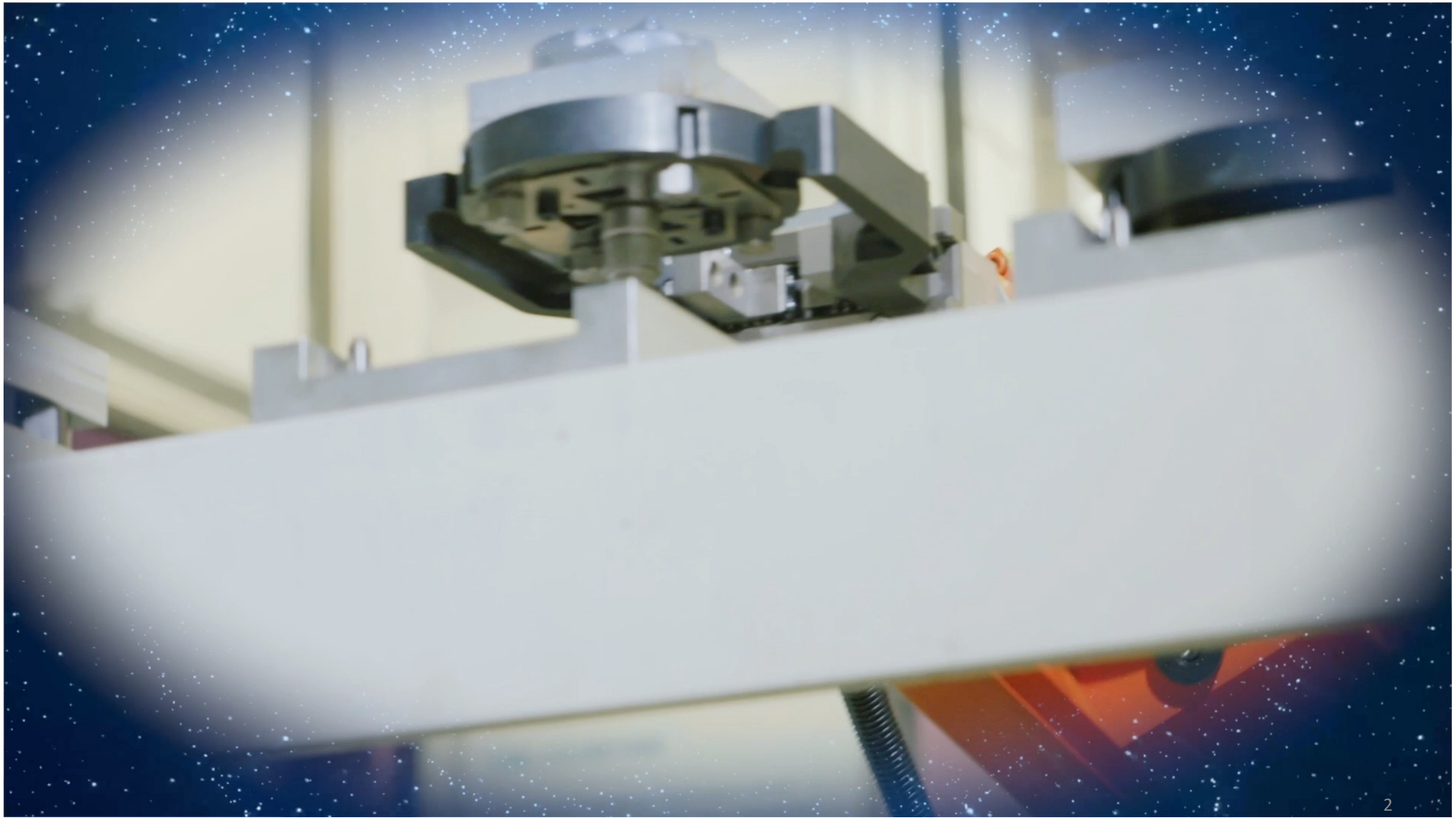
Advanced Network Management
a.k.a.
AIOps: **Autonoumous IT Operations** or **AI for IT Operations**
course#: 80240663
Spring 2024

群聊: ANM2024 Spring

Instructor: Dan Pei
peidan@tsinghua.edu.cn

Teaching Assistant:
Zhe Xie
xiez22(at)mails(dot)tsinghua(dot)edu(dot)cn



Roadmap

- **Self-Introduction**
- Course Logistics
- What is AIOps?
- Course Coverage (Website)
- Project

How to pronounce “Pei”?

Just pronounce “Pei” as “Pay”

In fact, just call me “Dan”

Louvre Museum

I-M Pei: the famous architect who designed the glass pyramid

About the Instructor

- Tenured Associate Professor.
- UCLA Ph.D. Best Ph.D. Thesis Award in UCLA CS in 2005.
- Joined Tsinghua CS Department in December 2012
- Homepage: **http://netman.aiops.org/~peidan**
- Previously a Principal Researcher at AT&T Research, a co-founder and founding CEO of a mobile health company in Beijing, before joining Tsinghua.
- Supervised interns from CMU, Cornell, Princeton, UCLA, GaTech, Michigan, Northwestern etc. Now @ Google, MSR, IBM, Purdue, Northeastern, HKUST
- ACM/IEEE Senior Member

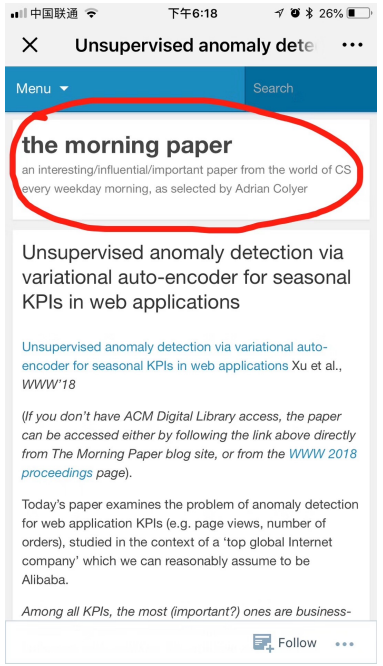
My Research Group @ Tsinghua: NetMan
http://netman.aiops.org

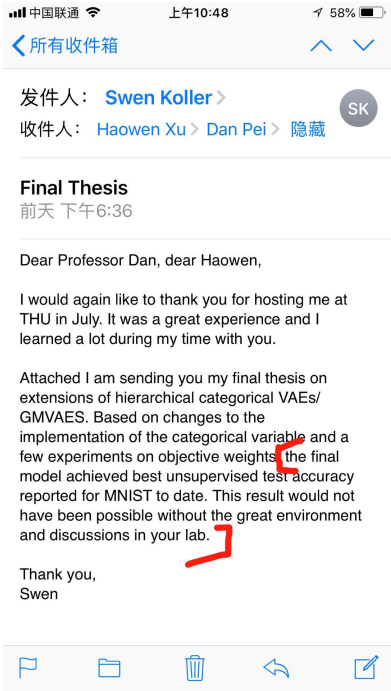
- Currently advising 7 Ph.D. and 3 M.S. students at Tsinghua.
- 3 affiliated associate professors
- 2 post-docs

- Graduated 22 PhDs & 9 Masters (4 Huawei “Young Genius”， 3 Kuaishou Star, 2 Ali Star, 4 went to MSRA, 5 became assistant/associate professors)

Industry Collaborators

Research results are covered by technology media such as MIT technology Review, Hacker News, Mother Board, Morning paper, and many Chinese media.



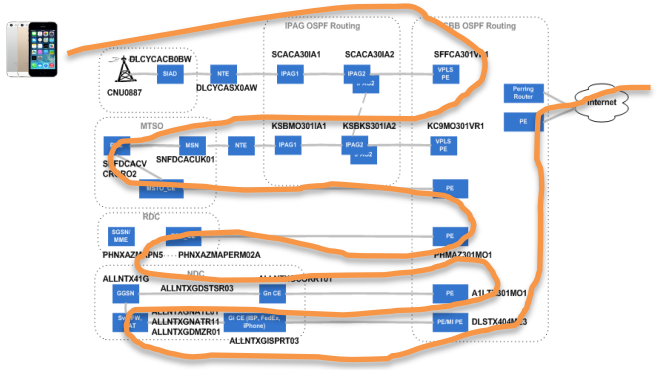


Publications:

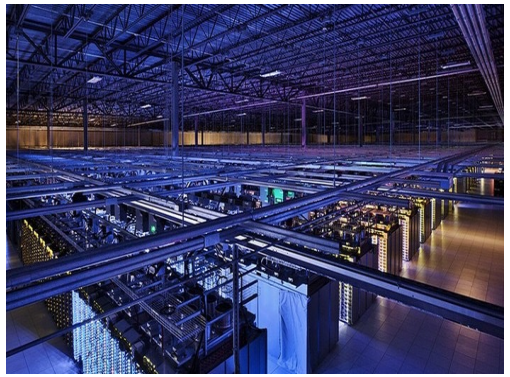
200+ AIOps papers and 20+ issued US Patents. Published in SIGCOMM、WWW、KDD、SIGMETRICS、TON、ATC、INFOCOM、FSE、ASE、IMC、CoNEXT、IJCAI etc.

AIOps: Autonomous IT Operations through Machine Learning

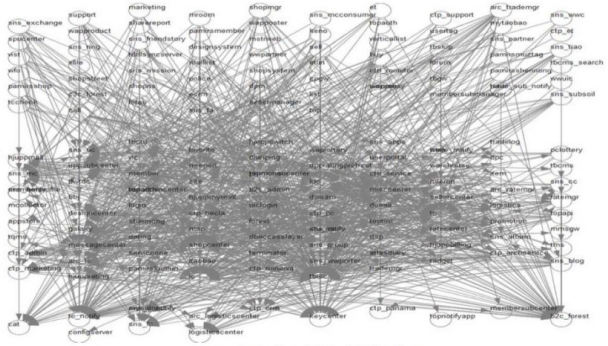
Large & complex access network



Large & complex data center



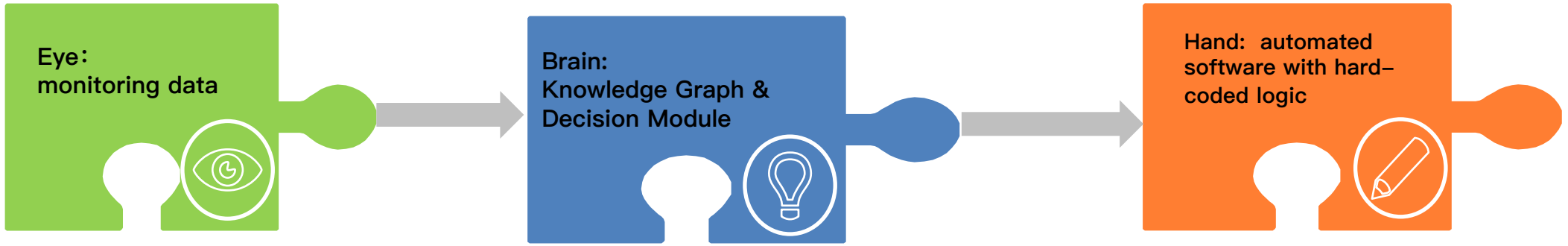
Large & complex application software



2012 淘宝核心链路应用拓扑图

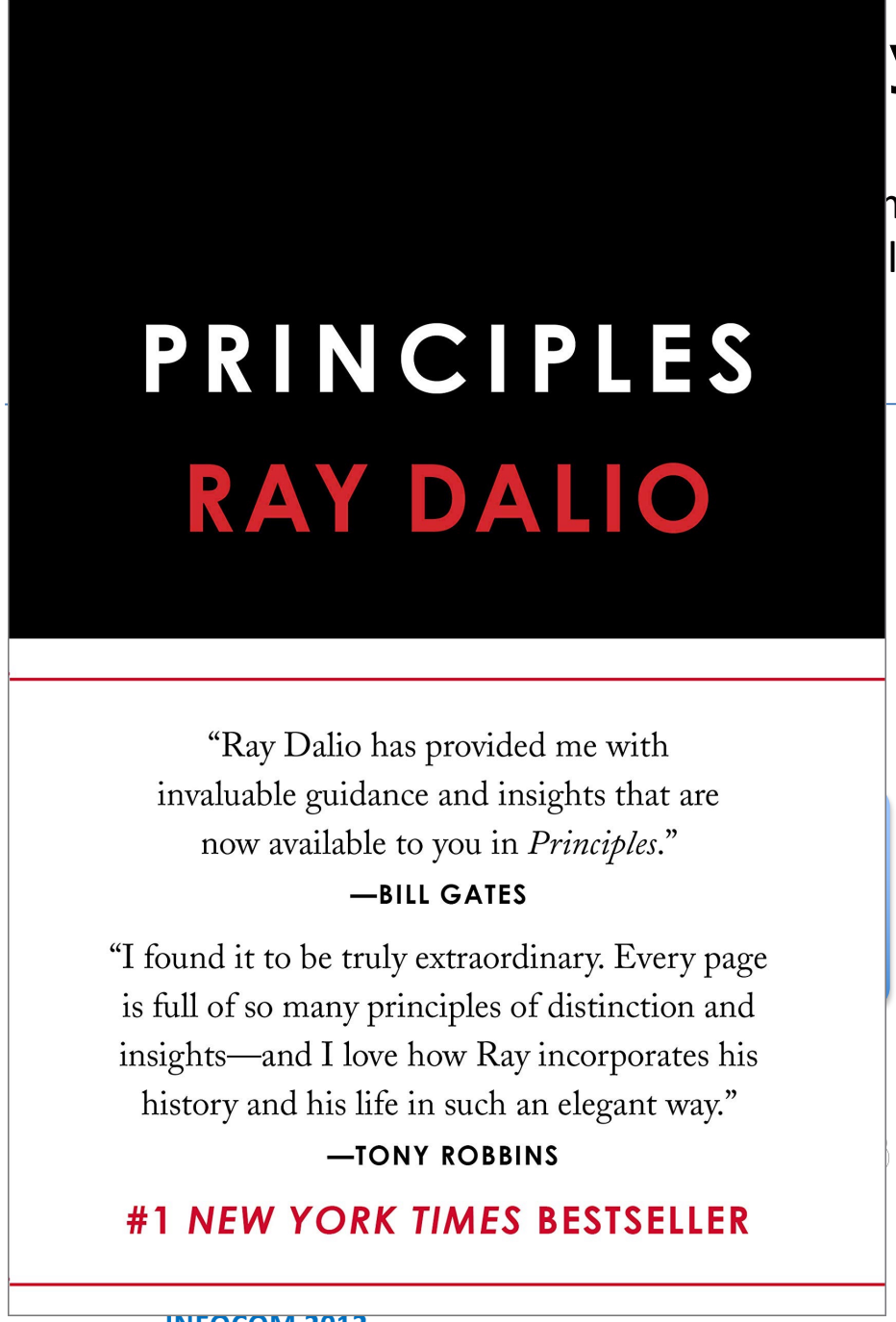
- Imagine that you are running an Internet-based service with hundreds of thousands of servers and many software modules, a large, complex, cross-layer, and rapidly evolving distributed system.
- You want to achieve 99.999% service reliability, but machine-generated monitoring data and hundreds of operators (IT operation engineers) alone won't get you there, because of the high complexity and sheer scale of the software/hardware system and the vast amount of machine-generated data.
- Machine learning is the direction to enable Autonomous IT Operations autonomous.

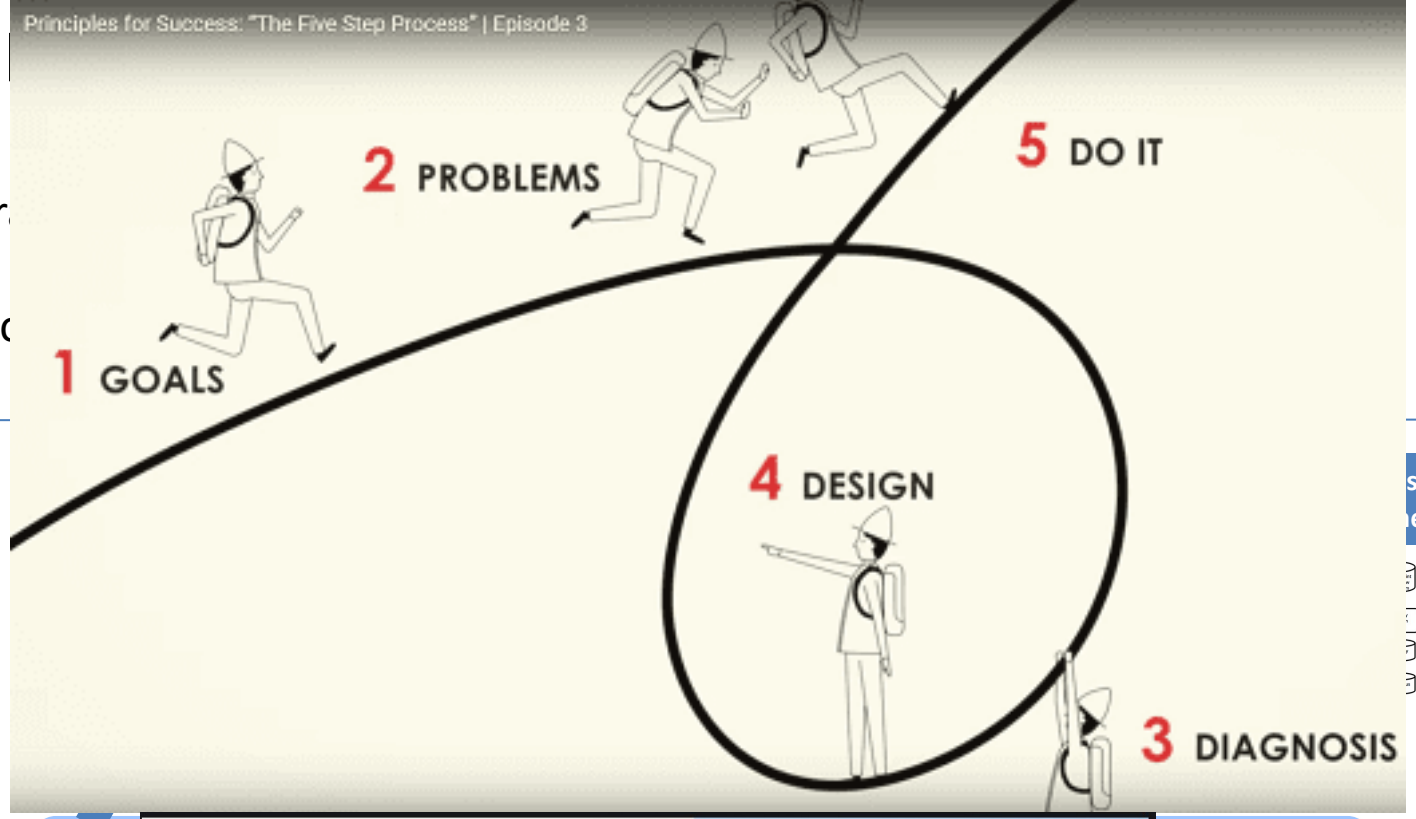
AIOps Architecture & Algorithms



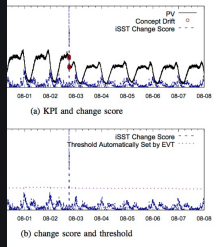
- The major topics of AIOps often coincide with its more general counterparts in Machine Learning:
  1. Anomaly Detection in Time Series, Logs (semi-structured text), Traces (program execution trace), and Graphs
  2. Anomaly Localization
  3. Failure/Event Prediction
  4. Causal Inference and its application in Root Cause Analysis

- State-of-art Machine Learning Algorithms are applied to solve the unique challenges in AIOps:
  1. Deep Neural Networks for Time Series or Sequence
  2. Deep Generative Model (VAE, GAN)
  3. Deep Reinforcement Learning
  4. Large Language Model
  5. Causal Inference





Transfer Learning for concept drift



INFOCOM 2012

ISSRE 2018 Best Paper

Roadmap

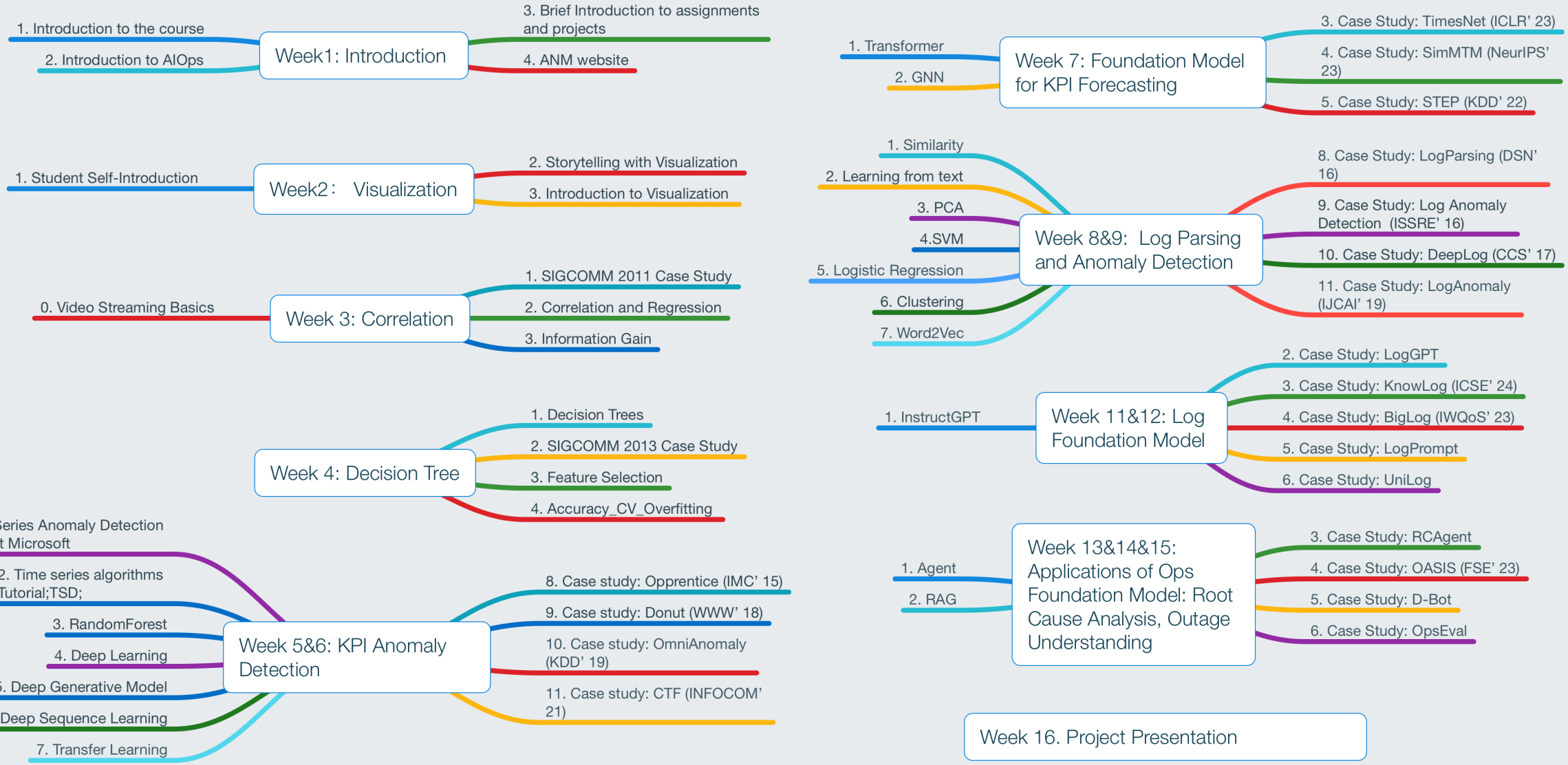
- Self-Introduction
- **Course Logistics**
- What is AIOps?
- Course Coverage (Website)
- Project

Summary

- AIOps is an **interdisciplinary** research field between **Machine Learning** and **Systems/Networking**, which is why this course had this historical title “Advanced Network Management”. If you are interested in learning how a large distributed system can be better run with the help of machine learning, this course is for you. If you want to learn **how machine learning can help solve challenging problems in a very complex system**, this course is for you.

- This course will cover the latest progress in major topics of AIOps using **case studies** from recent research papers in top conferences in all major computer science fields, including Machine Learning, Data Mining, System/Networking, Software Engineering, Database, Multimedia, etc.

- Through these case studies, we will show how the latest Machine Learning Algorithms are applied to solve the unique challenges in AIOps. The **basics of these Machine Learning** algorithms will be briefly **reviewed in an easy-to-understand way**, without going through the detailed theory behind them. Thus by the end of the course, you will be able to learn **roughly how these algorithms work, and how it can be applied to solve real-world problems.**


Week1: Introduction
1. Introduction to the course
2. Introduction to AIOps
3. Brief Introduction to assignments and projects
4. ANM website
Week2： Visualization
1. Student Self-Introduction
2. Storytelling with Visualization
3. Introduction to Visualization
Week 3: Correlation
0. Video Streaming Basics
1. SIGCOMM 2011 Case Study
2. Correlation and Regression
3. Information Gain
Week 4: Decision Tree
1. Decision Trees
2. SIGCOMM 2013 Case Study
3. Feature Selection
4. Accuracy_CV_Overfitting
Week 5&6: KPI Anomaly Detection
1. Time-Series Anomaly Detection Service at Microsoft
2. Time series algorithms Tutorial;TSD;
3. RandomForest
4. Deep Learning
5. Deep Generative Model
6. Deep Sequence Learning
7. Transfer Learning
8. Case study: Opprentice (IMC' 15)
9. Case study: Donut (WWW' 18)
10. Case study: OmniAnomaly (KDD' 19)
11. Case study: CTF (INFOCOM' 21)
Week 7: Foundation Model for KPI Forecasting
1. Transformer
2. GNN
3. Case Study: TimesNet (ICLR' 23)
4. Case Study: SimMTM (NeurIPS' 23)
5. Case Study: STEP (KDD' 22)
Week 8&9: Log Parsing and Anomaly Detection
1. Similarity
2. Learning from text
3. PCA
4.SVM
5. Logistic Regression
6. Clustering
7. Word2Vec
8. Case Study: LogParsing (DSN' 16)
9. Case Study: Log Anomaly Detection (ISSRE' 16)
10. Case Study: DeepLog (CCS' 17)
11. Case Study: LogAnomaly (IJCAI' 19)
Week 11&12: Log Foundation Model
1. InstructGPT
2. Case Study: LogGPT
3. Case Study: KnowLog (ICSE' 24)
4. Case Study: BigLog (IWQoS' 23)
5. Case Study: LogPrompt
6. Case Study: UniLog
Week 13&14&15: Applications of Ops Foundation Model: Root Cause Analysis, Outage Understanding
1. Agent
2. RAG
3. Case Study: RCAgent
4. Case Study: OASIS (FSE' 23)
5. Case Study: D-Bot
6. Case Study: OpsEval
Week 16. Project Presentation

Course Requirements

- Course website: http://course.aiops.org/

- **Prerequisites:**
  - You are expected to be familiar with at least one programming language, preferably Python.
    - If not, please quickly learn one.
  - You don't have to have previous machine learning knowledge.

- Encourage interaction and discussion
  - stop me and ask questions at any time!
  - You get credits for interaction (Yes, our TA notices your interaction)

Course Info

- Time: Wednesday **9:50am-12:15pm**
  - 15 classes, each has three 45-minute sessions.
- Session 1: Machine Learning Basics
- Sessions 2 & 3: AIOps case studies
- Office Hours:
  - After every class.

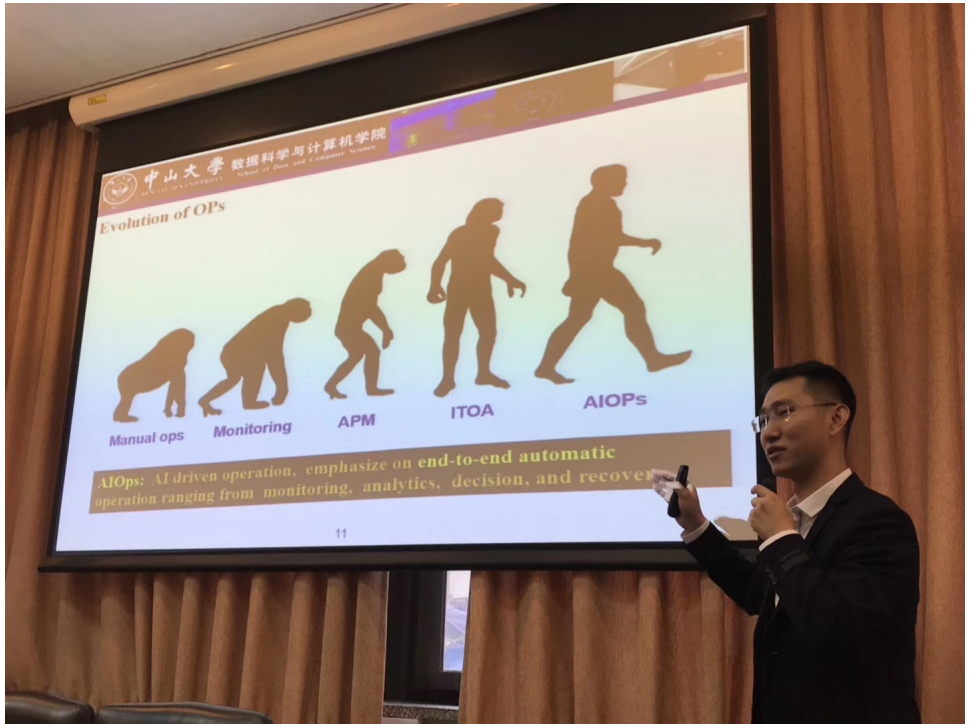
Grading:

- Attendance & Interaction: 10%
  - Presence: 0.5% for each lecture in Weeks 3-15. 6% in total
  - 1% for each question asked or answered. 4% maximum.

- Project:
  - Task 1: Paper Reading and Presentation (20%)
  - Task 2: Reproduction (40%)
  - Task 3: Report Writing (30%)

- The final grade will be in letter grading scale (e.g., A,B,C,D)

Invited Talks



Roadmap

- Self-Introduction
- Course Logistics
- **What is AIOps?**
- Course Coverage (Website)
- Project

Roadmap

- Self-Introduction
- Course Logistics
- What is AIOps?
- **Course Coverage (Website)**
  - **http://course.aiops.org**
- Project

Case Studies: Papers and Slides

- Already Posted on the course website
  - Maybe updated before each class, so please check the website weekly.
- Lecture Coverage
  - You were supposed to read these papers carefully
- Reading List
  - Relevant papers, read them if you have time

Why reading papers/case studies?

- Purpose: background knowledge, algorithm, methodology, writing, design principle, solution inspiration, evaluation methodology for the problem you are working on, and (sometimes) vision.

- My strongly biased personal opinion: Papers are not for finding topics for your next AIOps paper.
  - Instead, if possible, “From Practice, into Practice”: E.g. try existing algorithms in papers in practice to discover where it does not work

What's a good paper

- A new and important problem, solid solution.

- Old and challenging problem, a new *simple but elegant* solution, with straightforward insight and intuition behind it.

Using Google Scholar to find relevant papers (demo)

- Browse latest proceedings of relevant conferences, find one relevant paper *p*. *Relevant paper set S={p}*

for (each new *p* in *S) {*

Browse *p*'s references, and put relevant ones into *S*;

Browse *p*'s citations in Google Scholar, and put relevant ones into *S*;

}

https://www.editage.com/insights/8-winning-hacks-to-use-google-scholar-for-your-research-paper

Roadmap

- Self-Introduction
- Course Logistics
- What is AIOps?
- Course Coverage (Website)
- **Project**

Project: Time Series Forecasting Paper Reading and Reproduction

- **The entire class collaboratively completes a Benchmark for Time Series Forecasting. There are 3 tasks in this project.**

- **Task 1 (Paper Reading and Presentation):** Each student will be assigned n (TBD) papers. After carefully reading the papers, students will be required to explain the key technologies and insightful ideas used in the papers during class in a 25-minutes presentation **(Week 11 to Week 15).**

Project: Time Series Forecasting Paper Reading and Reproduction

- **Task 2 (Reproduction): We will provide a list of selected papers with open-sourced algorithm code and a list of datasets.** Students are required to **reproduce the algorithms in the papers and perform evaluation on the datasets.** Improvements to the algorithm are encouraged (bonus). Students will give presentation on the results of their reproduction and the problems they encountered in **Week 16.**

- **More details will be provided at the start of the Project.**

Project: Time Series Forecasting Paper Reading and Reproduction

- **Task 3 (Report Writing):** Students are required to summarize based on the papers and the results of the reproduction, and compile the content and the **observation** into a part of a **Report.** The whole class will collaboratively complete the writing of the benchmark report. **(Week 16)**
- We will provide a Latex template (in Overleaf) and the outline.
- Upon approval, the benchmark report may be uploaded to arXiv and submitted to a journal or conference

- **More details (paper list, etc) will be provided at the start of the Project.**

Enjoy the course!

Spring 2017

Spring 2018

Fall 2018

Fall 2019

Fall 2020

Spring 2023