NMRG @ NOMS2018

Toward Data-Driven Management

Kohei Shiomoto

Tokyo City University

24.04.2018

Agenda

- Motivation
- Data-driven management
- Building block: machine learning
- Applications
  - Anomaly detection
  - Root cause analysis
  - Resource management
  - Knowledge discovery (Workflow, Policy, Intent, ⋯)

Motivation

- Softwarization –Disaggregation-
  - SDN
    - Switch
    - Separation of control plane
  - NFV
    - Middlebox
    - Data plane as software

- How to deal with complex system?

Data-driven management

- Conventional Mechanism-driven approach
  - Given understanding precise mechanisms of components, build up a model of entire system.
- Towards holistic Data-driven approach
  - Mining data of inputs and outputs of Black-box
    - Given data, infer the relationship between inputs and outputs.
  - Machine learning is a key.

Data-driven management

- Mining data of inputs and outputs of Black-box
  - Machine learning & deep Learning
- Expectations
  - Anomaly detection
  - Root cause analysis/Correlation and causality inference
  - Traffic prediction
  - Resource management
  - Knowledge discovery
  - …

Building block: machine learning

- Big data accumulated in cloud [1]
- Scalable computing: Map reduce, GFS, Haddop,
- Framework: Spark, TensorFlow, Caffe, PyTorch, etc.
- Public cloud: AWS, Azure, etc.

[1] I. Stoica, et al., “A berkeley view of systems challenges for ai,” UCB/EECS-2017-159.

Building block: machine learning

- Tasks:
  - Inference, Recognition, Prediction, Learning,
- Tools:
  - Supervised
    - Regression, Classification (Naïve Bayes, Neural Network, SVM,⋯), ⋯
  - Unsupervised
    - Clustering (k-means, DBSCAN, Random forest, ⋯), Model estimation, ⋯
  - ⋯

Applications

- Anomaly detection
- Root cause analysis
- Resource management
- Knowledge discovery (Workflow, Policy, Intent, …)

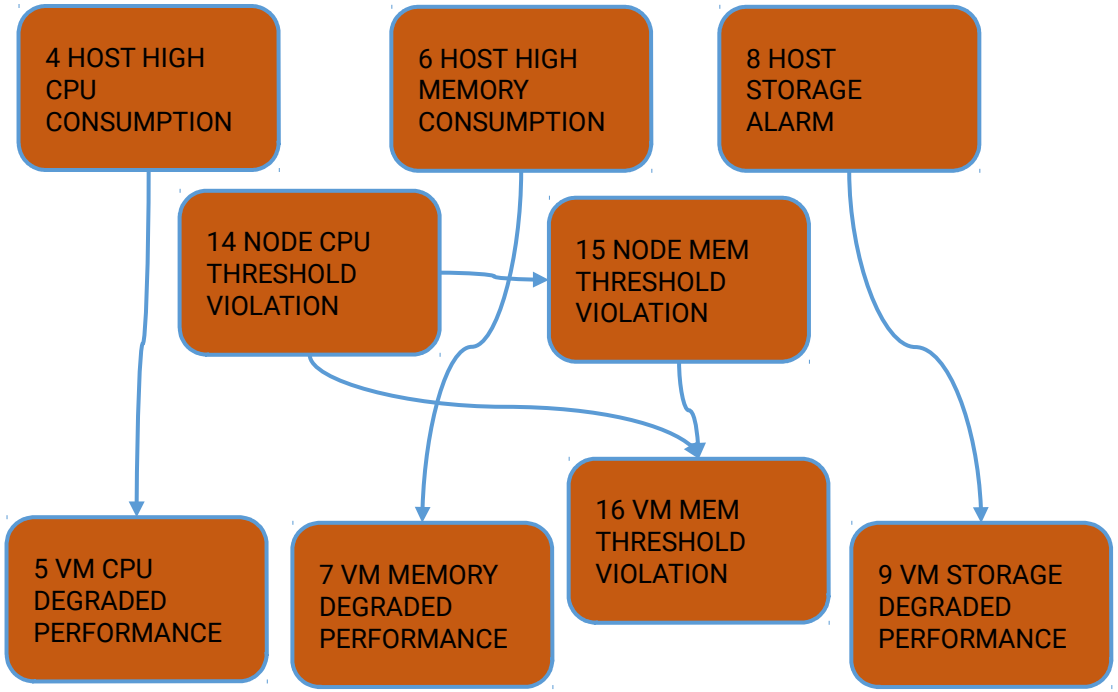
Applications
Anomaly detection

- Correlation detection: NICE [2], G-RCA[3]
- Syslog analytics: SyslogDigest [4], Spatio-Temporal [5]

[2] Ajay Mahimkar, et al., “Troubleshooting chronic conditions in large ip networks,” CoNEXT ‘08.
[3] H. Yan, et al., “G-rca: A generic root cause analysis platform for service quality management in large ip networks,” CoNEXT ’10.
[4] T. Qiu, et al., “What happened in my network: Mining network events from router syslogs,” IMC ’10.
[5] T. Kimura, et al., “Spatio-temporal factorization of log data for understanding network events,” IEEE INFOCOM 2014.

Applications
Root cause analysis

- Inference graph & Bayesian network
  - How to build
  - How to solve



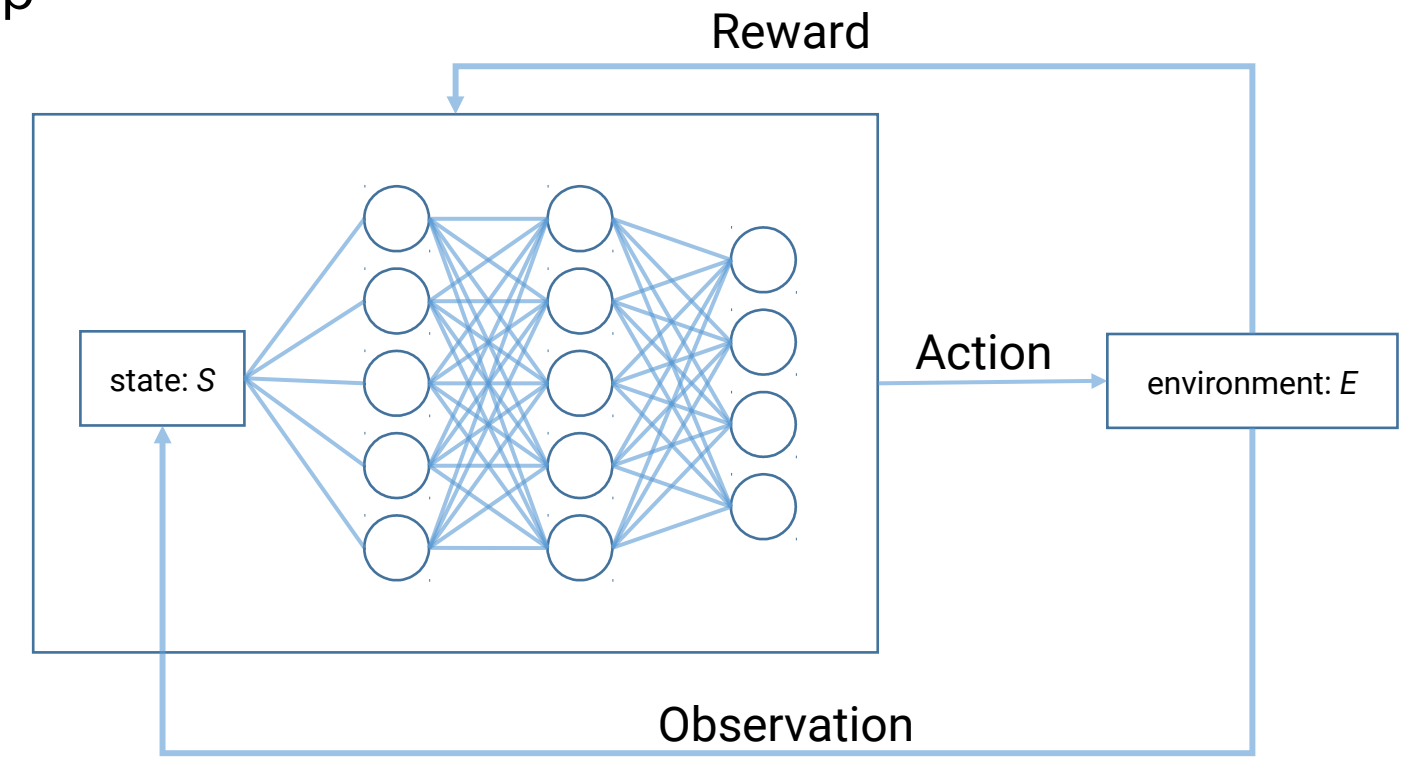
Applications
Root cause analysis

- IP over Fiber: SCORE [6], Srhink [7], BayesNet [8]
- Enterprise network: Sherlock [9]
- CDN: WISE [10]

[6] R.R. Kompella, et al., “Ip fault localization via risk modeling,” NSDI’05.
[7] S. Kandula, et al., “Shrink: A tool for failure diagnosis in ip networks,” MineNet ’05.
[8] M. Steinder and A.S. Sethi, “Increasing robustness of fault localization through analysis of lost, spurious, and positive symptoms,” Infocom 2002.
[9] P. Bahl, et al., “Towards highly reliable enterprise network services via inference of multi-level dependencies,” SIGCOMM ’07.
[10] M. Tariq, et al., “Answering what-if deployment and configuration questions with wise,” SIGCOMM ’08.

Applications
Resource Management

- Deep Reinforcement Learning
  - Job scheduling in cluster: Deep RM [11]
  - Bandwidth calendaring
  - Adaptive streaming



[11] H. Mao, et al., "Resource management with deep reinforcement learning," HotNets '16.

Applications
Knowledge Discovery

- Trouble ticket analysis [12]
- Router config error detection: Mineral [13]
- Mobile network eNodeB: AESOP [14]

[12] A. Watanabe, et al., “Workflow extraction for service operation using multiple unstructured trouble tickets,”. NOMS 2016.
[13] F. Le, et al., “Minerals: Using data mining to detect router misconfigurations,” MineNet ’06.
[14] S. Deb, et al., “Aesop: Automatic policy learning for predicting and mitigating network service impairments,” KDD ’17.

Closing Remarks

- Disaggregate vertically integrated system into components to achieve sustainable healthy growth.

- Machine learning & deep Learning
  - Mining data of inputs and outputs of Black-box

- How to collect data?
- Architecture for monitor, analytics, and actuator?
- Standard?