjunction with IEEE SmartComp and had a panel on Security and Trustworthiness in SCC in 2018
- BITS Workshop will be organized in 2019 and beyond, expanding its scope with smart mobility and energy topics

Dissemination:

- Outcomes will be disseminated through a dedicated website, keynote talks, seminars and tutorials
- Results will be published in top-tier conferences (IEEE INFOCOM, NDSS, Security and Privacy; ACM CCS, SenSys, BuildSys, CPSWeek) and high quality journals (IEEE TMC, ToN, TDSC; ACM TCPS, TOSN)
- Data generated from experimental testbeds (smart energy simulation with real smart meter data and smart mobility experiments with traffic data) will be made available to research community in anonymized fashion

Collaboration and Management Plan

Coordination plan:

September 2018 meeting at MST

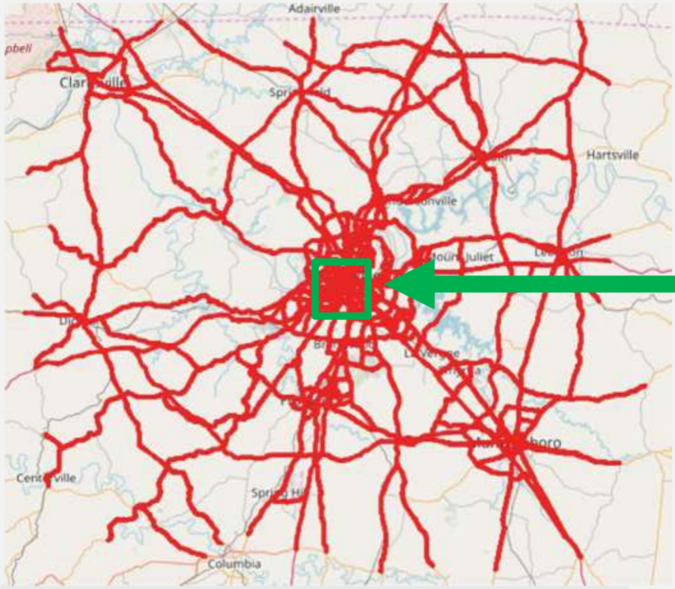
- Weekly Skype meeting
- Reciprocal faculty visits and student exchanges across partner institutions, and PhD student co-advising
- High quality publications coauthored by PIs
- Initial kickoff meeting on September 14-15, 2018 at MST
- PIs meeting at Vanderbilt Univ (June 2019) and at Osaka / Waseda Univ (June 2020)
- Offering joint short courses, project wiki, organizing workshops and special issues

Inter-institutional Student Visits:

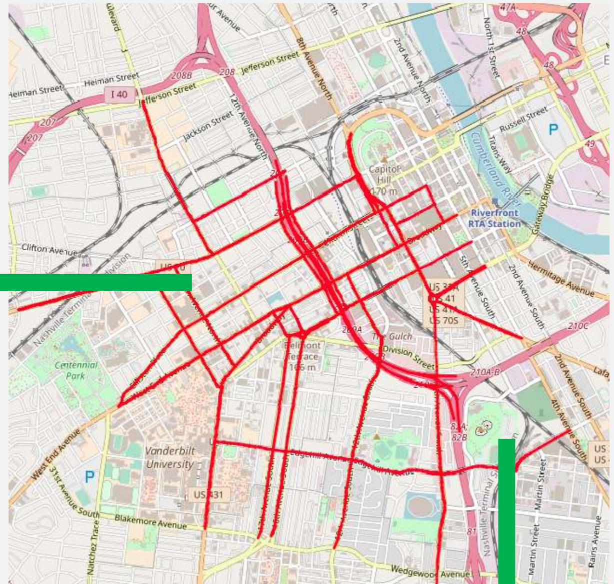
Students working on the STEAM project will visit the partner institutions for one week (during summer) for immersive collaboration experience.

- ❖ Design of security and trust models in SCC applications (e.g., smart grid and mobility) at MST
- ❖ Hands-on experience on the smart grid and smart mobility testbeds at Vanderbilt Univ.
- ❖ Learn about FHE based security schemes at Waseda Univ.
- ❖ Implement on smart transportation testbed at Osaka Univ. and NAIST

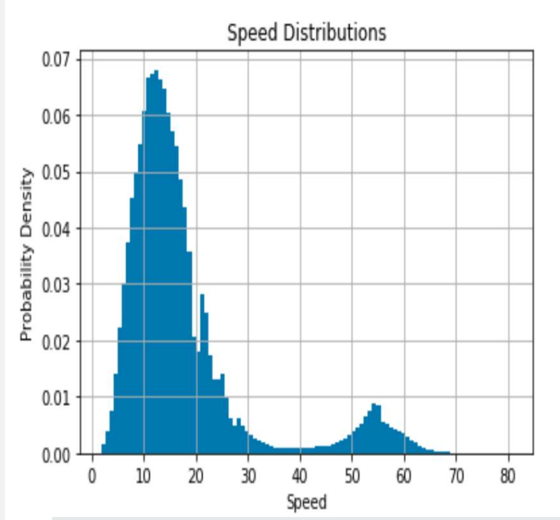
Current Progress



3 years of multivariate traffic data @1m from Nashville

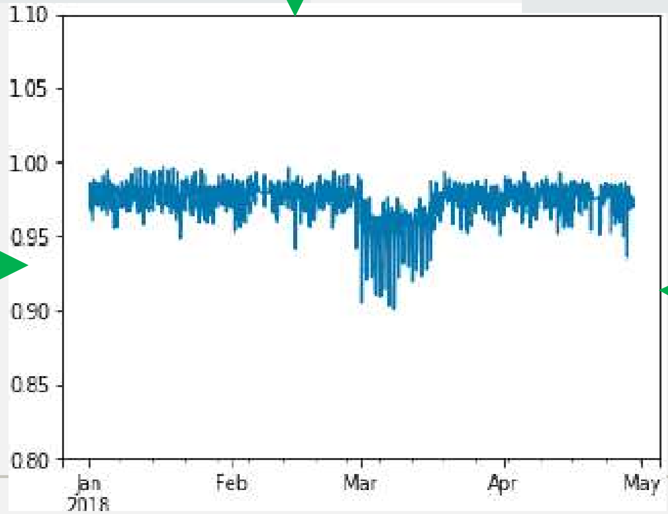


Focusing on Zones: Downtown Nashville > 100 sensor



Transform to Gaussian Distribution. Compute HM/AM ratio for samples in the zone @1hour interval

- We have developed deductive attack signature using Pythagorean mean metric
- We have built a large scale Matsim based simulation platform to study the impact of false data injection
- A distributed query and analysis framework has been developed to handle the 3 TB of sensor data.



All Hands Meeting + Planned Activities

- Missouri S&T: Sept 14-15, 2018 (Two full days)
- Tokyo: Oct 26, 2018 morning and Oct 27 morning
- Kyoto: March 11-15, 2019 (during IEEE PerCom 2019)
- Vanderbilt Univ: June 17-18, 2019 (after IEEE SmartComp 2019)
- Waseda Univ / Osaka Univ: 2020 (TBD)
- PhD Committees: PIs will serve on PhD committees across institutions
- Vision (Magazine) Paper: on integrated energy and transportation in IEEE Computer, IEEE Pervasive Computing or IEEE Communications
- Publication Venues: ACM ICCPS, IEEE PerCom + Workshops, ISORC, IEEE SmartComp, ACM Middleware, ASIA(CCS)
- Special Issue Organization (Magazine and/or Journal)

JUNO2: US-Japan Collaborative Project

STEAM: Secure and Trustworthy Framework for Integrated Energy and Mobility in Smart Connected Communities

PI Meeting – Chicago, October 11, 2019

Our US-Japan Team

MISSOURI S&T

Sajal K. Das

WESTERN MICHIGAN UNIVERSITY

Shameek Bhattacharjee

VANDERBILT

Abhishek Dubey

Hayato Yamana

Waseda University

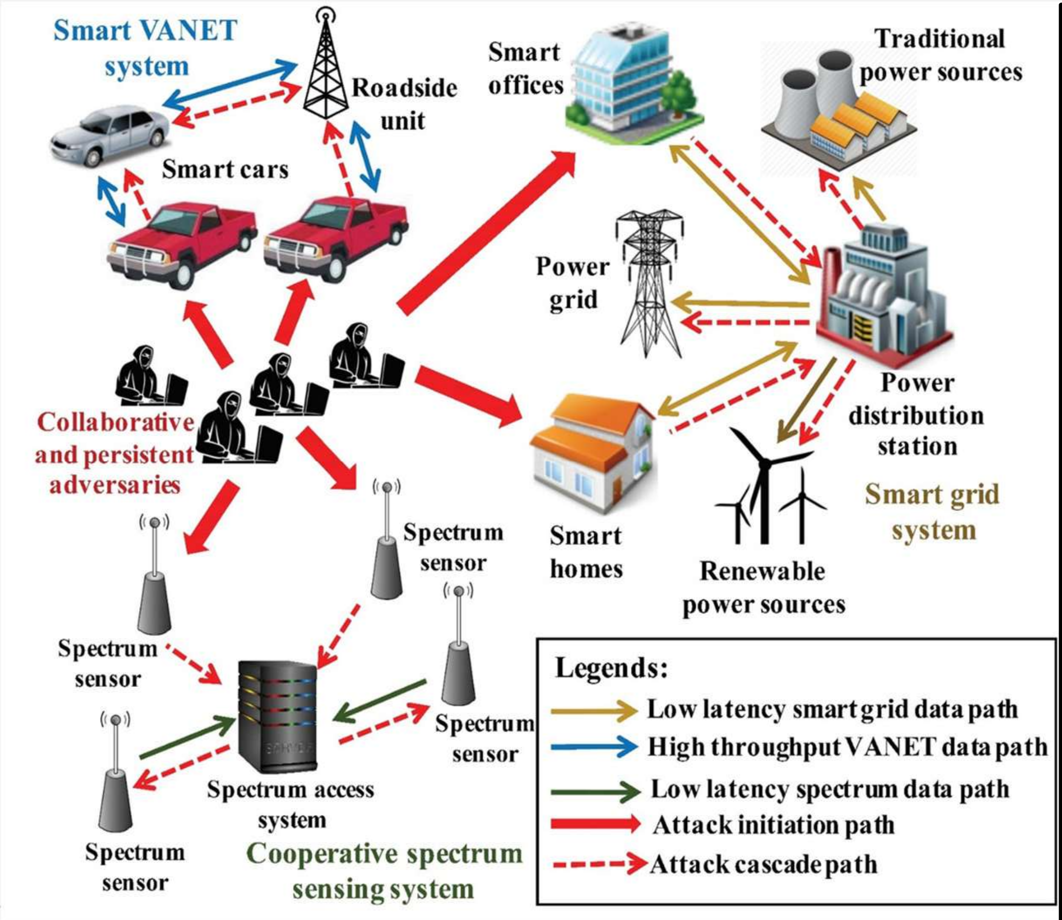
Hirozumi Yamaguchi

大阪大学 OSAKA UNIVERSITY

Keiichi Yasumoto

Nara Institute of Science and Technology NAIST

Security in a Smart City Scenario



Smart Mobility and Smart Energy:

- Interdependent, societally critical CPS networks
- Ensuring safety, resiliency and privacy preservation

Unique Challenges in Securing S & CC

- Large variations in individual end point data due to behavioral and activity differences.
- No strict stationarity over time or space.
- Heterogeneous time granularities of sensing and network sizes.

Goals and Novelty of STEAM Project:

- Develop integrated frameworks, algorithms and models to address security, dependability and trustworthiness challenges in mobility and energy under various threats.
- Design lightweight resilient anomaly detection and privacy preserving encryption schemes & middleware architecture.
- Trust building in S & CC applications; efficient mechanisms to handle conflicting goals of identifying anomalies; trade-off between security, privacy and integrity at scale.
- Efficient co-design and calibration of encryption and robust anomaly detection schemes.

STEAM Project: Year 1 Progress Report

Thrust 1: Secure and Trustworthy Decision Making under Uncertainty

- Fast, hierarchical, efficient, and accurate decentralized anomaly detection methodology for streaming transportation sensors using Pythagorean means and long short-term memory (LSTM) networks
- Fast and accurate detection of compromised smart meters under temporally distributed stealthy attacks for smart energy networks
- Algorithms for detecting components with attack margins much below the standard deviation of data with high accuracy.

Thrust 2: Privacy Preserving Computations using Fully Homomorphic Encryption (FHE)

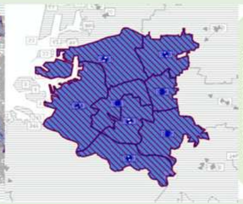
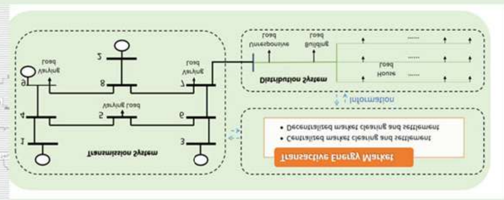
- Developed efficient approaches for enhancing FHE to execute privacy preserving decisions that require complex calculations.
- **Approach 1:** Table lookup with a non-colluding server to adopt any kinds of calculations at aggregators for higher speed-up
- **Approach 2:** Approximate Homomorphic Encryption scheme (HEAAN) to leverage floating-point arithmetic (e.g., log computation) over encrypted data

Thrust 4: Developing a Secure and Trustworthy Middleware Architecture

- A novel middleware architecture to assign tasks over IoT devices (e.g., RSUs, smart meters) taking into account the required quality of service (QoS) levels
- Implementation and evaluation of a prototype middleware using Docker technology for easy deployment
- Development of a Smart Transportation Service Emulation Testbed
- Defining an optimal task assignment problem and developing an algorithm with federated learning

Thrust 5: Validation With Real Datasets for Smart Mobility and Energy

- Large scale road traffic data collection from Osaka and Nashville
- An integrated energy simulation testbed development for experimenting with integrated mobility and energy scenarios

Progress on Thrust 1:

(Shameek Bhattacharjee, Sajal Das, Abhishek Dubey)

Secure and Trustworthy Decision Making under Uncertainty

Tasks:

1.1 Lightweight Anomaly Detection
1.2 Stochastic Trust Models
1.3 Dependable Decision Making

Trustworthy and Efficient Anomaly Detection

Threat Model: Orchestrated attacks on a collection of sensors to maximize the effect of attack on the global transportation system

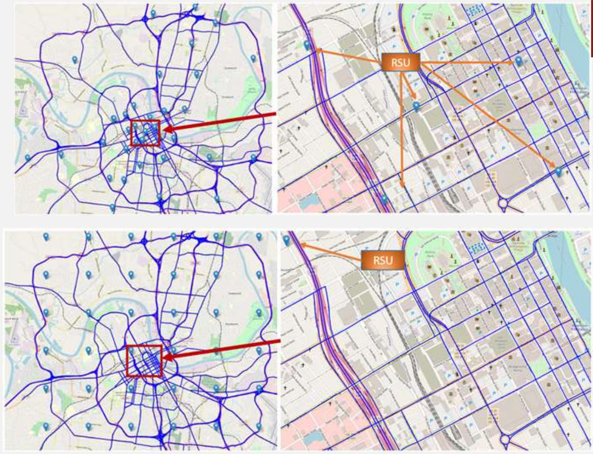
Tried on real data from Nashville, TN

Optimal RSU Placement: designed for optimal anomaly detection using ROC curves

Macro Model: designed for large scale decentralized anomaly detection in real time

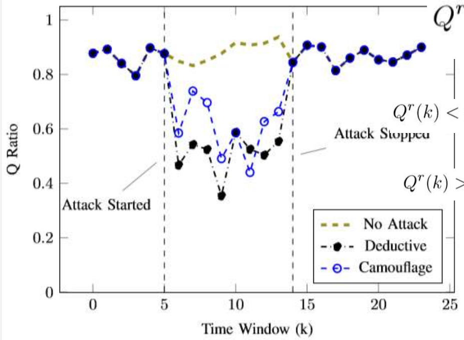
Micro Model: highly accurate and fine grained-anomaly detection. Computationally intensive

Congestion Progression: how the effects of anomalies propagate through the network



Optimized (for detection) RSU Deployment

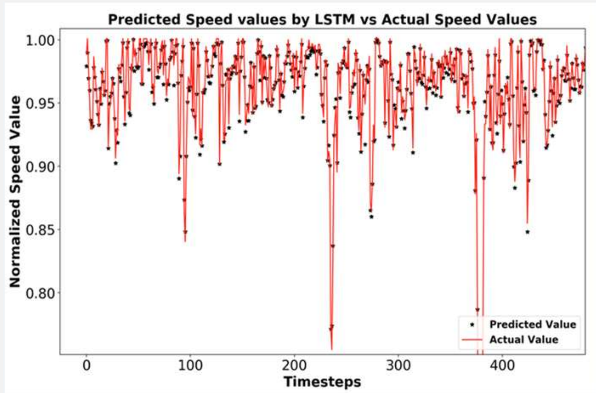
Data Collection	Zone Level Detection	Sensor Level Detection	Network Level State Estimation
Speed data sent from sensor to its corresponding RSU.	Light-weight anomaly detection run at the RSU level. Statistical means approach. Purpose: identify orchestrated data-integrity attacks.	LSTM based detection run at the centrally located cloud on high-capacity detection nodes. Identify which sensors are compromised by data-integrity attack. Only runs when attack is identified at the zone level.	If a real-incident is detected then then congestion propagation framework is used to identify the future affects in short-term. Then the state of the transportation network can be used for optimal routing. LSTM Networks are used to mitigate the anomalous information if no physical incident was identified



$$Q^r(k) = \frac{HM^r(k)}{AM^r(k)}$$

$$Q^r(k) < Q^r_{ave}(k) - \epsilon^r * Q^r_{std}(k)$$

$$Q^r(k) > Q^r_{ave}(k) + \epsilon^r * Q^r_{std}(k)$$



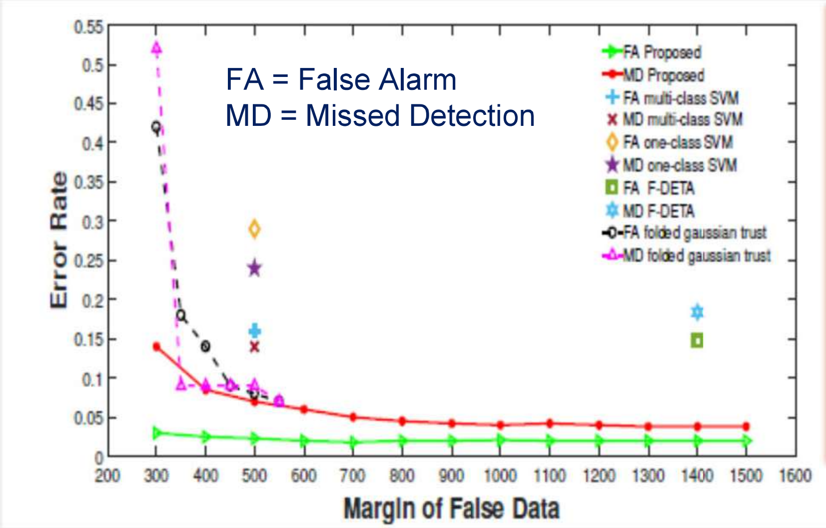
Macro-Level Detection: Built efficient streaming statistical algorithm

Micro-Level Detection: Efficient long short-term memory (LSTM) based traffic predictor by modeling each road segment in large scale traffic network as a function of neighboring roads.

Attack Context Embedded Trust Scoring Models

Challenge: *Fast* and *accurate* detection of Compromised Smart Meters under temporally distributed stealthy data falsification attacks.

- **Introduction of Attack Responses as Robust Statistical Measures**
  - Pythagorean Means and Real Analysis
  - Median Absolute Deviation
  - Location Parameter Correction
  - Attack Probability Time Ratio
- **Embedding of Responses** → magnify divergence in probability space for information theoretic detection
- **Magnified Divergence** → high detection accuracy, reduced false alarms, decreased convergence time under stealthy attacks
- **Multi-granular anomaly based attack detector → across temporal scales → better unsupervised threshold design**



Key improvement over existing works

- *Detects meters with attack margin > 300W*
- *FA < 10% for large sized grid*
- *Micro-grid level detection of attack presence ~ 100W*

Key Theory → Schur Ostrowski Criterion

$(x_i - y_i)\left(\frac{\partial f}{dx_i} - \frac{\partial f}{dy_i}\right) \geq 0$ → *Schur Convexity*

Products: (1) *ACM Trans. on Privacy and Security,* minor revision, Oct 2019
(2) *IEEE Trans. on Dependable and Secure Computing,* to appear

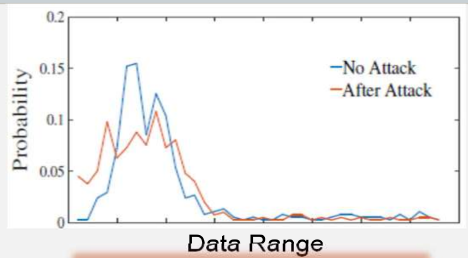
Broader Impacts:

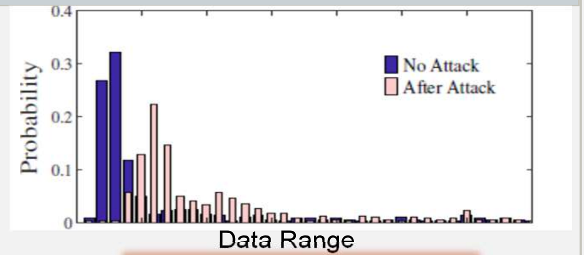
- Includes closed form approximations and performance limits of robust statistics under various attacks
- Validated across big datasets from Texas (800 meters) and Ireland (5000 meters)

Diversity Index Trust Scoring

Problem and Challenges: Detect Components with attack margins much below the standard deviation of data and minimal shape parameter change with high accuracy

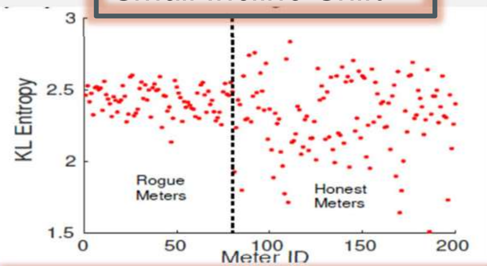
- **Introduced a new approach towards information theoretic trust scores using**
  - *- Modified Hill's Diversity Index*
  - *- Weighted Version of Renyi Entropy*
- **Captured horizontal, vertical, and incline shifts in statistical distributions**
- **Introduced the notion of Expected Temporal Self Similarity**
- **Optimized Model Parameters**
- **Higher Scores means more dishonest**

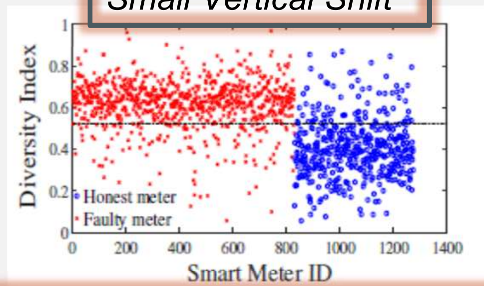




Small Incline Shift

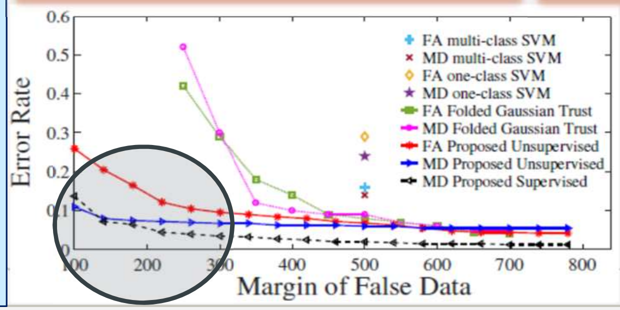
Small Vertical Shift





KL Divergence fails to classify

Our method classifies the same

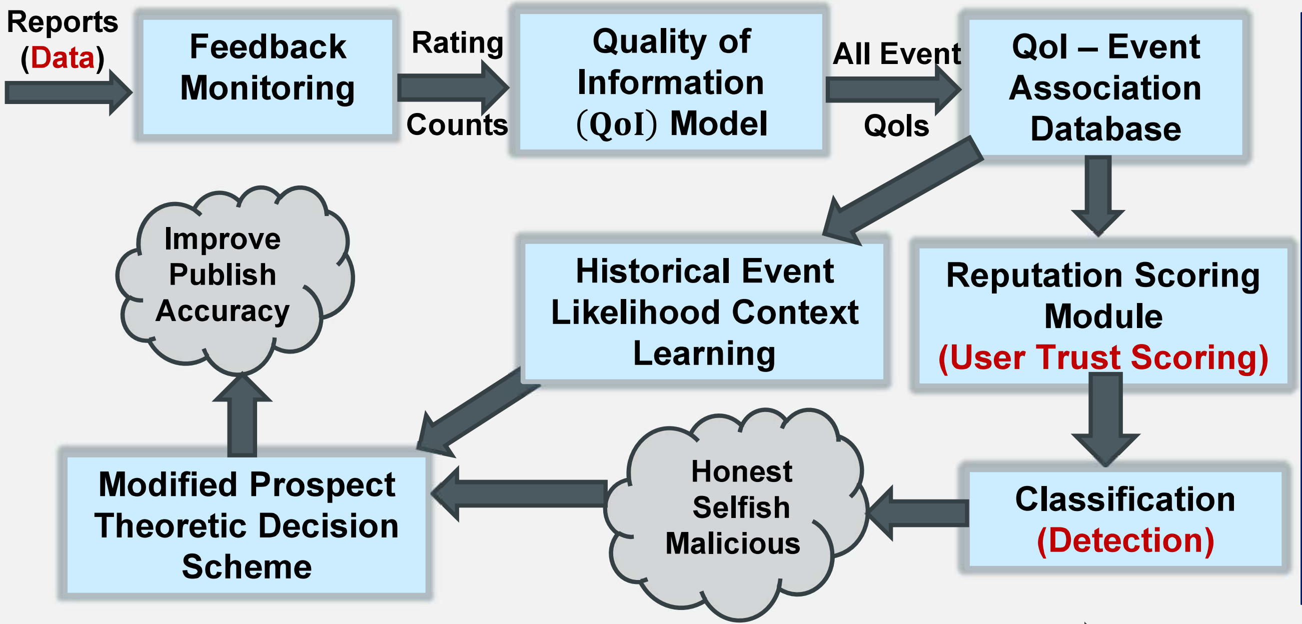


Broader Impacts:

- Applied to mobility data
- Validated by real datasets

(Product: Under double blind review in a conference)

Robust Decision Making under Attacks and Uncertainty



Problem and Challenges: Improve publish decision accuracy in vehicular social sensing under attacks and observation uncertainty

Two Level Decision Tree Formulation
- Which event type ?
- What event confidence ?

Key Theory for Tree Design
- Modified Prospect Theory (CPT)
- Tversky Kahneman Function
- Dual Prob. Weighing Function

Compare with classical decision tree with expected utility maximization

Modified Tversky-Kahneman Utility Function

$$v(\mathcal{C}_j) = \begin{cases} (\mathcal{C}_j)^{\theta_2}, & \text{if } \mathcal{C}_j \geq 0.5 \\ -\lambda_2.(0.5 - \mathcal{C}_j)^{\phi_2}, & \text{if } \mathcal{C}_j < 0.5 \end{cases}$$

Modified Dual Prob. Weighing Function

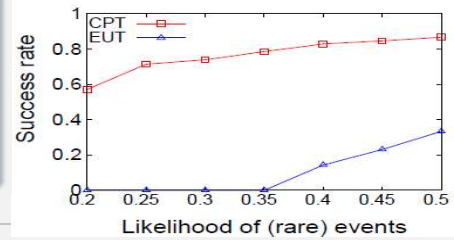
$$\pi^+(p_j) = \frac{p_j^{\delta_1}}{(p_j^{\delta_1} + (1 - p_j)^{\delta_1})^{\frac{1}{\delta_1}}},$$

$$\pi^-(\overline{p_j}) = \frac{(\overline{p_j})^{\delta_2}}{((\overline{p_j})^{\delta_2} + (1 - \overline{p_j})^{\delta_2})^{\frac{1}{\delta_2}}}$$

$$util(\mathcal{P}) = g_1 * v(\mathcal{C}_j) * \pi^+(p_j) + l_1 * v(\mathcal{C}_j) * \pi^-(\overline{p_j})$$

g_1 = publish given event j occurred (gain)

l_1 = publish given event j occurred (loss)



(**Products:** IEEE Transactions on Mobile Computing, and ACM Transactions on CPS)

Progress on Thrust 2:

(Shameek Bhattacharjee, Sajal Das, Hayato Yamana)

Privacy-preserving Computations using Fully Homomorphic Encryption (FHE)

Tasks:

2.1 FHE Calculations with Table Search
2.2 Handling Range Search
2.3 Applying FHE to Secure Decisions

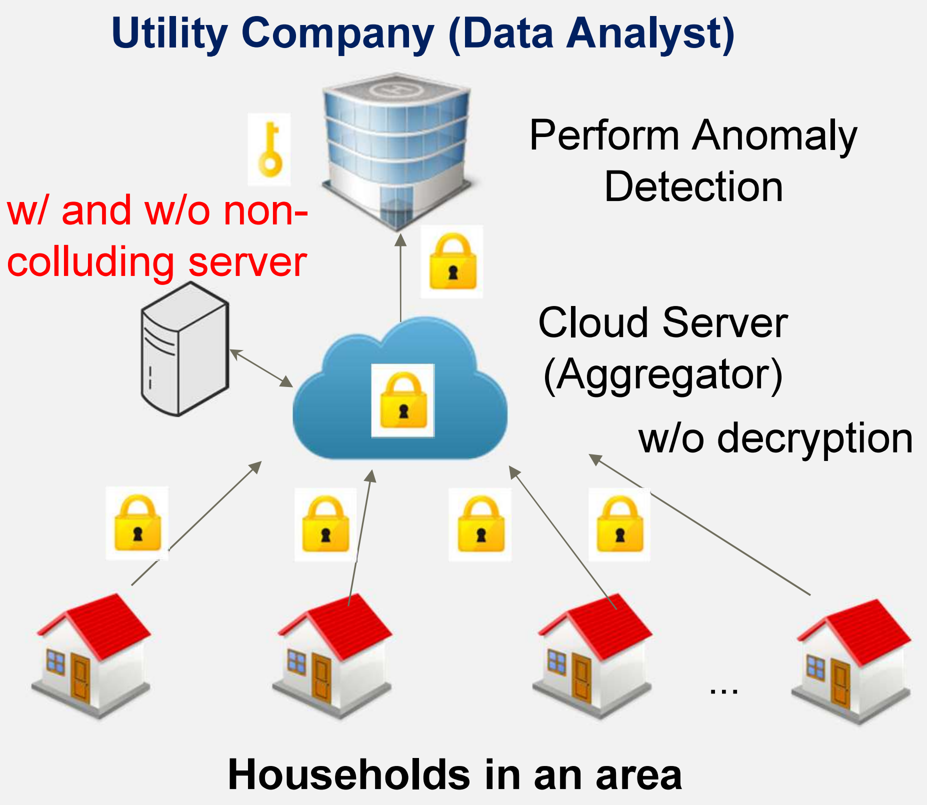
Preserving Privacy

Motivation

- Ensure secure and trustworthy decisions across integrated domains of smart energy and smart mobility networks
- Preserve privacy of each data output by IoT

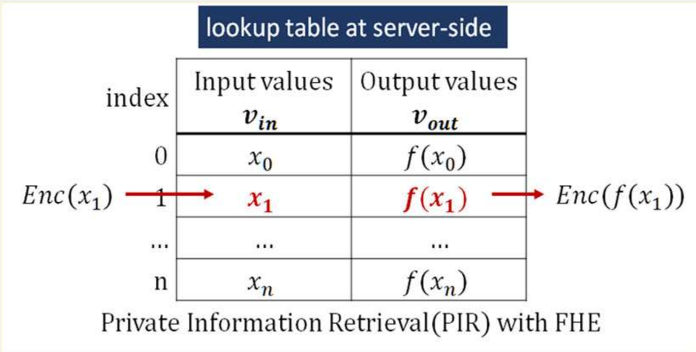
Objectives

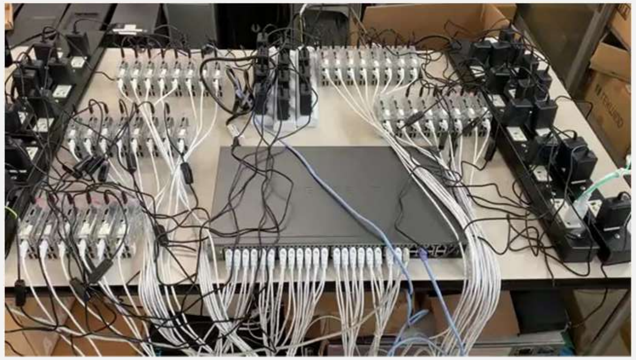
- Enhance *Fully Homomorphic Encryption* (FHE) to execute privacy preserving decisions that require complex calculations - TWO TECHNICAL APPROACHES
- Speed-up the FHE execution to satisfy the required performance



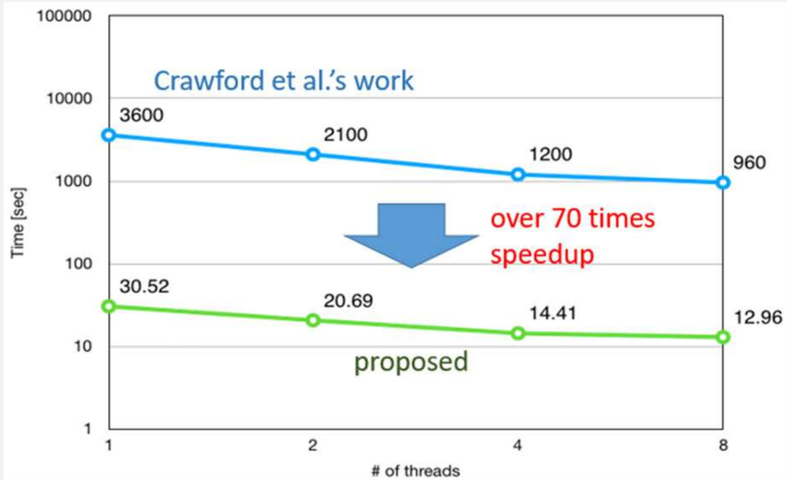
Preserving Privacy – Approach 1

- Table lookup with a non-colluding server to adopt any kinds of calculations at aggregators and to speed-up.





35 working Raspberry Pi's to emulate power meters



Execution time of one table lookup (Intel Core i7-8700 @3.2 GHz)

Novelty

- Any kinds of calculations can be adopted
- Integer-based encoding for further speedup in comparison with state-of-the-art research [Crawford et al. 2018] adopting bitwise-based encoding.

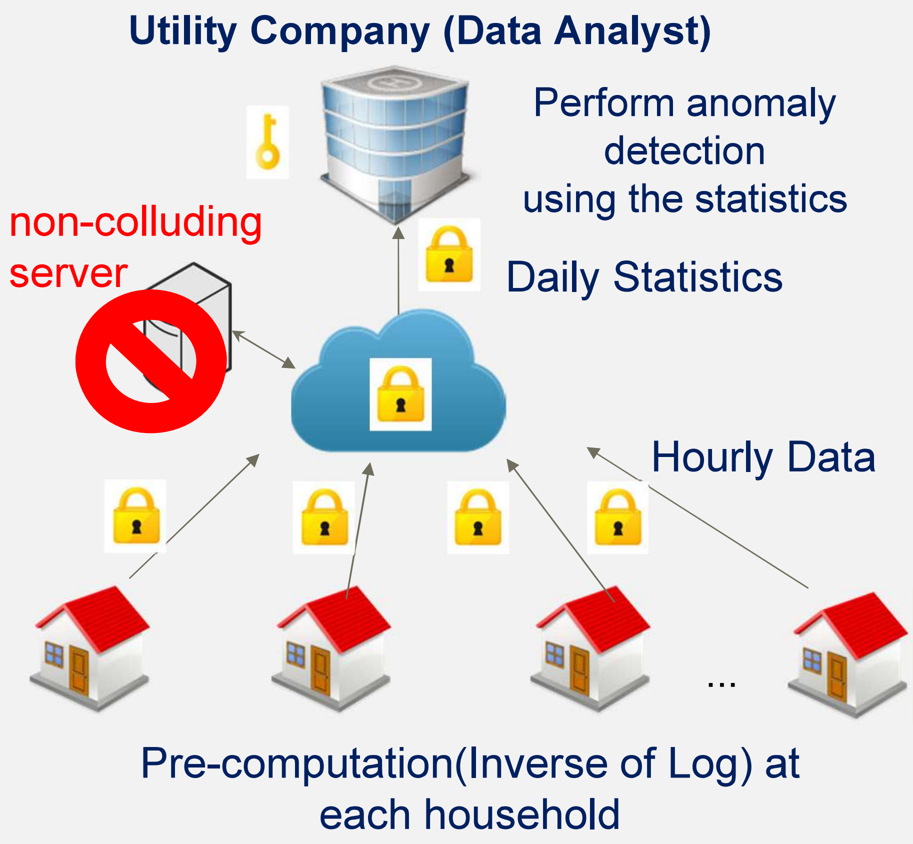
R. Li, Y. Ishimaki and H. Yamana, "Fully Homomorphic Encryption with Table Lookup for Privacy-Preserving Smart Grid," *3rd IEEE International Workshop on Big Data and IoT Security in Smart Computing* (BITS), June 2019.

Preserving Privacy – Approach 2

- Adopted Approximate Homomorphic Encryption scheme (HEAAN) to leverage floating-point arithmetic (log computation) over encrypted data

Novelty

- Anomaly detection algorithm over encrypted data w/o using non-colluding servers (more secure than Approach 1)
- Pre-computations of logarithm and its inverse at each household
- Optimized for FHE-friendly anomaly detection
- Homomorphic evaluation up to daily statistics that hide individual power consumption



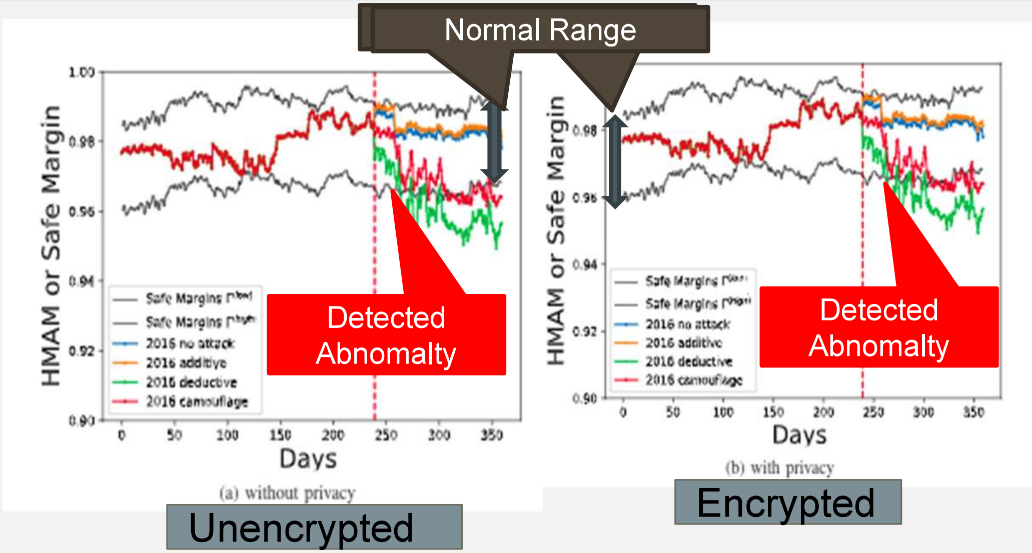
Preserving Privacy: Results and Conclusions

Findings

- Almost same accuracy over both encrypted and unencrypted methods (possible to mitigate small accuracy error via post-processing)
- Server-side computation is feasible: 3.303 s/hour (each hourly time-slot)

Next Challenges

- Enhance table lookup method to adopt multiple values, and propose less-than comparison for input values to handle wide range of inputs
- Balance Privacy-Performance trade-off with quantification (e.g., FHE with Differential Privacy)



Ciphertext Size:
Household → Server: 2,270KB/hour
Server → Utility: 224KB/date

All experiments are with 136 households.
Experiment is done with Intel Xeon CPU E5-1620 v4 @ 3.50GHz in single-threaded mode.

Progress on Thrust 4:

(Abhishek Dubey, Keiichi Yasumoto)

Developing Secure and Trustworthy Middleware Architecture

Tasks:

4.1 Distributed Aggregation
4.2 Secure Anonymization
4.3 Decision Making under Trade-offs

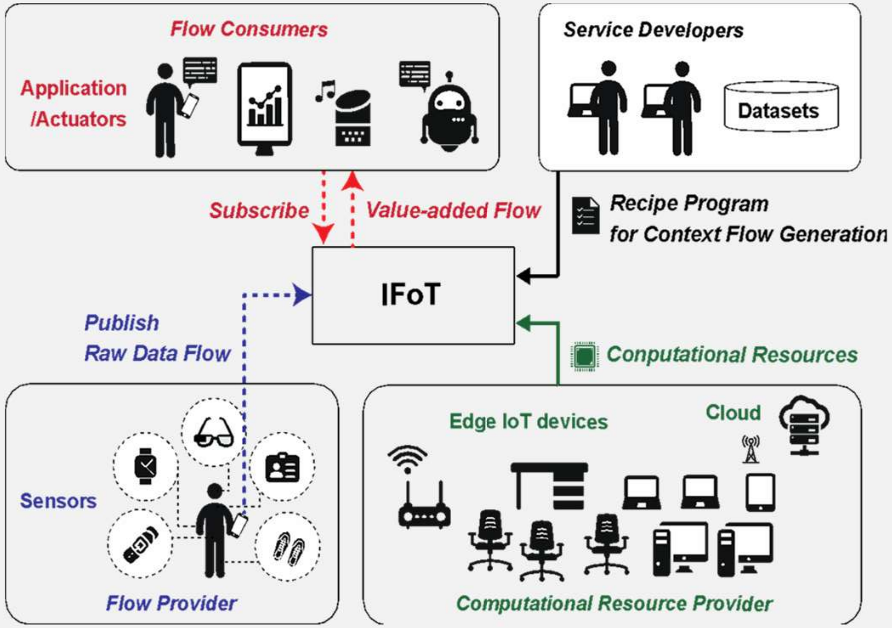
Middleware for Smart and Connected Communities

Challenges

- How to build multi-domain architecture for smart mobility and smart energy?
- How and where to implement computations related to privacy, security, and trust?
- What are computational/resource challenges for scalability?

Goal

Propose a novel middleware framework that distributes security features across multiple tasks and incorporates privacy, trustworthiness, resource constraints, and distributed decision support.



We have designed and developed IFoT middleware [SEC2018]

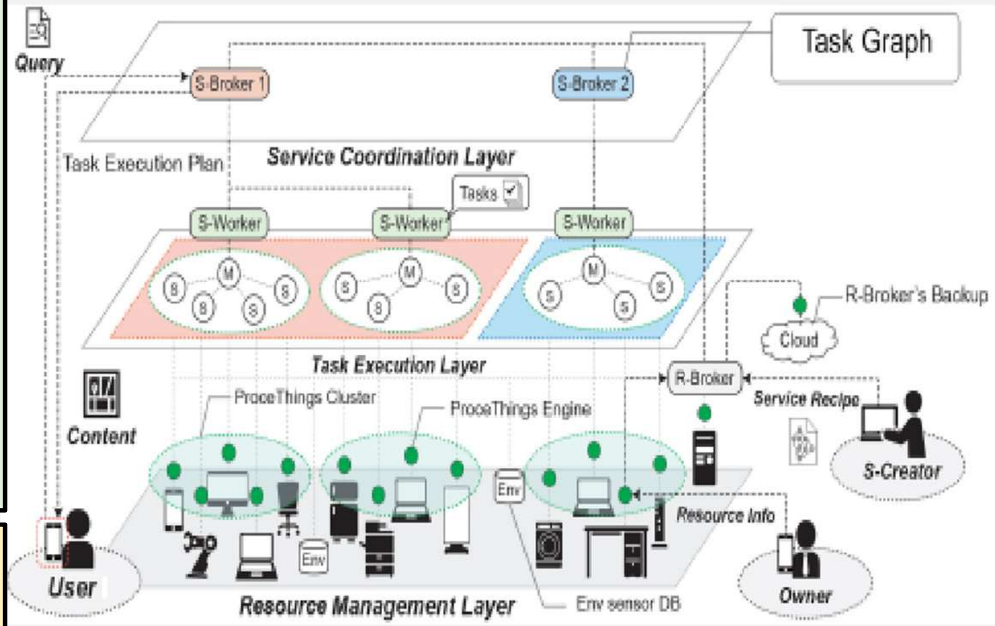
In-situ Distributed Computation and Aggregation

Innovation:

1. Designed a middleware architecture to assign tasks over IoT devices (e.g., RSUs, smart meters) taking into account required QoS level
2. Implemented/ evaluated a prototype middleware
   - Used Docker technology for easy deployment
   - Implemented a smart workspace use case (occupancy level of multiple rooms is queried)
3. Developed Smart Transportation Service Emulation Testbed

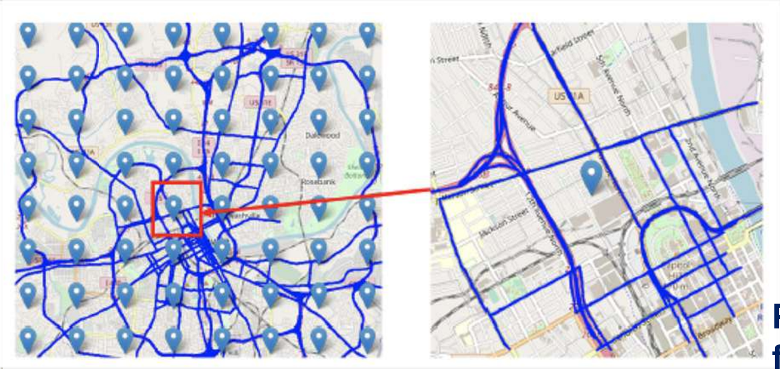
Findings:

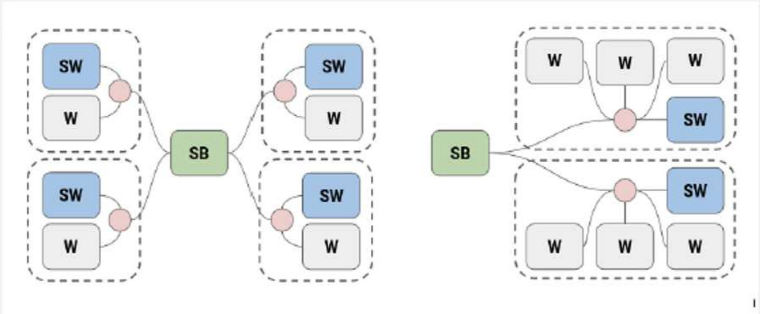
- Smart and Connected Community services can be realized by distributed execution of tasks assigned over IoT devices
- Query response time can be reduced by distributing tasks over IoT devices



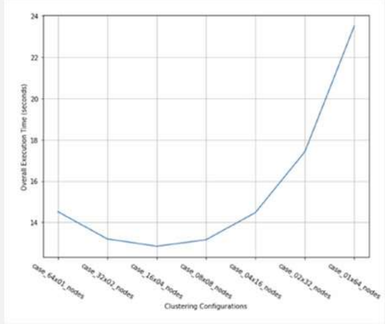
J. P. Talusan, K. Yasumoto, et al., "Evaluating Performance of In-Situ Distributed Processing on IoT Devices by Developing a Workspace Context Recognition Service," *IEEE PerCom Workshops,* 2019.

Smart Transportation Middleware





Processing time of a query (travel time b/w points) for different network configurations are tested



Processing time

80 km^2 map of Nashville, TN is divided into 8 x 8 grids, where each grid deploys an RSU which receives/stores traffic data (vehicle speed data) within the grid.

J. P. Talusan, K. Yasumoto, A. Dubey, S. Bhattacharjee, et al: Smart Transportation Delay and Resiliency Testbed based on Information Flow of Things Middleware. *IEEE BITS* 2019.

Ongoing work

- Develop federated learning on RSUs
- Optimal task assignment problem and algorithm with federated learning
- Develop and test anomaly detection algorithm (Target: submit to IoTDI 2020)

Progress on Thrust 5:

(Abhishek Dubay, Hirozumi Yamaguchi)

Validation with Real Datasets

Tasks:

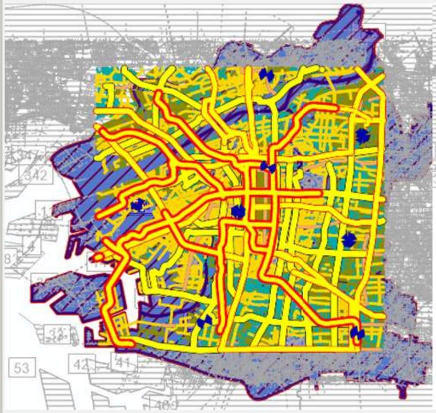
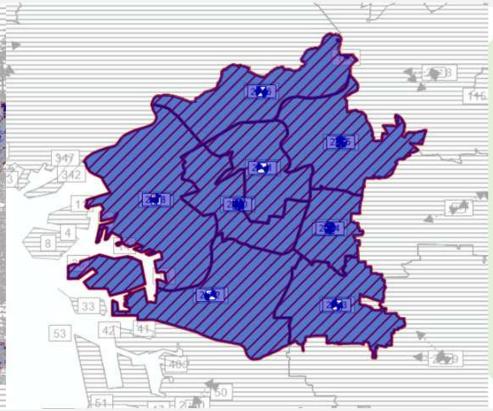
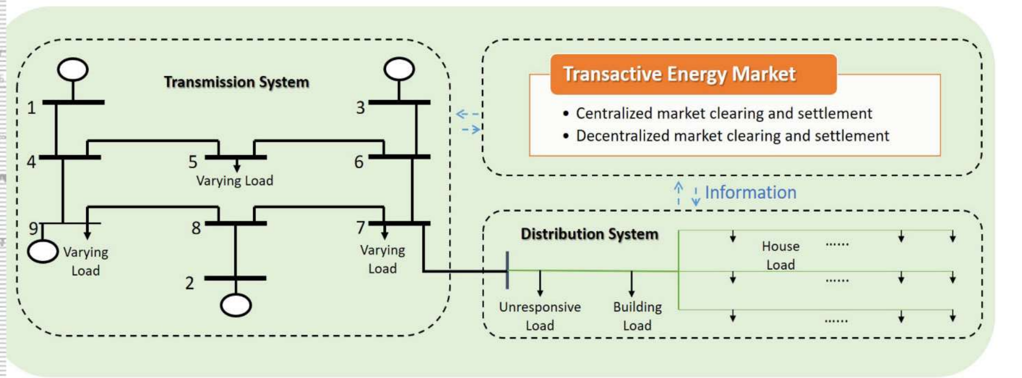
5.1 Smart Transportation Application
5.2 Smart Energy Application

Validation with Real Datasets

Objective: Validate the proposed models and approaches using smart mobility and smart energy distribution / consumption scenarios with real-world datasets

Approach:

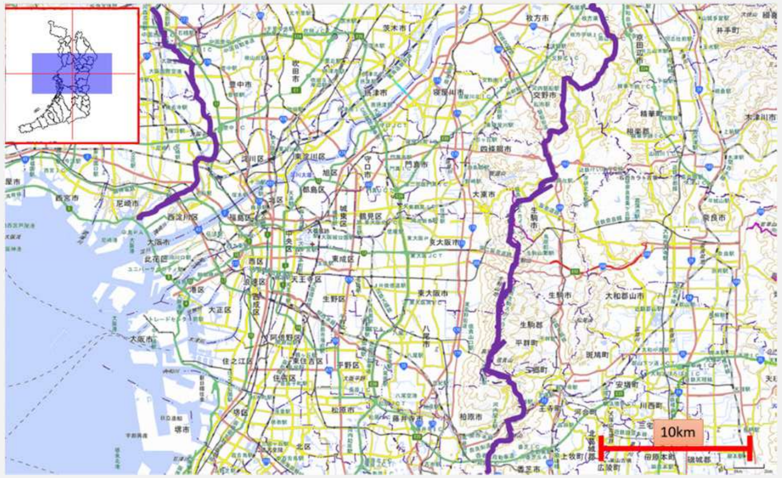
- Generate large-scale mobility data from real datasets in Osaka
- Design a transactive energy testbed that can integrate energy market data

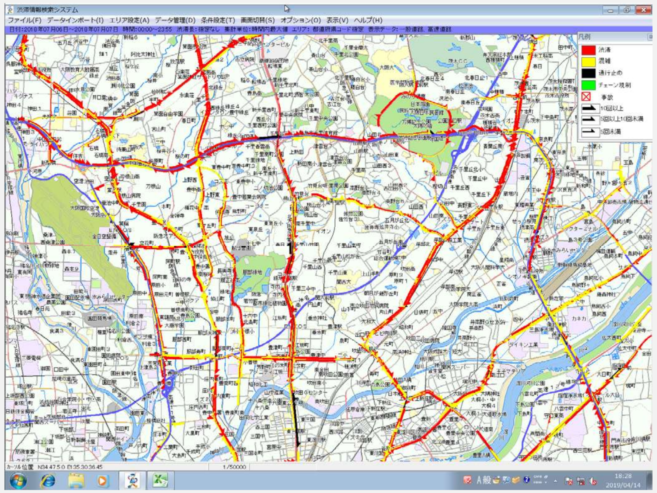




Real World Data Sets

Road Traffic Census data (obtained by nation-wide survey by Ministry Road Bureau) : contains traffic volume & velocity and OD of a particular day

VICS (obtained via IR beacons) : contains queue length of major city roads and highways for 3months (Osaka prefecture whole region)

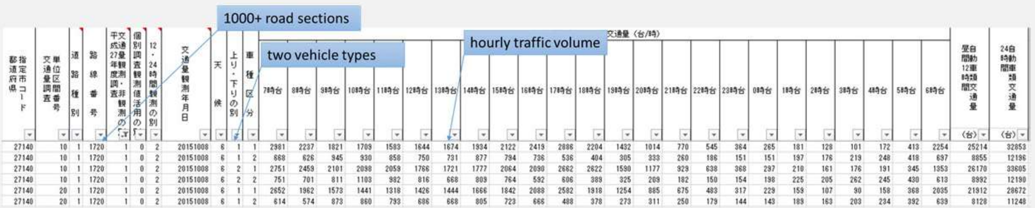




Infrared beacons
(Ordinary trunk roads)

Infrared beacons are installed on the ordinary trunk roads and provide information covering about 30 km in the forward direction and about 1 km in the rear direction.

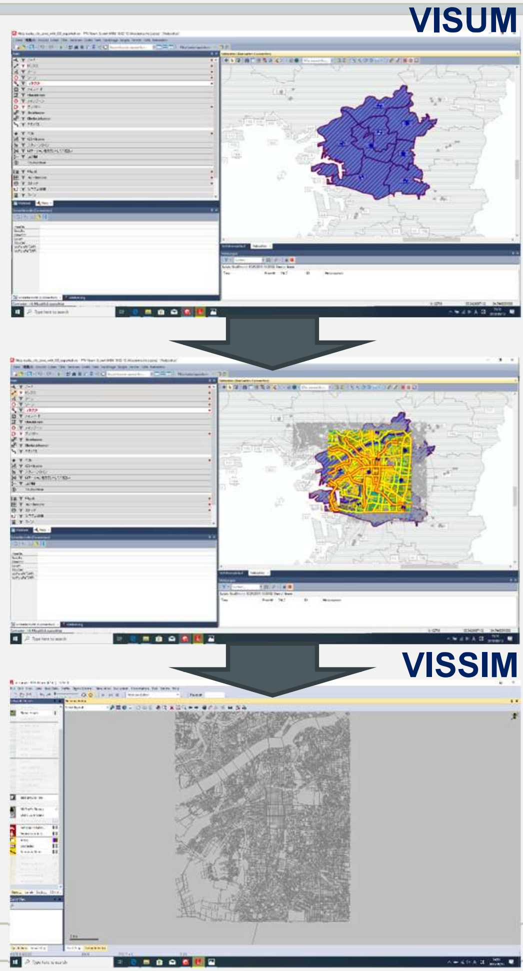
- Traffic congestion and travel time information.
- Information on restrictions due to accidents, construction, disasters, and weather conditions.
- Parking availability.



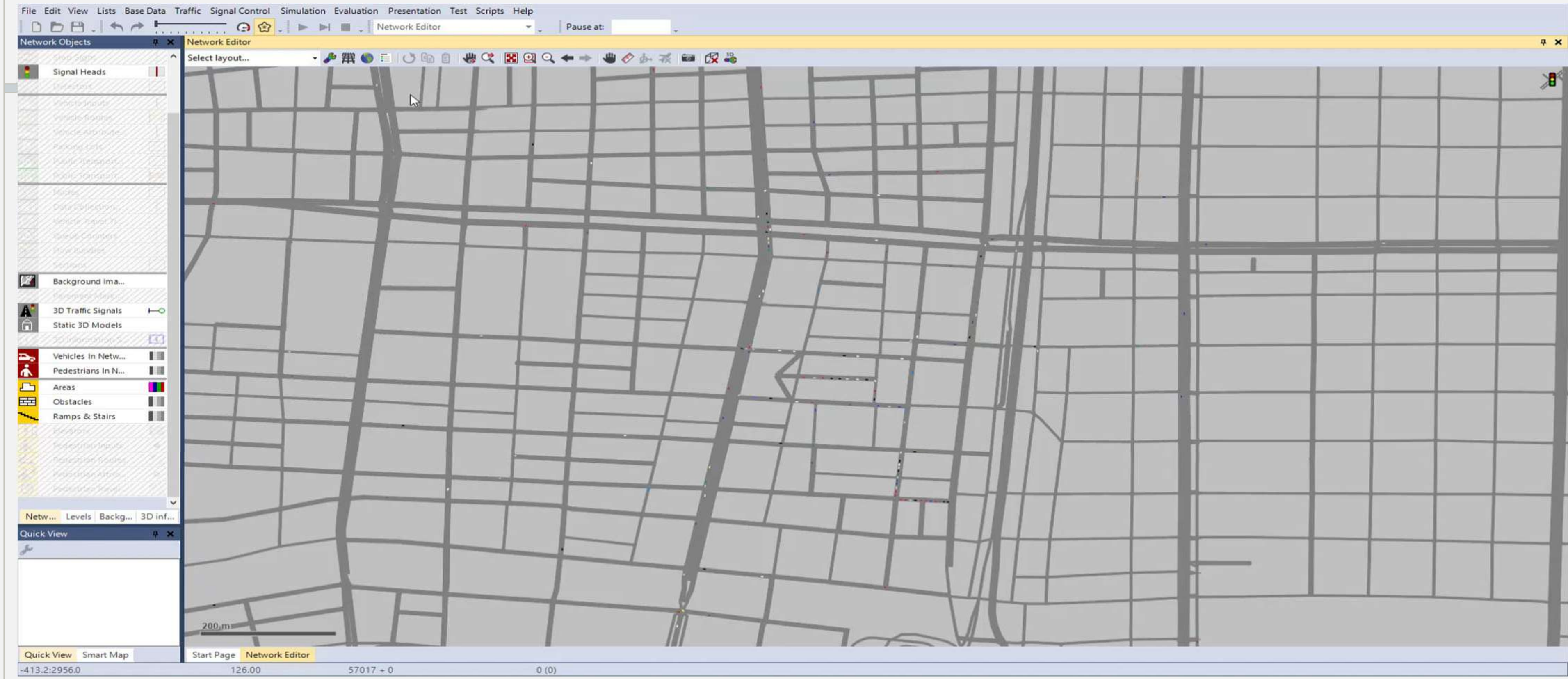
都道府県・指定市コード	交通量調査単位区間番号	道路種別	路線番号	平成27年度調査 交通量観測・非観測の別	個別調査観測値活用の別	12・24時間観測の別	交通量観測年月日	天候	上り・下りの別	車種区分	交通量（台/時） 7時台	8時台	9時台	10時台	11時台	12時台	13時台	14時台	15時台	16時台	17時台	18時台	19時台	20時台	21時台	22時台	23時台	0時台	1時台	2時台	3時台	4時台	5時台	6時台	昼間12時間自動車類交通量（台）	24時間自動車類交通量（台）
27140	10	1	1720	1	0	2	20151008	6	1	1	2981	2237	1821	1709	1583	1644	1674	1934	2122	2419	2886	2204	1432	1014	770	545	364	265	181	128	101	172	413	2254	25214	32853
27140	10	1	1720	1	0	2	20151008	6	1	2	668	626	945	930	858	750	731	877	794	736	536	404	305	333	260	186	151	151	197	176	219	248	418	697	8855	12196
27140	10	1	1720	1	0	2	20151008	6	2	1	2751	2459	2101	2098	2059	1766	1721	1777	2064	2090	2662	2622	1590	1177	929	638	368	297	210	161	176	191	345	1353	26170	33605
27140	10	1	1720	1	0	2	20151008	6	2	2	751	701	811	1103	982	816	668	809	764	592	606	389	325	209	182	150	154	198	225	205	262	245	430	613	8992	12190
27140	20	1	1720	1	0	2	20151008	6	1	1	2652	1962	1573	1441	1318	1426	1444	1666	1842	2088	2582	1918	1254	885	675	483	317	229	159	107	90	158	368	2035	21912	28672
27140	20	1	1720	1	0	2	20151008	6	1	2	614	574	873	860	793	686	668	805	723	666	488	378	273	311	250	179	144	143	189	163	203	234	392	639	8128	11248

Osaka City Vehicle Mobility Generation

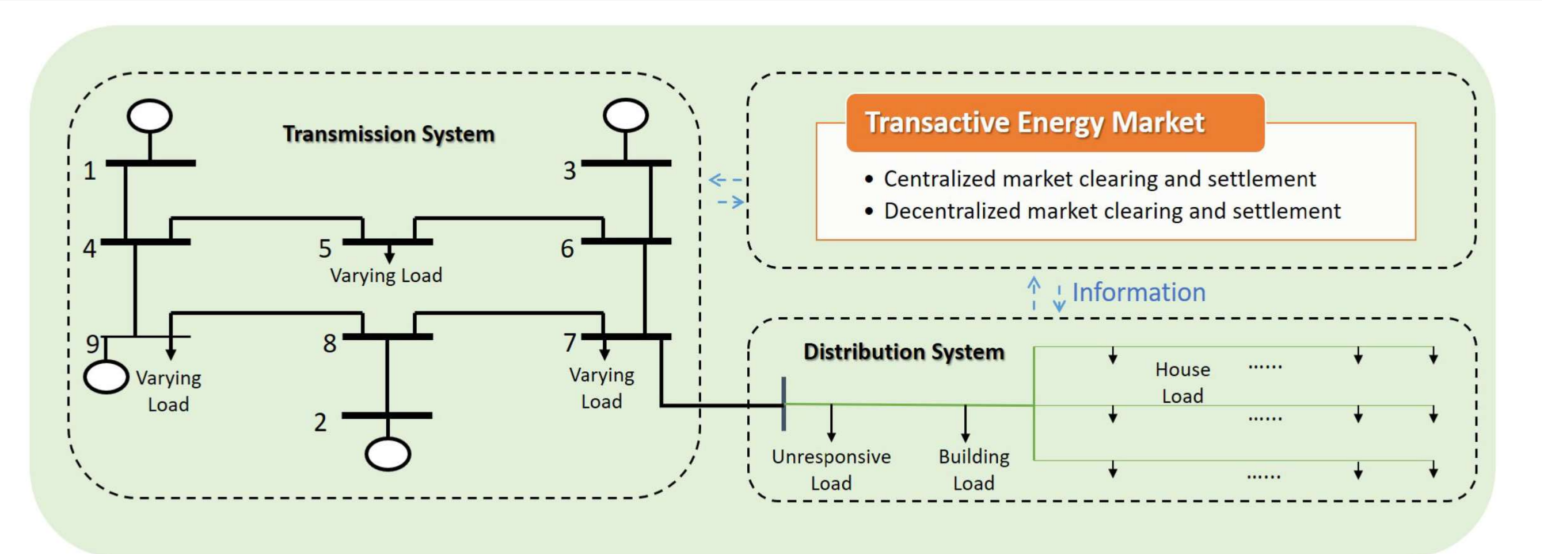
- **Procedure**
  - Arrange map (Open Street Map)
  - Set ODs (Census Survey Data)
  - Calculate route for each OD (by Traffic Simulator)
  - Determine OD volumes (Census Survey Data)
  - Adjust link-level traffic volume (VICS)
- **The generated mobility data includes**
  - 10K – 60K vehicles in Osaka city region
  - 90 days (will include data of the day of “big rain disaster” in July 2018)
- **Next Challenges:**
  - RSU placement design in 5G realistic situations
  - Evaluation of the impact of anomaly data on decision making using Osaka and Nashville data



Simulator Video



Transactive Energy Testbed



Architecture of TESST

Physical System | Transactive Energy Market | Network Simulator

Centralized Market Option | Decentralized Market Option

STEAM Project: Broader Impacts

➢ Interdisciplinary Education and Experiential Learning for Students:

- Praveen Madhavarapu; Prithwiraj Roy (Missouri S&T and Western Michigan Univ., USA)
- Michael Wilbur; Geoff Pettet (Vanderbilt Univ., USA)
- Yu Ishimaki; Ruixiao Li (Waseda Univ., Japan)
- Jose Paolo Talusan; Francis Tiausas (NAIST, Japan)
- S. Choochotkaew; Yuki Akura (Osaka Univ. Japan)

➢ Student Visit Exchanges:

- Y. Ishimaki (Waseda) visited MST in Aug-Sep 2018 for one month, and WMU for 2 weeks in June and Oct 2019
- P. Madhavarapu and P. Roy (Missouri S&T) respectively visited WMU for 4 weeks in July 2019 and Aug 2019
- J.P. Talusan (NAIST) visited Vanderbilt for 3 weeks in June 2019.
- M. Wilbur (Vanderbilt) visited WMU for 1 week.

➢ Integration of Research into Courses:

- Dubey (Vanderbilt Univ.) integrated anomaly detection module in his course on *Reliable Distributed Systems*, fall 2019.
- Shameek (WMU) incorporated CPS security challenges and smart grid security solutions in his course *Science of Cybersecurity*, spring 2019.

➢ Outreach (Workshop Organization):

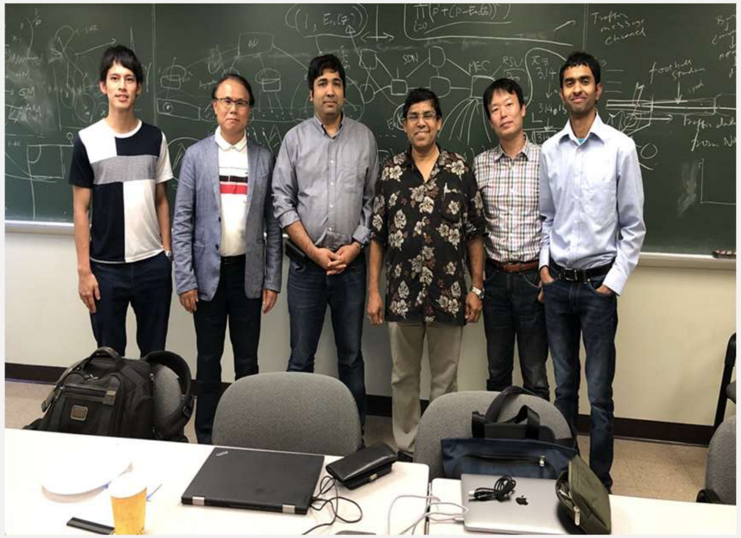
- Big Data and IoT Security (BITS), in conjunction with IEEE SmartComp 2019, Washington, DC, June 2019.
- Science of Smart City Operations and Platforms Engineering (SCOPE), during CPS-IoT Week, Montreal, Canada, April 2019.

Coordination and Collaboration

- Weekly Skype meeting; Very coherent group
- Joint publications by PIs and their students
- Co-organization of BITS and SCOPE workshops in 2019
- Planned Vision Paper: Security in Integrated Energy and Mobility
- Planned Special Issue Editing: Magazine and/or Journal

All Hands Meeting:

- Missouri S&T: Sept 14-15, 2018
- Tokyo: Oct 26-27, 2018 (JUNO2 Kick-off)
- Kyoto: March 11-14, 2019 (IEEE PerCom)
- Washington, DC: June 12-14, 2019 (IEEE SmartComp)
- Chicago: Oct 11, 2019 (JUNO2 PI Meeting)
- Bologna: June 20-23, 2020 (IEEE SmartComp)
- Nara: January 5-8, 2021 (ACM ICDCN)



September 14-15, 2018 meeting (Missouri S&T)

Collaborative Publications

1. S. Roy, N. Ghosh, and S. K. Das, "bioSmartSense: A Bio-inspired Data Collection Framework for Energy-efficient, QoI-aware Smart City Applications," *17th Annual IEEE International Conference on Pervasive Computing and Communications* (PerCom), Kyoto, Mar 2019.
2. Y. Nishimura, A. Fujita, A. Hiromori, H. Yamaguchi, T. Higashino, A. Suwa, H. Urayama, S. Takeshima and M. Takai, "A Study on Behavior of Autonomous Vehicles Cooperating with Manually-Driven Vehicles," *17th Annual IEEE PerCom*, pp. 212-219, Kyoto, Mar 2019.
3. J. P. Talusan, K. Yasumoto, et al, "Evaluating Performance of In-Situ Distributed Processing on IoT Devices by Developing a Workspace Context Recognition Service," *IEEE PerCom Workshop*, Kyoto, Mar 2019.
4. H. Yamaguchi, "Toward Urban Vehicle Mobility Modeling in Japan," *4th International Science of Smart City Operations and Platforms Engineering Workshop* (SCOPE), pp. 1-6, Apr 2019.
5. R. Li, Y. Ishimaki and H. Yamana, "Fully Homomorphic Encryption with Table Lookup for Privacy-Preserving Smart Grid," *IEEE BITS2019 Workshop*, pp. 19-24, June 2019.
6. M. Wilbur, A. Dubey, B. Leão and S. Bhattacharjee, "A Decentralized Approach for Real Time Anomaly Detection in Transportation Networks," *4th IEEE International Conference on Smart Computing* (SMARTCOMP), Washington, DC, pp. 274-282, June 2019.
7. J. P. Talusan, K. Yasumoto, A. Dubey, and S. Bhattacharjee, "Smart Transportation Delay and Resiliency Testbed based on Information Flow of Things Middleware," *IEEE BITS Workshop*, June 2019.
8. Y. Ishimaki, H. Yamana, "Non-Interactive and Fully Output Expressive Private Comparison," *INDOCRYPT*: 355-374, 2018.
9. S. Bhattacharjee and S. K. Das, "Detection and Forensics against Stealthy Data Falsification in Smart Metering Infrastructure," *IEEE Transactions on Dependable and Secure Computing*, to appear, 2019.
10. R. P. Barnwal, N. Ghosh, S. K. Ghosh, and S. K. Das, "Publish or Drop Traffic Event Alerts? Quality-aware Decision Making in Participatory Sensing-based Vehicular CPS," *ACM Transactions on Cyber-Physical Systems*, to appear, 2019.
11. S. Bhattacharjee, N. Ghosh, V. K. Shah, S. K. Das, "QnQ: A Quality and Quantity Unified Approach for Secure and Trustworthy Crowdsensing, *IEEE Transactions on Mobile Computing*, to appear, 2019.
12. A. Sturaro, S. Silvestri, M. Conti, and S. K. Das, "A Realistic Model for Failure Propagation in Interdependent Cyber-Physical Systems," *IEEE Transactions on Network Science and Engineering*, to appear, 2019.
13. S. Bhattacharjee, V. P. Madhavarapu, S. Silvestri, and S. K. Das, "Attach Context Embedded Data Driven Trust Diagnostics in Smart Metering Infrastructure," *ACM Transactions on Privacy and Security*, under minor revision, Oct 2019.

JUNO2: US-Japan Collaborative Project

STEAM: Secure and Trustworthy Framework for Integrated Energy and Mobility in Smart Connected Communities

PI Meeting – Chicago, October 11, 2019

Thank You

JUNO2: US-Japan Collaborative Project

STEAM: Secure and Trustworthy Framework for Integrated Energy and Mobility in Smart Connected Communities

PI Meeting (August 2021)

Our Super Team

MISSOURI S&T

Sajal K. Das

WESTERN MICHIGAN UNIVERSITY

Shameek Bhattacharjee

VANDERBILT

Abhishek Dubey

Hayato Yamana

Waseda University

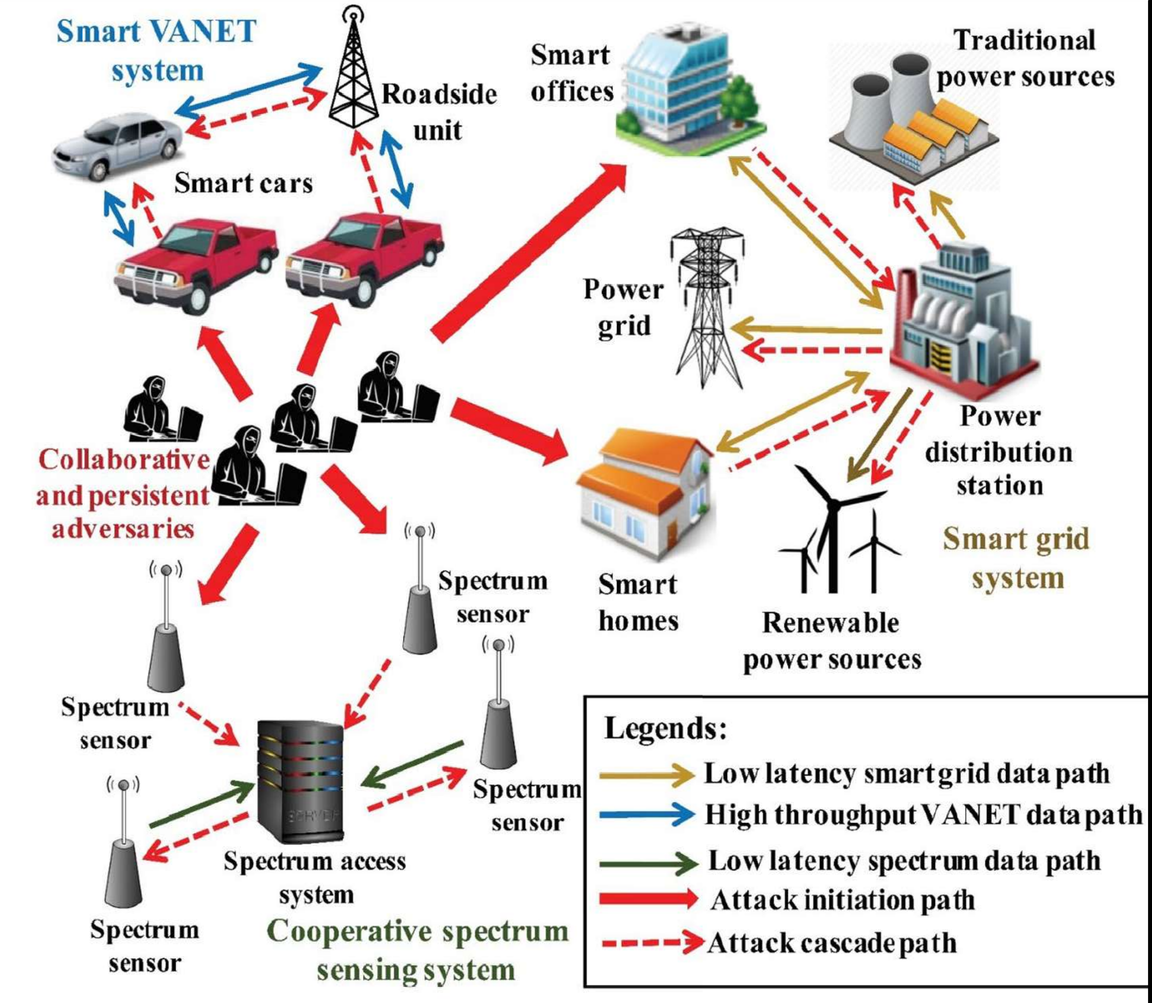
Hirozumi Yamaguchi

大阪大学 OSAKA UNIVERSITY

Keiichi Yasumoto

Nara Institute of Science and Technology NAIST

Security in a Smart City Scenario



Smart Mobility and Smart Energy:

- Interdependent, societally critical CPS networks
- Ensuring safety, resilience and privacy preservation

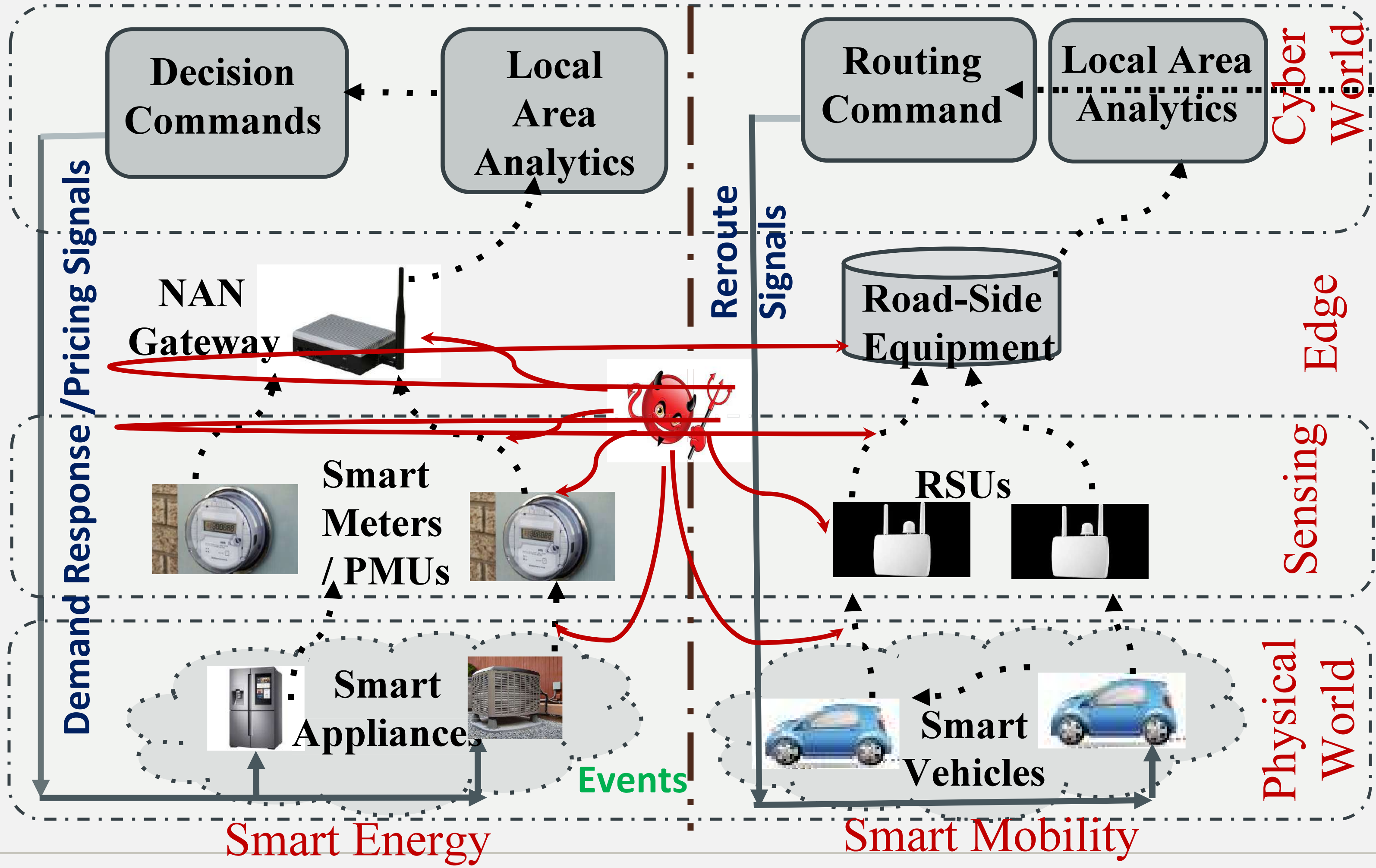
Unique Challenges in Securing S & CC

- Large variations in individual endpoint data due to human behavior and activity differences.
- No strict stationarity over time or space.
- Heterogeneous time granularities of sensing and network sizes.

Goals and Novelty of STEAM Project:

- Develop integrated frameworks, algorithms and models to address security, dependability and trustworthiness challenges in mobility and energy under various threats.
- Design lightweight resilient anomaly detection and privacy preserving encryption schemes & middleware architecture.
- Trust building in S & CC applications; efficient mechanisms to handle conflicting goals of identifying anomalies; trade-off between security, privacy and integrity at scale.
- Efficient co-design and calibration of encryption and robust anomaly detection schemes.

Energy and Mobility: Integrated CPS View



STEAM Project: Intellectual Merit

Thrust 1: Secure and Trustworthy Decision Making under Uncertainty (Bhattacharjee, Das)

- **Task 1.1: Lightweight Anomaly Detection:** Fast, efficient and accurate decentralized anomaly detection for compromised smart meters and transportation sensors under stealthy attacks, using Pythagorean means and long short-term memory (LSTM) networks.
- **Task 1.2 Trust Models:** Trust scoring models for diagnosing device compromisation with attack margins much below standard deviation.
- **Task 1.3: Trustworthy Decision Making:** Dependable decisions

Thrust 2: Privacy-Preserving Computations using FHE (Yamana, Bhattacharjee, Das)

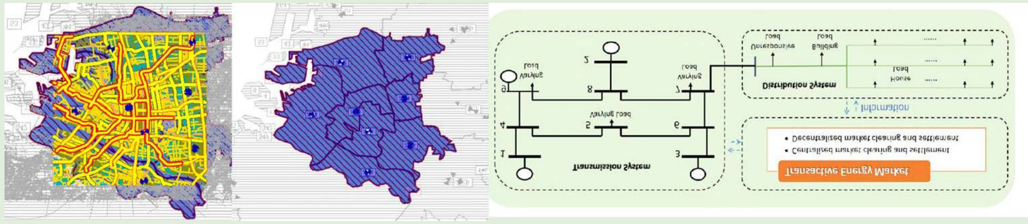
- **Task 2.1: FHE (Complex) Calculations:** Efficient schemes to compute FHE for privacy preserving decisions
- **Task 2.2: Handling Range Search:** Table lookup with non-colluding server for calculations at aggregator for higher speed-up;
- **Task 2.3: Applying FHE to Secure Decisions:** Approximate Homomorphic Encryption (HE) to leverage floating-point arithmetic (e.g., log computation) over encrypted data.

Thrust 4: Developing a Secure and Trustworthy Middleware Architecture (Dubey, Yasumoto)

- **Task 4.1: Distributed Aggregation:** Developed a middleware platform for SCC apps via distributed processing at edge nodes
- **Task 4.2: Secure Anonymization:** Developed a secure anonymization mechanism for smart mobility apps
- **Task 4.3: Decision Making under Trade-offs:** Balance query throughput, route accuracy, and privacy protection level

Thrust 5: Validation With Real Datasets for Smart Mobility and Energy (Yamaguchi, Dubey)

- **Task 5.1: Smart Transportation Application:** Large scale road traffic data collected from Osaka, Japan and Nashville, USA.
- **Task 5.2: Smart Energy Application:** Transactive energy testbed.



Thrust 1 Results:

(Shameek Bhattacharjee, S. K. Das)

Secure and Trustworthy Decision Making under Uncertainty

Tasks:

1.1 Lightweight Anomaly Detection
1.2 Stochastic Trust Models
1.3 Dependable Decision Making

Task 1.1 Anomaly Detection

Threat Model:

- Attacks from sensing layer affect operations in smart energy and mobility systems
- Special Events such as
  data omission □ energy
  traffic accidents □ mobility

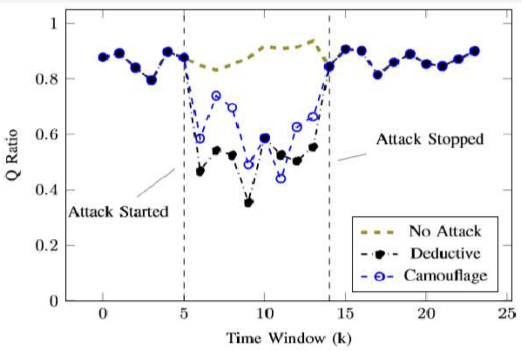
Macro Level: designed for large scale decentralized anomaly detection in real time for energy and transportation □ *Pythagorean Mean*

Micro Model: highly accurate and fine grained-anomaly detection □ *LSTM, Folded Gaussian Trust Models*

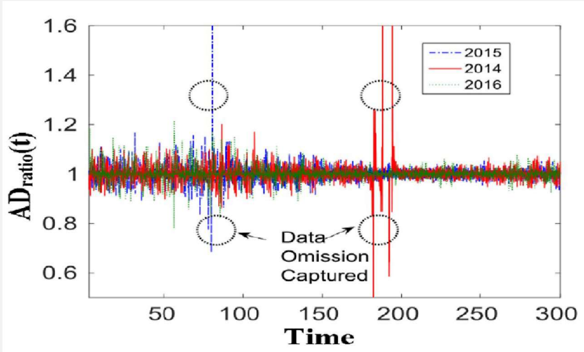
Key Theory → Schur Ostrowski Criterion

$(x_i - y_i)\,(\frac{\partial f}{dx_i} - \frac{\partial f}{dy_i}) \geq 0$ *→ Schur Convexity*

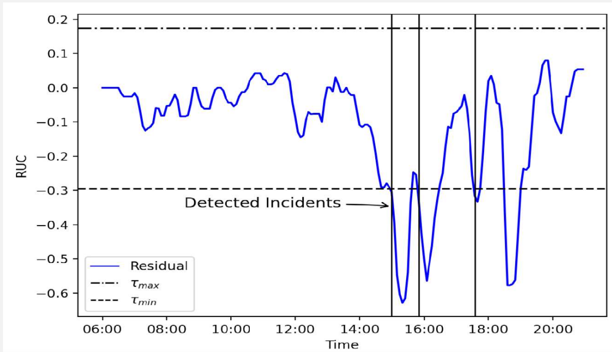
Validation:
Mobility: real data from Nashville, TN
Energy: real data from TX and Ireland



Anomaly under attacks



Anomaly under data omission



Anomaly under accidents

Metrics:

- Q ratio
- AD_{ratio}
- RUC

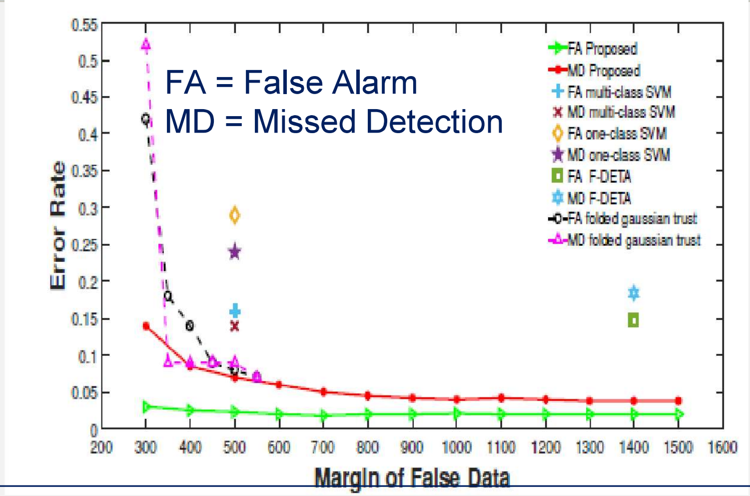
Derived from Pythagorean Means

Products:
(1) IEEE Trans. on Dependable and Secure Computing 2021
(2) ACM Trans. on Privacy and Security 2021
(3) IEEE Smart Computing Conference 2019
(4) IEEE Big Data and IoT Security Workshop 2019
(5) IEEE Big Data Conference (in preparation)

Task 1.2: Trust Scoring Models

Challenge: *Fast* and *accurate* detection of Compromised Smart Meters, RSUs, users of distributed data falsification attacks.

- **Introduction of Attack Responses as Robust Statistical Measures**
  - Pythagorean Means and Real Analysis
  - Location Parameter Correction
  - Attack Probability Time Ratio
- **Embedding of Responses** magnify divergence in probability space for information theoretic detection
- **Magnified Divergence** high detection accuracy, reduced false alarms, decreased convergence time under stealthy attacks
- **Multi-granular anomaly-based attack detector across temporal scales better threshold design indicates attack responses**



Key Theory:

- *Folded Gaussian Trust Model (Density based)*
- *Response Enhanced KL Divergence (Distance based)*
- *Neuro-cognitive models (Behavioral AI based)*

Key improvement over existing works

- *Detects meters with attack margin > 300W*
- *FA < 10% for large sized grid*
- *Compared with existing works*
- *Detects malicious users in vehicular crowdsensing applications*

Products:

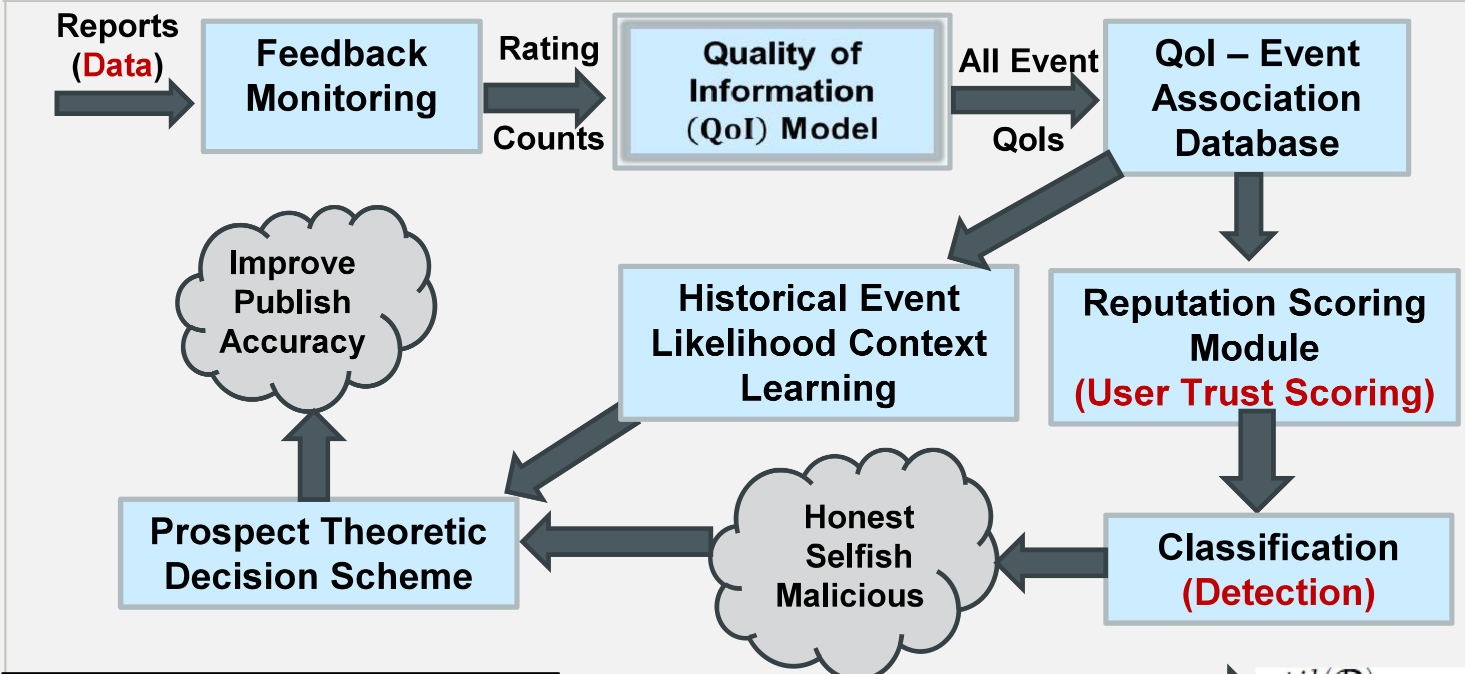
(1) *ACM Trans. on Privacy and Security, '21* ***(2)*** *IEEE Trans. on Mobile Computing, '21*

(3) *Journal (in preparation for IEEE TIFS)* ***(4)*** *IEEE MASS '20*

Broader Impacts:

- Includes closed form approximations and performance limits under attacks
- Validated across big datasets from Texas (800 meters) and Ireland (5000 meters)
- Preliminary efforts show success with other IoT domains (e.g., smart home)

Task 1.3: Trustworthy Decision Making



Problem and Challenges: Improve publish decision accuracy in vehicular social sensing under attacks and observation uncertainty

Two Level Decision Tree Formulation

- Which event type ?
- What event confidence ?

Key Theory for Tree Design

- Cumulative Prospect Theory (CPT)
- Tversky Kahneman Function
- Dual Prob. Weighing Function

Compare with classical decision tree with expected utility maximization

Modified Tversky-Kahneman Utility Function

$$v(\mathcal{C}_j) = \begin{cases} (\mathcal{C}_j)^{\theta_2}, & \text{if } \mathcal{C}_j \geq 0.5 \\ -\lambda_2.(0.5 - \mathcal{C}_j)^{\phi_2}, & \text{if } \mathcal{C}_j < 0.5 \end{cases}$$

Modified Dual Prob. Weighing Function

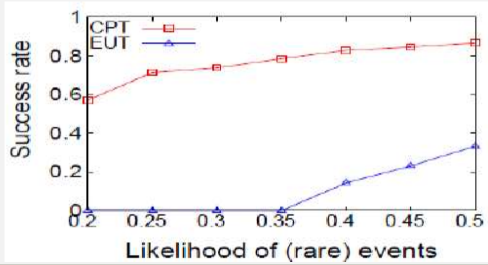
$$\pi^+(p_j) = \frac{p_j^{\delta_1}}{(p_j^{\delta_1} + (1-p_j)^{\delta_1})^{\frac{1}{\delta_1}}},$$

$$\pi^-(\overline{p_j}) = \frac{(\overline{p_j})^{\delta_2}}{((\overline{p_j})^{\delta_2} + (1-\overline{p_j})^{\delta_2})^{\frac{1}{\delta_2}}}$$

$$util(\mathcal{P}) = g_1 * v(\mathcal{C}_j) * \pi^+(p_j) + l_1 * v(\mathcal{C}_j) * \pi^-(\overline{p_j})$$

g_1 = publish given event j occurred (gain)

l_1 = publish given event j occurred (loss)



Products: **(1)** *IEEE Trans. on Mobile Computing '20* **(2)** *ACM Trans. on CPS '20*, **(3)** *IEEE WoWMoM '21*

Result on Thrust 2:

(S. Bhattacharjee, S. K. Das, H. Yamana)

Privacy-preserving Computations using Fully Homomorphic Encryption (FHE)

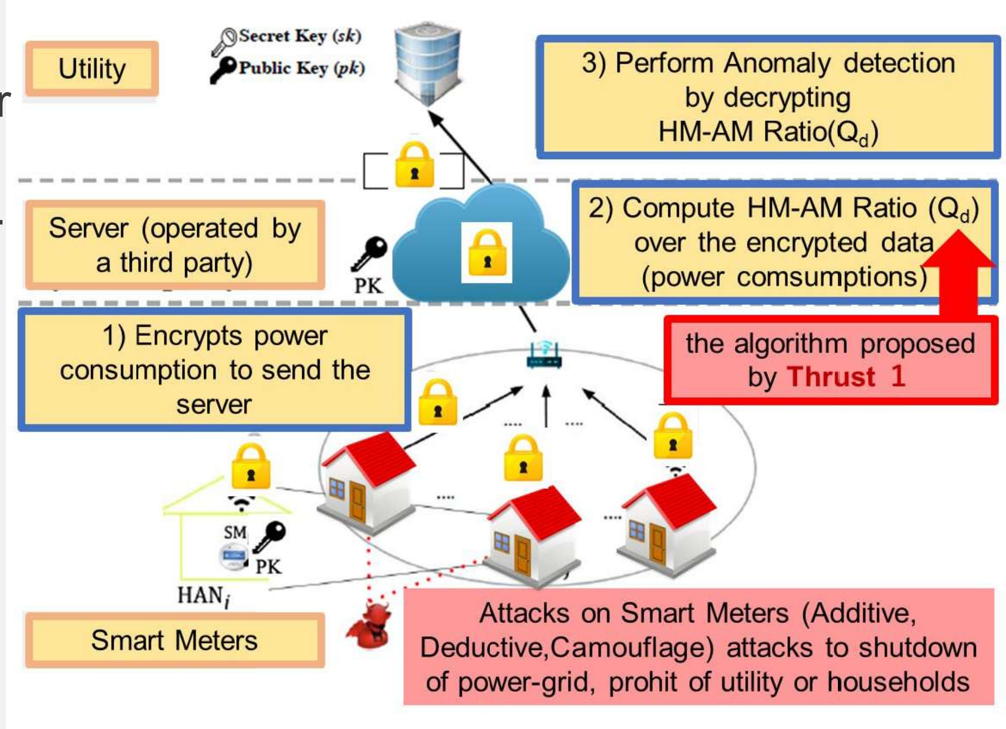
Tasks:

2.1 FHE Calculations with Table Search
2.2 Handling Range Search
2.3 Applying FHE to Secure Decisions

Preserving Privacy – Goal and Results

- Goal
  - Establishing a **privacy-preserving anomaly detection** method by adopting both Thrust 1 and "fully homomorphic encryption(FHE)," which has not been possible in the past.

- Results
  - Enabling **10 sec** anomaly detection for by our proposed special optimization (**Method 1**)
  - Enabling **5 min** anomaly detection for power-grid by adopting "table search mechanism," which is applicable in various anomaly detection algorithms. (**Method 2**)



[Note]
- Used 2014 to 2016 power consumption data of 200 households in Texas.
- The requirement for power-grid is every 10 min detection.
- Experiments on single core execution on the server (Intel Xeon E5-1620v4)

Preserving Privacy – Method 1

Calculation for the anomaly detection

$$\text{timeslot}\quad \begin{pmatrix} p_1^{(1)} & p_1^{(2)} & \cdots & p_1^{(N)} \\ p_2^{(1)} & p_2^{(2)} & \cdots & p_2^{(N)} \\ \vdots & \cdots & \cdots & \vdots \\ p_T^{(1)} & p_T^{(2)} & \cdots & p_T^{(N)} \end{pmatrix} \longrightarrow$$

Daily HM-AM ratio

$$Q_d = \frac{\sum_{t=1}^{T} \mathrm{HM}_t}{\sum_{t=1}^{T} \mathrm{AM}_t}$$

$$\mathrm{HM}_t = \frac{N}{\sum_{i=1}^{N}(1/\log(\boxed{p_t^{(i)}}+2))}$$

$$\mathrm{AM}_t = \frac{\sum_{i=1}^{N}\log(\boxed{p_t^{(i)}}+2)}{N}$$

HM_t: harmonic mean of all households' power consumption at time t

AM_t: arithmetic mean of all households' power consumption at time t

$p_i^{(j)}$: **power consumption** of household **i** at time **j**

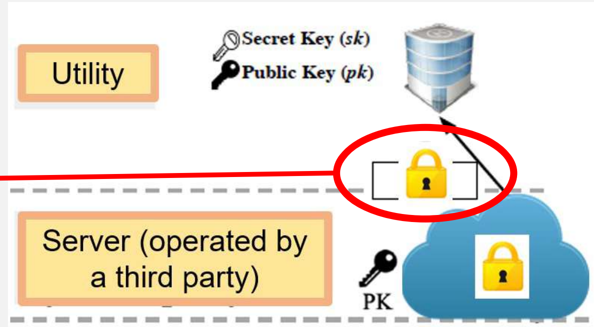
Problems to adopt FHE（HE）

1) Logarithm cannot be implemented with FHE (only addition/multiplication are adopted)
2) Division with a variable cannot be implemented (cannot calculate inverse)

before → after

$p_i^{(j)}$ is sent to the server

 Logarithms and divisions are required to calculate Q_d



idea 1) eliminating the calculation of logarithms at the server

Besides $p_t^{(i)}$, $\log(p_t^{(i)}+2)$ and $1/\log(p_t^{(i)}+2)$ are sent to the server

idea 2) eliminating the calculation of divisions at the server

Instead of sending $\mathrm{Enc}(Q_d)$, $\mathrm{Enc}(\mathrm{AM}_t)$ and $\mathrm{Enc}(\mathrm{HM}_t)$ from the server to the utility, separately. Then, the utility will decrypt them to calculate Q_d

[1]Y. Ishimaki, S. Bhattacharjee, H. Yamana and S. K. Das, "**Towards Privacy-preserving Anomaly-based Attack Detection against Data Falsification in Smart Grid**," *2020 IEEE Int'l Conf. on Communications, Control, and Computing Tech. for Smart Grids*, pp. 1-6, 2020. Int'l co-authorship Method1 Proposal

[2] Y. Ishimaki and H. Yamana, "**Faster Homomorphic Trace-Type Function Evaluation**," in IEEE Access, vol. 9, pp. 53061-53077, 2021 Speedup methods

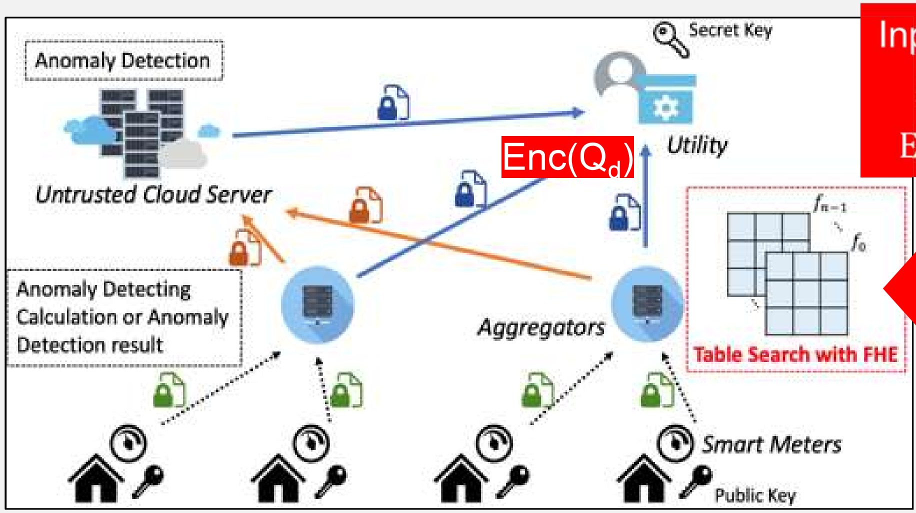
Preserving Privacy – Method 2

Problems to adopt FHE（HE）

1) Logarithm cannot be implemented with FHE (only addition/multiplication are adopted)
2) Division with a variable cannot be implemented (cannot calculate inverse)

idea : replacing $f(x, y, ...)$ to calculate Q_d to table lookup

Table lookup with a non-colluding server to adopt any kinds of calculations at the server (aggregators).



Inputs:

$Enc(\log(p_t^{(i)} + 2))$

$\mathrm{Enc}(1/\log(p_t^{(i)} + 2))$

Output :Enc(Q_d)

Server

index	Input values v_{in}	Output values v_{out}
0	x_0, y_0	$f(x_0, y_0)$
1	x_0, y_1	$f(x_0, y_1)$
...	...	...
$(n+1)^2$-1	x_n, y_n	$f(x_n, y_n)$

$Enc(x_0, y_1) \rightarrow 1 \rightarrow x_0, y_1$; $f(x_0, y_1) \rightarrow Enc(x_0, y_1)$

Private Information Retrieval(PIR) with FHE

Division is implemented by table lookup

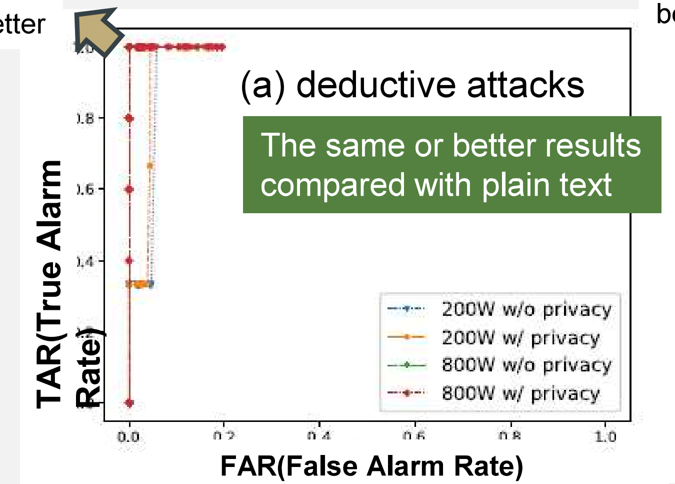
[3] Ruixiao Li, Yu Ishimaki and Hayato Yamana, "**Fully Homomorphic Encryption with Table Lookup for Privacy-Preserving Smart Grid**," Proc. of the 3rd IEEE International Workshop on Big Data and IoT Security in Smart Computing, pp.19-24 (2019.6) supported one input version

[4] Ruixiao Li, Yu Ishimaki, Hayato Yamana, "***Privacy Preserving Calculation in Cloud using Fully Homomorphic Encryption with Table Lookup***," Proc. of the 5th IEEE International Conference on Big Data Analytics (ICBDA2020), pp.315-322 (2020.05) supported any number of inputs version

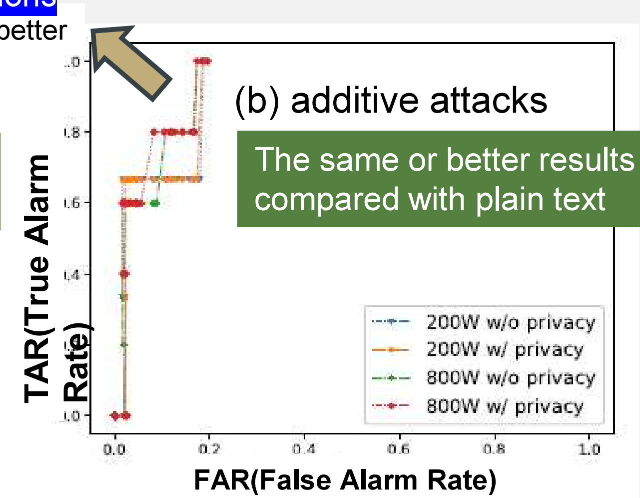
[5] Ruixiao Li and Hayato Yamana, ***Fast and Accurate Function Evaluation with LUT over Integer-based Fully Homomorphic Encryption***, Proc. of the 35th International Conference on Advanced Information Networking and Applications (AINA-2021), pp.620-633 (2021.5) supported any input value version / adopting approximation for inputs to select the best entry in the table

Preserving Privacy: Evaluation

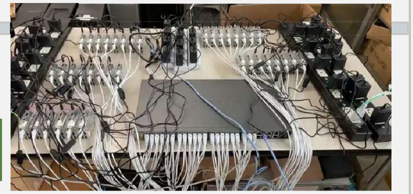
Method 1) eliminating FHE unfriendly calculations



The same or better results compared with plain text



The same or better results compared with plain text



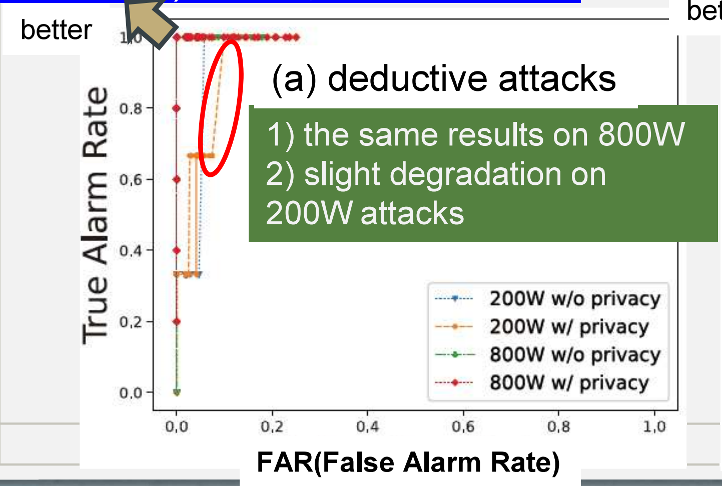
Execution time

	time [sec]
1) Aggregator	9.935
2) Utility	0.022
Total time	9.957

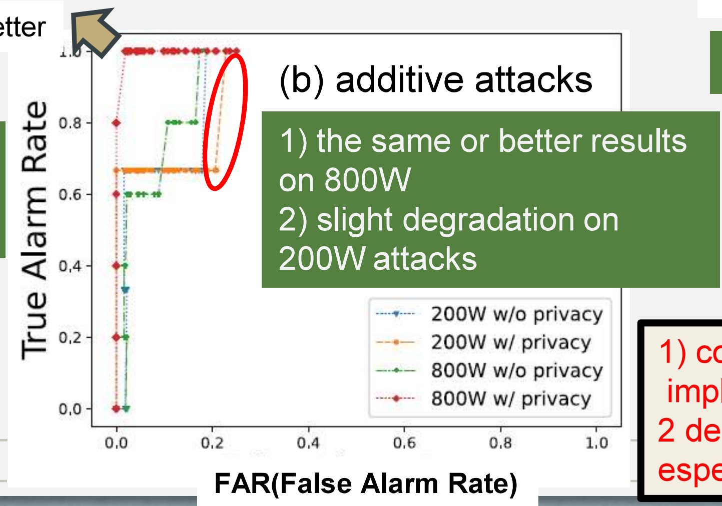
Aggregator：Intel Xeon E5-1620v4 3.5GHz(single thread)
Encrypted data size: 11,784KB(household→aggregator)
288KB(aggregator→utility)

※execution time at power meter is 0.112 sec/data

Method 2) adopting FHE table lookup



1) the same results on 800W
2) slight degradation on 200W attacks



1) the same or better results on 800W
2) slight degradation on 200W attacks

Execution time

	time
1) Aggregator	5 min

Input table size: 22,795 x 663
Output table size： 15,113,085
average power consumption is 1kWh/household)

1) confirmed **the same results** with plain text implementation w/ **800W attacks**
2 depending on table size, TAR/FAR will vary especially for small attacks.

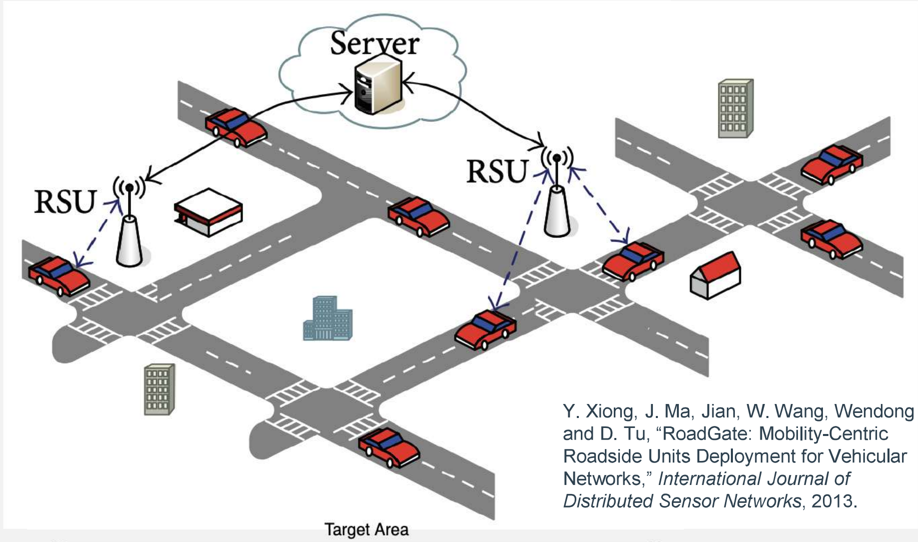
Thrust 3 Results:

(Abhishek Dubey, S. K. Das, S. Bhattacharjee, K. Yasumoto)

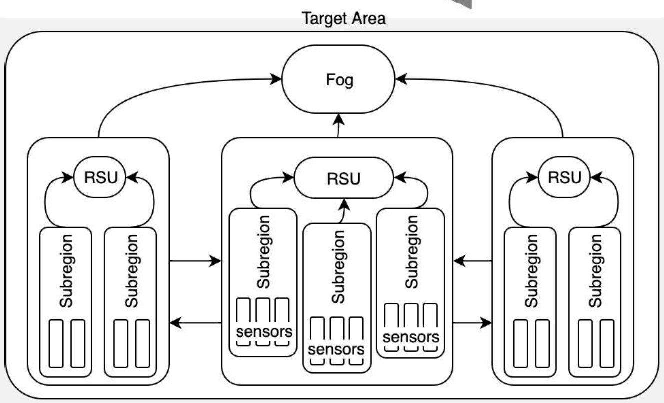
Security and Performance Tradeoff

Thrust 3 : Anomaly Detection & Performance Tradeoff

- Detect anomalies in real-time in transportation and energy data collected from urban areas.
- Correlated anomalies with incidents in urban transportation network.
- Data used:
  - ❖ Road segment and traffic mobility data
  - ❖ Traffic incidents: Nashville Police, Fire department data and Waze data
  - ❖ Weather data



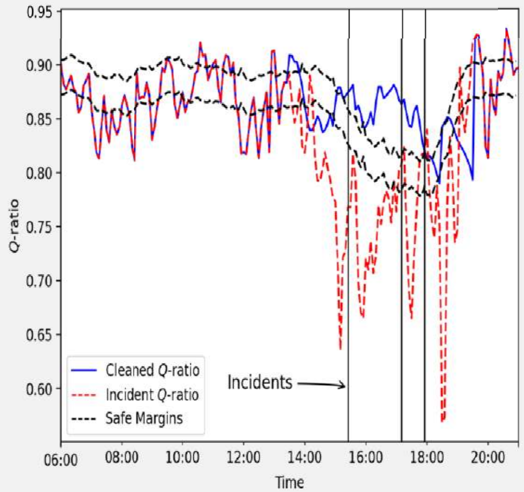
Y. Xiong, J. Ma, Jian, W. Wang, Wendong , and D. Tu, "RoadGate: Mobility-Centric Roadside Units Deployment for Vehicular Networks," *International Journal of Distributed Sensor Networks*, 2013.

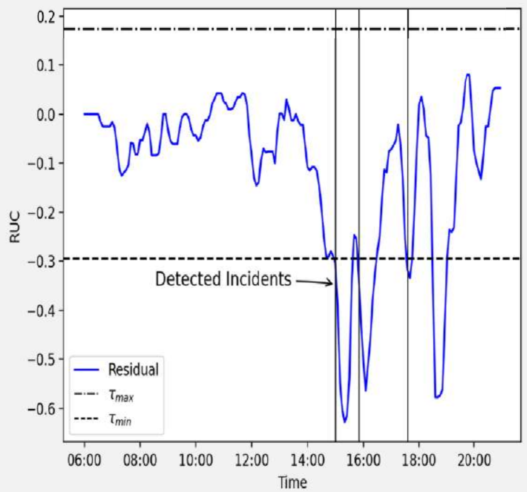


Proposed architecture for distributed anomaly detection

Thrust 3: Clustering Approach

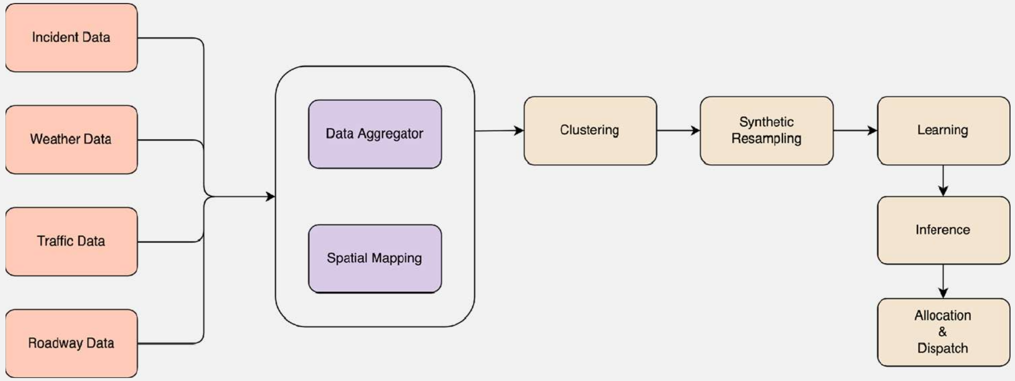
- Use clustering and statistical ratios to identify sudden changes.
- Averaged 2019 speed data into 7 day of week data and used only weekends
- Cluster segments using different spatio-temporal granularities
  - *Spatial*: Max distance between segments
  - *Temporal*: Resampling speed data

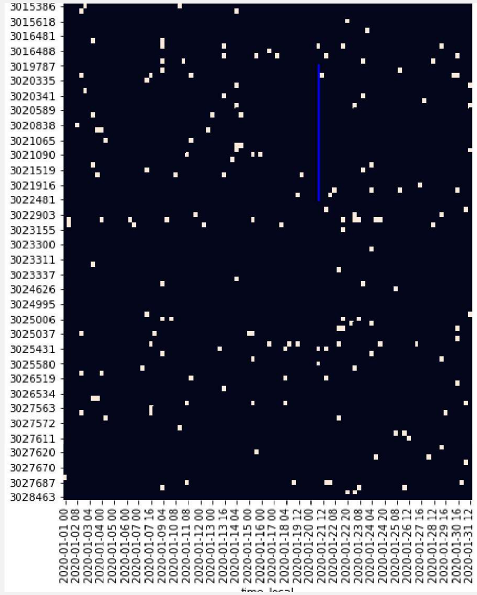




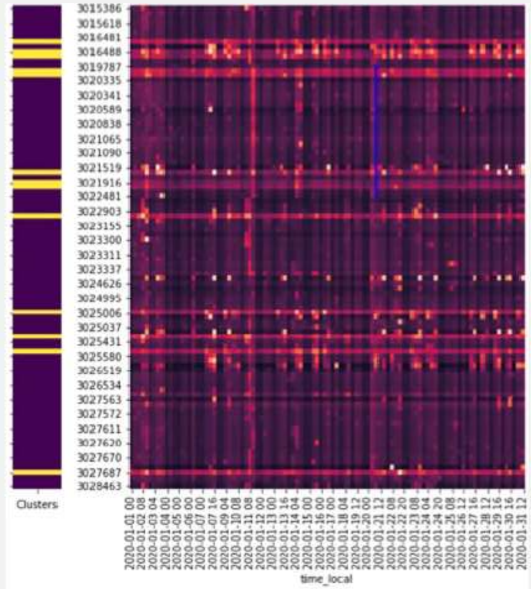
Thrust 3: Predicting Anomalies

- **Challenges**
  - **Big Data:** Many factors are involved in road accidents, which requires collecting various types of data from assorted resources with different resolution and quality.
  - **Sparsity:** Although frequency of road accidents is high, when viewed from the perspective of total time and space, incidents are rare events.
  - **Irregularity:** Accidents are random in nature, especially in high spatial-temporal resolution





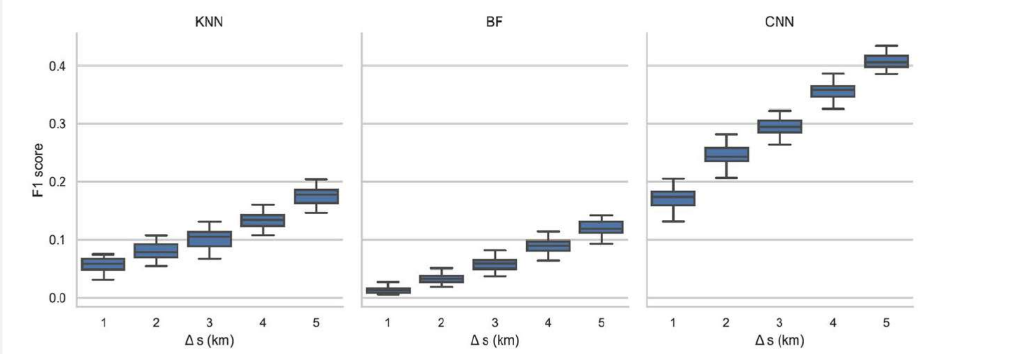
Test data over 4-hour window

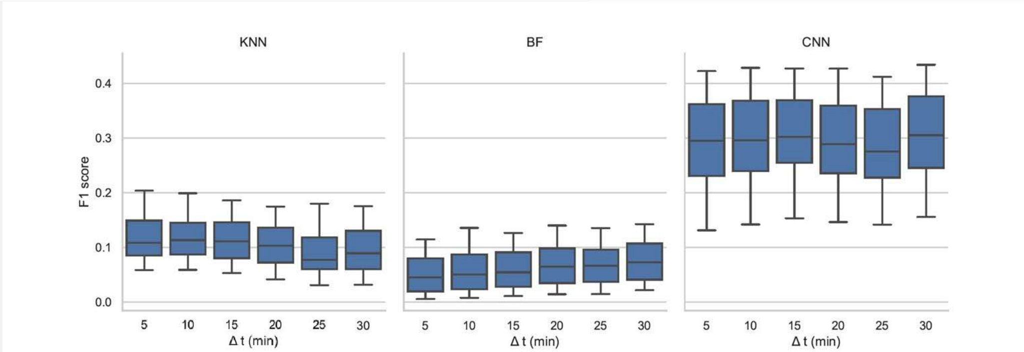


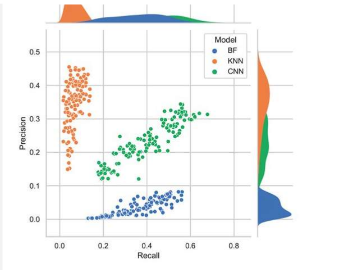
Prediction over 4-hour window

Thrust 3: Uncertainty Handling

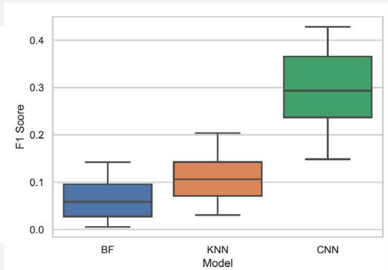
- **Challenge** – Uncertainty of data. This requires fine tuning of the detection thresholds.
- **Spatial discretization** – Reduces precision but improves robustness
- **Temporal discretization** – Reduces usefulness of detection metric. But high temporal resolution reduces recall.
- **Solution** – Uses pareto optimization with CNNs (convolutional neural networks)







CNN model can balance precision and
recall to achieve higher F-1 scores.



Thrust 4 Results:

(Abhishek Dubey, Keiichi Yasumoto)

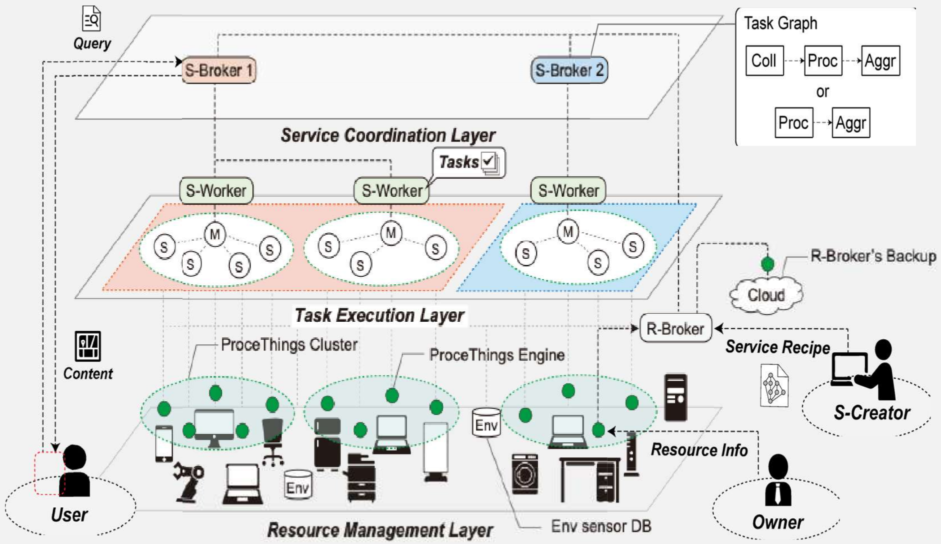
Developing Secure and Trustworthy Middleware Architecture

Tasks:

4.1 Distributed Aggregation
4.2 Secure Anonymization
4.3 Decision Making under Trade-offs

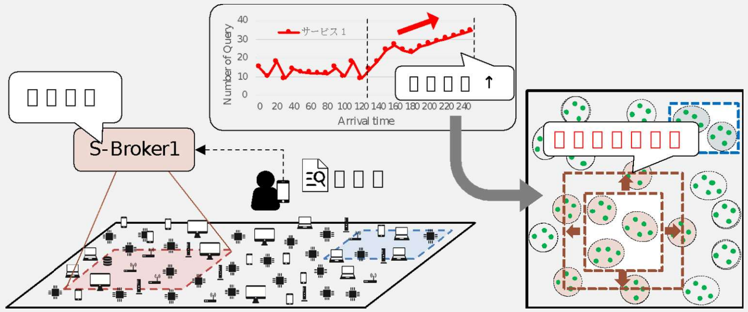
Task 4.1: Distributed Aggregation

Proposed a novel middleware framework for Smart and Connected Communities that distributes security features across proposed tasks and incorporates privacy, trustworthiness, resource constraints, and distributed decision support.



Middleware Platform [MUSICAL2021]

Developed a middleware platform for SCC apps through distributed processing among edge nodes [MUSICAL 2021]



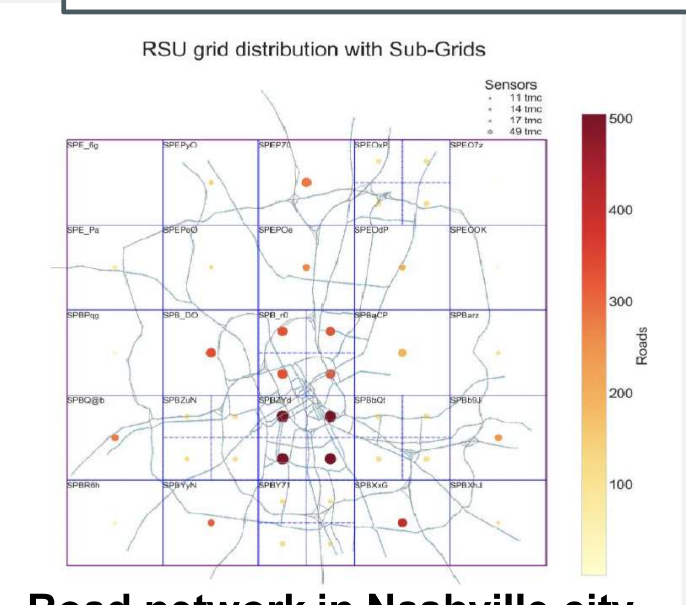
Adaptive scale-out mechanism

the number of edge nodes is increased as necessary

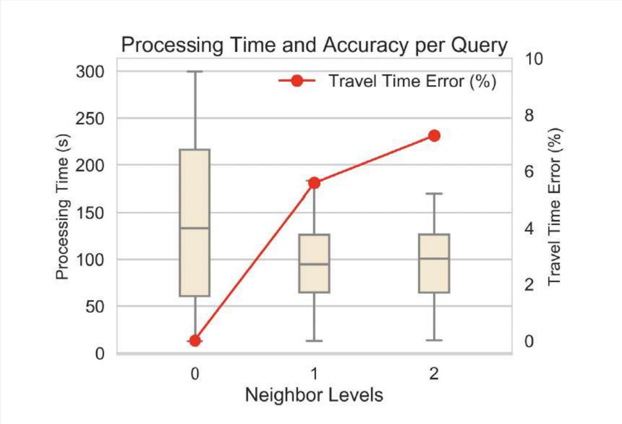
Smart Mobility App Developed on Middleware

Distributed route planning was implemented on the middleware [IEEE Access 2021]

For evaluation, emulation environment was constructed
- 49 Docker containers corr. to RSUs were virtually deployed over the 25 grids
- Each RSU running the middleware uses real traffic data to compute the shortest time paths
- “adaptive scale-out” is applied to the heavy-loaded grids (tasks are off-loaded to neighboring grids)

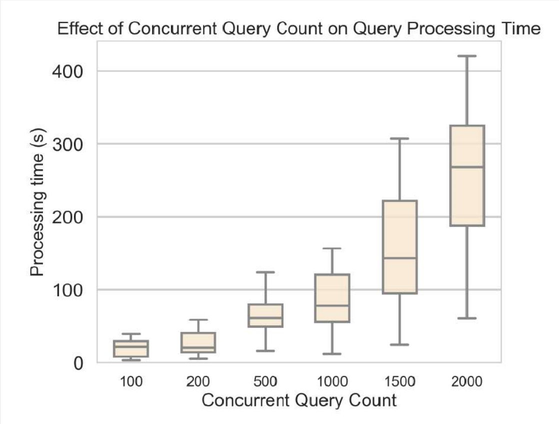


Road network in Nashville city
divided into 25 grids



Query response time for 1000 queries
(neighbor level means # hops for offloading tasks)

Offloading (adaptive scale-out) enabled a great reduction of resp. time

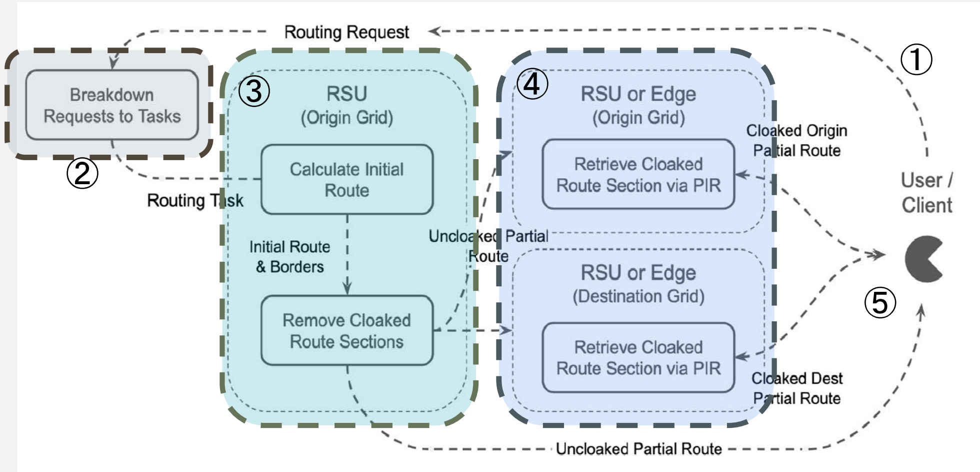


Response time for 100 to 2000 queries
(Neighbor level = 1)

Up to 2000 queries can be processed in practical time

Task 4.2: Secure Anonymization

Developed a secure anonymization mechanism based on APSP/PIR for smart mobility apps that allows users to get a query result securely without revealing origin/dest. points [SmartComp2021]



(i) User sends a query with blurred origin & dest. points (in grid level) to S-broker

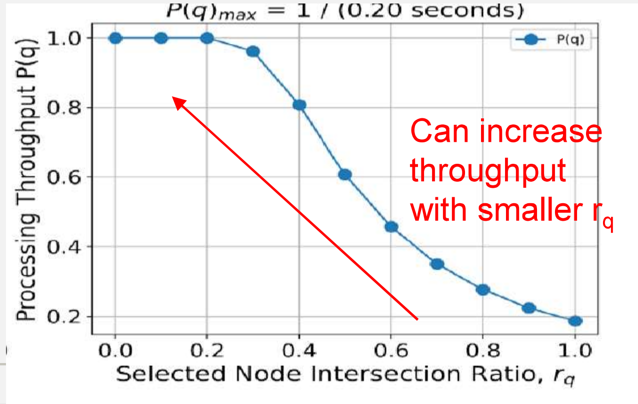
(ii) S-broker assign tasks for computing sub-shortest-paths to grids (RSUs) along the grid sequence

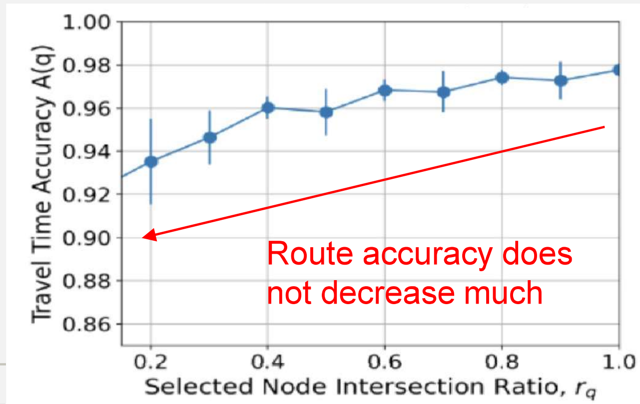
(iii) Each intermediate grid computes the shortest paths between border intersections

(iv) Source & dest. grids compute APSP (all-pairs-shortest-paths) between all intersections

(v) User gets sub-shortest paths (PIR is used to retrieve the paths for the source and destination grids) and concatenate them

- Applying APSP/PIR to all queries could cause some edge nodes (e.g., center grids) to be heavily loaded
- Reducing num. intersections can improve the throughput while keeping accuracy high enough.





Task4.3: Decision Making under Trade-offs

Developed a tradeoff mechanism that balances query throughput, route accuracy and privacy protection level [SmartComp2021]

We formulated multi-objective optimization problem and developed NSGA-II based algorithm

Set of queries at time slot t

$$Maximize\ (P(Q_t), V(Q_t), A(Q_t))\ s.t.\ (3)$$

Query throughput | Privacy level | Route accuracy

$$\forall g \in G, |\{q \mid q \in Q_t, src(q) = g \vee dst(q) = g\}| \leq Cap(g) \quad (3)$$

#queries processed at each RSU is limited

$$\mathcal{P}(q) = \frac{1}{\mathbf{1^T} \cdot [R(\mathbf{g_q}) + I(\mathbf{g_q}) + C(\mathbf{g_q})]}$$

Throughput = 1÷(route calc. + PIR calc. + delay)

$$\mathcal{V}(q) = \mathbf{1^T} \cdot [(\mathbf{a_q}^{\gamma} \odot \mathbf{v'_q}) \cdot C_{lp}]$$

Privacy level = area size x cover rate x coefficient

$$\mathcal{A}(q) = \mathbf{1^T} \cdot M(\mathbf{v_q}, \mathbf{r_q})$$

Accuracy = computed by #intersections and cover rate

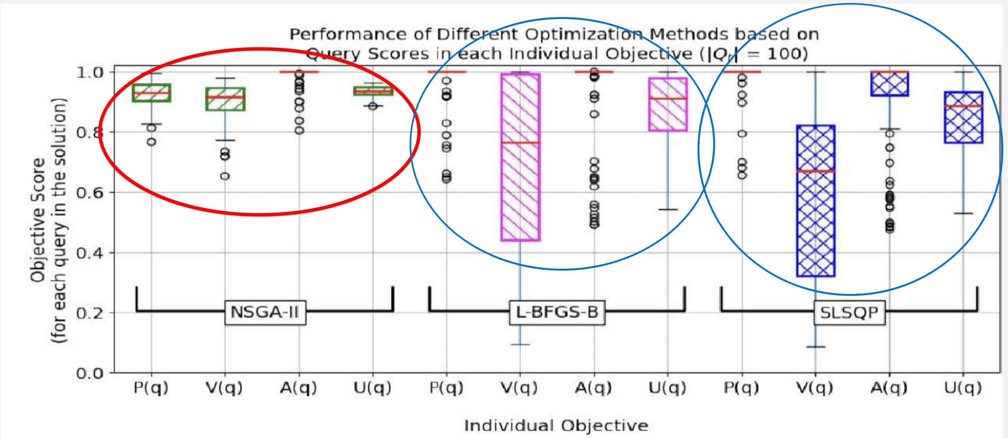
$$P(Q_t) = \frac{\sum_{q \in Q_t} \mathcal{P}(q) \cdot H_P(q)}{|Q_t|}$$

$$V(Q_t) = \frac{\sum_{q \in Q_t} \mathcal{V}(q) \cdot H_V(q)}{|Q_t|}$$

$$A(Q_t) = \frac{\sum_{q \in Q_t} \mathcal{A}(q) \cdot H_A(q)}{|Q_t|}$$

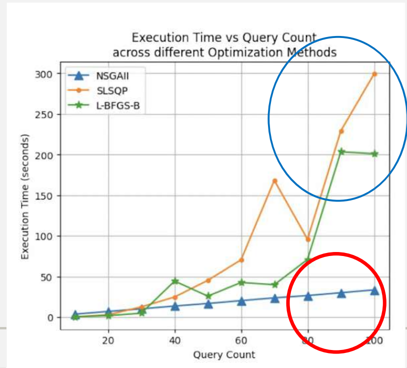
Aggregate by multiplying User preference Hx

Conducted simulations using Osaka city traffic data from T5.1



Ours (NSGA-II) outperforms others in all objectives: P, V, A, U

P: throughput, V: privacy level, A: accuracy, U: average utility



Our method outperforms others in execution time

Thrust 5 Results:

(A. Dubay, Hirozumi Yamaguchi)

Validation with Real Datasets

Tasks:

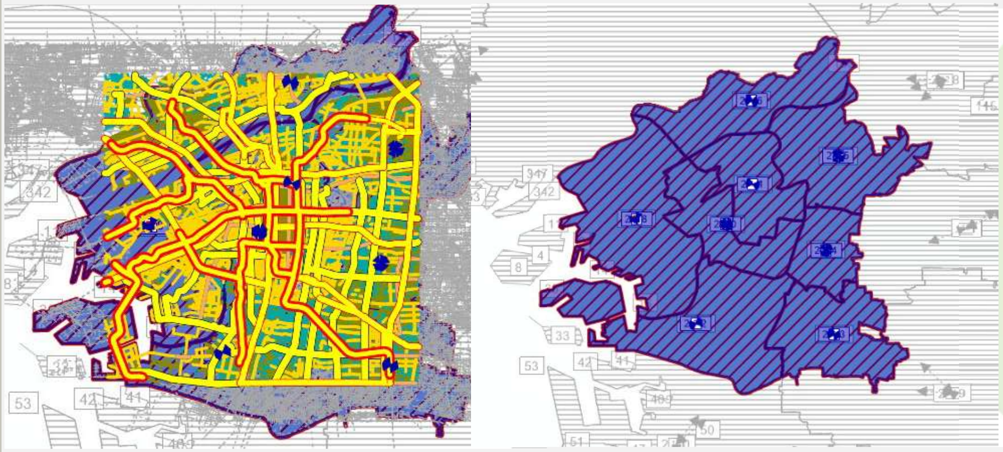
5.1 Smart Transportation Application
5.2 Smart Energy Application

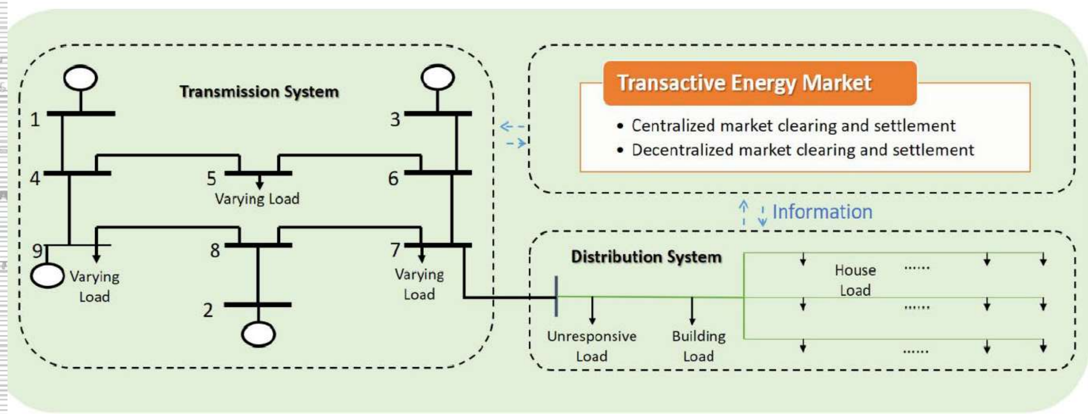
Validation with Real Datasets

Objective : Validate the proposed models and approaches using smart mobility and smart energy distribution / consumption scenarios with real-world datasets

Approach :

- Generate large-scale mobility data from real datasets in Osaka
- Design a transactive energy testbed that can integrate energy market data





Task 5.1: Smart Transportation Application

VICS (obtained via IR beacons) : contains queue length of major city roads and highways (Osaka prefecture whole region)

is simulated using real dataset

Infrared beacons (Ordinary trunk roads)

Infrared beacons are installed on the ordinary trunk roads and provide information covering about 30 km in the forward direction and about 1 km in the rear direction.

- Traffic congestion and travel time information.
- Information on restrictions due to accidents, construction, disasters, and weather conditions.
- Parking availability.

Link Traffic

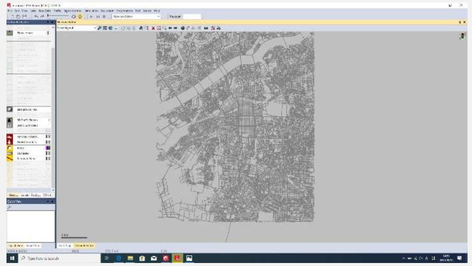
time / road	5:00	6:00	・・・	4:00
road a	a1	a2		a24
road b	b1	b2		b24
・・・				
road z	z1	z2		z24

estimation

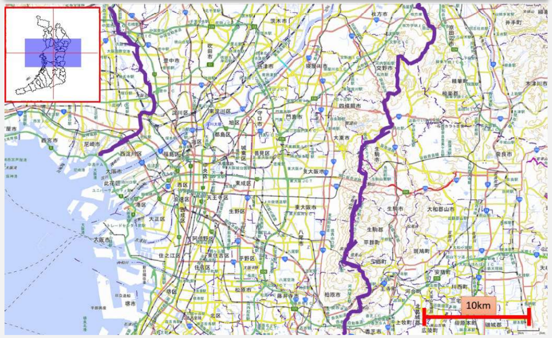
OD Matrix

OD	①	②	③
①	AA	BB	CC
②	DD	EE	FF
③	GG	HH	II

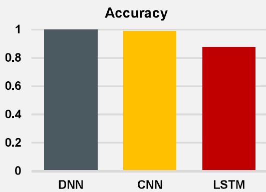
simulation

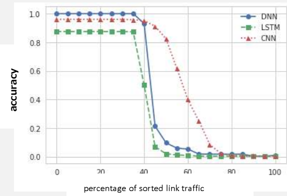


Road Traffic Census data (obtained by nation-wide survey by Ministry Road Bureau)



1. generate possible OD patterns
2. for each OD matrix, simulate vehicles and measure the link traffic to annotate OD matrix label to link traffic data
3. train DNN, LSTM and CNN using this data that predicts OD matrix from link traffic (5,378 links for 24 hours in case of Osaka)

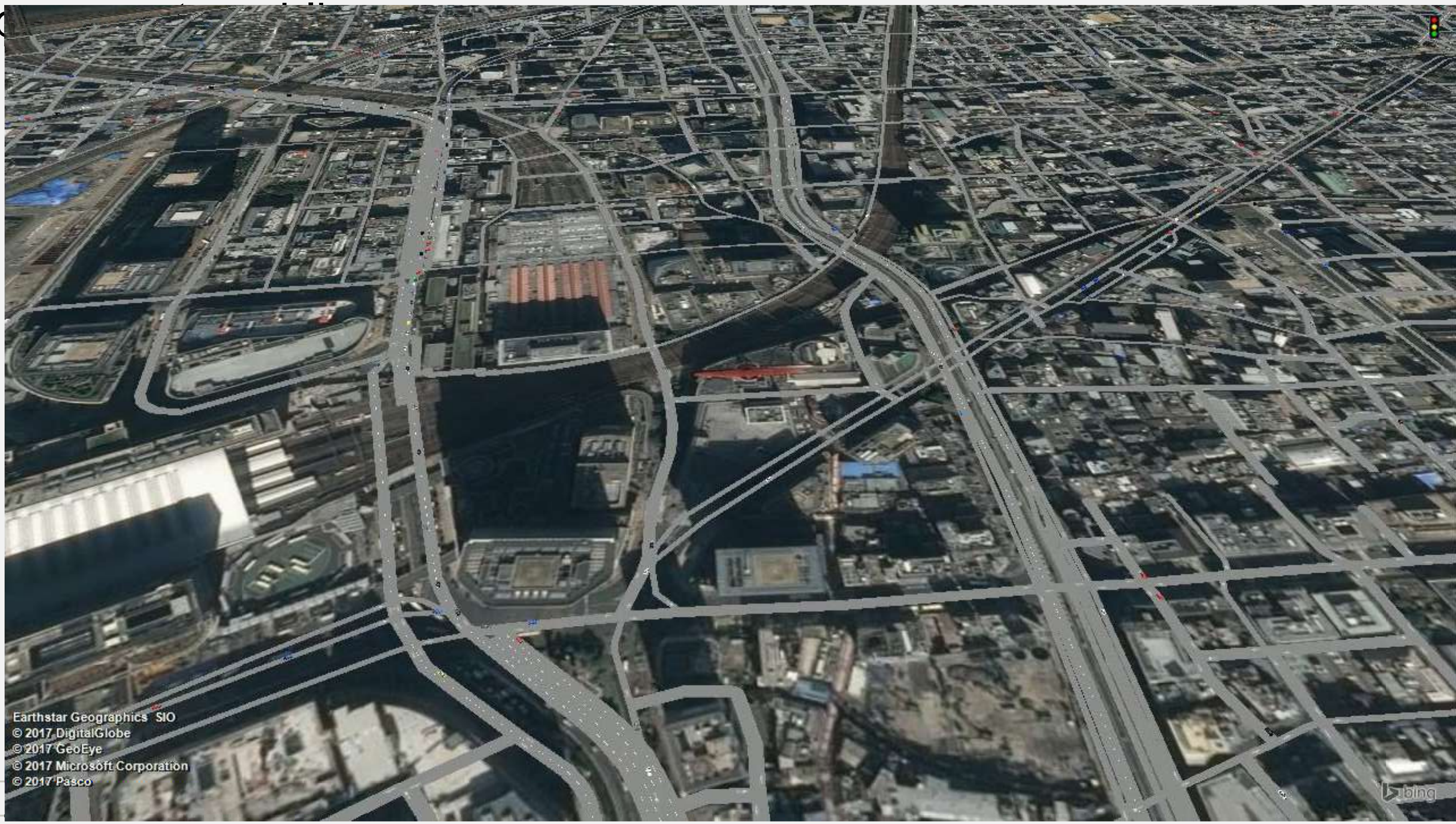




100% accuracy (DNN with all links traffic)
80% accuracy (CNN with 50% link traffic)

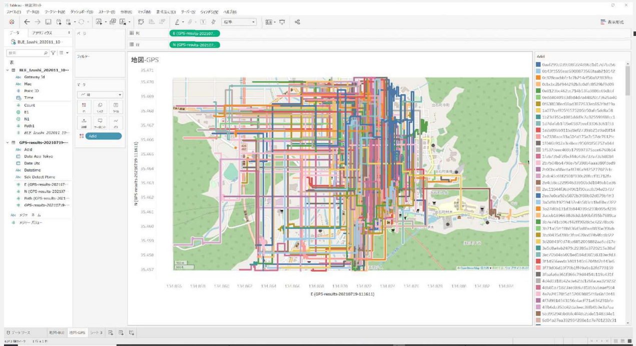
Simulated Vehicles in Osaka Downtown

- will b

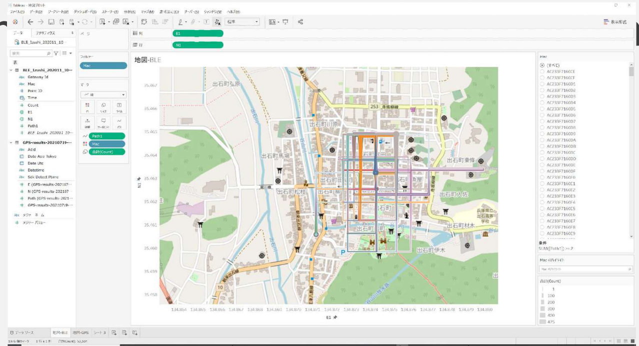


Simulation with Vehicles and Pedestrians

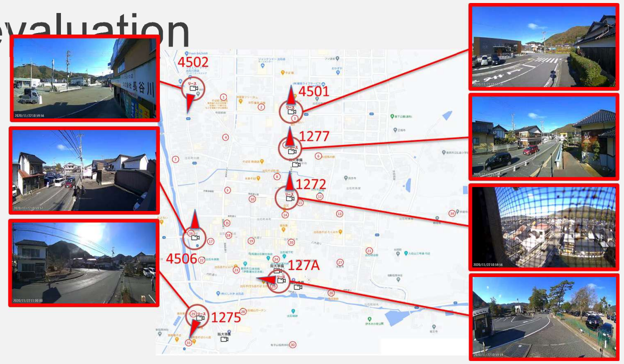
- generating vehicle/human mobility data of Toyooka city (Hyogo prefecture)
  - using anonymized GPS/BLE traces as well as link traffic volume data from the city
  - -m … n evaluation



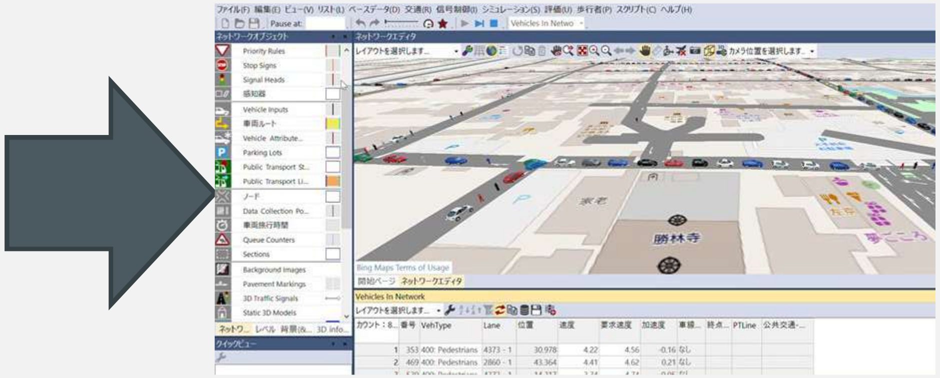
500+ GPS Traces (grid)



170 BLE Traces (grid)

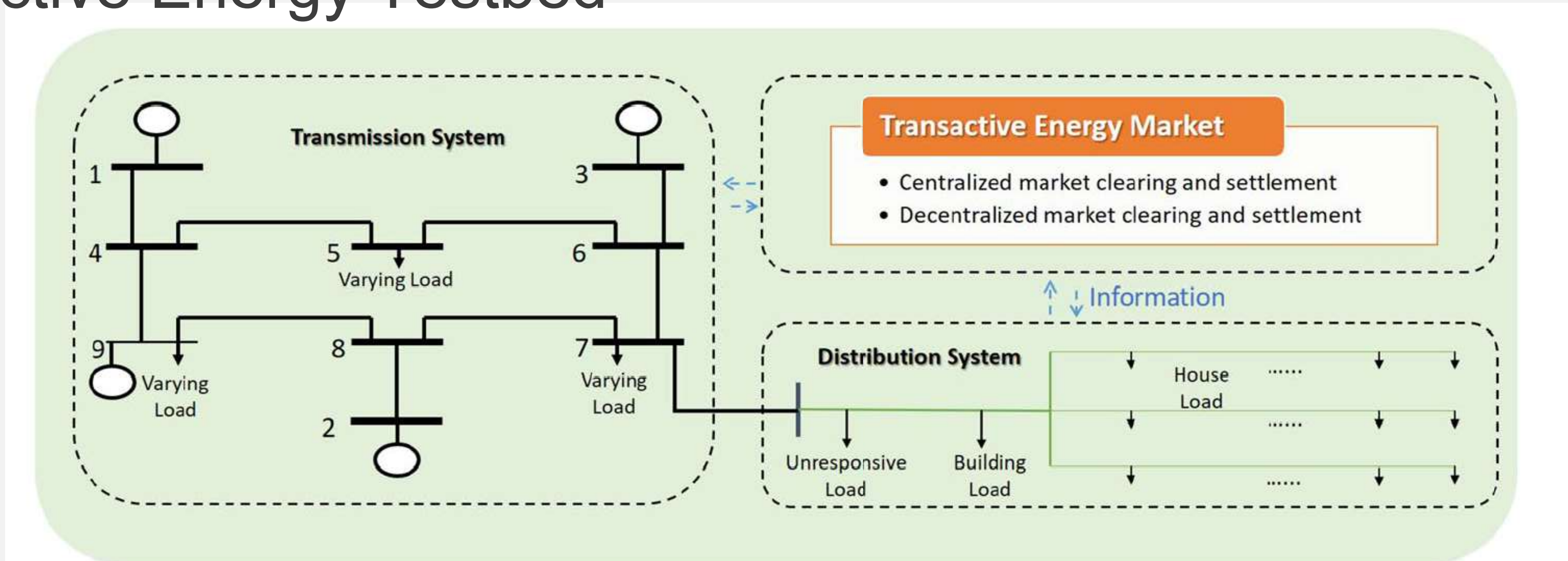


Vehicle/Pedestrian Counting via Roadside Cameras



Task 5.2 Smart Energy Application

Transactive Energy Testbed



Architecture of TESST

Physical System | Transactive Energy Market | Network Simulator

Centralized Market Option | Decentralized Market Option

STEAM Project: Broader Impacts

➢ Interdisciplinary Education and Experiential Learning for Students:

- Prithwiraj Roy and Venkata Praveen Madhavarapu (Missouri S&T), PhD - graduating in fall 2021
- Michael Wilbur; Geoff Pettet (Vanderbilt Univ.)
- Yu Ishimaki; Ruixiao Li (Waseda Univ.) - graduated
- Jose Paolo Talusan; Francis Tiausas (NAIST)
- S. Choochotkaew; Yuki Akura (Osaka Univ.)

➢ Student Visit Exchanges:

- Y. Ishimaki (Waseda) visited MST in Aug-Sep 2018 for one month, and WMU for 2 weeks in June and Oct 2019.
- V. P. Madhavarapu and P. Roy (Missouri S&T) visited WMU for 4 weeks in July and Aug 2019, respectively.
- J. P. Talusan (NAIST) visited Vanderbilt for 3 weeks in June 2019.
- M. Wilbur (Vanderbilt) visited WMU for 1 week in 2019.

➢ Integration of Research into Courses:

- Missouri S&T: Developed and taught a new course on *Advances in CPS Security*, spring 2021.
- Vanderbilt Univ.: Incorporated anomaly detection module in *Reliable Distributed Systems*, fall 2019.
- WMU: Covered CPS and smart grid security in *Science of Cybersecurity*, spring 2019.

➢ Outreach Activities:

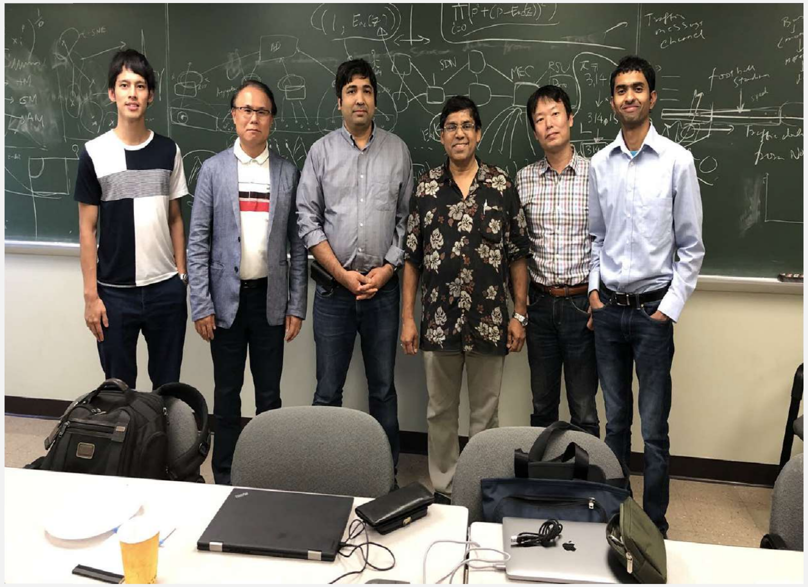
- **Organized** Big Data and IoT Security (BITS) workshop in conjunction with IEEE SmartComp 2019-2021.
- Organized Science of Smart City Operations and Platforms Engineering (SCOPE) workshop during CPS-IoT Week, 2019.
- Led to two NSF projects from CNS and SaTC programs (MST, WMU).
- Supported UG students in research.; mentored high school students.
- Das delivered Keynote Talks at various conferences.

Coordination and Collaboration

- Bi-weekly Skype/Zoom meeting; Very coherent group
- NUmerous joint publications by PIs and their students
- Co-organization of BITS and SCOPE workshops
- Planned Vision Paper: Security in Integrated Energy and Mobility
- Planned Special Issue Editing: Magazine and/or Journal

All Hands Meeting:

- Missouri S&T: Sept 14-15, 2018
- Tokyo, Japan: Oct 26-27, 2018 (JUNO2 Kick-off Meeting)
- Kyoto, Japan: March 11-14, 2019 (IEEE PerCom)
- Washington, DC: June 12-14, 2019 (IEEE SmartComp)
- Chicago: Oct 11, 2019 (JUNO2 PI Meeting)
- Bologna, Italy: June 20-23, 2020 (IEEE SmartComp) - NO
- Nara, Japan: January 5-8, 2021 (ACM ICDCN) – NO due to Covid-1



September 2018 (Missouri S&T)

(Collaborative) Publications

1. S. Roy, N. Ghosh, and S. K. Das, "bioSmartSense: A Bio-inspired Data Collection Framework for Energy-efficient, QoI-aware Smart City Applications," *17th Annual IEEE International Conference on Pervasive Computing and Communications* (PerCom), Kyoto, Mar 2019.
2. Y. Nishimura, A. Fujita, A. Hiromori, H. Yamaguchi, T. Higashino, A. Suwa, H. Urayama, S. Takeshima and M. Takai, "A Study on Behavior of Autonomous Vehicles Cooperating with Manually-Driven Vehicles," *17th Annual IEEE PerCom*, pp. 212-219, Kyoto, Mar 2019.
3. J. P. Talusan, K. Yasumoto, et al, "Evaluating Performance of In-Situ Distributed Processing on IoT Devices by Developing a Workspace Context Recognition Service," *IEEE PerCom Workshop, Kyoto, Mar 2019.*
4. H. Yamaguchi, "Toward Urban Vehicle Mobility Modeling in Japan," *4th International Science of Smart City Operations and Platforms Engineering Workshop* (SCOPE), pp. 1-6, 2019.
5. R. Li, Y. Ishimaki and H. Yamana, "Fully Homomorphic Encryption with Table Lookup for Privacy-Preserving Smart Grid," *IEEE BITS2019 Workshop*, pp. 19-24, June 2019.
6. M. Wilbur, A. Dubey, B. Leão and S. Bhattacharjee, "A Decentralized Approach for Real Time Anomaly Detection in Transportation Networks," *4th IEEE International Conference on Smart Computing* (SMARTCOMP), Washington, DC, pp. 274-282, June 2019.
7. J. P. Talusan, K. Yasumoto, A. Dubey, S. Bhattacharjee, "Smart Transportation Delay/Resilience Testbed using Information Flow of Things Middleware," *IEEE BITS Workshop*, 2019.
8. Y. Ishimaki, H. Yamana, "Non-Interactive and Fully Output Expressive Private Comparison," *INDOCRYPT*: 355-374, 2018.
9. S. Bhattacharjee and S. K. Das, "Detection and Forensics against Stealthy Data Falsification in Smart Metering Infrastructure," *IEEE Transactions on Dependable and Secure Computing*, 18(1): 356-371, Jan/Feb 2021.
10. S. Bhattacharjee, V. K. P. Madhavarapu, S. Silvestri, and S. K. Das, "Attack Context Embedded Data Driven Trust Model in Smart Metering Infrastructure," *ACM Transactions on Privacy and Security*, 24(2): 9:1-9:36, Apr 2021.
11. R. P. Barnwal, N. Ghosh, S. K. Ghosh, and S. K. Das, "Publish or Drop Traffic Event Alerts? Quality-aware Decision Making in Participatory Sensing-based Vehicular CPS," *ACM Transactions on Cyber-Physical Systems*, to appear, 2019.
12. S. Bhattacharjee, N. Ghosh, V. K. Shah, S. K. Das, "QnQ: A Quality and Quantity Unified Approach for Secure and Trustworthy Crowdsensing, *IEEE Transactions on Mobile Computing*, 19(1): 200-216, Jan 2020.
13. J. P. Talusan, M. Wilbur, A. Dubey, K. Yasumoto, "On Decentralized Route Planning Using the Road Side Units as Computing Resources," *IEEE International Conference on Fog Computing (ICFC 2020)*, pp. 1-8, Apr. 2020.
14. A. Sturaro, S. Silvestri, M. Conti, and S. K. Das, "A Realistic Model for Failure Propagation in Interdependent Cyber-Physical Systems," *IEEE Transactions on Network Science and Engineering*, (Special Issue on Network Science for High-Confidence Cyber-Physical Systems), 7(2): 817-831, Apr-Jun 2020.
15. M. Wilbur, C. Samal, J. P. Talusan, K. Yasumoto, A. Dubey, "Time-dependent Decentralized Routing using Federated Learning," *23rd IEEE International Conference on Real-Time Distributed Computing (ISORC 2020)*, pp. 56-64, May 2020.
16. T. Limbasiya, D. Das, and S. K. Das, "MComIoV: Secure and Energy-Efficient Message Communication Protocols for Internet of Vehicles," *IEEE/ACM Transactions on Networking*, 29(3): 1349-1361, June 2021.
17. H. Vasudev, V. Deshpande, D. Das, and S. K. Das, "A Lightweight Mutual Authentication Protocol for V2V Communication in Internet of Vehicles," *IEEE Transactions on Vehicular Technology*, 69(6): 6709-6717, June 2020.
18. J. P. Talusan, M. Wilbur, A. Dubey, K. Yasumoto, "Route Planning Through Distributed Computing by Road Side Units," *IEEE Access*, Vol. 8, pp. 176134-176148, 2020.
19. S. Bhattacharjee, V. P. Madhavarapu, and S. K. Das, "A Diversity Index based Scoring Framework for Identifying Smart Meters Launching Stealthy Data Falsification Attacks," *ACM Asia Conference on Computer and Communications Security* (ASIACCS), Hong Kong, pp. 26-39, June 2021.
20. P. Roy, S. Bhattacharjee, and S. K. Das, "Real Time Stream Mining based Attack Detection in Distribution Level PMUs for Smart Grids," *IEEE Global Communications Conference (GlobeCom) – Symposium on Smart Grid Communications and Power Line Communications*, Taipei, Taiwan, Dec 2020.
21. Y. Ishimaki, S. Bhattacharjee, H. Yamana, and S. K. Das, "Towards Privacy-Preserving Anomaly-based Attack Detection against Data Falsification in Smart Grid," *IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids* (SmartGridComm) – *Cyber Security and Privacy Symposium*, Nov 2020.
22. F. Tiausas, J. Talusan, Y. Ishimaki, H. Yamana, H. Yamaguchi, S. Bhattacharjee, A. Dubey, K. Yasumoto, and S. K. Das, "User-centric Distributed Route Planning in Smart Cities based on Multi-objective Optimization," *IEEE International Conference on Smart Computing* (SMARTCOMP), Irvine, California, Aug 2021.

JUNO2: US-Japan Collaborative Project

STEAM: Secure and Trustworthy Framework for Integrated Energy and Mobility in Smart Connected Communities

PI Meeting: Aug 18 -19, 2021

Thank You