SK telecom

IoT 센서 데이터 자동 분석 소프트웨어 아키텍처

스마트 빌딩 적용 사례

Deokwoo Jung, PhD.

SK 텔레콤 종합기술원

Partner for New Possibilities

목차

Data-Driven Energy Analytics (DDEA)

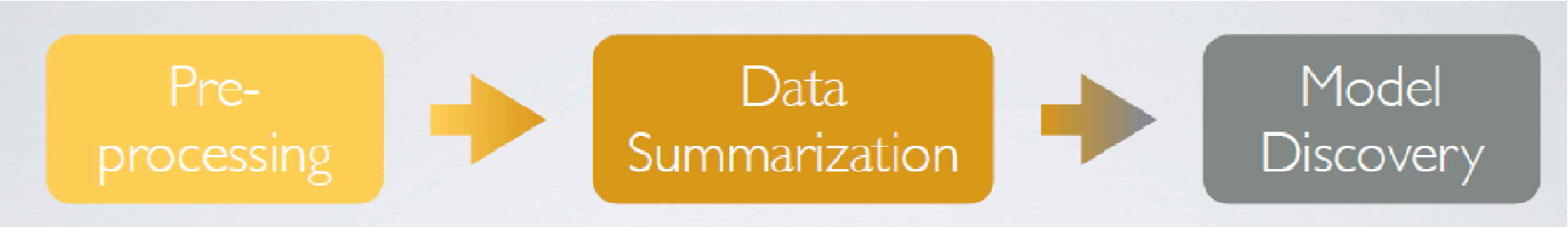
DATA DRIVEN ENERGY ANALYTICS

Industry-Scale Demand-Response Resource Finder

WHAT DDEA IS AIMING

- Demand-Response Resource Finder
- Unsupervised-Learning with Sensor Data
- Industrial-Scale DR Service Provider (AaaS)

HOW DDEA WORKS



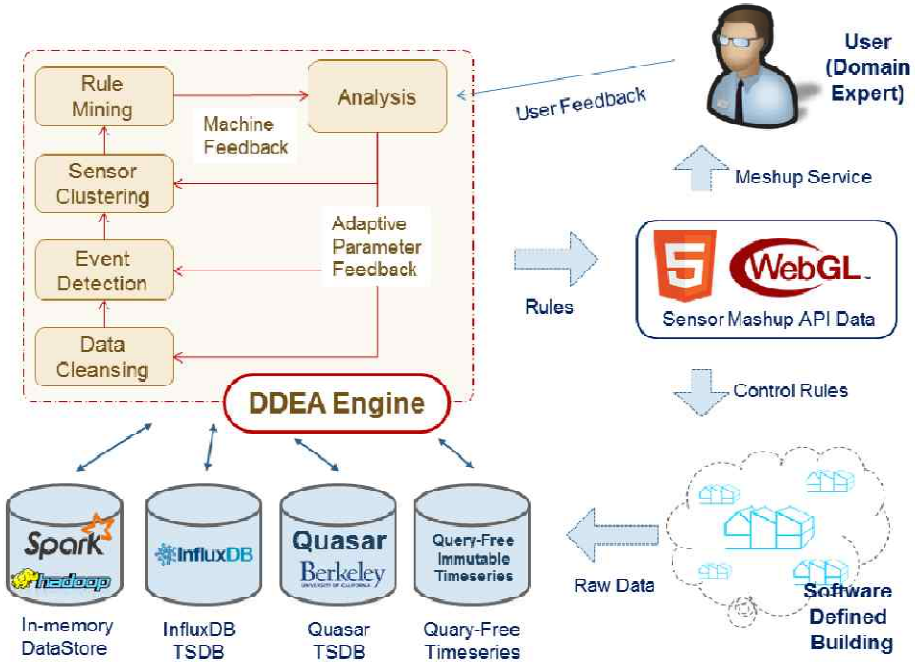
- Pre-processing Sensor & Environmental Data
- Data Summarization (Feature Extraction & Clustering)
- Model Discovery by Learning Algorithm

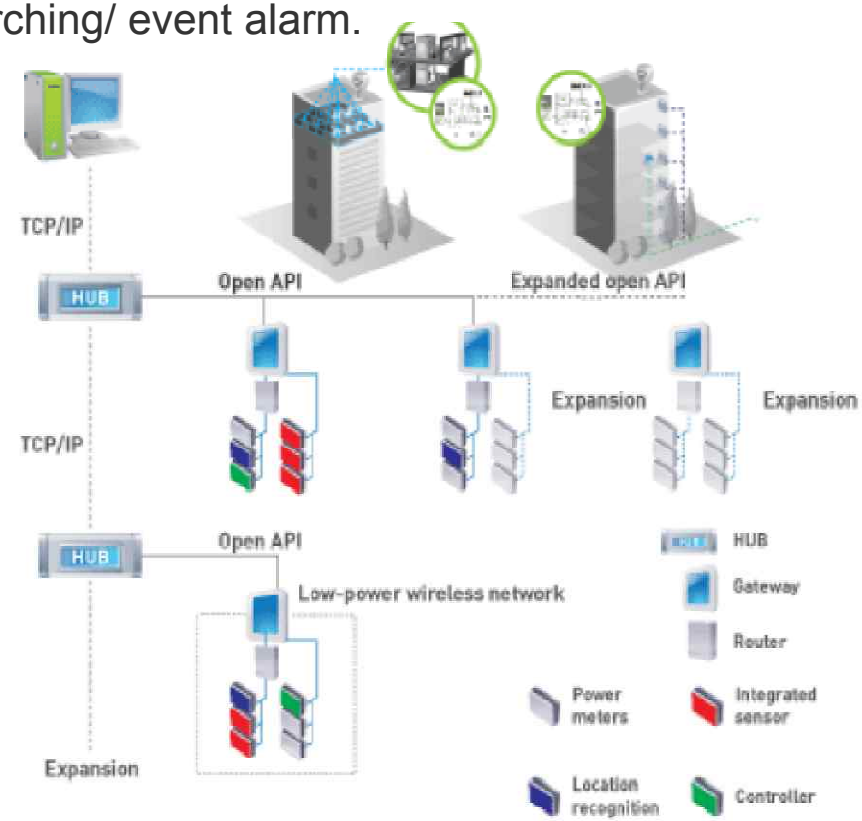
The Data-Driven Energy Analytics (DDEA)

Scalable Energy Use Analysis Platform for Smart building

Project Goals

- **Make DDEA scalable to >100x buildings :**
  - data query/ loading / processing time optimization.
- **Make DDEA integratable to Internet of Things (IoT).**
  - inter-operating with WSN, meta-data transformation,
- **Increase DDEA usability for Layman**
  - IFTTT style rule setting / checking/ searching/ event alarm.
  - Intuituve Data Analysis Visualization.





The Data-Driven Energy Analytics (DDEA)

Scalable Energy Use Analysis Platform for Smart building

●Data Management & Processing Methodology

- ○ An approach to process BEMS Bigdata utilizing Open Source Python Scikit-Learn (http://scikit-learn.org/) and R bnlearn (http://www.bnlearn.com/).
- ○ A strategy to locally save, to simplify, and to process data in parallel.
- ○ A data model applicable to any type of BEMS and climate data.

● Data Modeling Methodology

- ○ A mechanism for Automatic BEMS data pattern recognition.
- ○ A topology for a data model without Naming Scheme information.
- ○ A correlation model for BEMS data pattern, climate data, and time.

●Analysis Methodology based on BEMS Data Model

- ○ Potential causation models among BEMS pattern, climate data, and time.
- ○ Algorithms to automatically compute the weight of data points utilizing statistical causation model.
- ○ Automatic computation and visualization methods for causation graph using Bayesian Network based on the weight of data points.

●Adaptive Operatiblity on Various Environment

- ○ Ability to load sensor data from various data infrastructures.
- ○ Operability on wide variety of hardwares and software environements.
- ○ Web-based Frontend to support connectivity from mobile to desktop

DDEA Architecture

DDEA architecture : Current implementation

Data Preprocessing

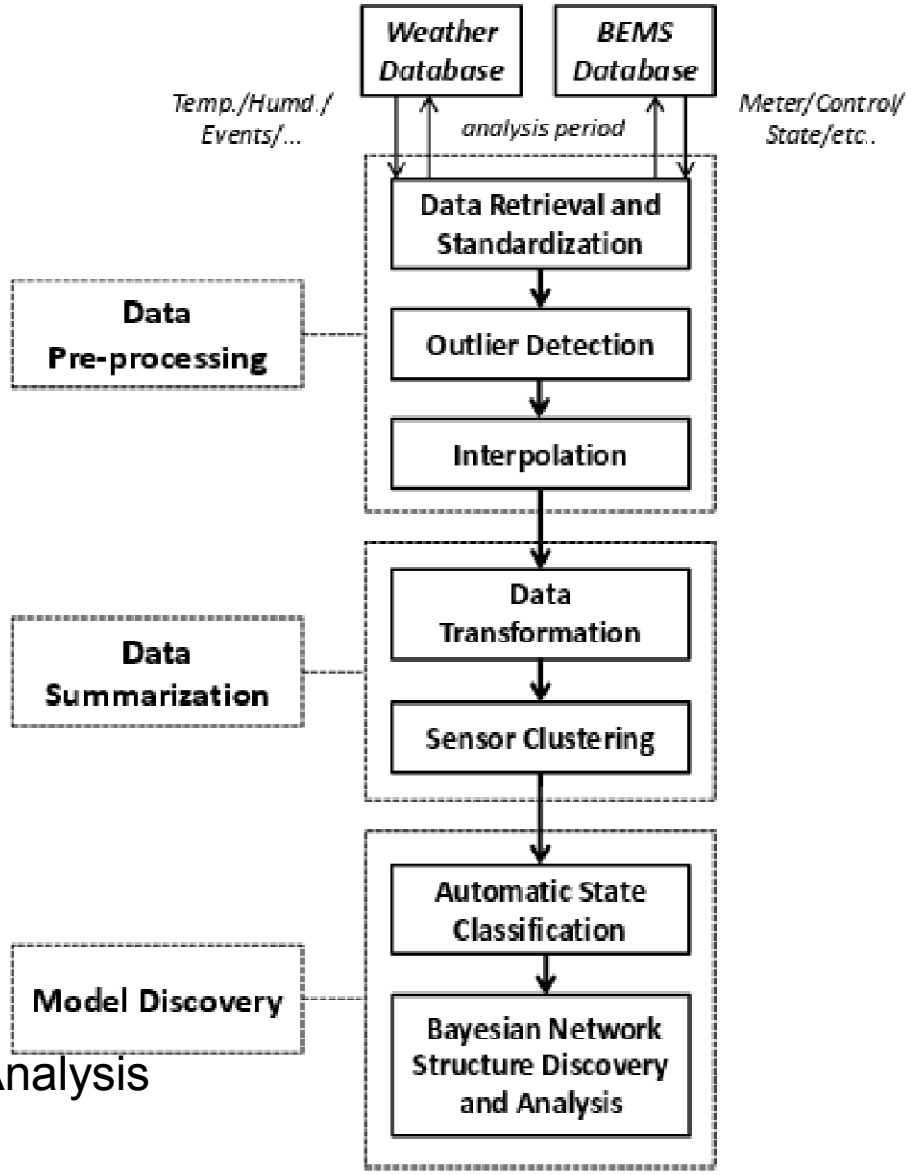
- Data Retrieval and Standardization
  - User-specified analysis period
  - Weather and BEMS database
  - Binary files in hard-disk
  - Single global time reference
- Interpolation / Outlier detection

Data Summarization

- Data Transformation
  - Regular/ Irregular Event
- Sensor Clustering

Model (or Rule) Discovery

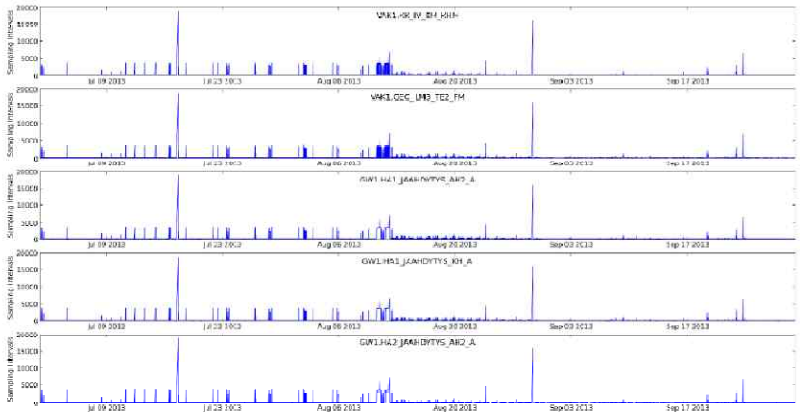
- Automatic State Classification
- Bayesian Network Discovery and Analysis



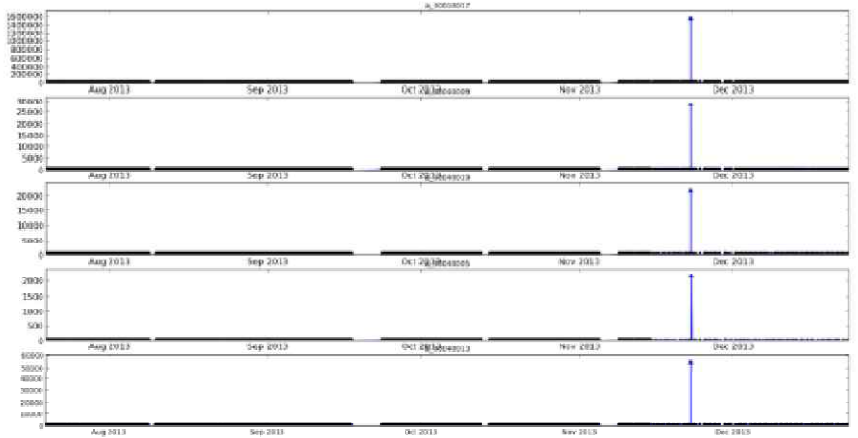
Data Preprocessing

- Data Retrieval and Standardization - Interpolation / Outlier detection

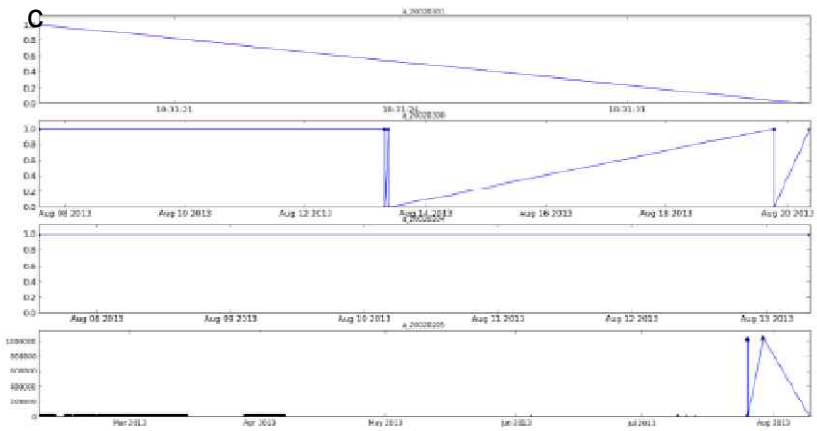
VTT Site: Large variations and correlated disruptions in sampling intervals among 5 sensors are observed



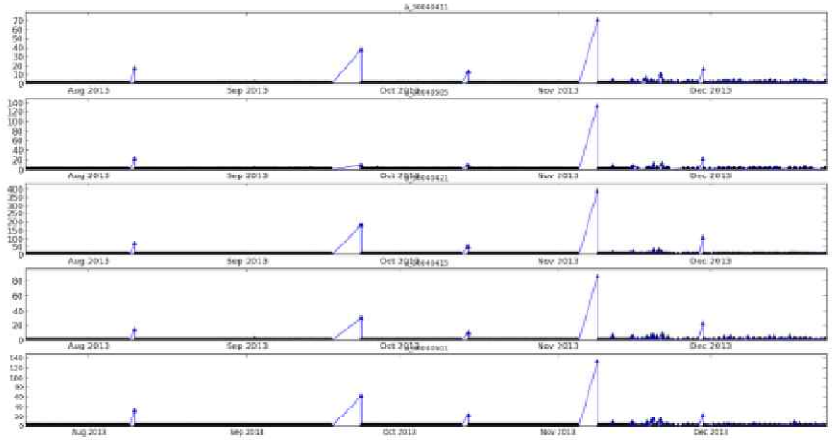
GSBC Site : Bad Outliers Power meter BMS data of ['30010017', '30040005', '30040009', '30040013', '30040019']



GSBC Site : Bad data acquisitions : E.g., Machine Room BMS Measurements of ['20020204', '20020205', '20020300', '20020301']



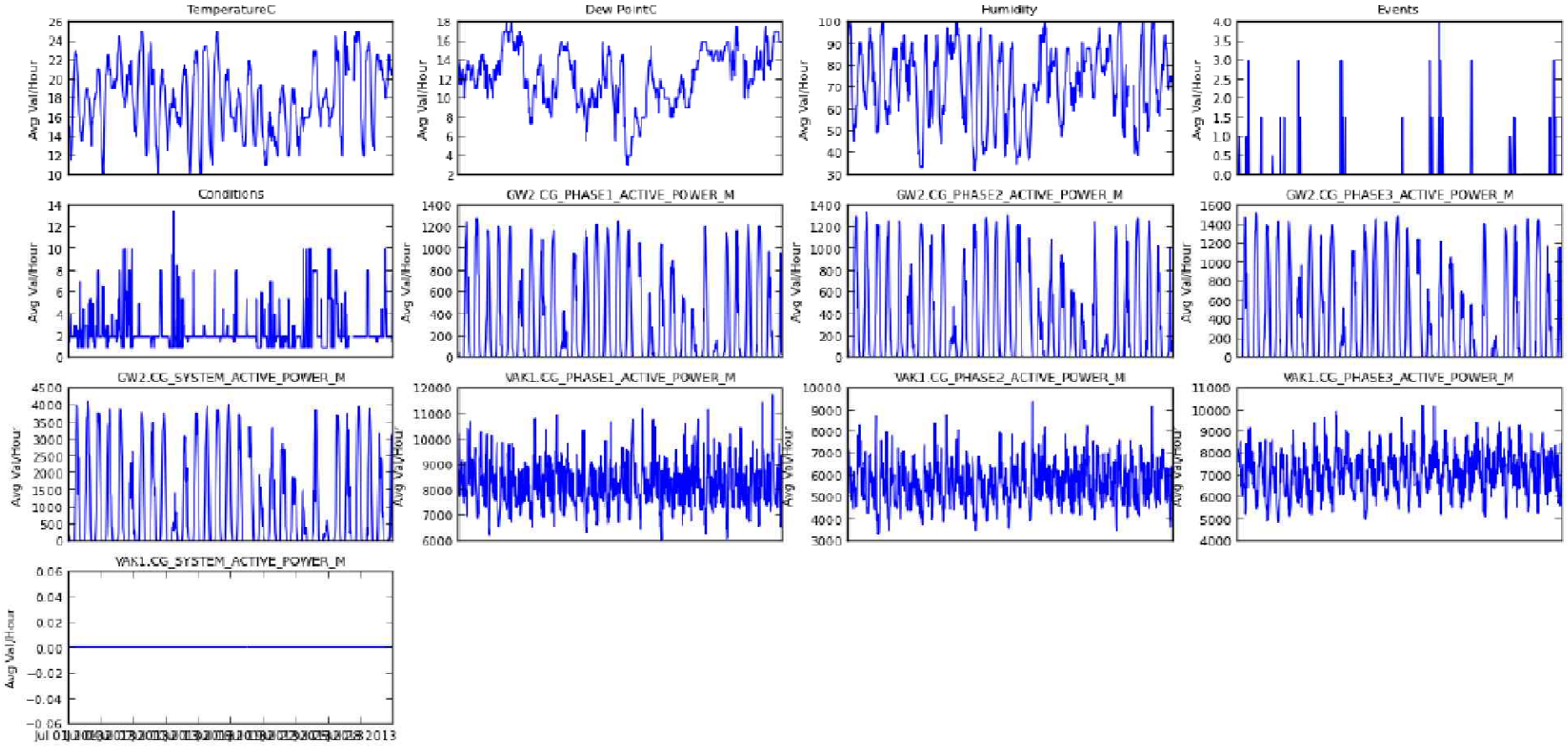
GSBC Site : Ambiguous Outliers Power meter BMS data of ['30040411', '30040415', '30040421', '30040501', '30040505']



Data Preprocessing

- Data Retrieval and Standardization - Interpolation / Outlier detection

13 sample data points retrieved from weather and BEMS database during 8 - 15 July, 2013



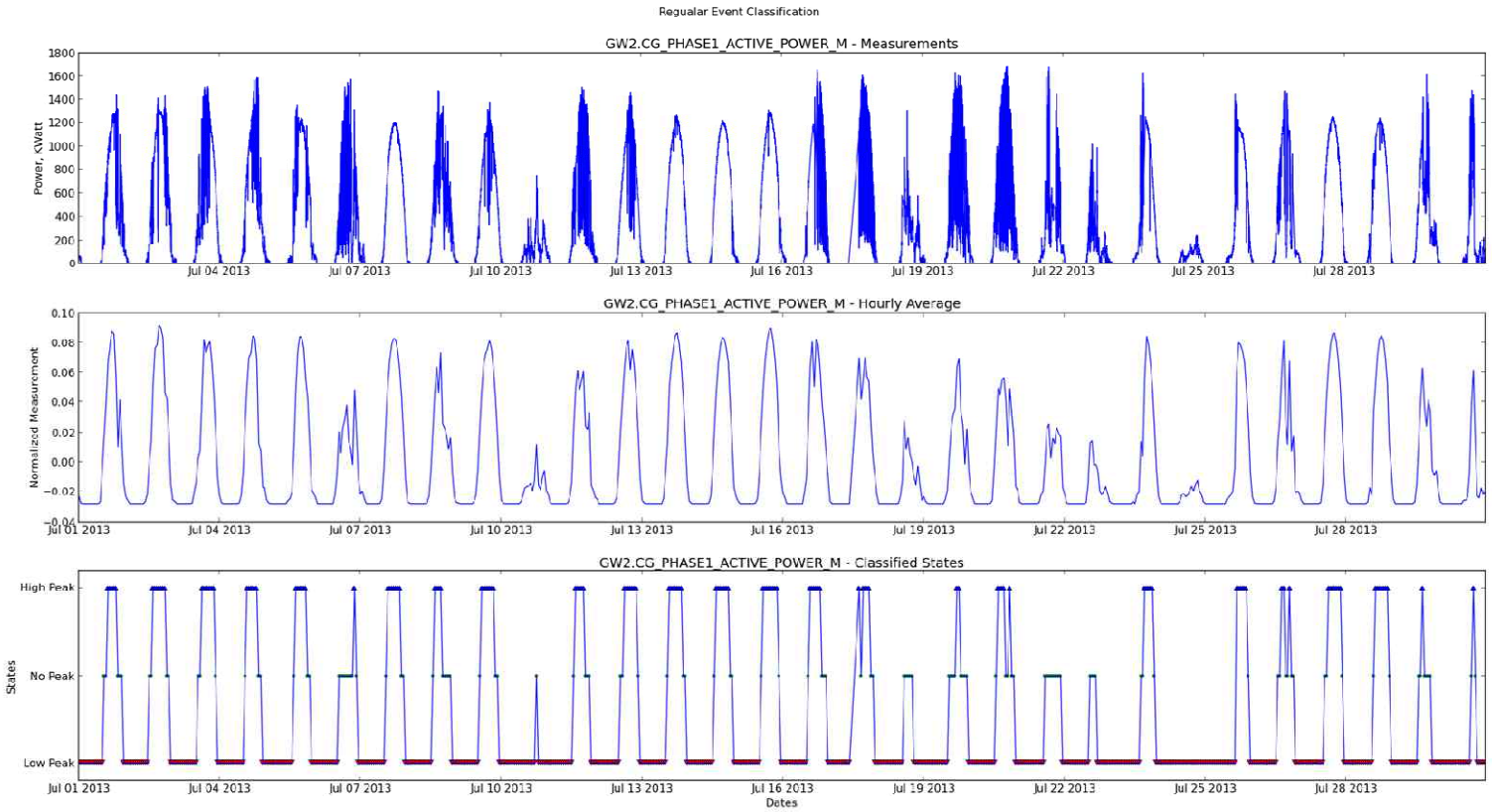
Data Summarization

- Data Transformation - Sensor measurement event classification

Proposed Signal Feature for Event Analysis	Regular Event (Average Value)	Irregular Event (Differential Value)
Data Characteristics	Periodically occurring measurement e.g., high and low peak in daily energy consumption	Non-Periodically occurring measurement . e.g., sudden drops in energy during daytime.
Time window	Set by users (default 15 min)	Set by users (default 15 min)
Data Feature	Extraction Average: Average value	Differential: Burstiness of Measurement Fluctuation
State Classification	Ternary States Low Peak, High Peak, and Non-Peak in hourly average of measurement	Binary States Regular or Irregular fluctuation in measurement differential.
Classification Method	K-Mean	K-Mean / GMM

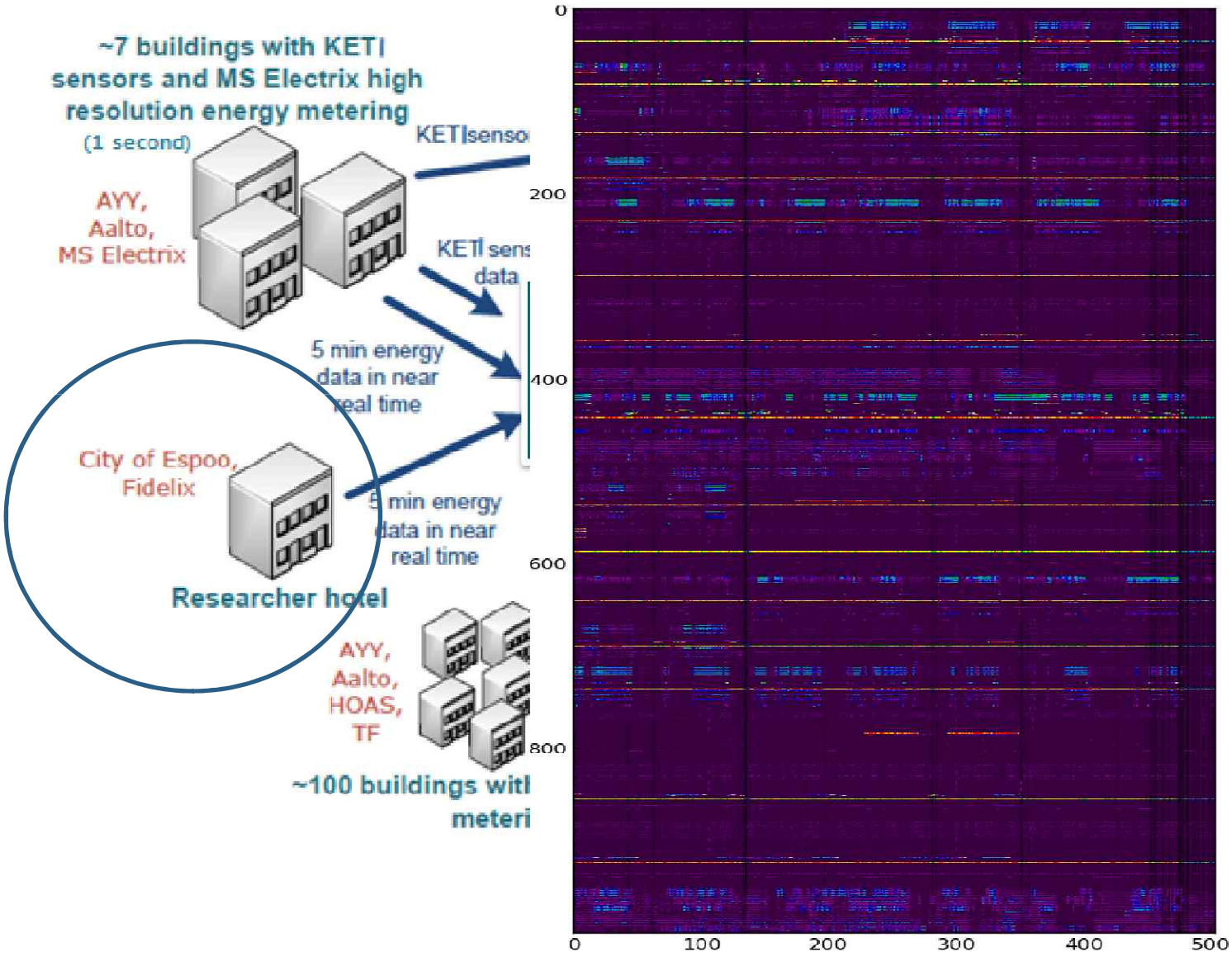
Data Summarization

- Data Transformation - Regular event classification
  - Automatically can transform all measurement data into hourly average and classify the hourly average into ternary states, ‘Low-Peak, High-Peak, and No-Peak.
  - The regular event classification greatly simplifies temporal characteristics of measurement points that allows us to build tractable Bayesian Network for regular events.



Finland VTT 빌딩 적용 사례

Sampling density of VTT testbed in one week (1/11/2013 - 8/11/2013)
analysis interval, t = 20 mins , x-axis: time index y-axis: sensor/data stream index
the brighter the color, the higher density



Analysis - VTT - Summary : Model Discovery

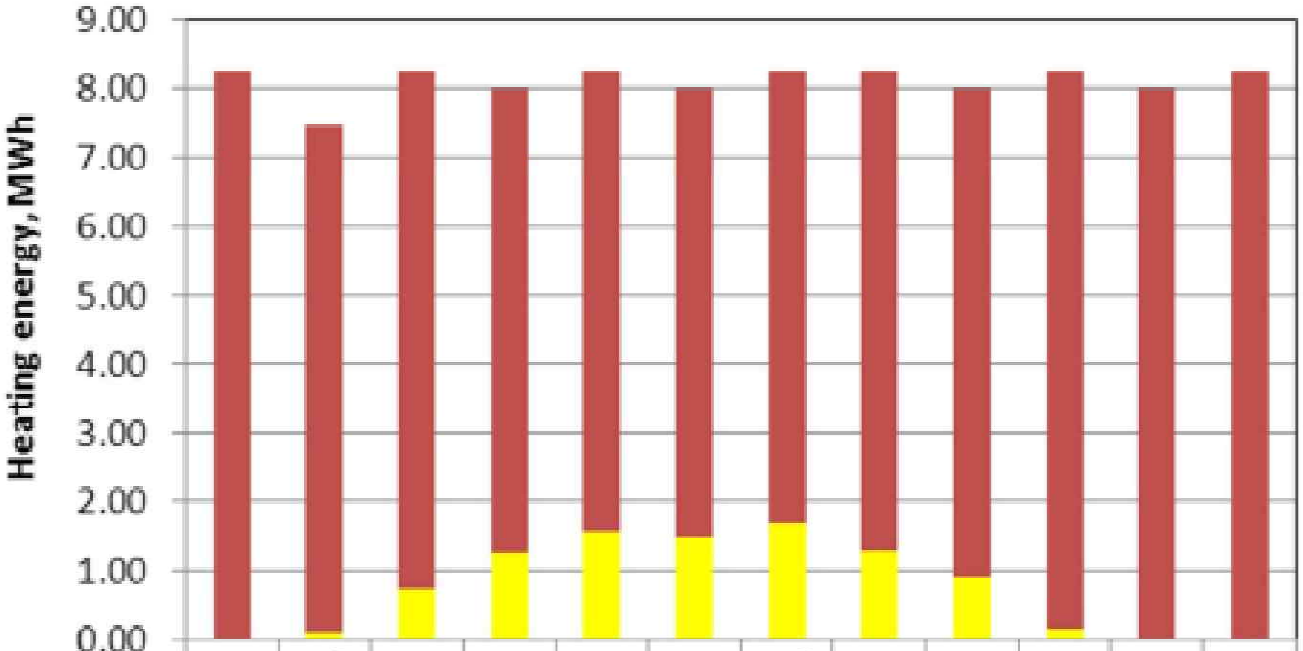
(Aalysis of GW2.CG_SYSTEM_ACTIVE_POWER_M)

- **Solar systems**

✓ Solar heat flat-plate collectors (collector area 16 m2 / estimated yearly heat production 575 kWh/ m2)

✓ PV-panels (panel area 35 m2 / peak power 5,2 kWp)

Figure 33. Module of 4 flat plate collectors of the solar heating system.

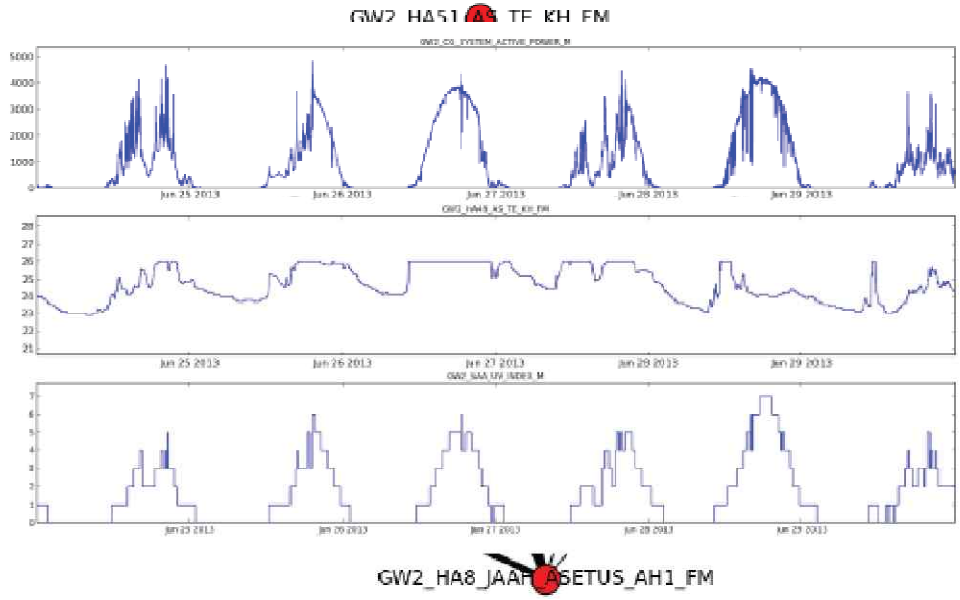


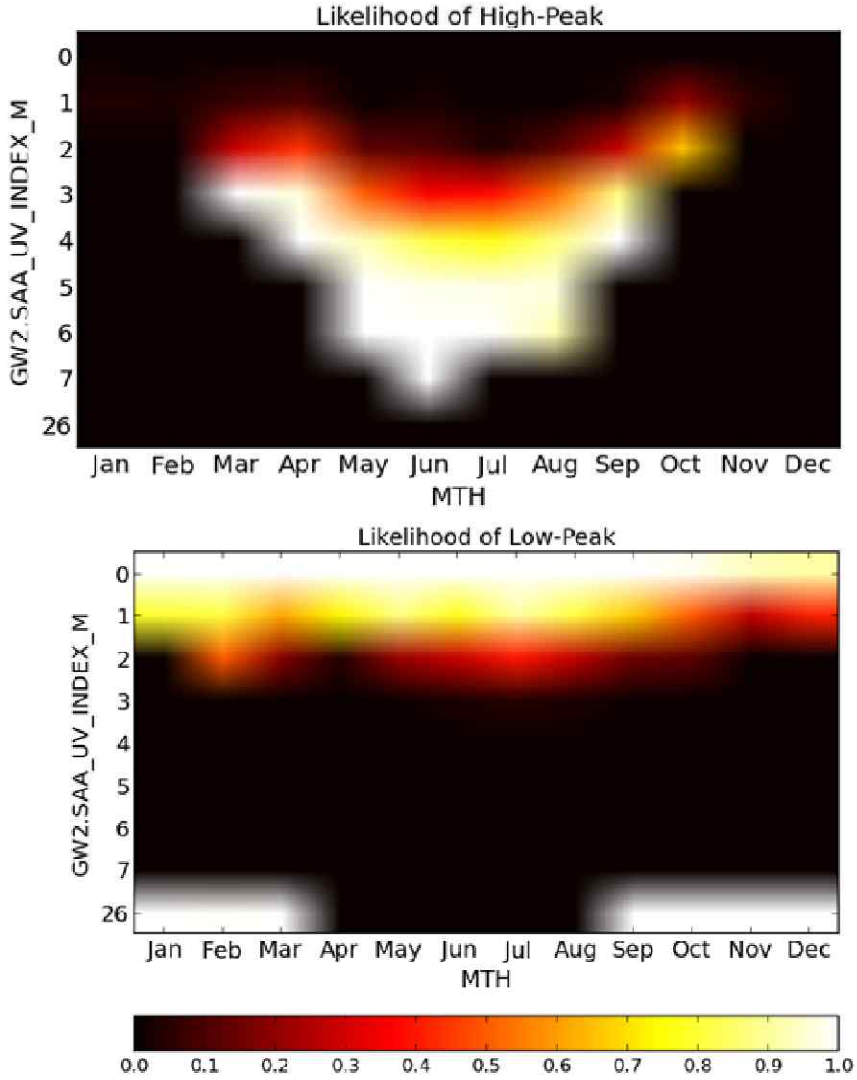
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
DHW demand	8.25	7.34	7.51	6.71	6.68	6.49	6.55	6.97	7.08	8.09	7.98	8.25
Solar heat	0.00	0.11	0.73	1.27	1.56	1.49	1.69	1.28	0.90	0.15	0.00	0.00

Analysis - VTT - Summary : Model Discovery

(Aalysis of GW2.CG_SYSTEM_ACTIVE_POWER_M)

- Sensor-Time-Weather Causality Analysis
  - Bayesian network structure that reveals the relationship between time, weather, and sensor measurements





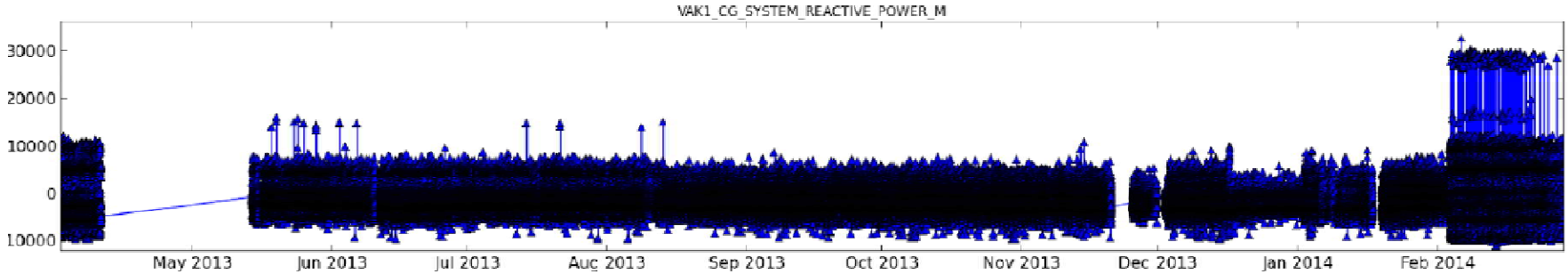
- Peak Demand : 6,7,8 월에 UV Measurement 영향
- UV 값이 높을수록 (26 제외) Peak Demand의 확률이 높음

Analysis - VTT - Summary : Model Discovery

VAK1_CG_SYSTEM_REACTIVE_POWER_M- Irregular (Differential)

- **Heating networks (GSHP 128/AHU+floor 200/320 kW)**

✓ AHU (Ground source heat pump, district heat, solar heating)
✓ Radiators (Ground source heat pump, district heat, solar heating)
✓ DHW (Ground source heat pump, district heat, solar heating, grey water heat recov)



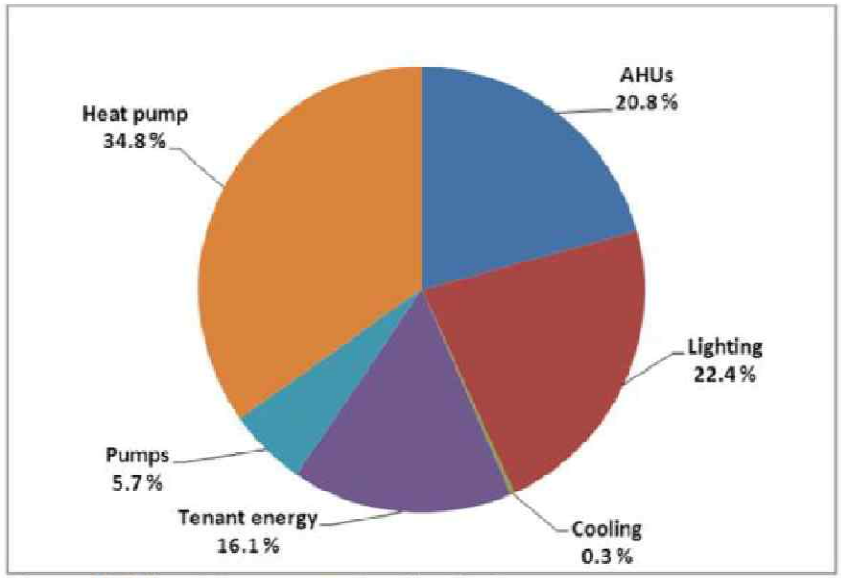


Figure 11. Electricity consumption distribution.

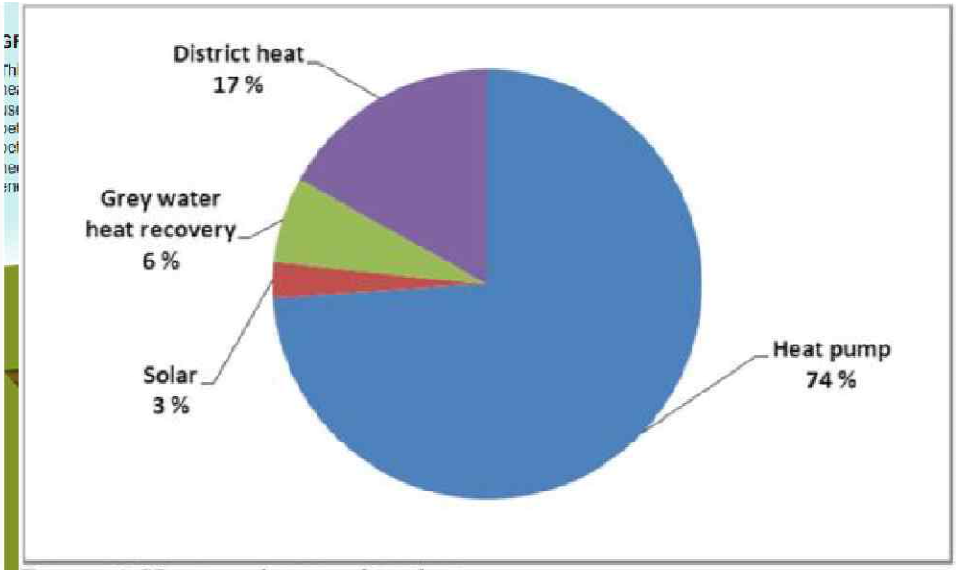
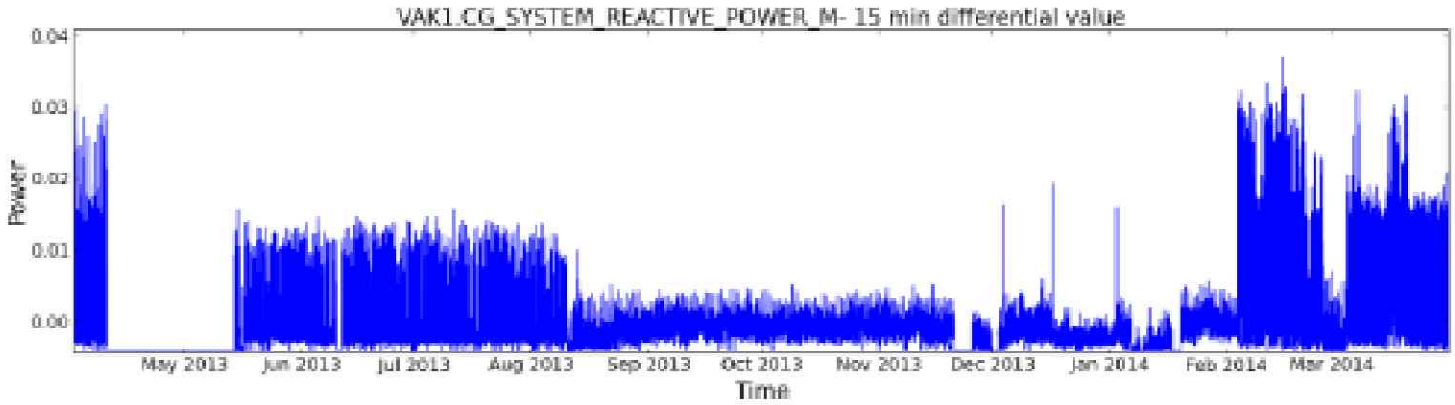


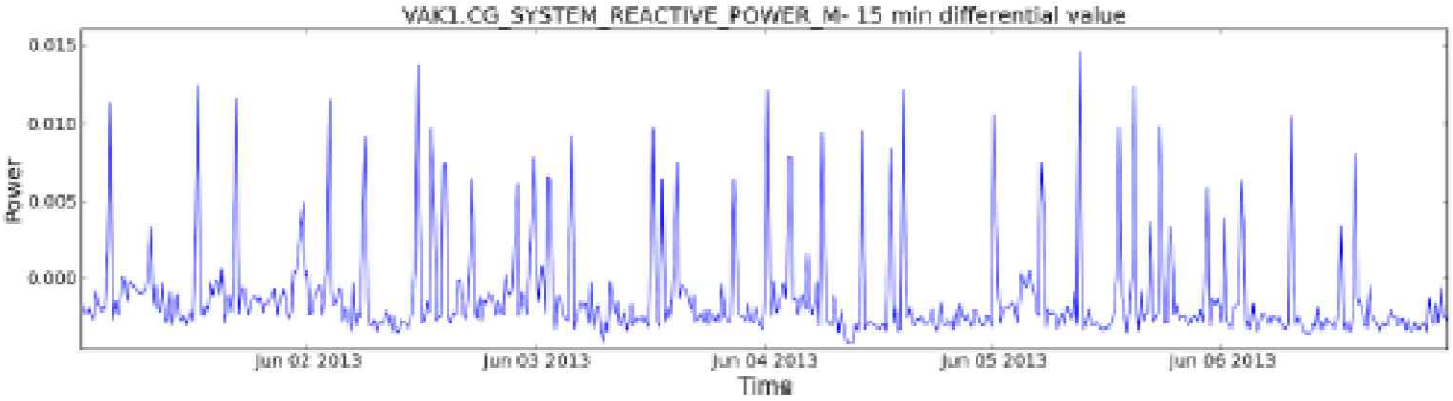
Figure 6. Heat production distribution.

Analysis - VTT - Summary : Model Discovery

VAK1_CG_SYSTEM_REACTIVE_POWER_M- Irregular (Differential)



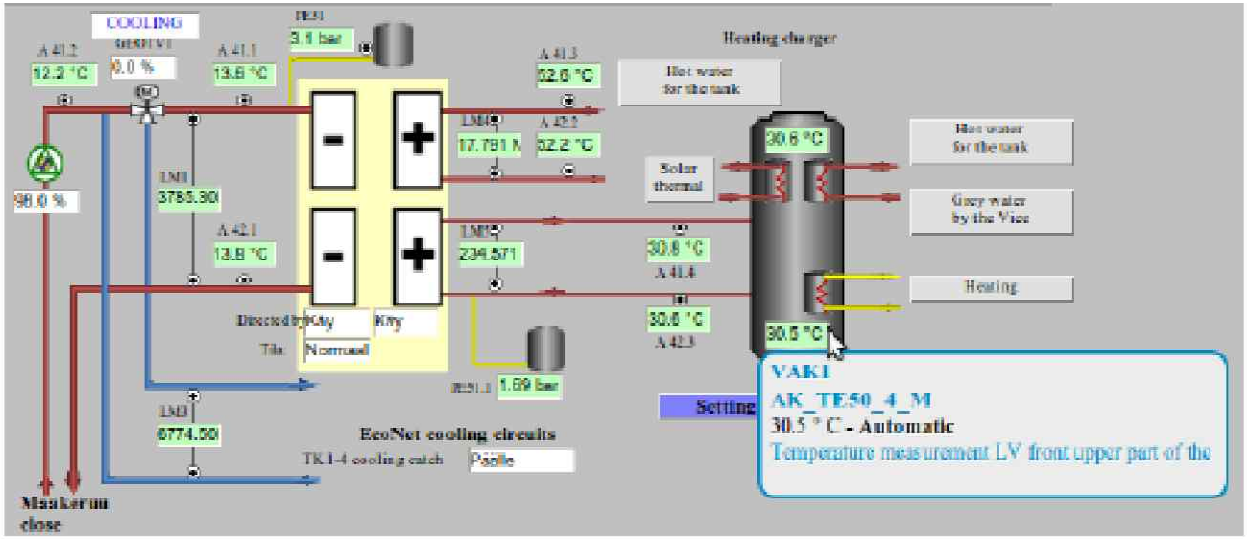
(a) One-year profile (Apr 2013 - Apr 2014)

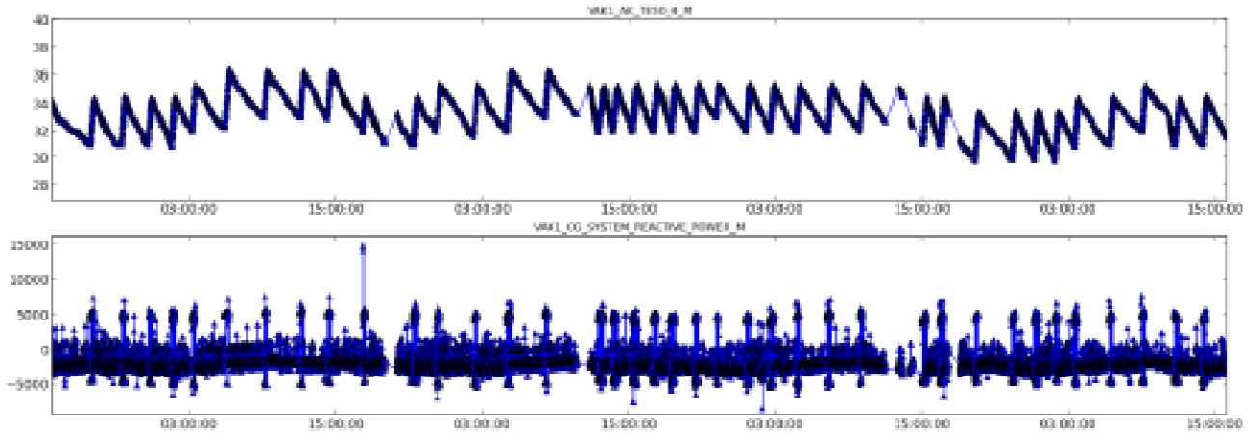


(b) One-week profile (June 1-7, 2013)

Analysis - VTT - Summary : Model Discovery

VAK1_CG_SYSTEM_REACTIVE_POWER_M- Irregular (Differential)





(b) Differential Measurements during Aug 20-24, 2013

Distributed DDEA

Work in Progress

CURRENT DDEA STACK

Display / UI Layer

Processing Layer

Scikit-Learn

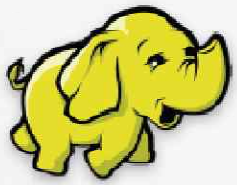
Numpy

R

Data Storage Layer

InfluxDB

Quasar



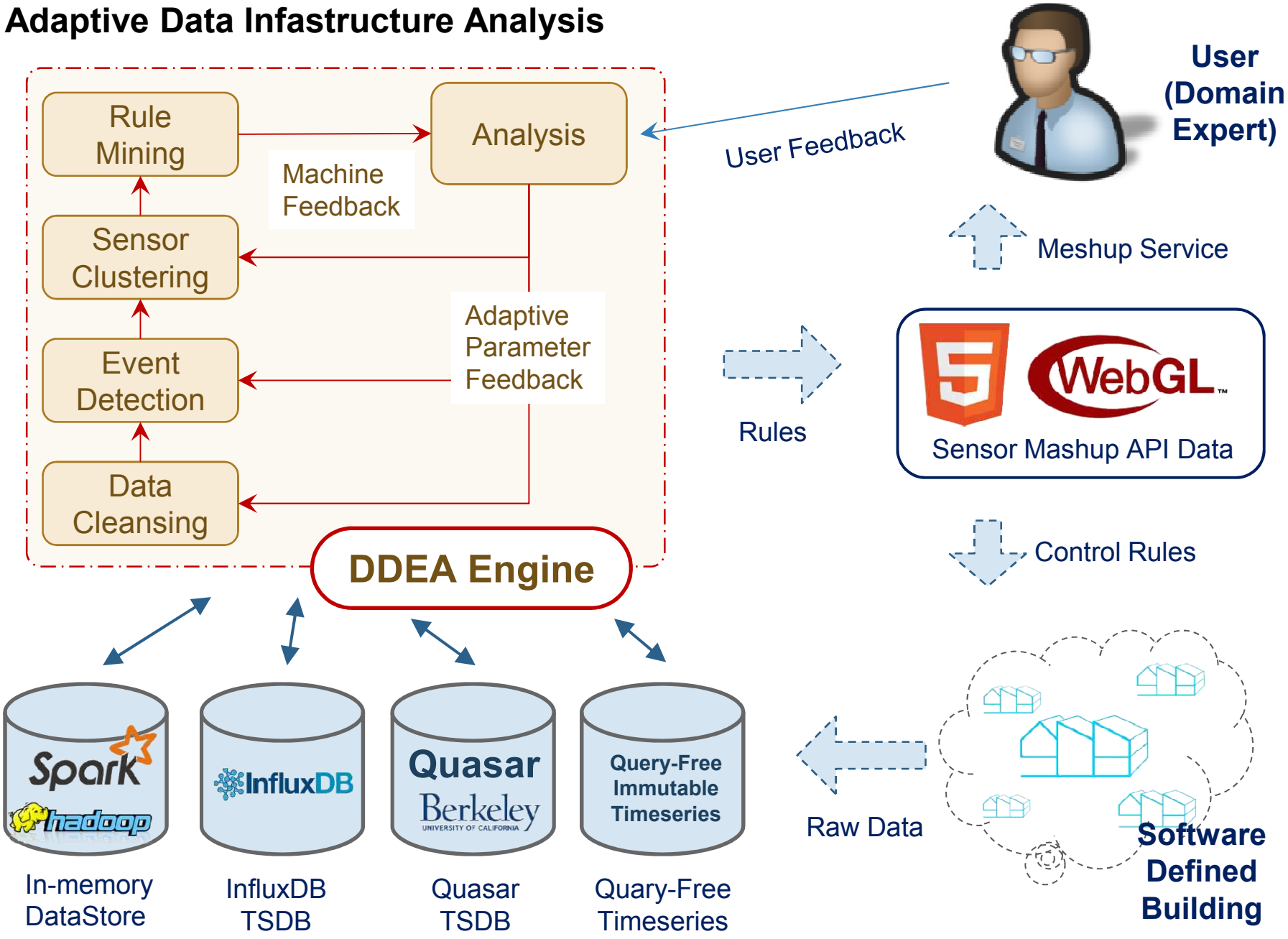
Hadoop

Query-less bin file

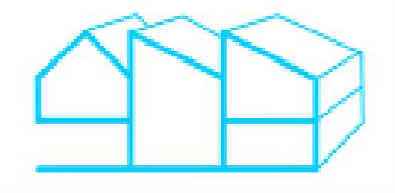
- Compact & Compatible
- Support Various Storage Types
- Single-Host Multicore Awareness

DDEA architecture

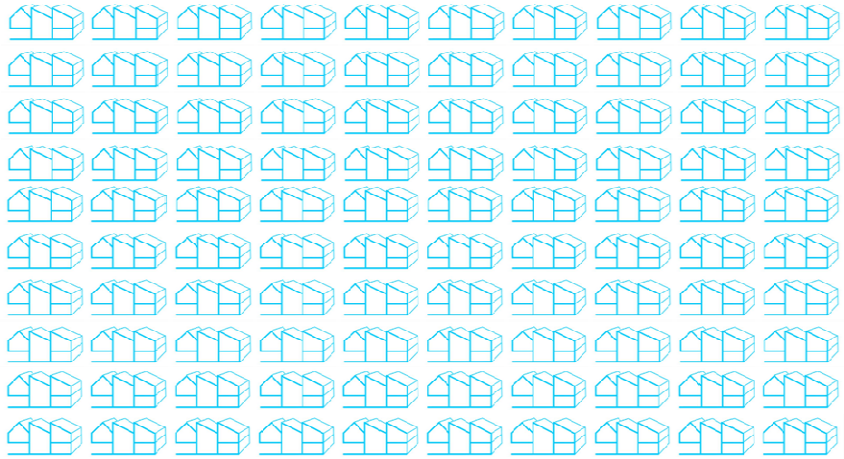
Adaptive Data Infastructure Analysis



Analyzing sMAP on Spark : Work-in-progress



Dent Meter Reading : ~50GB
BACnet Reading : ~200GB



x100 sMAP : 5~20TB

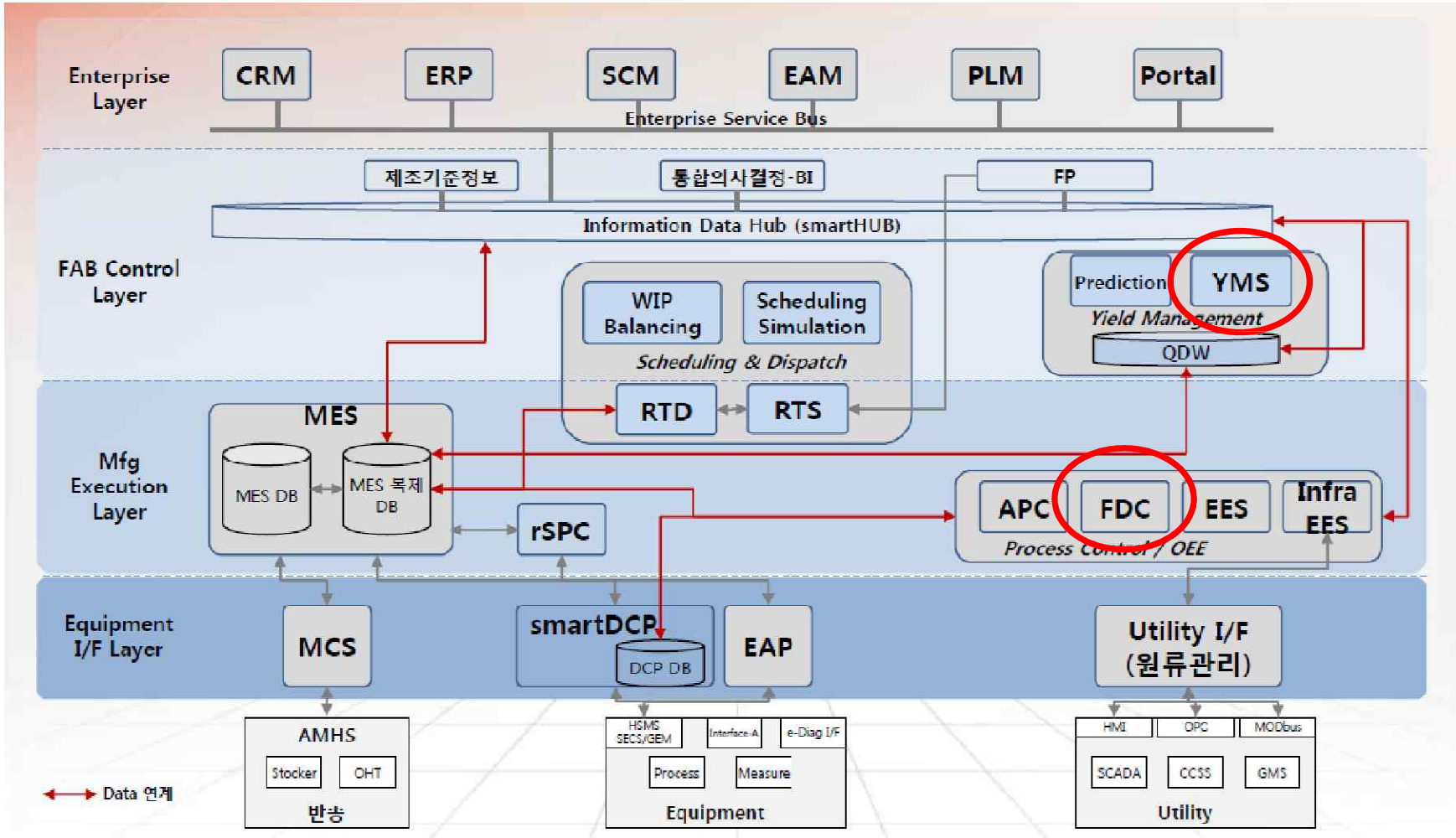
DISTRIBUTED DDEA

- Industry-Scale Time-Series Data Collection & Retrieval
- Batch & Realtime Data Processing Capacity
- Horizontal Scalability over Cluster / Cloud
- Centralized Data Collection & Processing

DDEA 의 타 vertical IoT 적용
- Industrial IoT -

반도체 공정 (FAB) 자동화 시스템 Overview

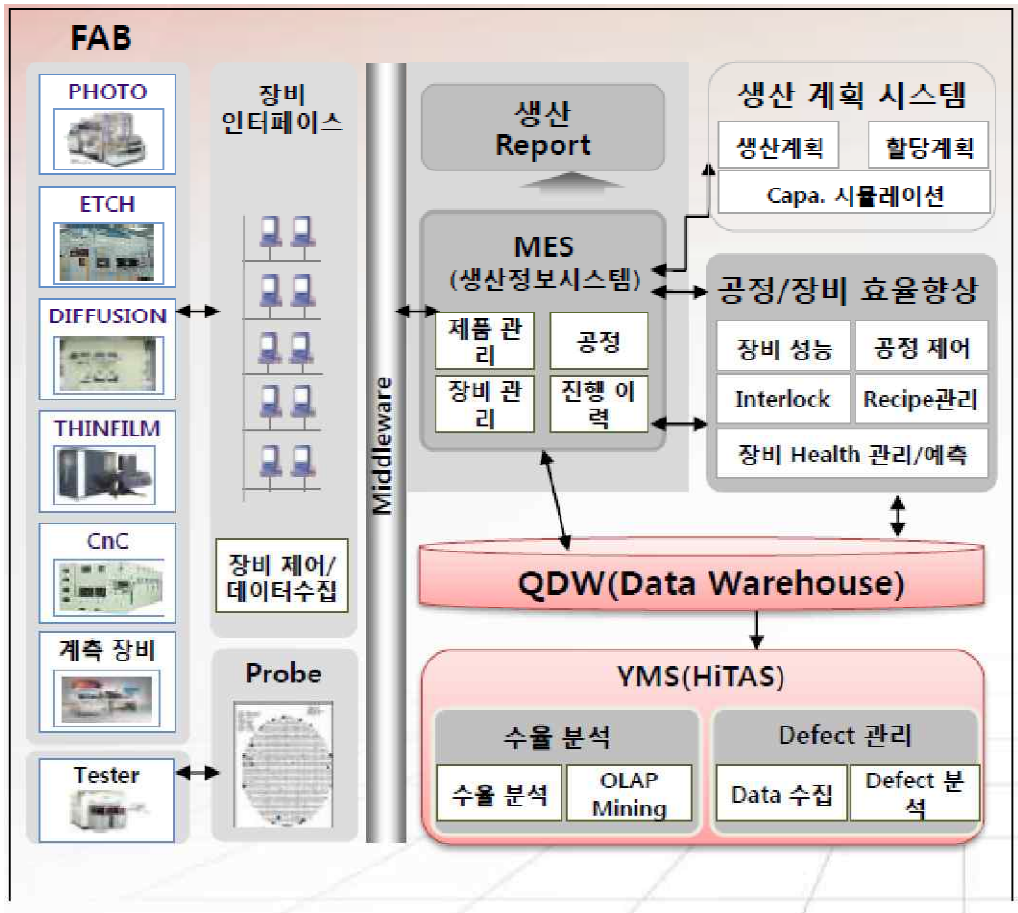
Discovering Hidden Factors is a key for success , yet challenging to automate it !!



Sources: Manufacturing Technology Center System Engineering PJT. [Kim.SH], SK Hynix, 2015 .

YMS (Yield Management System)

반도체 제조 공정에서 발생하는 모든 Data를 수집하여 Data Warehouse를 구축하고 공정 이상 발생 모니터링 / 원인 분석 및 예방을 위한 분석 기능 제공



- **전 공정 Data Trend Analysis**
  - 공정 Data의 이상 발생, 중심/산포의 변화 감지

- **Data Visualization**
  - 공정 Data를 Sheet, Chart, Map, Image의 형태로 분석 Eng'r에게 제공

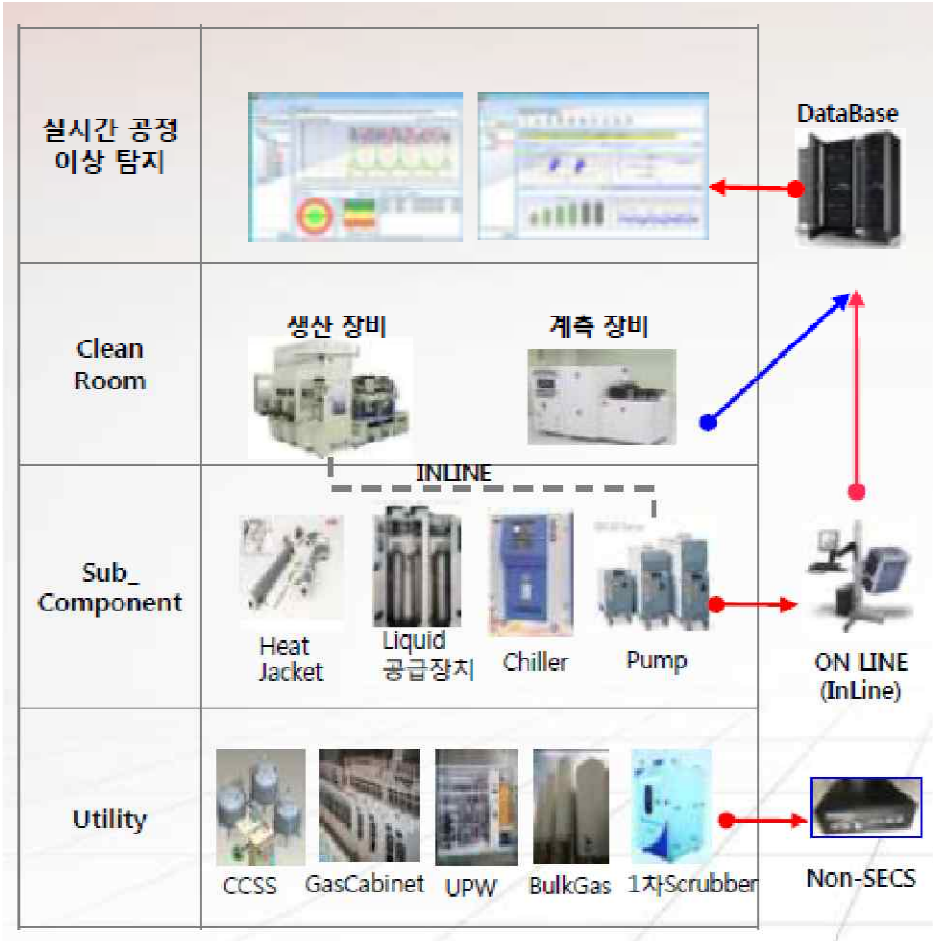
- **Data 통계 분석**
  - SPC : 통계적 공정 관리
  - Data Mining : 공정 이상 원인 탐색 및 검증
  - Data Pattern 탐색 : 이상 Data Pattern 자동 분류
  - OLAP : Online Analytical Processing Data Pivot, Data Drilling 분석

- **분석 Process 표준화**
  - Workflow 를 통한 사용자 분석 Process 표준화

Sources: Manufacturing Technology Center System Engineering PJT. [Kim.SH], SK Hynix, 2015 .

FDC (Fault Detection &Classification)

- 반도체 공정 Process 진행 중 설비에서 발생하는 미세 데이터의 수집/ 분류 / 분석 을 통한 설비 이상 유무 모니터링 및 진단

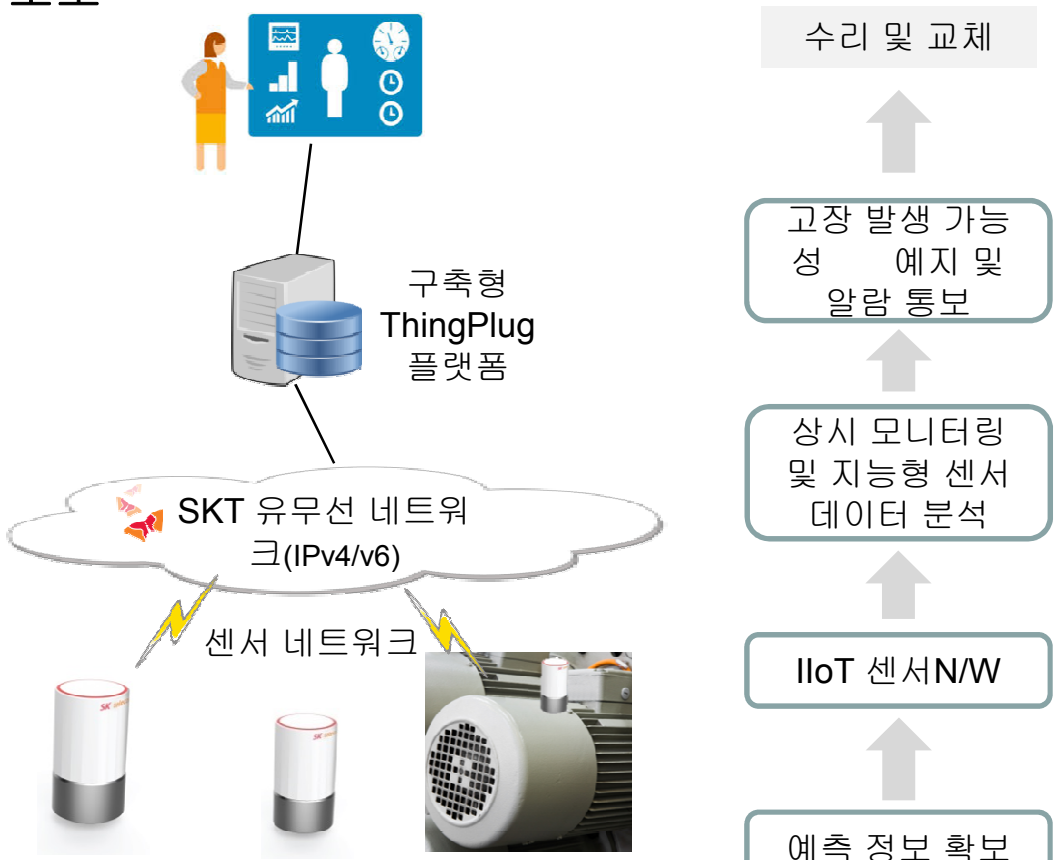


주요 기능	설명
실시간 장비 Data 수집	• Lot/Wafer별 Parameter Data의 수집 및 보관 • 동일 공정에 대한 장비간/ 장비 내 Chamber간 차이 분석
Data 통계 분석	• 다중 Parameter간 상호 연관성 비교 분석 • Source Parameter와 Response Parameter간 상관관계 분 석
자동 Fault Detection & Response	• 관리 한계치 설정에 따른 자동 Fault Detection 및 Notification • 자동화 시스템과 연계한 Process Interlock 구현
실시간 Pattern 분석	• Ideal Pattern 설정을 통한 실시간 장비 Health Check
Report 기능	• 장비의 상세 Process 이력 및 특성에 대한 관리 • 장비 별, Lot/Wafer별 Summary Trend 관리 • 장비 및 공정 관리를 위한 다양한 Report 제공
Classif-ication	• 축적된 Knowledge Base를 이용하여 Troubleshooting 지침으로 활용

Sources: Manufacturing Technology Center System Engineering PJT. [Kim.SH], SK Hynix, 2015 .

IIoT (Industrial IoT) /PdM(Predictive Maintenance) 솔루션

IIoT 센서 N/W 기반 PdM 솔루션 구조도



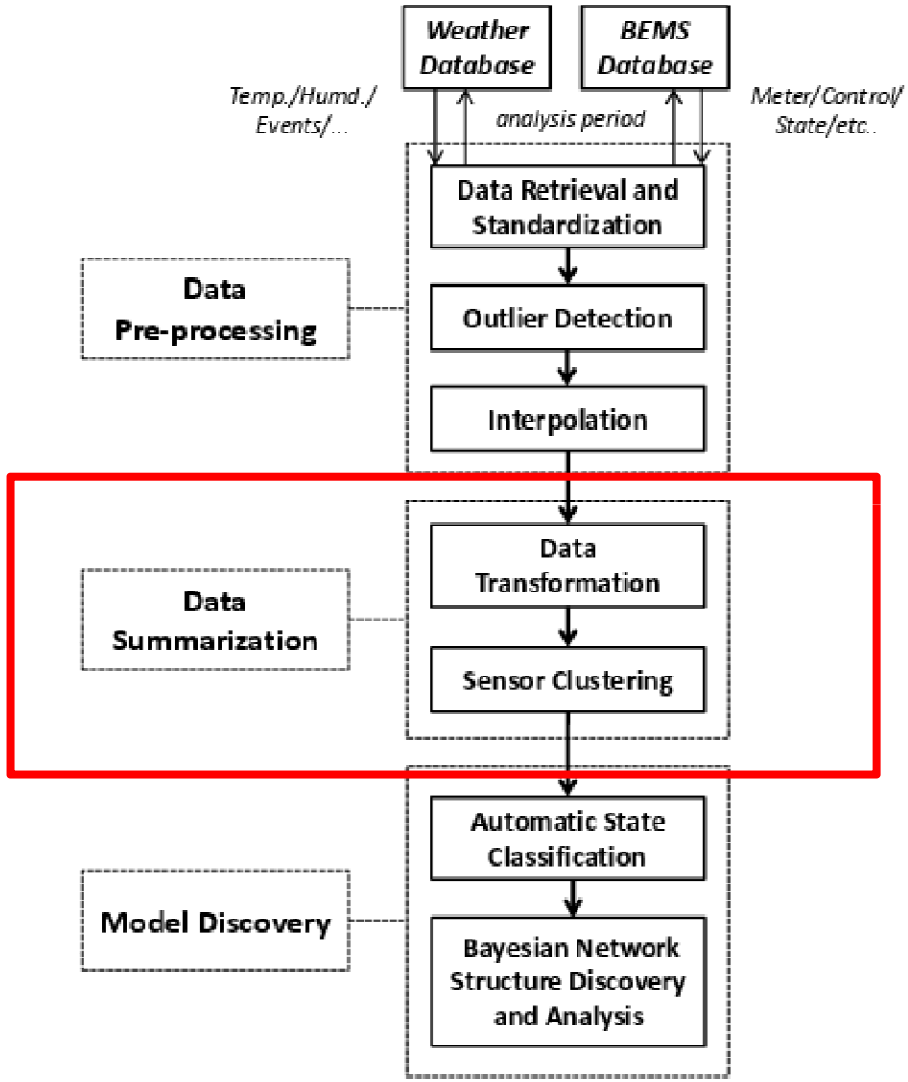
Stick & Play IIoT 센서 (고정밀 진동, 계측 등)

주요 적용 기술

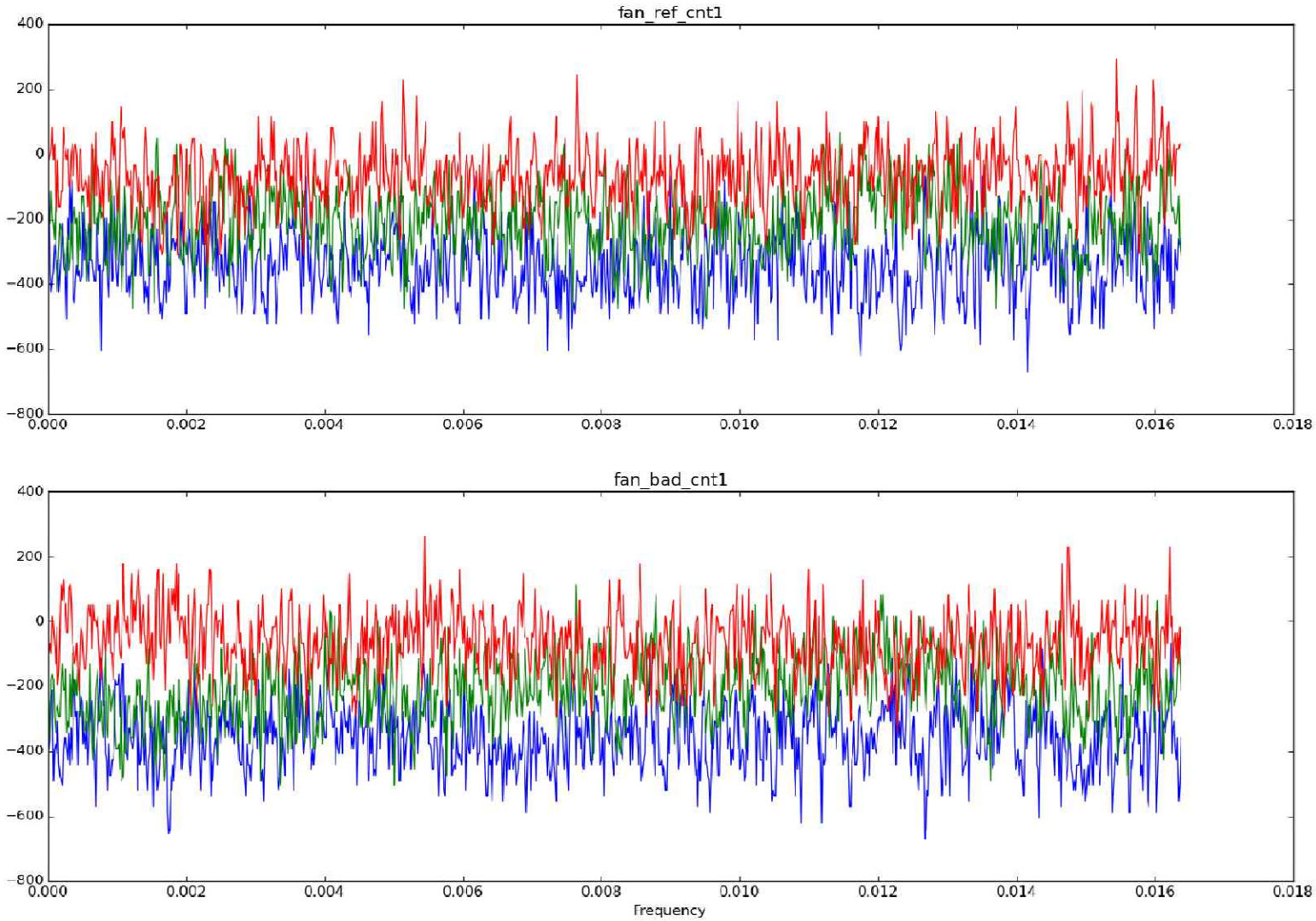
- 고정밀 진동 및 전류 계측 기술
- 고정밀 센서 데이터 전송 네트워크
- 빅 데이터 클라우드 운영 기술 (ThingPlug 구축형)
- 빅 데이터 분석 기술
  - 기계 학습 기반 고장 진단 알고리즘
  - 분석 데이터 시각화

Same Framework

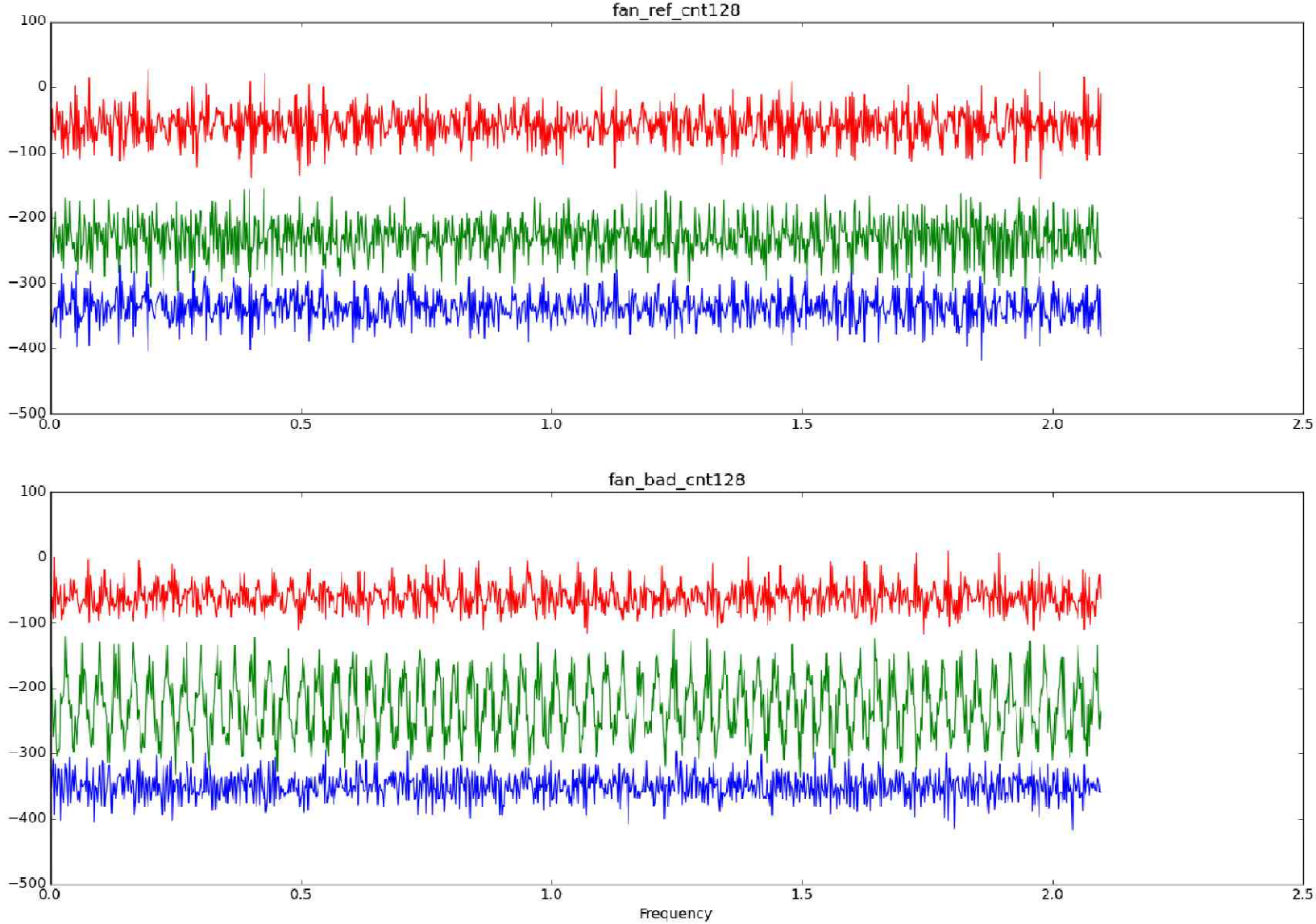
- Find Optimal Transform
- Exploit Data Sparsity
- Simplify your model.
- Robust Analysis



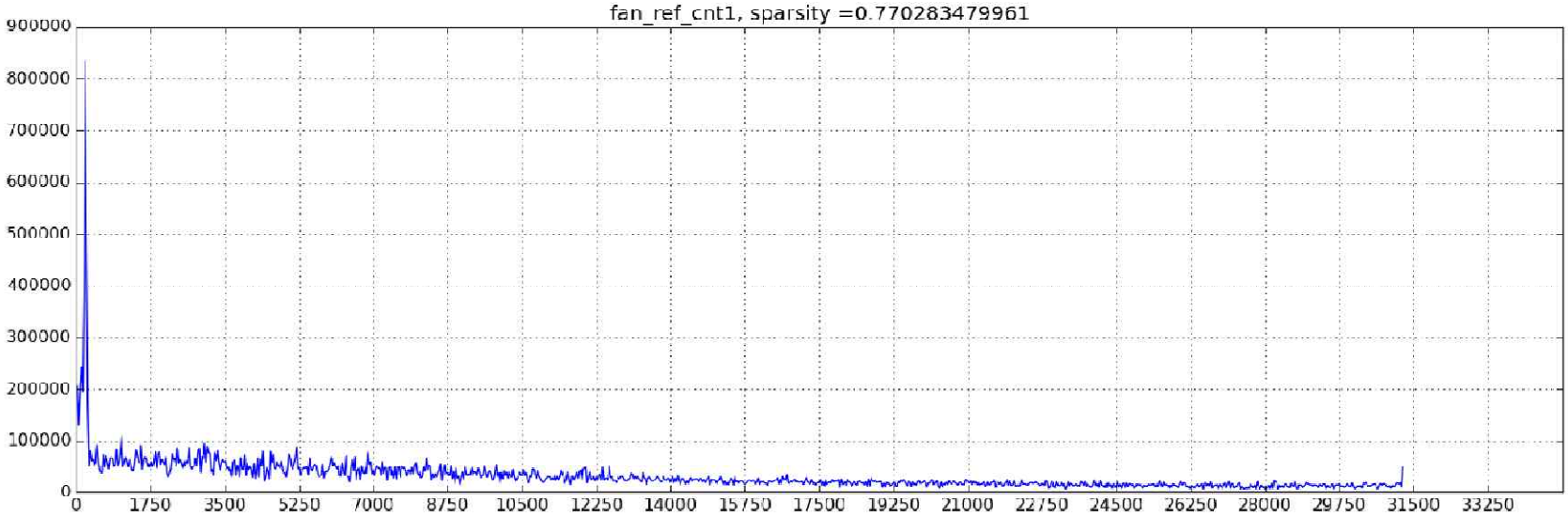
Vibration raw data - 37.5kHz

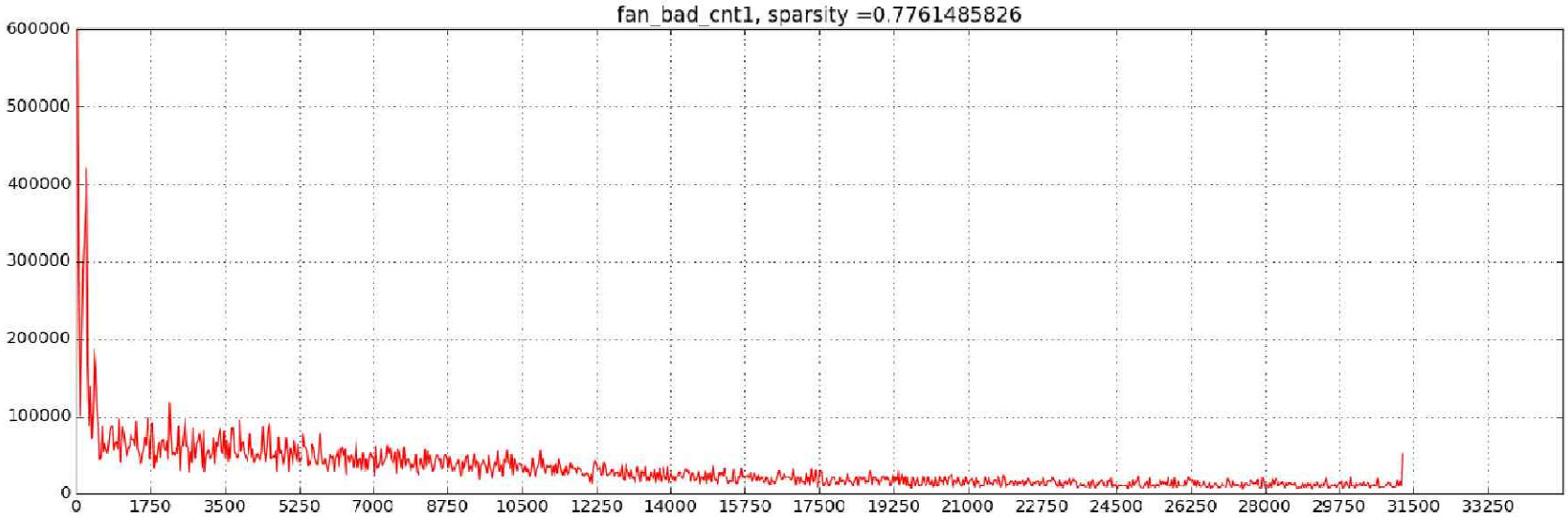


Vibration raw data - 300Hz

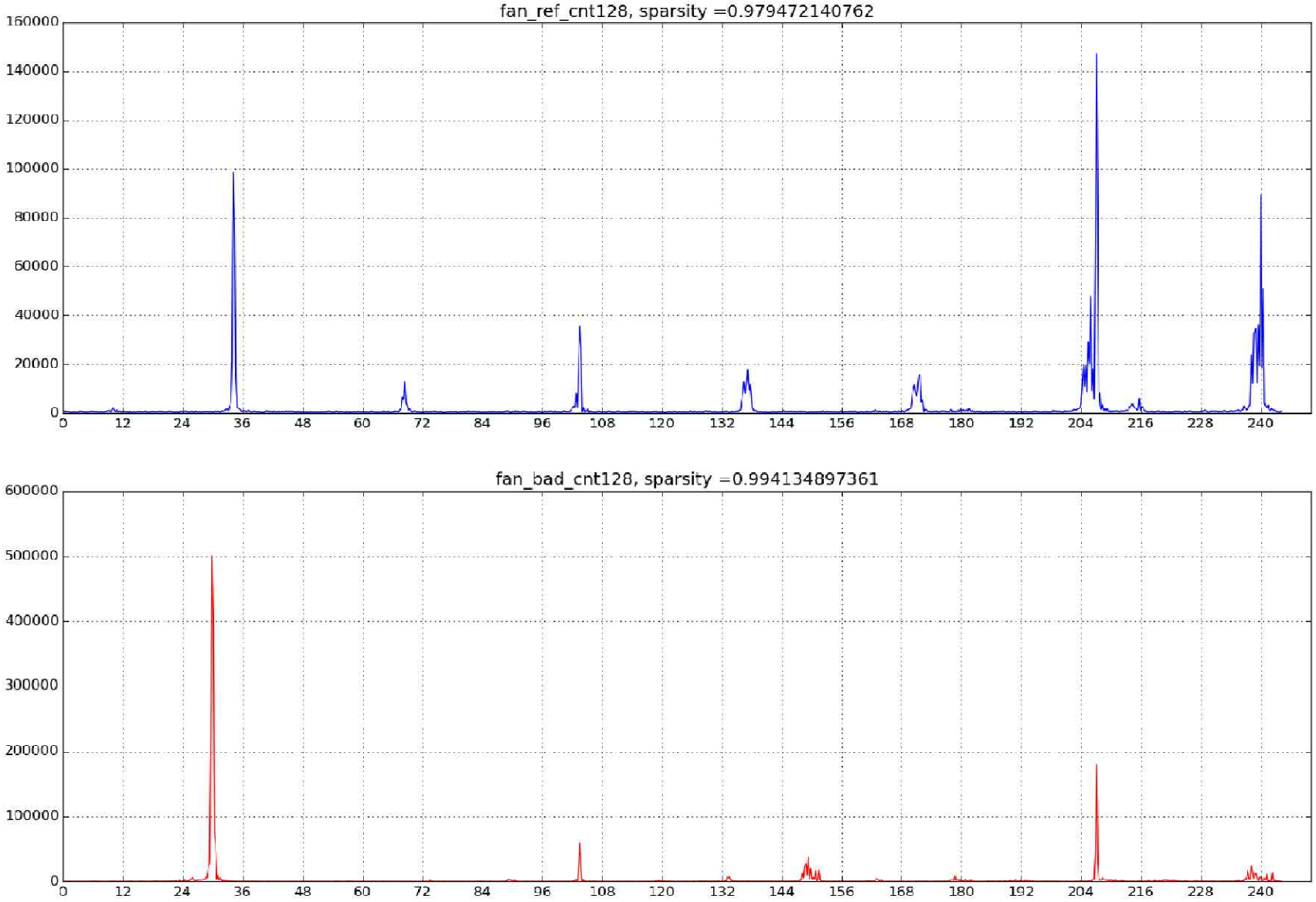


Discrete Cosine Transform- 37.5kHz

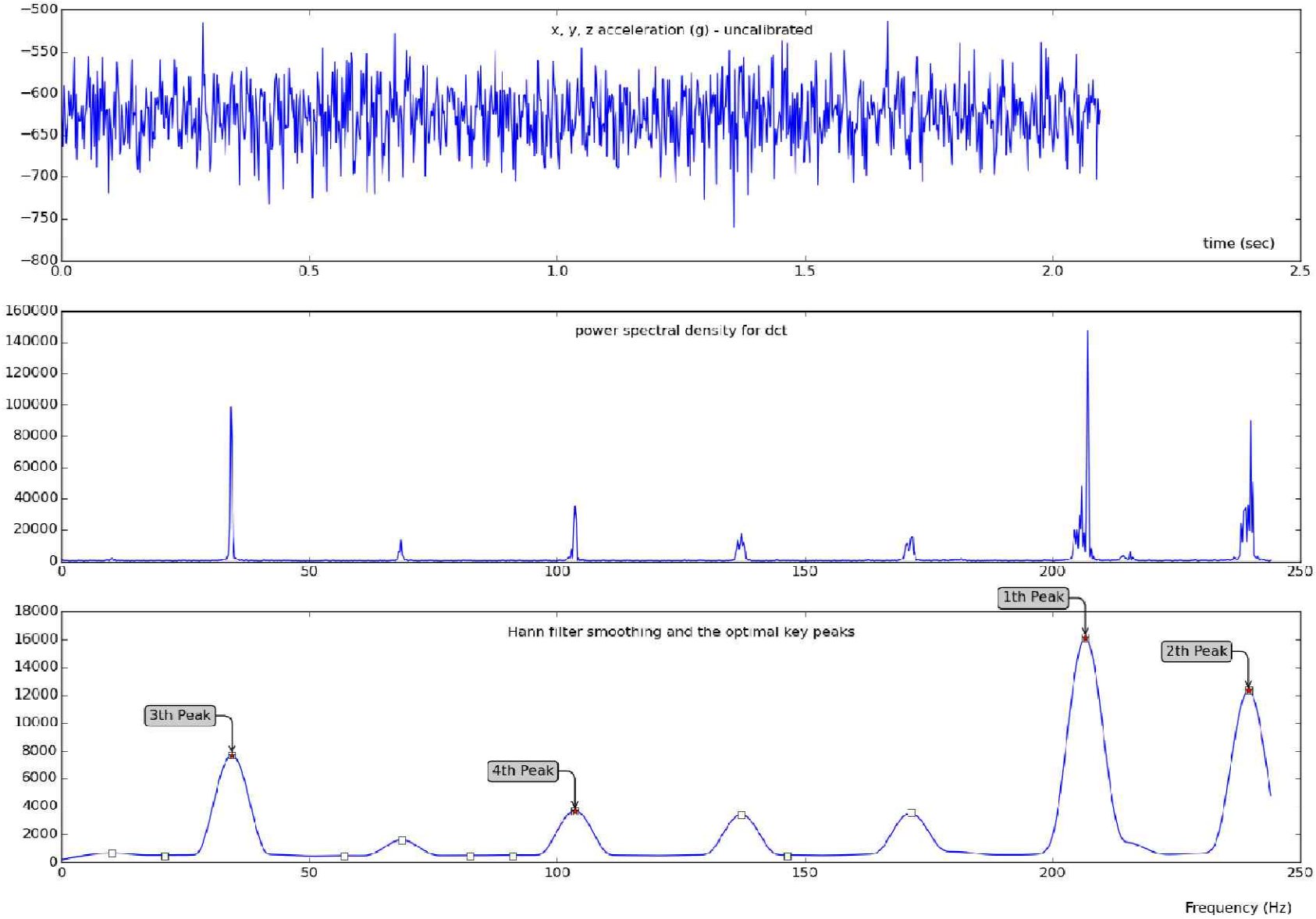




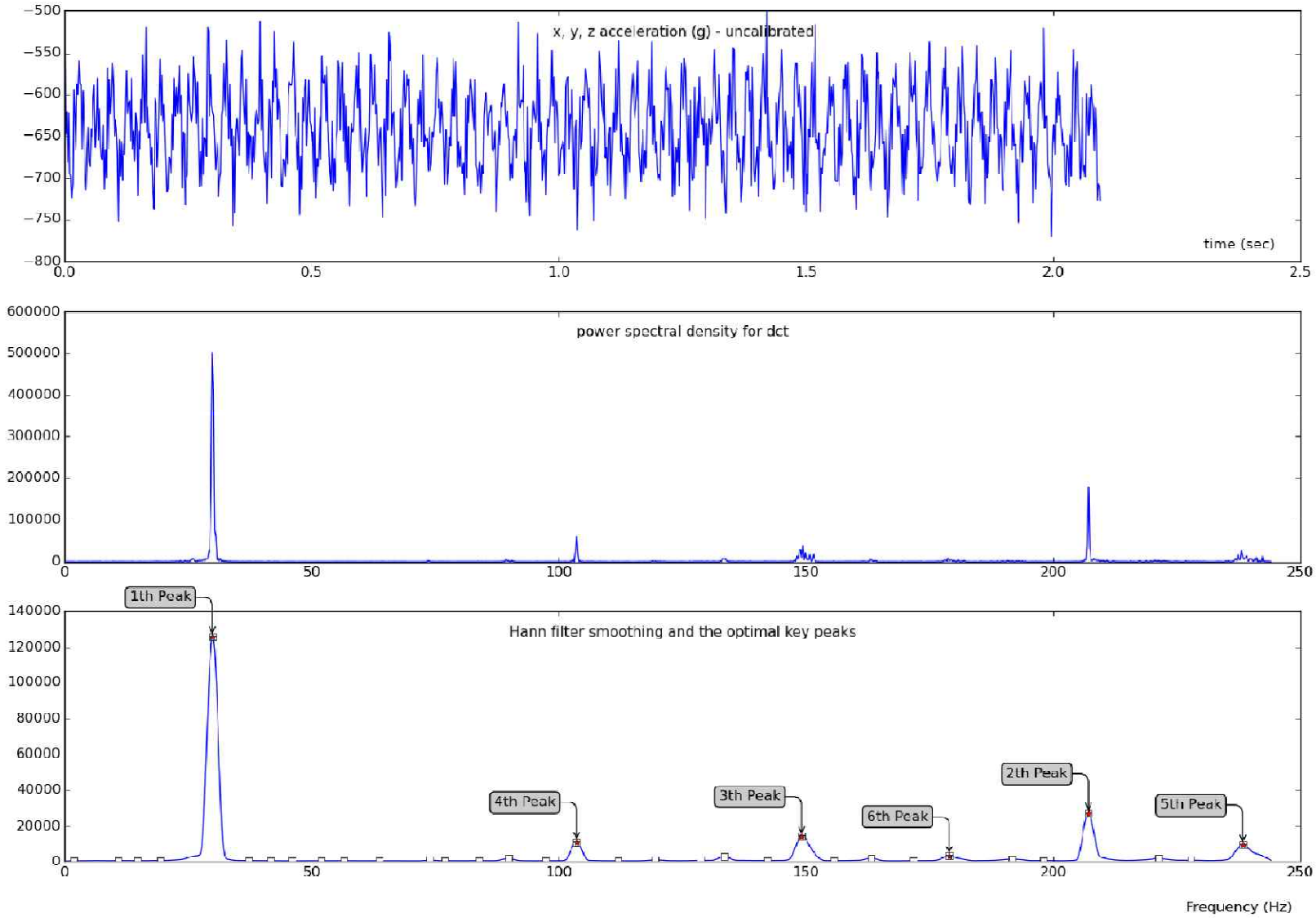
Discrete Cosine Transform- 300Hz



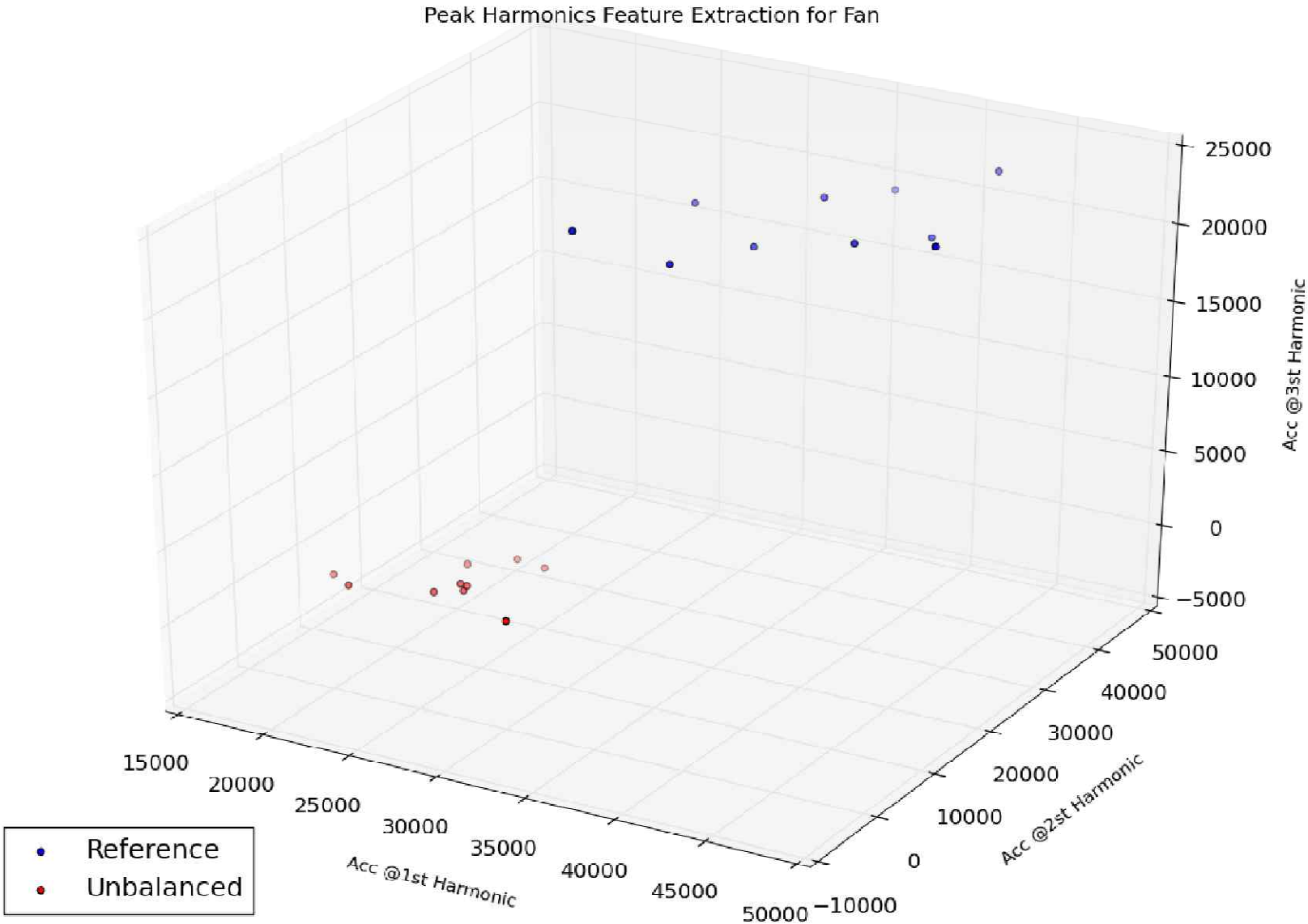
Automatic Feature Extraction - Reference



Automatic Feature Extraction - Unbalanced



Clustering for Reference & Unbalanced Fan



Questions ?