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)

Shameek Bhattacharjee

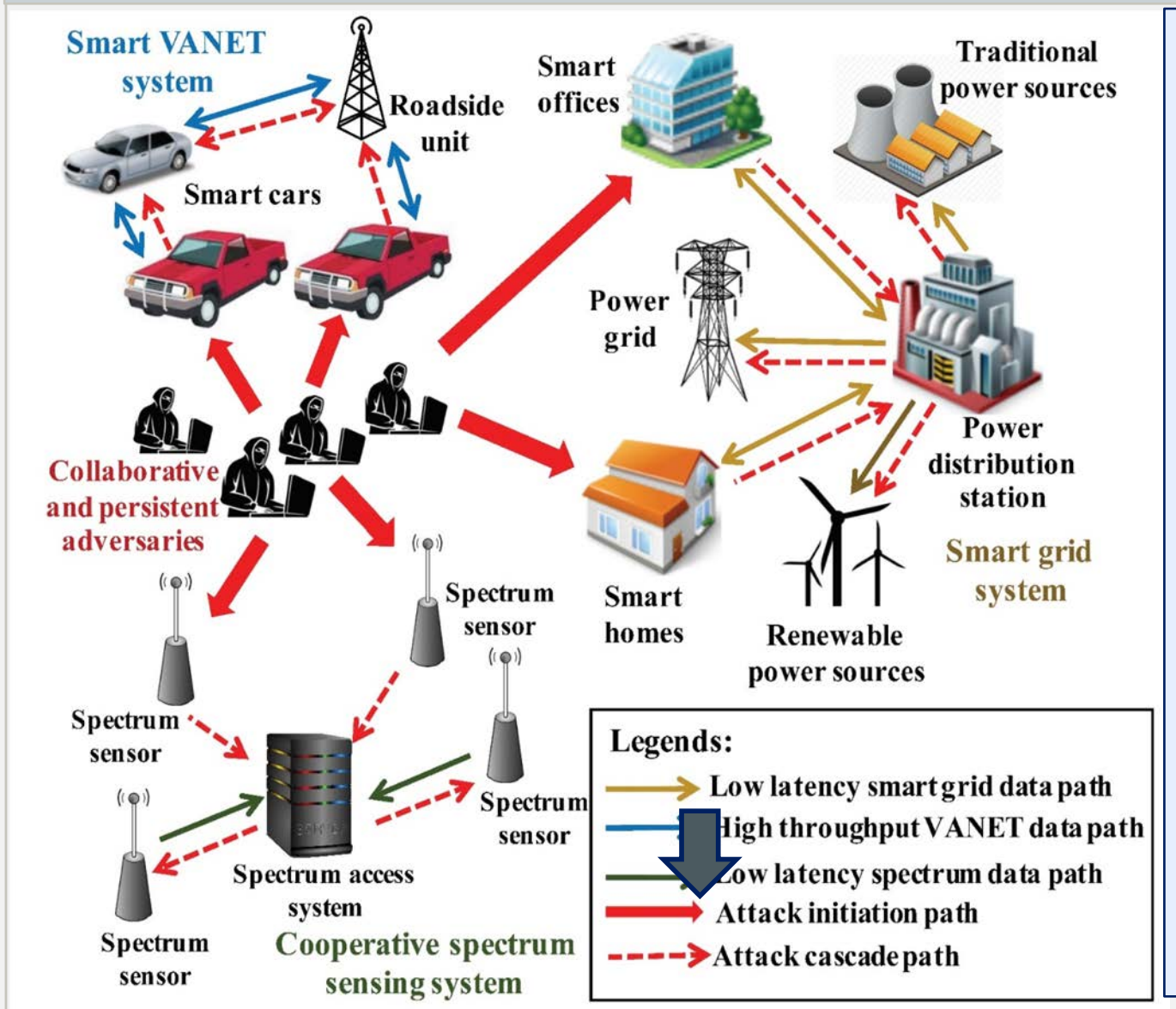
Abhishek Dubey

Hayato Yamana

Hirozumi Yamaguchi

Keiichi Yasumoto

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.

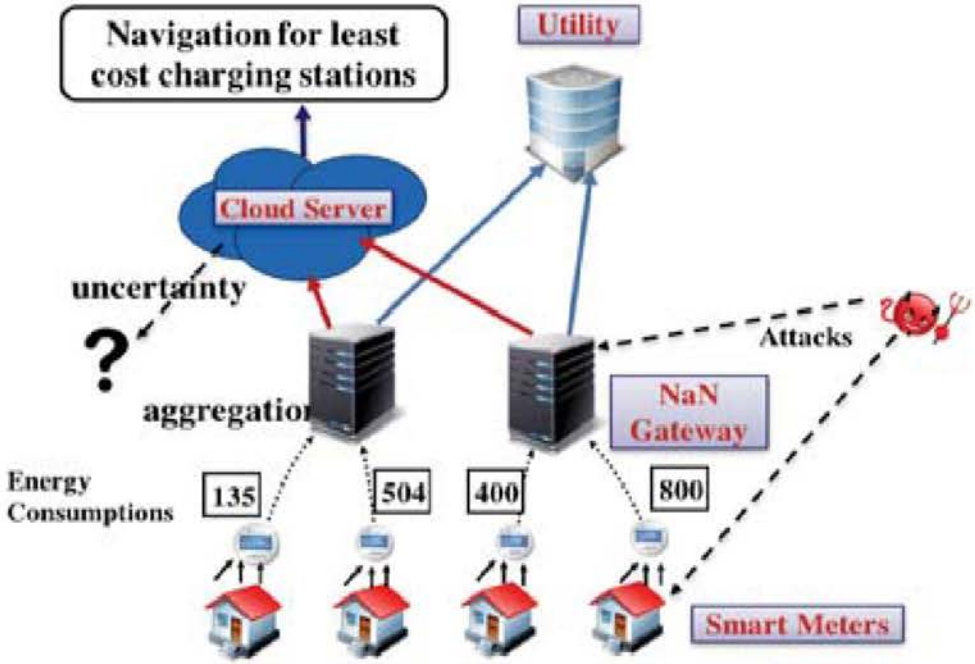
A integrated treatment of smart energy and smart mobility is important.

Integrated Treatment of Energy and Mobility

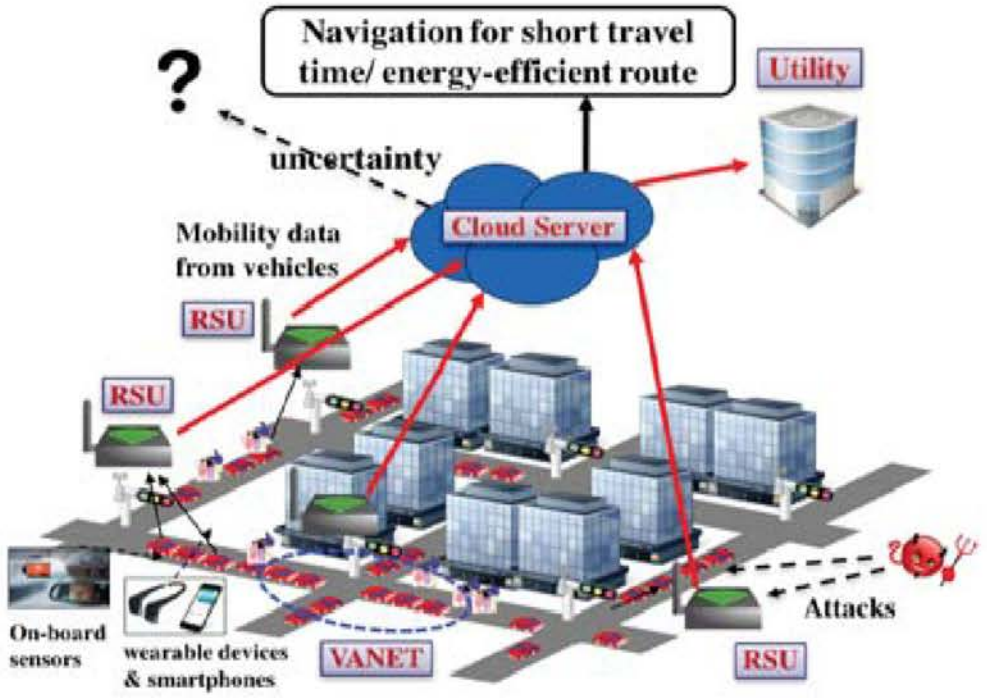
Similarity: Hierarchical structure, spatiotemporal relationships

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

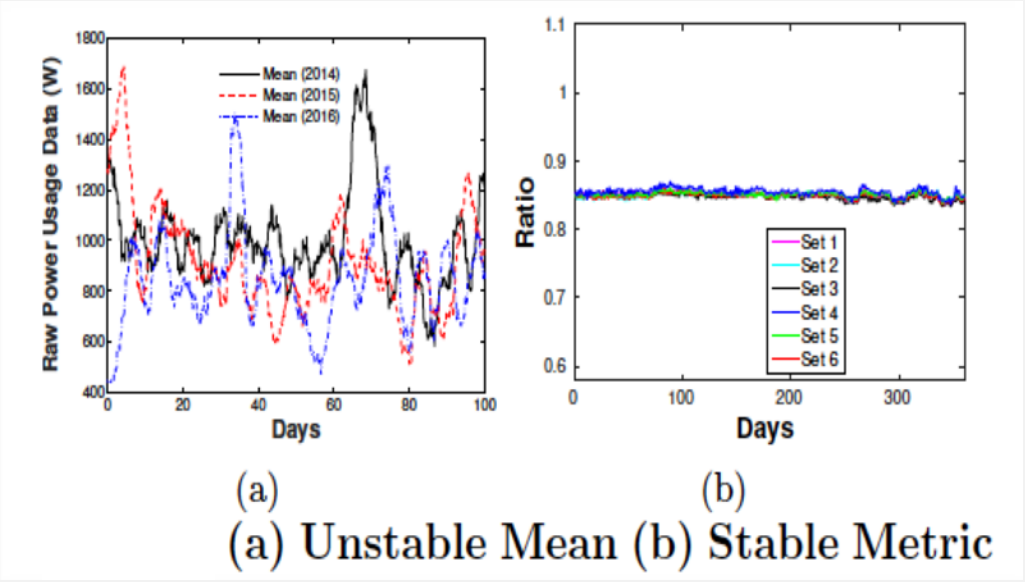
Task 1.1 Anomaly Detection

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].

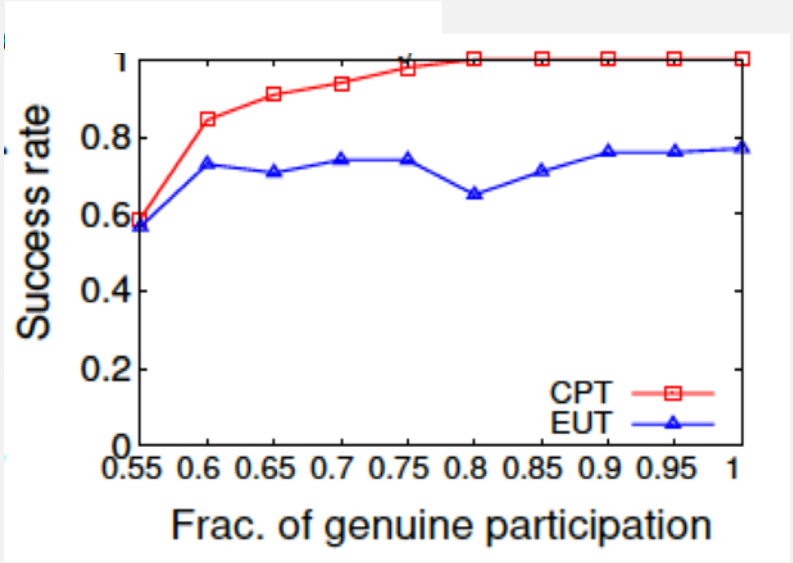
Task 1.3 Dependable Decision Models

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:

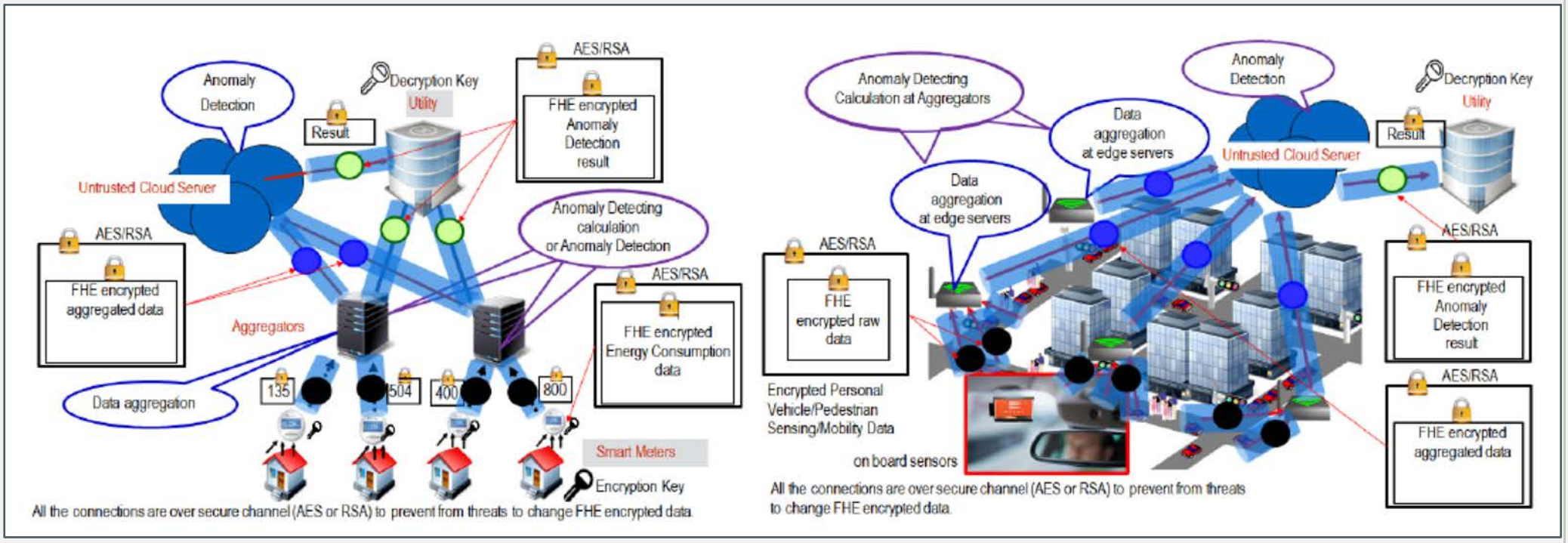
- Among all encryption schemes, only FHE enables additions and multiplications over encrypted data. By adopting FHE for calculation in Thrust 1, the individual raw data are not exposed to others.
- Drawbacks of existing FHE that prevent their adoption in SCC:
  - Too high complexity (space and time) → 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

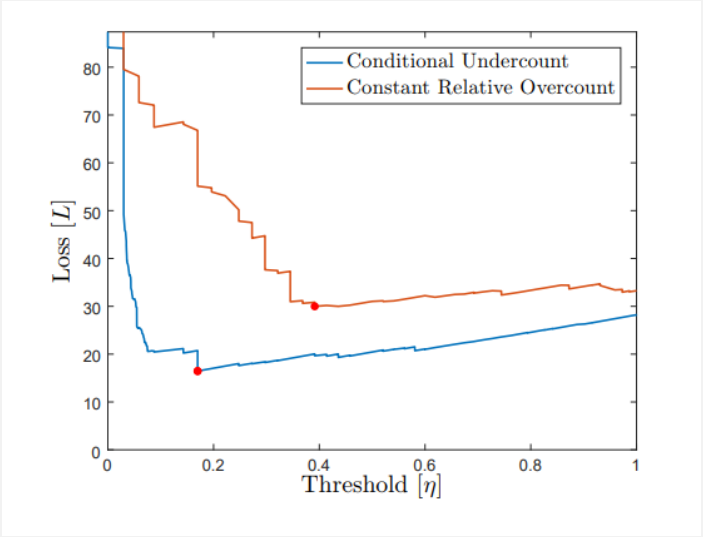
Thrust 3: Current Status

State of the Art:

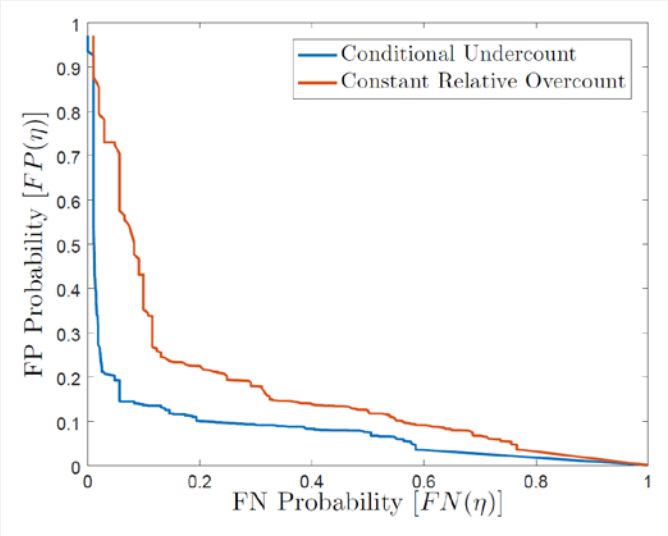
- SCC applications are typically data-driven, where sensors collect data at unprecedented scale and granularity. Since data privacy and integrity are two major concerns with such data, addressing their trade-off is crucial.
- Performance impact of state of the art methods on SCC depends on the choice 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

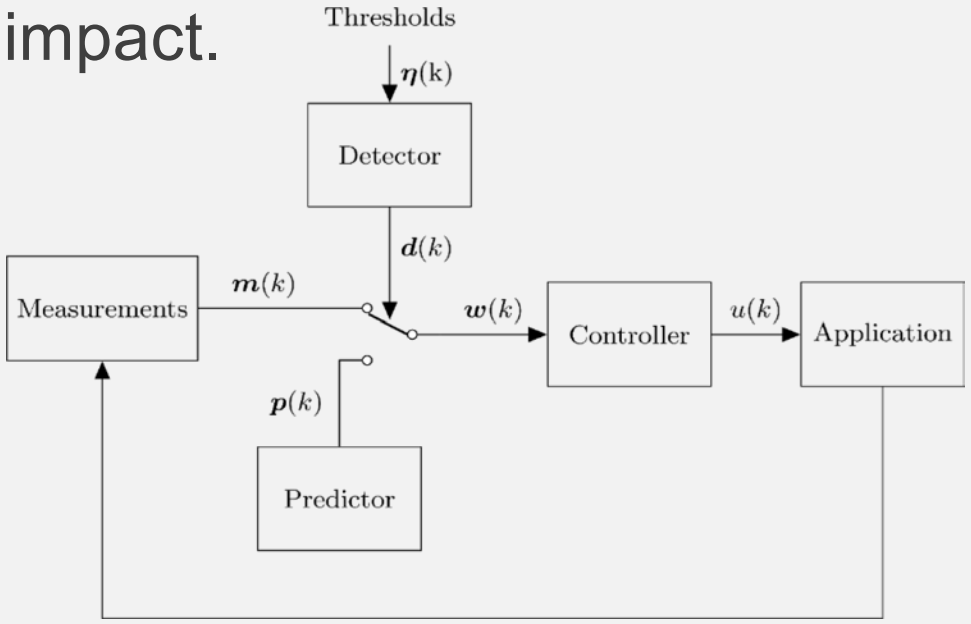


ROC curve – Precision vs Recall

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.

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



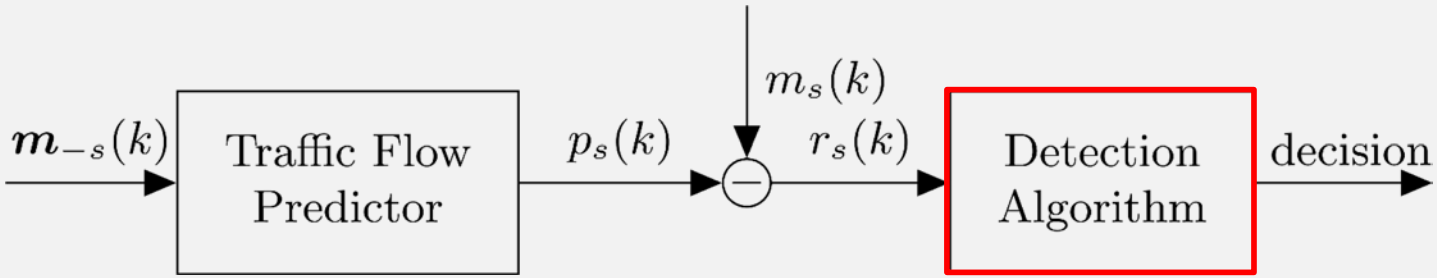
Example: Application Aware Anomaly Detection Thresholds

- 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: *Prof. Keiichi Yasumoto* (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:

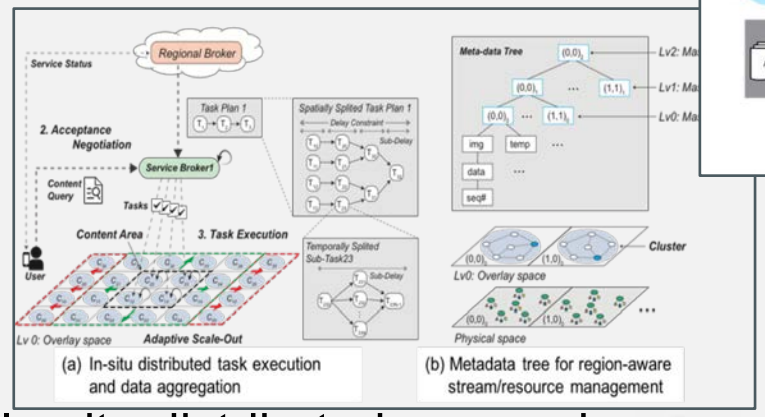
- Delay-awareness is essential to provide up-to-date mobility and energy information in a smart city for navigation services (e.g., movement and charging of PEVs).
- Edge/fog computing-based distributed architecture can minimize delay and wastage of bandwidth/resources for cloud servers.
- No attempt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

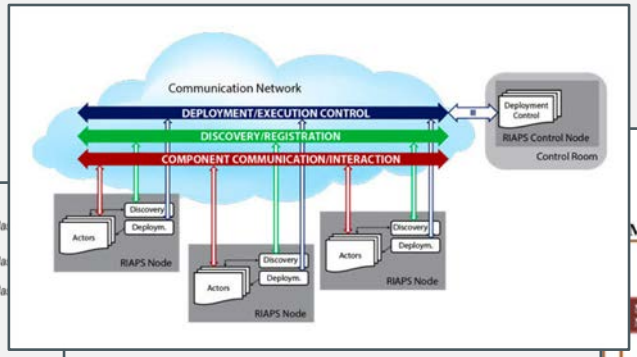
Preliminary Work:

- 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/)

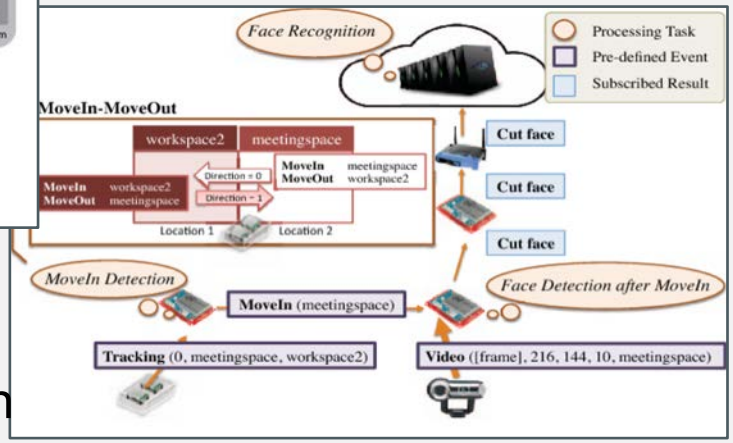
Thrust 4: Proposed Approach

- 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:

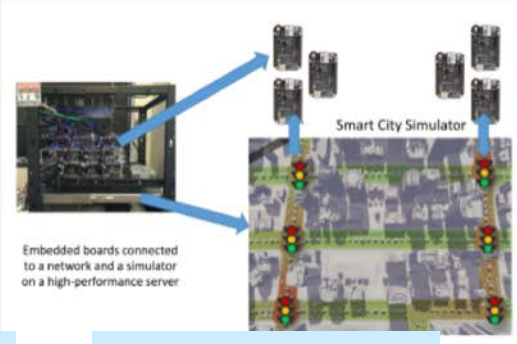
5.1 Smart Transportation Application

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

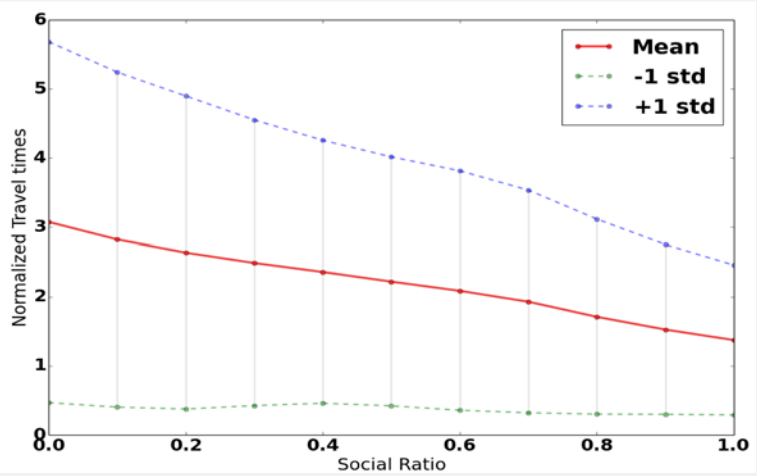
Available Testbeds and Datasets:

- **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

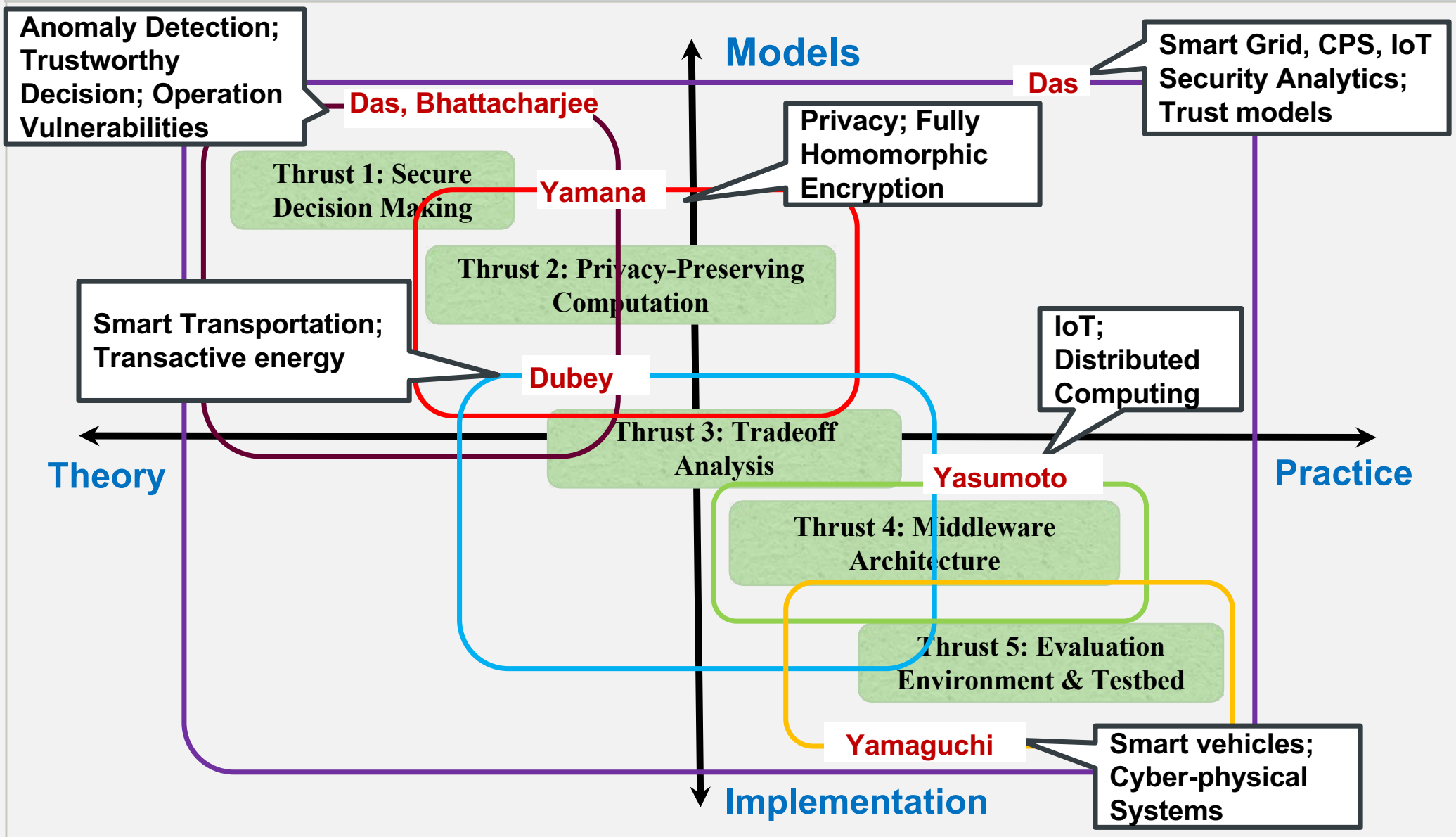
Smart Energy Application

- 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:

- 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

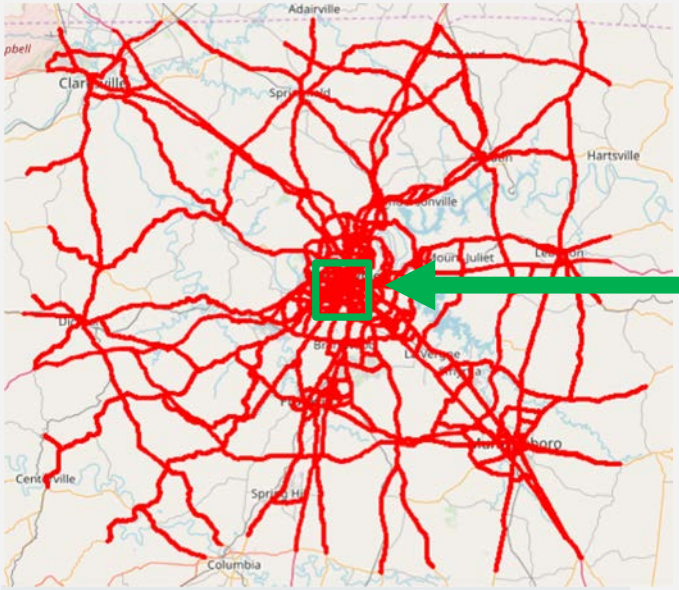
September 2018 meeting at MST

Inter-institutional Student Visits:

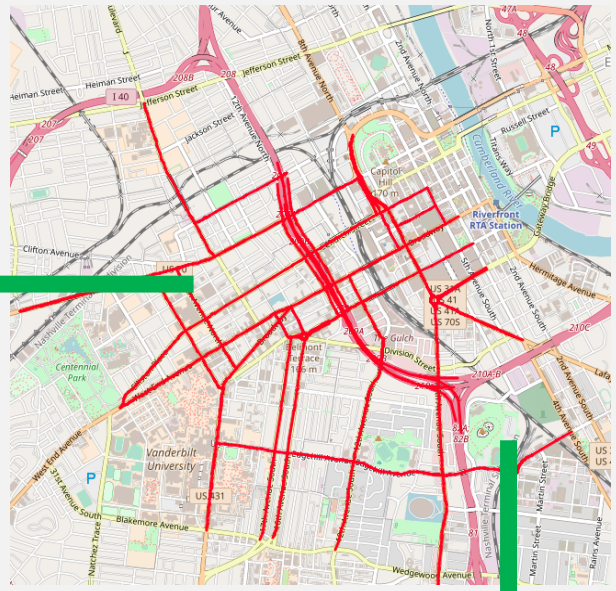
Students working on the STEAM project will visit the partner institutions for one week (during summer) for immersive collaboration experience.

- 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

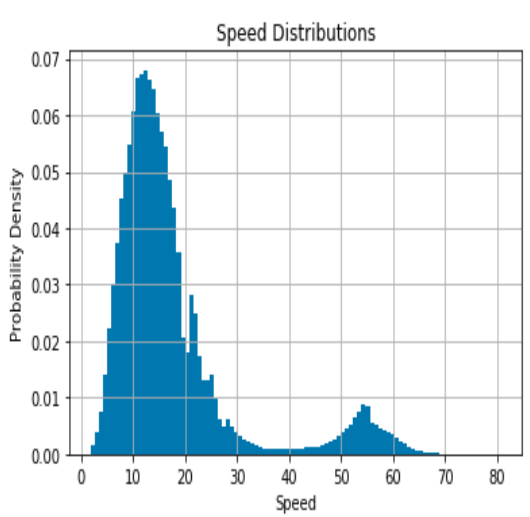
Current Progress



3 years of multivariate traffic data @1m from Nashville

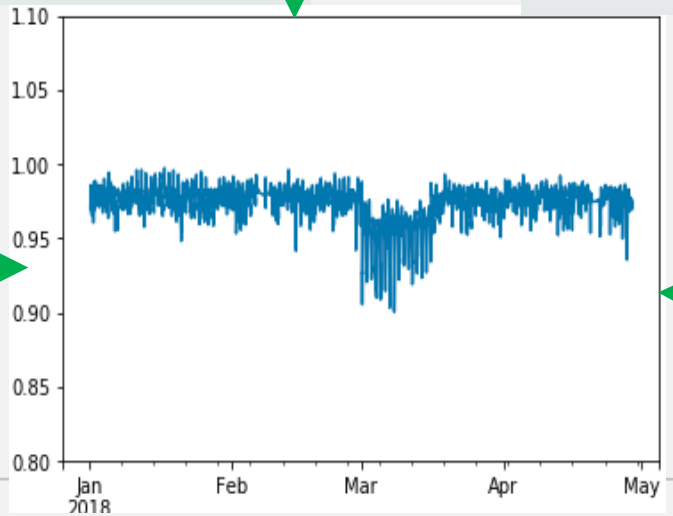


Focusing on Zones: Downtown Nashville > 100 sensor



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)