Emerging Trends in AI/ML and Implications for Networking Research

2021/05/19 @ theNetworkingChannel

Dan Pei

Tsinghua University

清華大學 | NetMan Tsinghua

Data Access

Algorithm competition, live testbed, etc. to bring together the community

Q: There is need for fair and broader access to data needed for developing AI models and addressing some of the challenges, but the data is currently controlled by a few operators? What can be done to address this challenge?

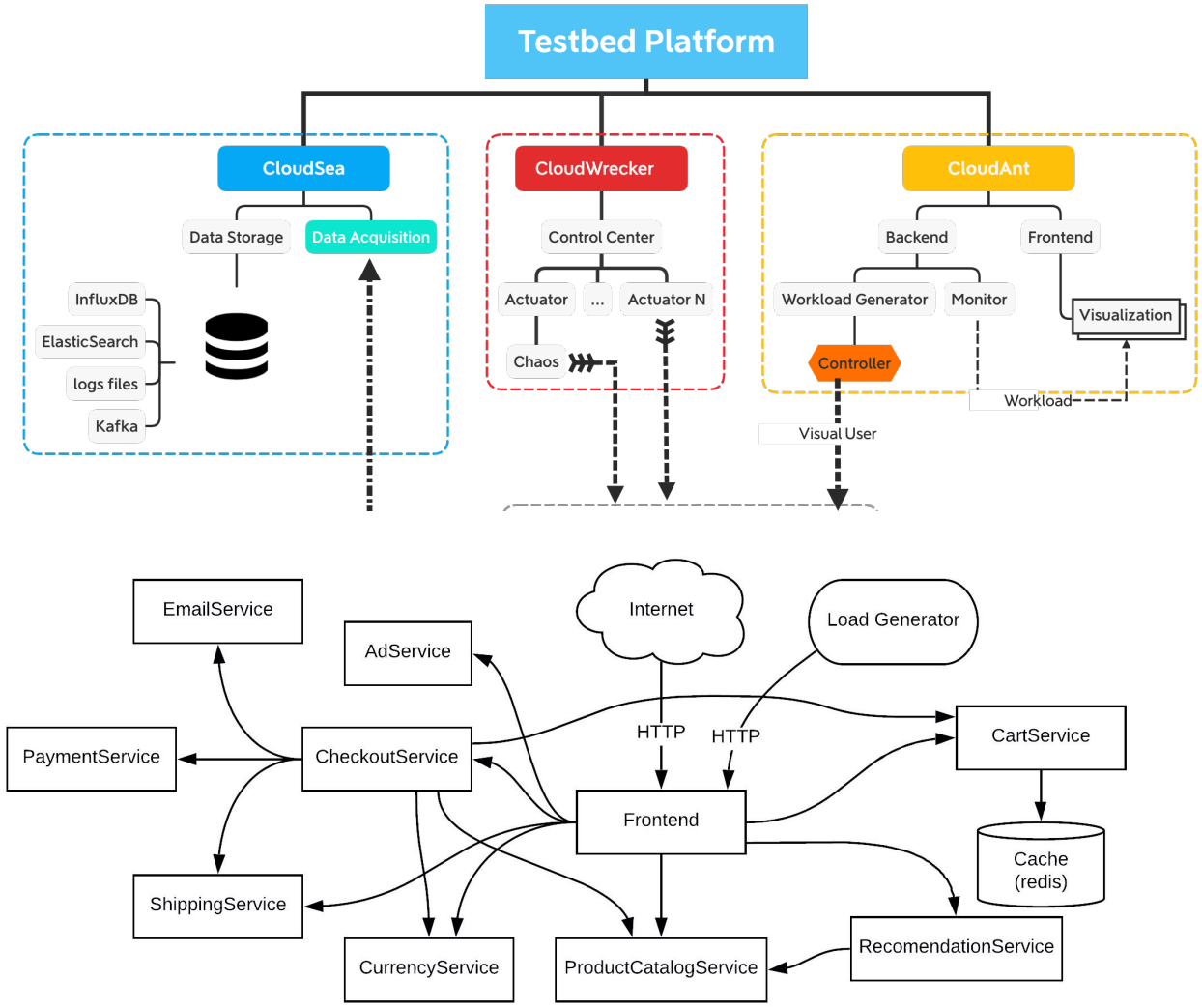
AIOps Challenge Algorithm Competitions

Datasets: https://github.com/netmanaiops

- 2018 AIOps Challenge: time series anomaly detection. Published labeled data from 5 Internet companies. More than 50 teams participated. Papers based on these data were published in KDD, IWQoS, etc.
  Data Downloadable @ https://github.com/NetManAIOps/KPI-Anomaly-Detection
- 2019 AIOps Challenge: multi-attribute time series anomaly localization. Published data from an Internet company. More than 60 teams participated.
  Data Downloadable @ https://github.com/NetManAIOps/MultiDimension-Localization
- 2020 AIOps Challenge: Anomaly detection and localization in a microservice system. Published data from a telecom company. More than 100 teams participated.
  Data Downloadable @ https://github.com/NetManAIOps/AIOps-Challenge-2020-Data
- 2021 AIOps Challenge: Anomaly detection and localization in banking systems. To be published data from two banks. More than 200 teams participated

A representative and live Net+AI testbed that the community can contribute and use

- New components can plug-and-play onto testbed
- Many missing pieces
  - Large-enough Industry-grade microservice based system
  - Realistic traffic
  - Failure patterns from industry
  - Failure injection systems
  - Realistic evaluation metrics



Use knowledge to glue all components (including AI-enabled ones)

Clearly define the properties and capabilities of AI-enabled component in the overall architecture

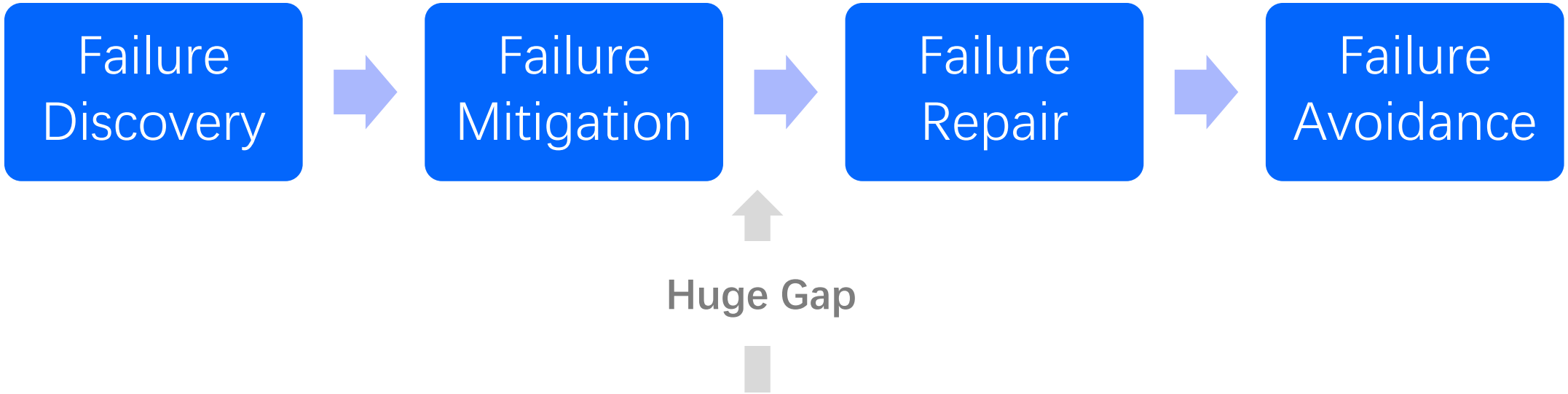
Q1: "It is hypothesized that as ML/AI solutions get infused in design of a range of networking functions, network architecture can be automated through simply optimizing generic AI models? What are your thoughts?"

Q2: "As networks are becoming increasingly mission critical; how must ML/AI strategies be adapted to operate in these environments?"

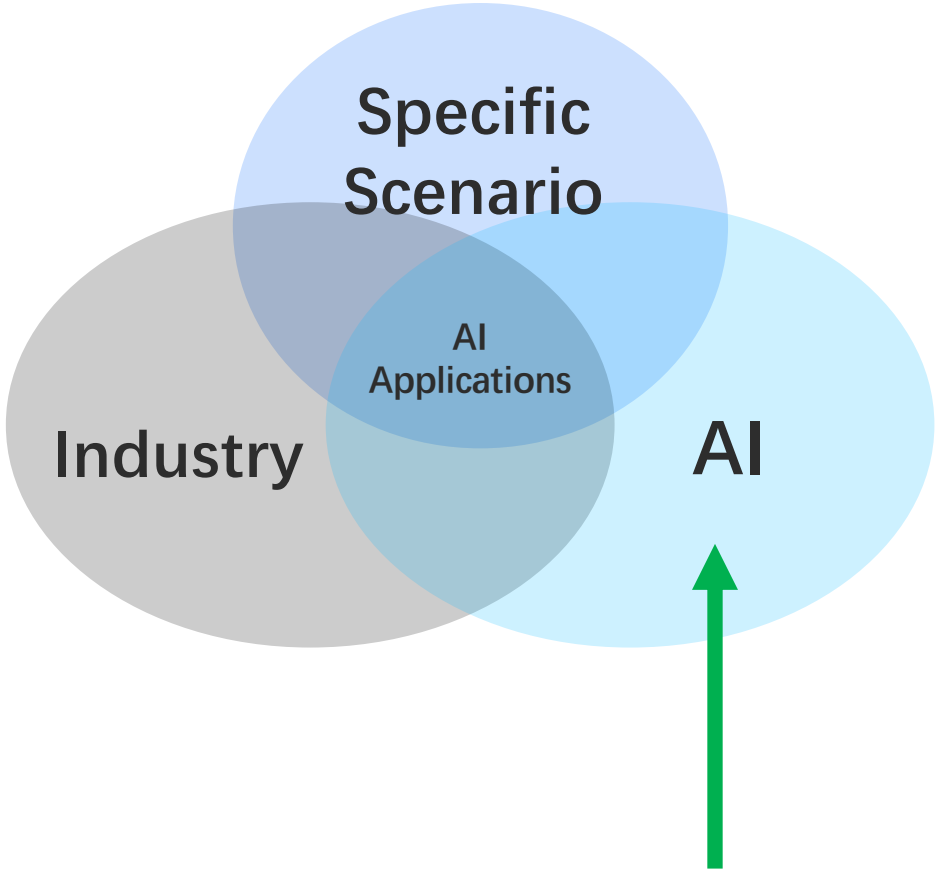
Q3: " AI/ML solutions have been shown to be brittle to adversarial perturbations or deviations from training data. Do you believe this will limit the use of AI/ML solutions in mission critical networks? How can network operators mitigate this threat?"

清華大學 | NetMan Tsinghua

Pitfalls: use general ML algorithms as blackbox to tackle Ops challenges



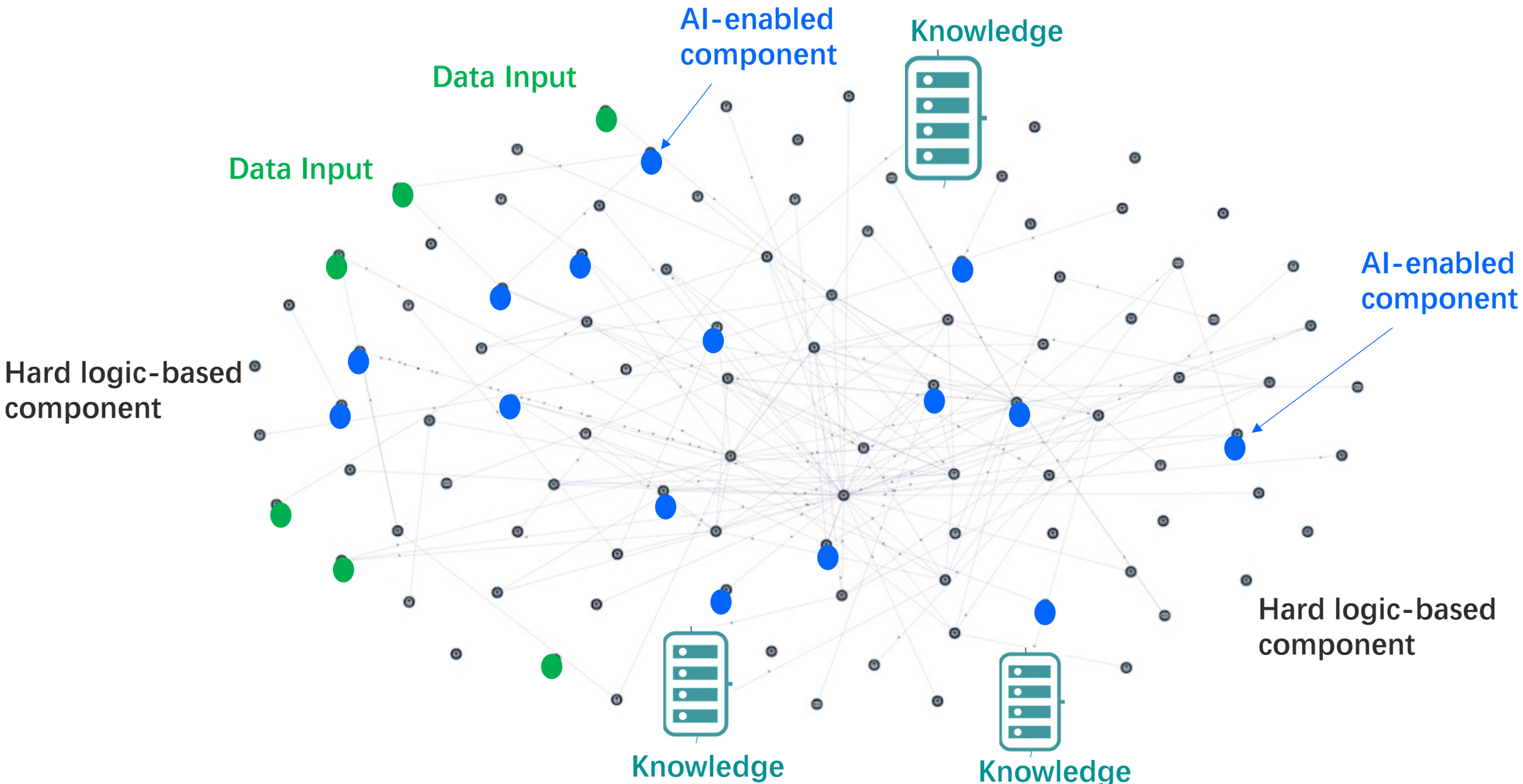
So far, AI succeeds only in specific application scenario in specific area in specific industry



Treat AI as a high-level programming language, to “code” some components

Output of AI-enabled components are probabilistic rather than deterministic

A Net+AI system: design the overall system around each component’s known capability and property, and “glue” the components using “knowledge”



Fully utilize latest AI technologies that enable better machine-human hybrid architecture

Active Learning, Transfer Learning, Ensemble Learning, Knowledge Graph, …

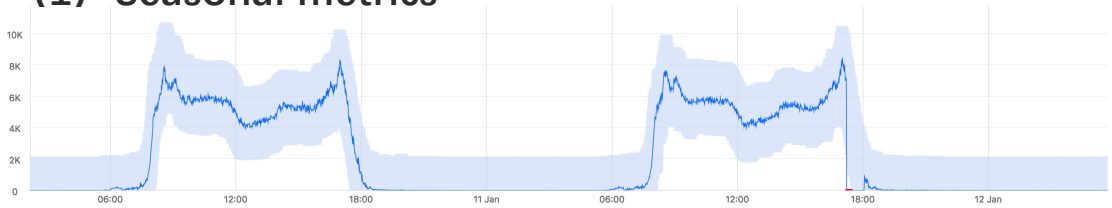
Q1: What are the expected breakthroughs in AI/ML technologies that will hold the most relevance for next generation communication networks and why?

Q2: What in your view are the most significant recent developments that hold promise for design and operation of communication networks?

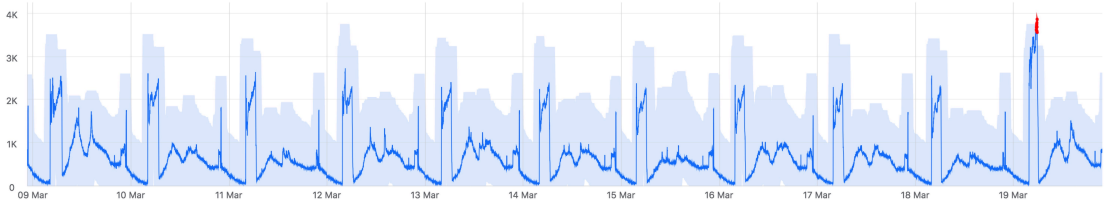
Q3: What breakthroughs do you expect in the next 5-10 years? Can we expect networks to operate autonomously in the next 10 years?

Time series anomaly detection: diverse metrics and their diverse anomalies

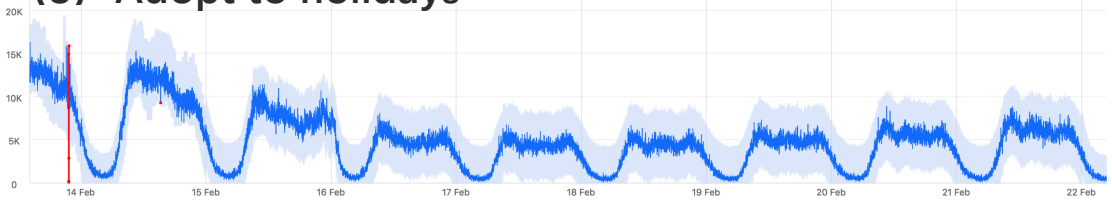
(1) Seasonal metrics



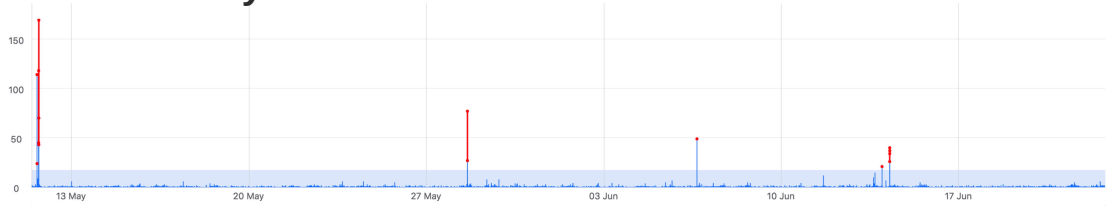
(2) Periodicity shift



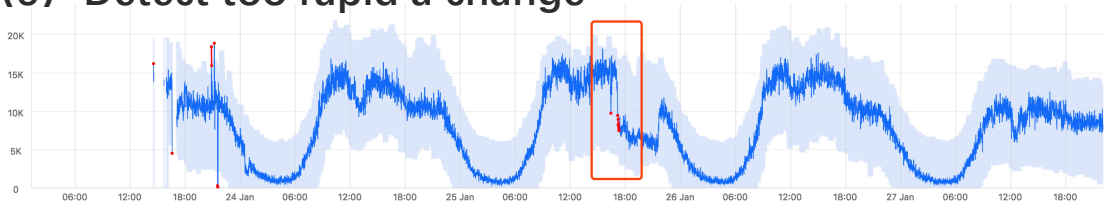
(3) Adopt to holidays



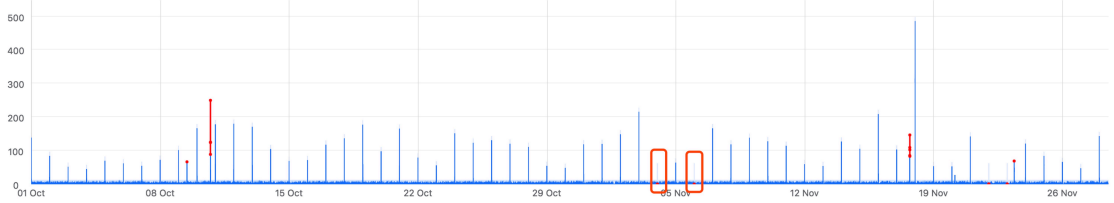
(4) Identify variable metrics and obtain extreme threshold



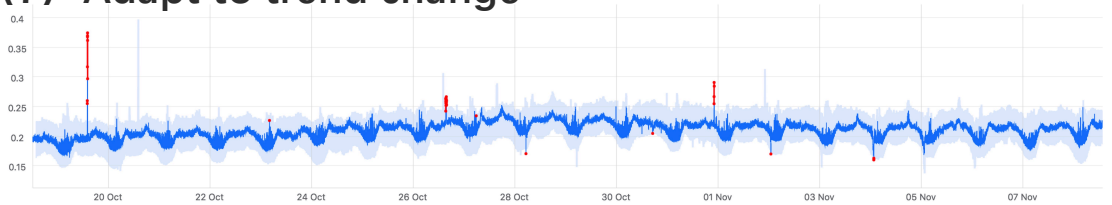
(5) Detect too rapid a change



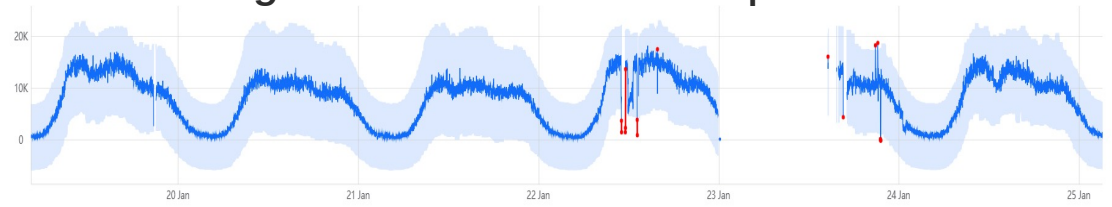
(6) Detect the lack of seasonality.



(7) Adapt to trend change



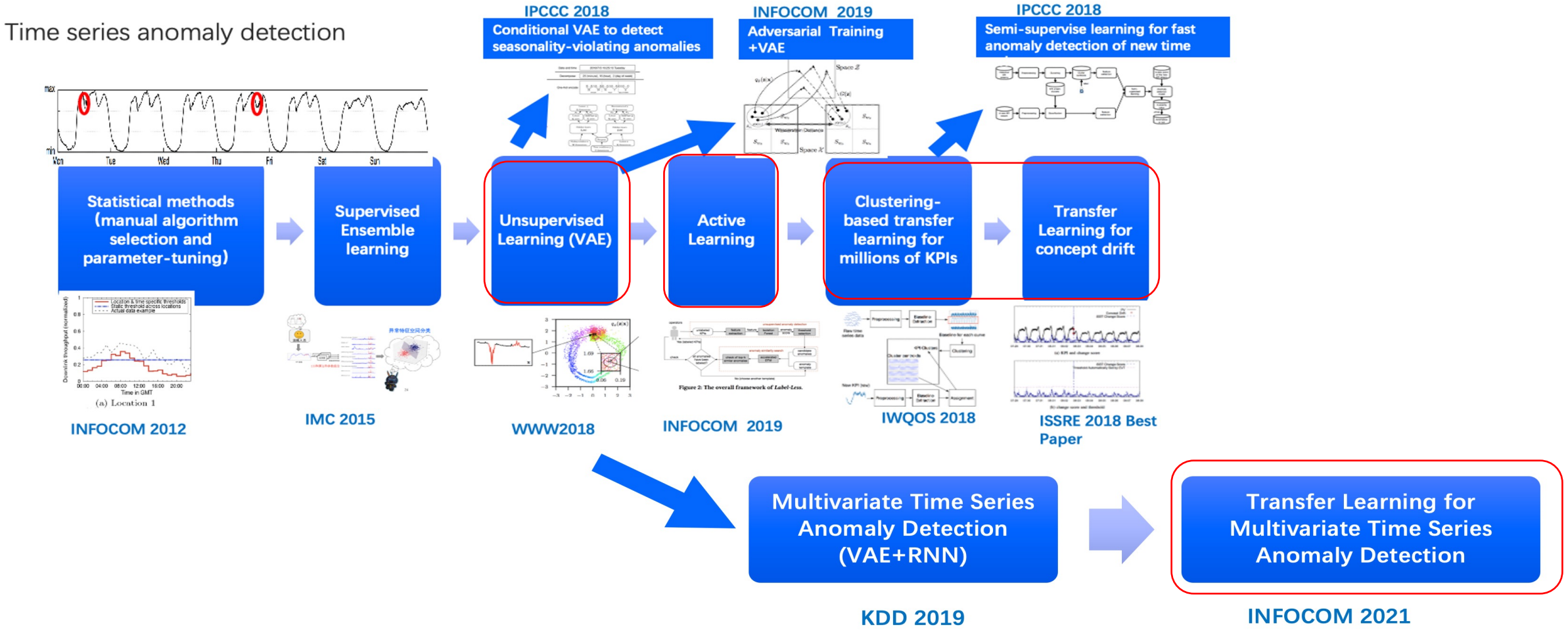
(8) Robust against data loss or interruption



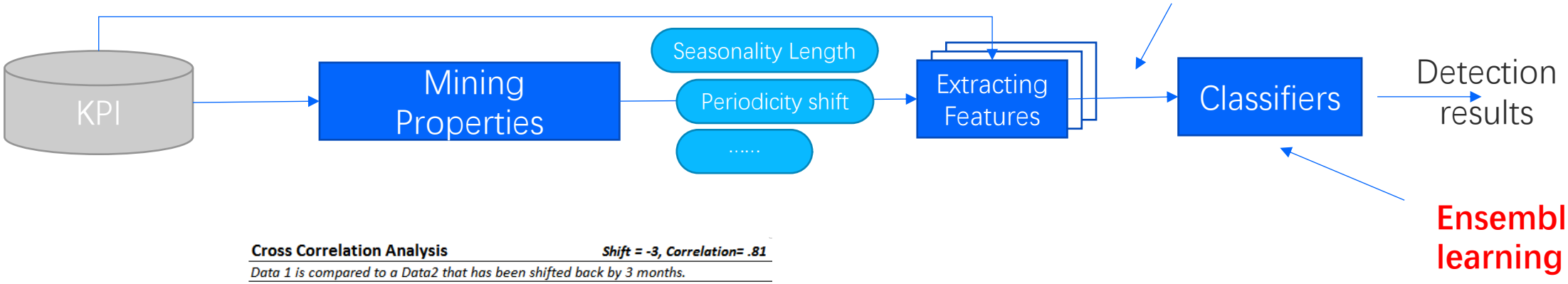
Labels are expensive and often unavailable → Unsupervised approaches

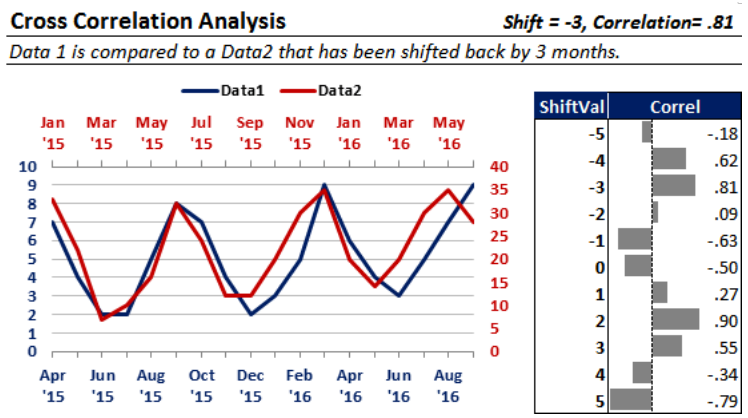
Hyperparameters affect unsupervised approaches' performance → active learning (human in the loop)

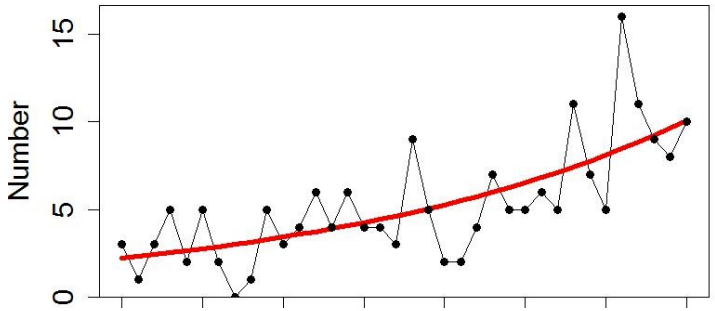
Training data can be sparse; data distribution can change → transfer learning



Knowledge graph from human experts for mapping profile to a set of classifiers







Other knowledge: causality rules, topology, etc.

Thanks!

Q&A

清華大學 | NetMan Tsinghua