On the Application of AI for Failure Management

AIOps for IT Operations, Networks and DevOps

18th Inter. Conf. on the Design of Reliable Communication Networks (DRCN)
Mar. 28 - 31, 2022, Spain

Prof. Jorge Cardoso
University of Coimbra
E-mail: jorge.cardoso@huawei.com
Chief Architect for AIOps
Munich/Dublin Huawei Research Center

2022.03.30

HUAWEI

Overview
Keynote

On the Application of AI for Failure Management: Problems, Solutions and Algorithms

Artificial Intelligence for IT Operations (AIOps) is a class of software which targets the automation of operational tasks through machine learning technologies. ML algorithms are typically used to support tasks such as anomaly detection, root-causes analysis, failure prevention, failure prediction, and system remediation. AIOps is gaining an increasing interest from the industry due to the exponential growth of IT operations and the complexity of new technology. Modern applications are assembled from hundreds of dependent microservices distributed across many cloud platforms, leading to extremely complex software systems. Studies show that cloud environments are now too complex to be managed solely by humans. This talk discusses various AIOps problems we have addressed over the years and gives a sketch of the solutions and algorithms we have implemented. Interesting problems include hypervisor anomaly detection, root-cause analysis of software service failures using application logs, multi-modal anomaly detection, root-cause analysis using distributed traces, and verification of virtual private cloud networks.

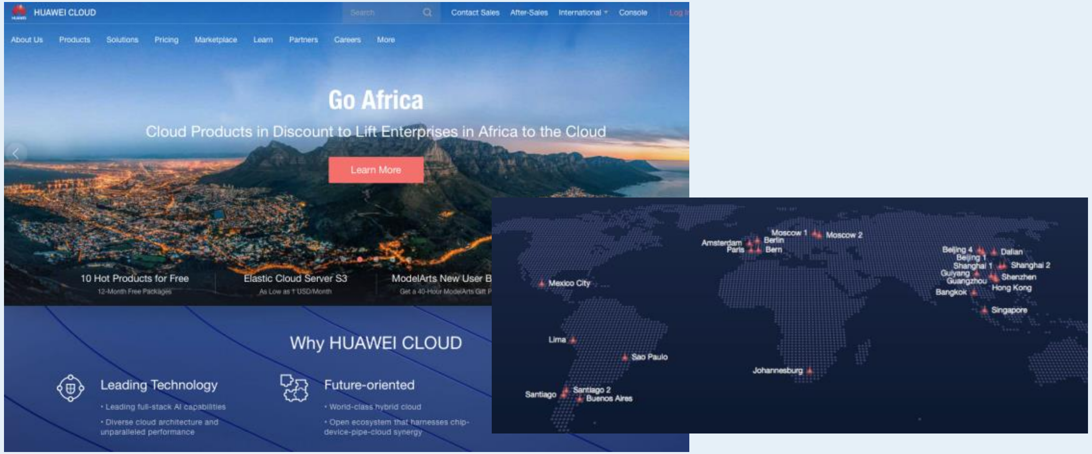
<table>
<tr><th>Time/Day</th><th>Monday 28</th><th>Tuesday 29</th><th>Wednesday 30</th><th>Thursday 31</th></tr>
<tr><td>9:00 - 9:30</td><td rowspan="4">Tutorial 1: Accurate and Practical Network Reliability Evaluation using Binary Decision Diagrams</td><td>Welcome</td><td rowspan="3">Keynote 3: On the Application of AI for Failure Management: Problems, Solutions and Algorithms</td><td></td></tr>
<tr><td>9:30 - 9:45</td><td rowspan="3">Keynote 1: Resiliency for Connected and Autonomous Vehicular Systems</td><td rowspan="7">Workshop 2: 1st International Workshop on Emerging Technologies to Deploy Secure and Reliable Edge Computing Networks, Systems and Services (Go2Edge)</td></tr>
<tr><td>9:45 - 10:00</td></tr>
<tr><td>10:00 - 10:30</td><td>Coffee Break</td></tr>
<tr><td>10:30 - 11:00</td><td>Coffee Break</td><td>Coffee Break</td><td rowspan="4">Session 3: Reliability and Network Robustness</td></tr>
<tr><td>11:00 - 11:30</td><td rowspan="3">Tutorial 2: Ultra-Reliability and Timing for Wireless Connectivity in 5G and Beyond</td><td rowspan="3">Session 1: Design, modelling and evaluation</td></tr>
<tr><td>11:30 - 12:00</td></tr>
<tr><td>12:00 - 12:30</td></tr>
<tr><td>12:30 - 13:00</td><td rowspan="2">Lunch</td><td rowspan="2">Lunch</td><td rowspan="2">Lunch</td><td rowspan="2">Lunch</td></tr>
<tr><td>13:00 - 13:30</td></tr>
<tr><td>13:30 - 13:45</td><td rowspan="8">Workshop 1: 1st International Workshop on Key challenges in global cybersecurity: Efforts and trends in EU (KCYEU)</td><td rowspan="2">Invited talk</td><td rowspan="4">Industrial Panel: Systems reliability: A challenge or a nightmare?</td><td rowspan="5">WIE: Preparing women for early career in science and engineering</td></tr>
<tr><td>13:45 - 14:00</td></tr>
<tr><td>14:00 - 14:30</td><td rowspan="2">Keynote 2: The Quantum Internet: Recent Advances and Challenges</td></tr>
<tr><td>14:30 - 15:00</td></tr>
<tr><td>15:00 - 15:30</td><td>Coffee break</td><td>Coffee break</td></tr>
<tr><td>15:30 - 16:00</td><td rowspan="3">Session 2: Protection and Resilience</td><td rowspan="4">Tutorial 3: Explainable Artificial Intelligence for Trusted Failure Management in Communication Networks</td><td rowspan="3">Tutorial 4: Aerospace Network Virtualization</td></tr>
<tr><td>16:00 - 16:30</td></tr>
<tr><td>16:30 - 17:00</td></tr>
<tr><td>17:00 - 17:30</td><td></td><td></td><td>Closing</td></tr>
<tr><td>17:30 - 18:00</td><td></td><td></td><td></td><td></td></tr>
</table>

HUAWEI

HUAWEI CLOUD
Site Reliability Engineering (SRE)

Reliability is an important feature of HAUWEI CLOUD. SRE is responsible for it …

HUAWEI CLOUD



… a few numbers…

- worldwide, Huawei Cloud has 45 availability zones across 23 regions (June 2019)
- more than 180 cloud services and 180 solutions for a wide range of sectors
- Costumers include European Organization for Nuclear Research (CERN), PSA Group, Shenzhen Airport, Port of Tianjin, …

SRE

- **Automation**
- Setup quantified **SLO**
- Run **Err Budget**
- Self-constraint, balanced Dev/Ops
- Fact oriented
- Four **golden signals**
- Altering from **time-series**

- Eliminating **Toil**
- Implement monitoring and develop handling processes
- Diagnosis, analysis, and detailed data
- Define contextual, **customer-focused SLOs**
- 70% of outages due to changes in a live system

- 50% software engineer / 50% system engineer
- Load slows down systems
- Elite forces, focusing on **high-ROI** work

- Training through brain games
- Perform regular **drills** in the environment

- Regular **review meetings**
- Humans add latency

- Addressing **cascading failures**
- Take turns to monitor and document solutions
- Distributed **consensus** for reliability

- Guarantee that there are few problems
- It's either a problem or a **quick recovery**
- The recovery tool is designed in the early stage of the process

- Continuous review and optimization

HUAWEI

Worldwide Trends
Cloud, transformation, edge, scale and complexity

Distributed Cloud

Constantly changing infrastructure that is heavily virtualized

O&M challenge will not be about replication, but about:
1. Automation
2. Integration
3. Reuse
4. Abstraction
5. Upgrade

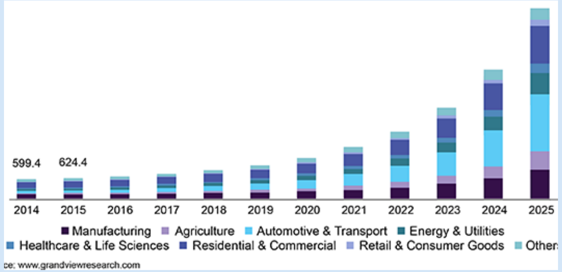
HC Public Cloud – HC Stack Online -- Edge cloud/CDN

Edge 2019	Edge 2025
Static processes	Adaptive processes
Hierarchic architectures	Fog/Mesh architectures
Static network topology	Dynamic network topology
Edge and cloud	Distributed cloud to the edge

Trend: 5 big clouds (GAAVI), 100+ industry clouds, 500+ regions, 5000+ edge sites. The average business runs 38% of workloads in public cloud and 41% in private cloud

Digital Transformation

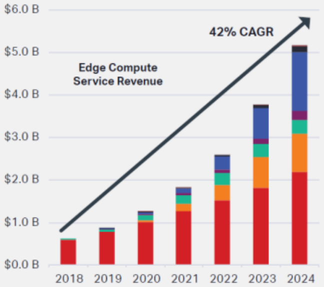
As digital transformation initiatives start to be implemented, IT infrastructure has become exponentially more complex.
1. Real-time monitoring
2. Prediction optimization
3. Control and decision making



Trend: **digital transformation** initiative is expected to growth **20%**/year until 2025. **Intelligent monitoring** market is expected to growth b/year until 2025

Cloud Edge Applications

Application	Description
Video/CDN	Content needed to be cached at the local edge to improve performance
Cloud Gaming	Replace gaming consoles by streaming games from low latency cloud servers at the Cloud Edge
Industrial IoT/ Automation	Edge computing reduces latency for mission critical remote applications
Smart Venues	Processing needed at the venues to reduce response times
Automotive	Infotainment, Autonomous Driving
AR /VR	Low latency to reduce motion sickness



Source: Mobile Experts, "Edge Computing for Enterprises 2019", July 2019.

Trend: Video/CDN, Cloud Gaming, Industrial IoT, Automotive, AR/VR

Ultra-scale and Complexity

Overwhelming number of alarms and monitoring data, makes it impossible to know where to focus during incident resolution.

Not only monitoring tools are important, the velocity of code deployments also becomes key
- Automation of 10k deployments/year
- >50 monitoring tools
- Trillions metrics/day
- Service availability?

	Stone Age	Last Century	Last Decade	This Decade	Today	Tomorrow
Technology Trends	**Mainframe**	**Client Server**	**Distributed**	**Virtualization**	**Cloud**	**Digital Business**
Server Count	1	10s	100s	1,000s	10,000s	100,000s
Deployments/Year	1	2	10s	100s	1,000s	10,000s
Monitoring Tools	1	3	5	10	25	50+
Events/Metrics/Day	**100s**	**1000s**	**100,000s**	**Millions**	**Billions**	**Trillions**
Organizational Silos	1	10	15	25	50	100
Humans Ability to Cope	Yep	Yep	Kind Of	Not Really	Nope	HELP!
Service Availability	100%	99.999%	99.99%	99.9%	99%	?

Trend: Digital Transformation increases the number of **managed servers 10x, 10k deployments/year, >50 monitoring tools, trillions metrics**

R&D Direction
AI-driven autonomous systems

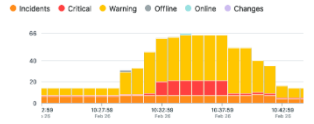
Business Driver: high reliability, high automation, low cost of IT operations

Objective: Use AI/ML to transform the cloud, IT operations and infrastructure by processing massive amounts of data to trigger **automated actions 24/7**, with **higher reliability**, **higher operational efficiency** and **cost savings**

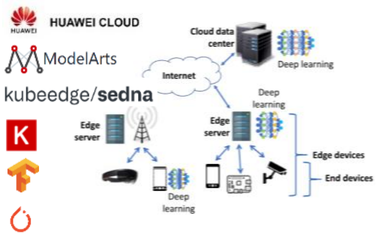
Research and Development

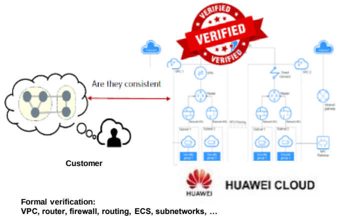
Research Fields	AI for IT Operations	Edge AI	AI for Network	AI for DevOps
Methods	Anomaly detection	Intelligent Container Tracking	Formal Verification	Log Recommendation
	Root-cause analysis	Computing Models	SmartNICs & Troubleshooting	Code Analysis
	Secure Operations	Software Framework	P4 Network Programming	Continuous Verification
	Failure Prediction	Federated Learning	Intent-Based Networks	Structured Logging

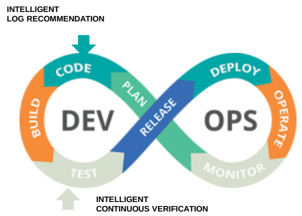
Scenarios



jupyter JAEGER OPENTSDB Spark







Fundamental Research: AIOps, DataOps, MLOps, federated learning, deep learning, formal verification methods

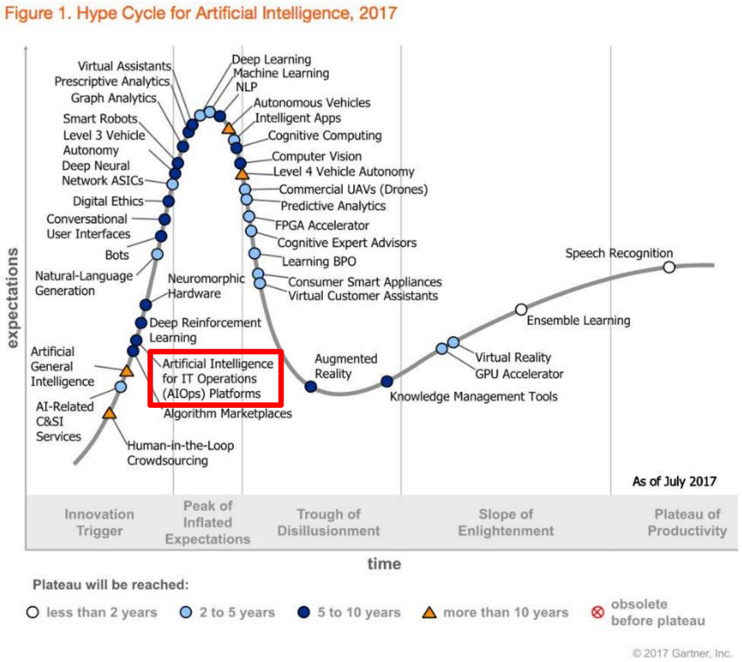
HUAWEI

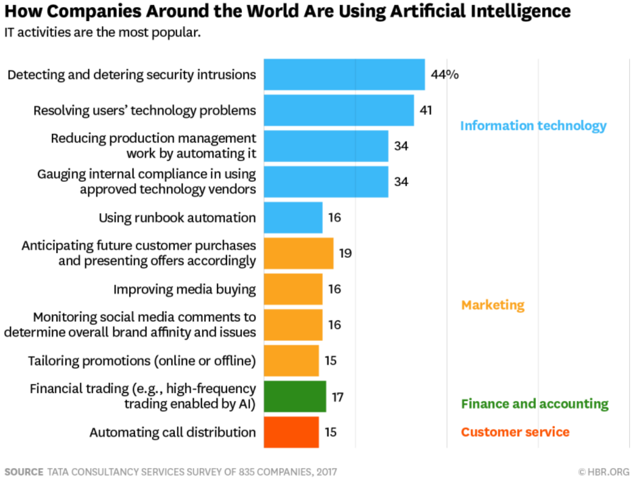
AI for IT Operations (AIOps)
Bringing AI to O&M

SRE / O&M Activities
- **System monitoring** and 24x7 technical support, Tier 1-3 support
- **Troubleshooting and resolution of operational issues**
- Backup, restoration, archival services
- Update, distribution and release of software
- Change, capacity, and configuration management
- …

“We began applying machine learning two years ago **(2016)** to **operate our data centers more efficiently**… over the past few months, DeepMind researchers began working with Google’s data center team to significantly improve the system’s utility. Using neural networks trained on different operating scenarios and parameters, we created a more efficient and adaptive framework to understand data center dynamics and optimize efficiency.” *Eric Schmidt, Dec. 2018*

“Virtyt Koshi, the EMEA general manager for virtualization vendor Mavenir, reckons Google is able to run all of its data centers in the Asia-Pacific with only about 30 people, and that a telco would typically have about 3,000 employees to manage infrastructure across the same area."





Harvard Business Review

Early indicators
Moogsoft AIOps, Amazon EC2 Predictive Scaling, Azure VM resiliency, Amazon S3 Intelligent Tiering

38.4% of organizations take more than 30 minutes to resolve IT incidents that impact consumer-facing services (PagerDuty)

https://www.lightreading.com/automation/google-has-intent-to-cut-humans-out-of-network/d/d-id/742158
https://www.lightreading.com/automation/automation-is-about-job-cuts---lr-poll/d/d-id/741989
https://www.lightreading.com/automation/the-zero-person-network-operations-center-is-here-(in-finland)/d/d-id/741695

Overview of AIOps Research
1990-2020

Results

- Majority of research (670 papers, 62.1%) are associated with failure management (FM)
  - Online failure prediction (26.4%)
  - Failure detection (33.7%)
  - Root cause analysis (26.7%)
- Most common problems in FM
  - Software defect prediction, system failure prediction, anomaly detection, fault localization and root cause diagnosis
- Failure detection has gained particular traction in recent years (71 publications for the 2018-2019 period)
- Root cause analysis (39) and online failure prediction (34)
- Failure prevention and remediation are the areas with least research

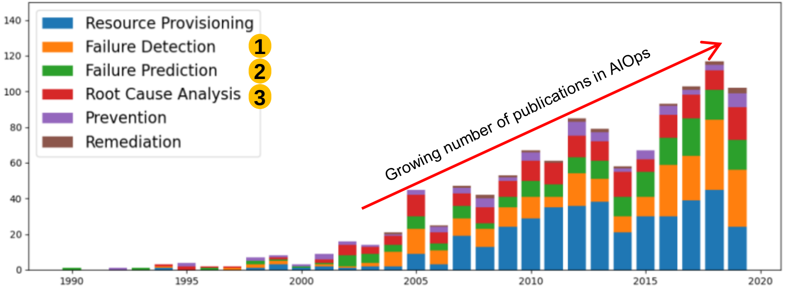


Fig. 4: Published papers in AIOps by year and categories from the described taxonomy.

Table 3: Selection of result papers grouped by data sources, targets and (sub)categories.

Ref.	Data Sources									Targets					Cat.
	Source Code	Testing Resources	System Metrics	KPIs/SLO data	Network Traffic	Topology	Incident Reports	Event Logs	Execution Traces	Source Code	Application	Hardware	Network	Datacenter	
[27]	•									•					1.1
[32]	•	•								•					1.2
[16]			•								•			•	1.3
[41]			•	•							•			•	1.3
[29]											•			•	1.4
[47]			•									•			2.1
[14]			•									•		•	2.1
[12]			•					•				•			2.1
[46]								•				•	•		2.1
[8]		•									•			•	2.2
[11]		•	•								•				2.2
[17]		•	•								•			•	2.2
[35]		•			•						•			•	2.2
[24]								•			•			•	2.2
[37]								•			•	•	•	•	2.2
[45]								•			•				2.2
[43]	•							•						•	3.1
[42]				•							•				3.1
[40]			•	•							•			•	3.1
[21]					•	•							•		3.1
[22]					•	•					•		•		3.1
[15]								•			•				3.1
[10]								•	•		•			•	3.1
[6]									•		•			•	3.1
[28]									•		•				3.1
[30]					•								•	•	3.2
[49]	•							•						•	3.3
[1]	•	•								•	•				4.1
[33]			•			•						•		•	4.1
[5]					•							•	•		4.1
[44]	•							•			•				4.2
[4]	•		•	•							•				4.2
[19]					•	•							•		4.2
[9]									•		•			•	4.2
[36]		•				•					•				4.3
[7]			•	•							•			•	4.3
[26]								•			•			•	4.3
[2]									•		•			•	4.3
[39]							•				•			•	5.1
[48]							•				•			•	5.2
[25]							•	•			•			•	5.2
[38]			•	•							•			•	5.3

(Sub)Category Legend					
1.1	Software Defect Prediction	2.2	System Failure Prediction	4.2	Root Cause Diagnosis
1.2	Fault Injection	3.1	Anomaly Detection	4.3	RCA - Others
1.3	Software Rejuvenation	3.2	Internet Traffic Classification	5.1	Incident Triage
1.4	Checkpointing	3.3	Log Enhancement	5.2	Solution Recommendation
2.1	Hardware Failure Prediction	4.1	Fault Localization	5.3	Recovery

A Systematic Mapping Study in AIOps. Notaro, P.; Cardoso, J. and Gerndt, M. In AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations, Springer, 2020.

A Survey of AIOps Methods for Failure Management. Notaro, P.; Cardoso, J. and Gerndt, M. In ACM Transactions on Intelligent Systems and Technology, 2021.

HUAWEI

AIOps Fields
Troubleshooting

SRE / O&M Activities

- System monitoring and 24x7/Tier 1-3 technical support
- **Troubleshooting and resolution of operational issues**
- Backup, restoration, archival services
- Update, distribution and release of software
- Change, capacity, and configuration management
- …

Complex System

- OBS. Object Storage Service
- EVS. Elastic Volume Service (block storage)
- VPC. Virtual Private (private virtual networks)
- ECS. Elastic Cloud Server (scalable computing)

Troubleshooting tasks

- **Anomaly detection.** Determine what constitutes normal system behavior, and then to discern departures from that normal system behavior
- **Fault diagnosis (root cause analysis).** Identify links of dependency that represent causal relationships to discover the true source of an anomaly
- **Fault Prediction.** Use of historical or streaming data to predict incidents with varying degrees of probability
- **Fault recovery.** Explore how decision support systems can manage and select recovery processes to repair failed systems

Anomaly Detection

- **Response Time Analysis**
  - A service started responding to requests more slowly than normal
  - The change happened suddenly as a consequence of regression in the latest deployment
- **System Load**
  - The demand placed on the system, e.g., REST API requests per second, increase since yesterday
- **Error Analysis**
  - The rate of requests that fail -- either explicitly (HTTP 5xx) or implicitly (HTTP 2xx with wrong content) -- is increasing slowly, but steadily
- **System Saturation**
  - The resources (e.g., memory, I/O, disk) used by key controller services is rapidly reaching threshold levels

Troubleshooting
Monitoring and its data sources

System's Components (e.g., OBS, EVS, VPC, ECS) are monitored and generate various types of data: Logs, Metrics, Traces, Events, Topologies

Logs. Service, microservices, and applications generate logs, composed of timestamped records with a structure and free-form text, which are stored in system files.

2017-01-18 15:54:00.467 32552 ERROR oslo_messaging.rpc.server [req-c0b38ace - default default] Exception during message handling

Metrics. Examples of metrics include CPU load, memory available, and the response time of a HTTP request.

{"tags": ["mem", "192.196.0.2", "AZ01"], "data": [2483, 2669, 2576, 2560, 2549, 2506, 2480, 2565, 3140, …, 2542, 2636, 2638, 2538, 2521, 2614, 2514, 2574, 2519]}

Traces. Traces records the workflow and tasks executed in response to, e.g., an HTTP request.

{"traceId": "72c53", "name": "get", "timestamp": 1529029301238, "id": "df332", "duration": 124957, "annotations": [{"key": "http.status_code", "value": "200"}, {"key": "http.url", "value": "https://v2/e5/servers/detail?limit=200"}, {"key": "protocol", "value": "HTTP"}, "endpoint": {"serviceName": "hss", "ipv4": "126.75.191.253"}]

Events. Major milestones which occur within a data center can be exposed as events. Examples include alarms, service upgrades, and software releases.

{"id": "dns_address_match", "timestamp": 1529029301238, …}
{"id": "ping_packet_loss", "timestamp": 152902933452, …}
{"id": "tcp_connection_time", "timestamp": 15290294516578, …}
{"id": "cpu_usage_average ", "timestamp": 1529023098976, …}

HUAWEI

Our Contribution to AIOps Research
2019-2022

Field	Layers	Tasks	Publication
General AIOps	Service Hypervisor Middleware OS Hardware Network	Anomaly Detection Root-cause Analysis Failure Prediction Fault Recovery	• **A Survey of AIOps Methods for Failure Management**. Notaro, P.; Cardoso, J. and Gerndt, M. In ACM Transactions on Intelligent Systems and Technology, 2021. • **A Systematic Mapping Study in AIOps**. Notaro, P.; Cardoso, J. and Gerndt, M. In AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations, Springer, 2020. • **Artificial Intelligence for IT Operations (AIOPS) Workshop White Paper**. Bogatinovski, J.; Nedelkoski, S.; Acker, A.; Schmidt, F.; Wittkopp, T.; Becker, S.; Cardoso, J. and Kao, O. In AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations, Springer, 2020.
Log Analysis	All		• **QuLog: Data-Driven Approach for Log Instruction Quality Assessment**. Bogatinovski, J.; Nedelkoski, S.; Acker, A.; Cardoso, J. and Kao, O. In 30th IEEE/ACM International Conference on Program Comprehension, 2022. • **Self-Supervised Log Parsing**. Nedelkoski, S.; Bogatinovski, J.; Acker, A.; Cardoso, J. and Kao, O. In European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 14-18 September, 2020, Belgium, 2020. • **Self-Attentive Classification-Based Anomaly Detection in Unstructured Logs.** Nedelkoski, S.; Bogatinovski, J.; Acker, A.; Cardoso, J. and Kao, O. In 20th IEEE International Conference on Data Mining (ICDM), Italy, 2020 • **Efficient Failure Diagnosis of OpenStack Using Tempest**. Bhatia, A.; Gerndt, M. and Cardoso, J. In IEEE Internet Computing, Vol. 22 (6): 61-70, 2018.
Trace Analysis	Service Middleware		• **Automated Analysis of Distributed Tracing: Challenges and Research Directions**. Bento, A.; Correia, J.; Filipe, R.; Araujo, F. and Cardoso, J. In Journal of Grid Computing, Vol. 19 (9), 2021. • **Self-Supervised Anomaly Detection from Distributed Traces**. Bogatinovski, J.; Nedelkoski, S.; Cardoso, J. and Kao, O. In IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC), 2020 • **Anomaly Detection and Classification using Distributed Tracing and Deep Learning**. Nedelkoski, S.; Cardoso, J. and Kao, O. In 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), 2019. • **Anomaly Detection from System Tracing Data using Multimodal Deep Learning**. Nedelkoski, S.; Cardoso, J. and Kao, O. In IEEE 12th International Conference on Cloud Computing (CLOUD), 2019.
Metric Analysis	All		• **IAD: Indirect Anomalous VMMs Detection in the Cloud-based Environment**. Jindal, A.; Shakhat, I.; Cardoso, J.; Gerndt, M. and Podolskiy, V. In AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations, Springer, 2021. • **Online Memory Leak Detection in the Cloud-based Infrastructures**. Jindal, A.; Staab, P.; Cardoso, J.; Gerndt, M. and Podolskiy, V. In AIOPS 2020 International Workshop on Artificial Intelligence for IT Operations, Springer, 2020.
Multi Source	All		• **Multi-source Distributed System Data for AI-Powered Analytics**. Nedelkoski, S.; Bogatinovski, J.; Mandapati, A. K.; Becker, S.; Cardoso, J. and Kao, O. In Service-Oriented and Cloud Computing (ESOCC 2020), 28-30 September, 2020, Crete, pages 161-176, 2020.

HUAWEI

Change Management
Intelligent Continuous Verification

BACKGROUND & MOTIVATION

Change Processes Cause Failures

Fig. Causes of failures [1]
- ***Upgrades*: 16%**
- *Bugs*: 15%
- *Config*: 10%
- …

Google SRE found +-70% outages are due to changes [2]

§	Root cause	#Sv	Cnt	%	Cnt '09-'15
	Unknown	29	355	-	M.M.M.M.M.M.M
5.1	Upgrade	18	54	16	7.4.M.5.M.4.7
5.2	Network	21	52	15	4.4.6.8.M.8.5
5.3	Bugs	18	51	15	M.4.9.8.9.9.2
5.4	Config	19	34	10	2.2.7.2.5.M.4
5.5	Load	18	31	9	2.5.5.5.4.8.2
5.6	Cross	14	28	8	-.2.4.M.5.3.4
5.7	Power	11	21	6	5.4.3.5.3.1.-
5.8	Security	9	17	5	7.-.2.1.3.4.-
5.9	Human	11	14	4	-.1.4.4.2.1.2
5.10	Storage	4	13	4	2.-.-.3.5.3.-
5.11	Server	6	11	3	-.3.-.2.2.4.-
5.12	NatDis	5	9	3	1.1.3.2.1.1.-
5.11	Hardware	4	5	1	1.-.-.3.1.-.-

Problem
- Many incidents are caused by service upgrades
- Manual verification of changes is expensive

INNOVATION

Intelligent Continuous Verification

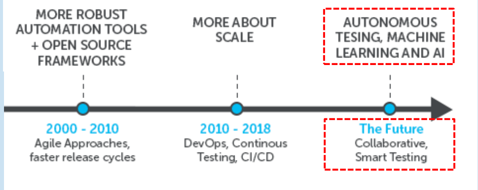


Fig. Verification, test, QA trends [3]

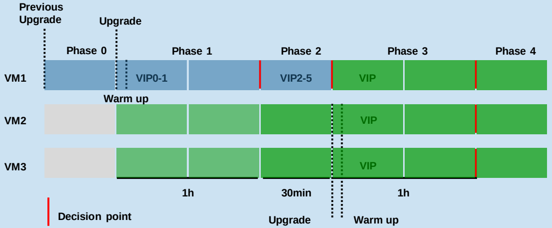
- Compare logs using ML approaches to detect changes in service upgrades or service reconfiguration
- Reason about metric and log comparisons to judge the correctness of service upgrades

DESCRIPTION

MAIN ACHIEVEMENT

Automated Service Change Verification
- Automatic validation of canary phases/gates during service deployments
- Rollback invalid service deployments to avoid failures in production

HOW IT WORKS
1. Collect service logs from release *n*-1. Divide logs into 4 phases. Train a ML model for each phase
2. Use a technique/algorithm such as NuLog [3], 2KDiff [4] or Drain [5] to evaluate the difference of two logs



3. Release *n* of service. Collect service log for phase *p* in [1, 2, …]. Use ML model of phase *p* to check validity of the service log *p*

ASSUMPTIONS & LIMITATIONS
- Only logs are used (traces and metrics are not analyzed)
- Commits involving a high number of modifications causes false positives

TRL 9: *Full operational system. Actual application of the technology in its final form and under real operating conditions*

ANTICIPATED IMPACT

Automated Change Management

Evolution. Analyze different versions of a system to highlight bugs or new/removed functionality.

Testing/Deployment. Differences of systems deployed in different environments, e.g., pre-production vs. production.

Malware Analysis. Differences between original system and a suspected infected one.

Fig. Verification results are pushed to Quality Gates after each service release (PoC)

Self-Supervised Log Parsing

Sasho Nedelkoski[1,3], Jasmin Bogatinovski[1,3], Alexander Acker[1], Jorge Cardoso[2], and Odej Kao[1]

[1] Distributed Systems, TU Berlin, Berlin, Germany
nedelkoski, jasmin.bogatinovski, alexander.acker, odej.kao@tu-berlin.de
[2] Department of Informatics Engineering/CISUC, University of Coimbra, Portugal
jcardoso@dei.uc.pt
[3] Equal contribution

Abstract. Logs are extensively used during the development and maintenance of software systems. They collect runtime events and allow tracking of code execution, which enables a variety of critical tasks such as troubleshooting and fault detection. However, large-scale software systems generate massive volumes of semi-structured log records, posing a major challenge for automated analysis. Parsing semi-structured records with free-form text log messages into structured templates is the first and crucial step that enables further analysis. Existing approaches rely on log-specific heuristics or manual rule extraction. These are often specialized in parsing certain log types, and thus, limit performance scores and generalization. We propose a novel parsing technique called NuLog that utilizes

Self-Supervised Log Parsing. Nedelkoski, S.; Bogatinovski, J.; Acker, A.; Cardoso, J. and Kao, O. In European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2020.

HUAWEI

[1] Why Does the Cloud Stop Computing? Lessons from Hundreds of Service Outages
[2] https://landing.google.com/sre/sre-book/chapters/introduction/
[4] Using finite-state models for log differencing. H. Amar, L. Bao, N. Busany, D. Lo, S. Maoz. ESEC/FSE 2018
[5] Drain: An online log parsing approach with fixed depth tree. He, P., Zhu, J., Zheng, Z., Lyu, M.R.,ICWS, 2017

Anomaly Detection
Detecting Faulty Hypervisors

BACKGROUND & MOTIVATION

Virtualization failures affect VMs but cannot be observed directly

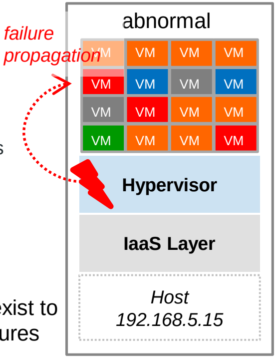


Fig. VMs exhibit problems when the hypervisor has technical issues

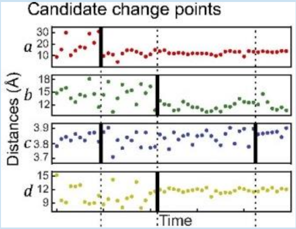
Problem

- No effective solution exist to detect hypervisors failures

INNOVATION

Indirect approach to detect hypervisor failures by monitoring VMs

Fig. Several time-series generated by several VMs running in the same hypervisor



- **Insight**. When an hypervisor is malfunctioning, resource saturation of VMs suddenly changes, within a window w

DESCRIPTION

APPROACH

Quorum change-point detection

- Analyzes individual time-series, and uses change points and voting to decide whether there is an hypervisor malfunction
- Key results: **F1 72%** (2 VMs); **80+%** (3+ VMs)

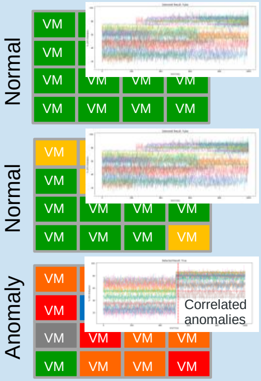
HOW IT WORKS

Method 1 (Change Points)
1. Treat time-series as univariate
2. Detect change points
3. Vote to decide global changes

Method 2 (Isolation Forest)
1. Treat time-series as features
2. Detect significant changes

Method 3 (ECP E.Divisive)
1. Treat time-series as multivariate
2. Detect multiple change points

Analyze VM resources to detect correlated anomalies



ASSUMPTIONS & LIMITATIONS

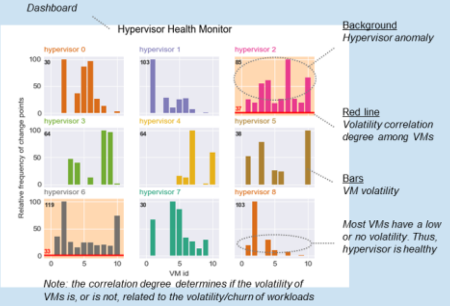
- Datasets used for evaluation were collected from simulation environment, synthetic data generator and public sources

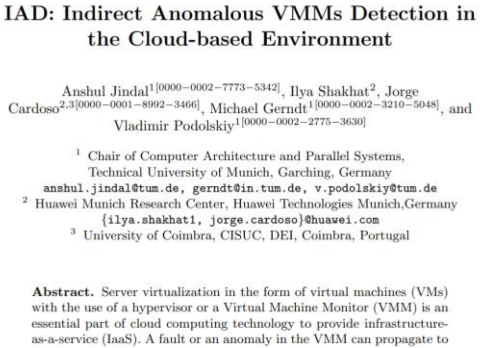
TRL 5. *Basic technological components are integrated with realistic supporting elements so it can be evaluated in testbed environment*

ANTICIPATED IMPACT

Predictive Maintenance

Migrate customers' VMs before hypervisors fail





IAD: Indirect Anomalous VMMs Detection in the Cloud-based Environment

Anshul Jindal[1][0000–0002–7773–5342], Ilya Shakhat[2], Jorge Cardoso[2,3][0000–0001–8992–3466], Michael Gerndt[1][0000–0002–3210–5048], and Vladimir Podolskiy[1][0000–0002–2775–3630]

[1] Chair of Computer Architecture and Parallel Systems, Technical University of Munich, Garching, Germany
anshul.jindal@tum.de, gerndt@in.tum.de, v.podolskiy@tum.de
[2] Huawei Munich Research Center, Huawei Technologies Munich,Germany
{ilya.shakhat1, jorge.cardoso}@huawei.com
[3] University of Coimbra, CISUC, DEI, Coimbra, Portugal

Abstract. Server virtualization in the form of virtual machines (VMs) with the use of a hypervisor or a Virtual Machine Monitor (VMM) is an essential part of cloud computing technology to provide infrastructure-as-a-service (IaaS). A fault or an anomaly in the VMM can propagate to

IAD: Indirect Anomalous VMMs Detection in the Cloud-based Environments, Jindal, A.; Shakhat, I.; Cardoso, J.; Gerndt, M. and Podolskiy, V. International Workshop on AIOPS 2021, Springer, 2021.

HUAWEI

Root Cause Analysis
Application Logs

BACKGROUND & MOTIVATION

Once an anomaly is detected, root cause analysis (RCA) is fundamental to resolve problems

Several forms of RCA exist

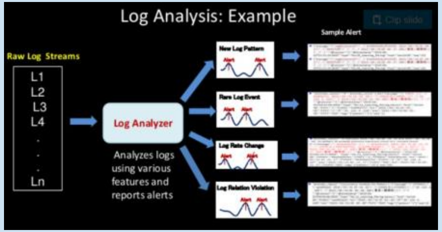
- App logs, metrics, traces, events, etc.

Problem

- Mainly log severity level has been used for AD & RCA
- High number of false positives

INNOVATION

Use a novel, fast algorithms for RCA using log analytics



- **Insight**. Recent research shows it is possible to model the underlying structure of application logs using machine learning [1, 2]

DESCRIPTION

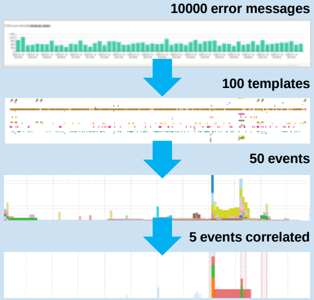
MAIN ACHIEVEMENT

Performs RCA based on application logs

- Anomaly detection in large volume of semi-structured logs
- Correlation between metric anomalies and alarms and logs
- Log summarization that 100x reduces amount of data a human has to process

HOW IT WORKS

1. **Template mining**. Fast log template reconstruction using Drain algorithm
2. **Natural Language Processing**. Language-aware log parsing and keyword extraction using NLP approaches (www.spacy.io)
3. **Dynamic Grouping**. Time-series classification using Poisson model Grouping using Pearson correlation coefficient Distance-aware correlation



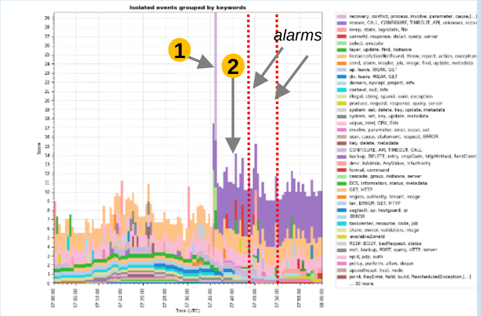
ASSUMPTIONS & LIMITATIONS

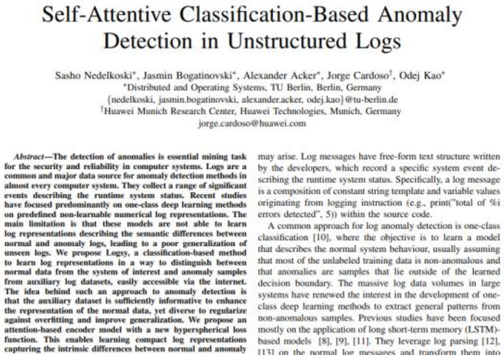
- On-demand processing requires a certain range of logs to learn normality
- Results depend on service logs quality

TRL 5. *Basic technological components are integrated with realistic supporting elements so it can be evaluated in testbed environment*

ANTICIPATED IMPACT

Lower troubleshooting time in 80%





Self-Attentive Classification-Based Anomaly Detection in Unstructured Logs

Sasho Nedelkoski*, Jasmin Bogatinovski*, Alexander Acker*, Jorge Cardoso†, Odej Kao*
*Distributed and Operating Systems, TU Berlin, Berlin, Germany
{nedelkoski, jasmin.bogatinovski, alexander.acker, odej.kao}@tu-berlin.de
†Huawei Munich Research Center, Huawei Technologies, Munich, Germany
jorge.cardoso@huawei.com

Abstract—The detection of anomalies is essential mining task for the security and reliability in computer systems. Logs are a common and major data source for anomaly detection methods in almost every computer system. They collect a range of significant events describing the runtime system status. Recent studies have focused predominantly on one-class deep learning methods on predefined non-learnable numerical log representations. The main limitation is that these models are not able to learn log representations describing the semantic differences between normal and anomaly logs, leading to a poor generalization of unseen logs. We propose Logsy, a classification-based method to learn log representations in a way to distinguish between normal data from the system of interest and anomaly samples from auxiliary log datasets, easily accessible via the internet. The idea behind such an approach to anomaly detection is that the auxiliary dataset is sufficiently informative to enhance the representation of the normal data, yet diverse to regularize against overfitting and improve generalization. We propose an attention-based encoder model with a new hyperspherical loss function. This enables learning compact log representations capturing the intrinsic differences between normal and anomaly

may arise. Log messages have free-form text structure written by the developers, which record a specific system event describing the runtime system status. Specifically, a log message is a composition of constant string template and variable values originating from logging instruction (e.g., print("total of %i errors detected", 5)) within the source code.

A common approach for log anomaly detection is one-class classification [10], where the objective is to learn a model that describes the normal system behaviour, usually assuming that most of the unlabeled training data is non-anomalous and that anomalies are samples that lie outside of the learned decision boundary. The massive log data volumes in large systems have renewed the interest in the development of one-class deep learning methods to extract general patterns from non-anomalous samples. Previous studies have been focused mostly on the application of long short-term memory (LSTM)-based models [8], [9], [11]. They leverage log parsing [12], [13] on the normal log messages and transform them into

Self-Attentive Classification-Based Anomaly Detection in Unstructured Logs. Nedelkoski, S.; Bogatinovski, J.; Acker, A.; Cardoso, J. and Kao, O. In 20th IEEE International Conference on Data Mining (ICDM), 17-20 November, 2020, Italy, 2020.

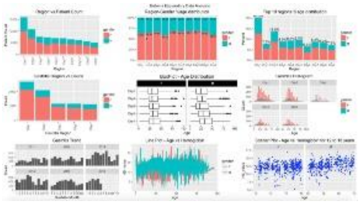
Technische Universität Berlin HUAWEI

Anomaly Detection
Multi-modal Anomaly Detection

BACKGROUND & MOTIVATION

Move from single source, single dimension to multi-source & dimensions

Fig. Metrics, logs, and traces are monitored by separated systems



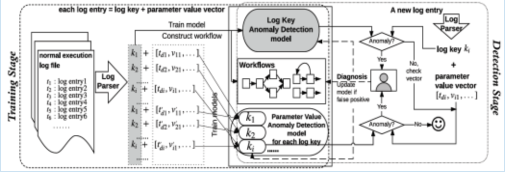
Problem

- High percentage of false positive alarms. Noisy signals requires new AD & RCA robust techniques

INNOVATION

Apply recent Sequence Learning approaches to AIOps

- State of the art results in many applications: image, video, translation and speech recognition to extract long-term dependencies



- e.g., unsupervised anomaly detection in log files (DeepLog)

DESCRIPTION

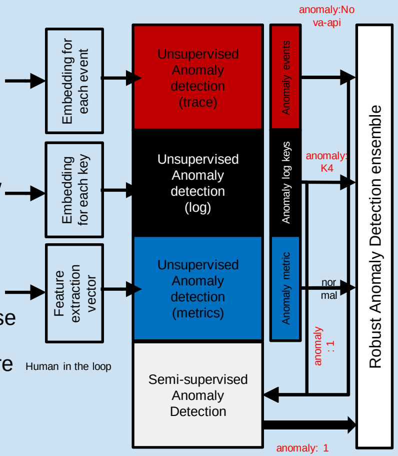
MAIN ACHIEVEMENT

New ensemble AI Algorithms to Detect Anomalies in Multi-source, Multi-dimension data

- Robust anomaly detection ensemble
- Extend approaches such as SkyWalking

HOW IT WORKS

1) Requests generate log events, traces, and metrics
2) Access and Data Transformation to provide an uniform view
3) Robust Anomaly Detection using an ensemble (multi-view)
4) Root Cause Analysis use the neural network and backward anomaly score propagation to identify the root of the problem



ASSUMPTIONS & LIMITATIONS

- Requires a special (not trivial) software development of recurrent neural networks, like LSTM
- Requires access to Topology Services

TRL 3: *Active research and development is initiated. Analytical studies and laboratory studies to validate analytical feasibility of the approach*

ANTICIPATED IMPACT

Lower false positive alarm rate

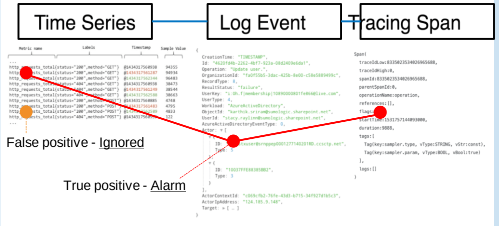
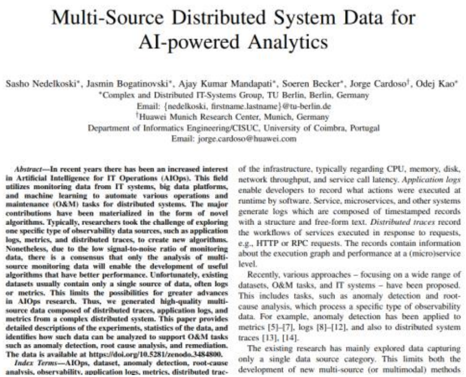


Fig. Multi-source analysis
Correlate single anomalies as a way to improve precision



Multi-Source Distributed System Data for AI-powered Analytics

Sasho Nedelkoski*, Jasmin Bogatinovski*, Ajay Kumar Mandapati*, Soeren Becker*, Jorge Cardoso†, Odej Kao*
*Complex and Distributed IT-Systems Group, TU Berlin, Berlin, Germany
Email: {nedelkoski, firstname.lastname}@tu-berlin.de
†Huawei Munich Research Center, Munich, Germany
Department of Informatics Engineering/CISUC, University of Coimbra, Portugal
Email: jorge.cardoso@huawei.com

***Abstract*—In recent years there has been an increased interest in Artificial Intelligence for IT Operations (AIOps). This field utilizes monitoring data from IT systems, big data platforms, and machine learning to automate various operations and maintenance (O&M) tasks for distributed systems. The major contributions have been materialized in the form of novel algorithms. Typically, researchers took the challenge of exploring one specific type of observability data sources, such as application logs, metrics, and distributed traces, to create new algorithms. Nonetheless, due to the low signal-to-noise ratio of monitoring data, there is a consensus that only the analysis of multi-source monitoring data will enable the development of useful algorithms that have better performance. Unfortunately, existing datasets usually contain only a single source of data, often logs or metrics. This limits the possibilities for greater advances in AIOps research. Thus, we generated high-quality multi-source data composed of distributed traces, application logs, and metrics from a complex distributed system. This paper provides detailed descriptions of the experiments, statistics of the data, and identifies how such data can be analyzed to support O&M tasks such as anomaly detection, root cause analysis, and remediation. The data is available at https://doi.org/10.5281/zenodo.3484800.**

***Index Terms*—AIOps, dataset, anomaly detection, root-cause analysis, observability, application logs, metrics, distributed trac-**

of the infrastructure, typically regarding CPU, memory, disk, network throughput, and service call latency. *Application logs* enable developers to record what actions were executed at runtime by software. Service, microservices, and other systems generate logs which are composed of timestamped records with a structure and free-form text. *Distributed traces* record the workflows of services executed in response to requests, e.g., HTTP or RPC requests. The records contain information about the execution graph and performance at a (micro)service level.

Recently, various approaches – focusing on a wide range of datasets, O&M tasks, and IT systems – have been proposed. This includes tasks, such as anomaly detection and root-cause analysis, which process a specific type of observability data. For example, anomaly detection has been applied to metrics [5]–[7], logs [8]–[12], and also to distributed system traces [13], [14].

The existing research has mainly explored data capturing only a single data source category. This limits both the development of new multi-source (or multimodal) methods

Multi-source Distributed System Data for AI-Powered Analytics. Nedelkoski, S.; Bogatinovski, J.; Mandapati, A. K.; Becker, S.; Cardoso, J. and Kao, O. In Service-Oriented and Cloud Computing (ESOCC), 2020.

Technische Universität Berlin
HUAWEI

Anomaly Detection & Root Cause Analysis
Distributed Traces

BACKGROUND & MOTIVATION

While popular, only visualization tools exist for trace management

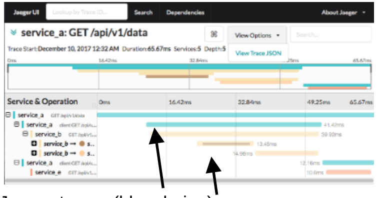


Fig. Jaeger traces (blue, beige)

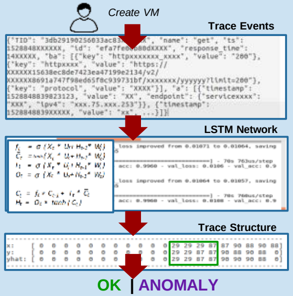
Current limitations

- Tracing tools only provides trace visualization
- Manually finding anomalies in traces is <u>error-prone</u> and <u>not scalable</u>

INNOVATION

Apply recent ML and statistical approaches to process sequential data

- Explore the use of Deep Learning: Long Short Term Memory (LSTM)
- Explore the use of attention networks
- Explore the use of association rules



DESCRIPTION

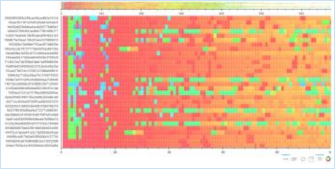
MAIN ACHIEVEMENT

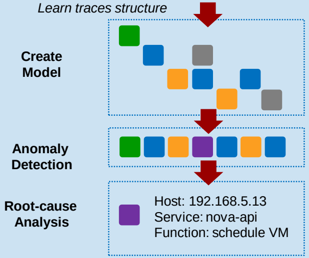
Trace anomaly detection and root-cause analysis using trace structure

- Deep Learning (LSTM, CNN), Machine Learning (Optiks), Sequence Analysis (LCS, Multiple sequence alignment, Needleman-Wunsch), etc.
- Attention networks yielded better results

HOW IT WORKS

1) *Learning*. For each service endpoint, learn the traces' structure it generates
2) *Modeling*. Aggregate all the traces into a behavior model
3) *Anomaly detection*. When a new trace is generate, compare its structure with the behavior model. If it was not seen before, an anomaly exists
4) *Root-cause analysis*. When an anomaly is detected, determine in which span it occurred and identify host, service, function





ASSUMPTIONS & LIMITATIONS

- Microservices are instrumented with tracing capabilities

TRL 4. *Small scale prototype. Basic technological components are integrated to establish that they will work together.*

ANTICIPATED IMPACT

Improve trace-based RCA in 90%

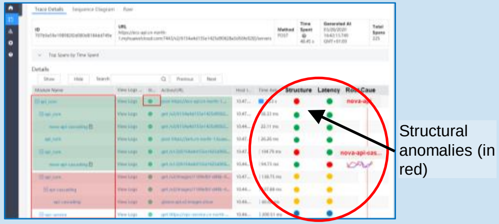
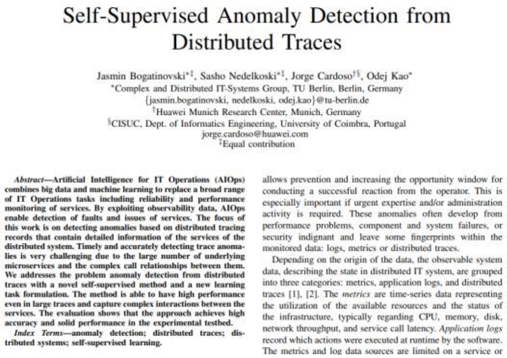


Fig. Trace management & trace analysis
Red circles show structural anomalies



Self-Supervised Anomaly Detection from Distributed Traces

Jasmin Bogatinovski*‡, Sasho Nedelkoski*‡, Jorge Cardoso†§, Odej Kao*
*Complex and Distributed IT-Systems Group, TU Berlin, Berlin, Germany
{jasmin.bogatinovski, nedelkoski, odej.kao}@tu-berlin.de
†Huawei Munich Research Center, Munich, Germany
§CISUC, Dept. of Informatics Engineering, University of Coimbra, Portugal
jorge.cardoso@huawei.com
‡Equal contribution

***Abstract*—Artificial Intelligence for IT Operations (AIOps) combines big data and machine learning to replace a broad range of IT Operations tasks including reliability and performance monitoring of services. By exploiting observability data, AIOps enable detection of faults and issues of services. The focus of this work is on detecting anomalies based on distributed tracing records that contain detailed information of the services of the distributed system. Timely and accurately detecting trace anomalies is very challenging due to the large number of underlying microservices and the complex call relationships between them. We addresses the problem anomaly detection from distributed traces with a novel self-supervised method and a new learning task formulation. The method is able to have high performance even in large traces and capture complex interactions between the services. The evaluation shows that the approach achieves high accuracy and solid performance in the experimental testbed.**

***Index Terms*—anomaly detection; distributed traces; distributed systems; self-supervised learning.**

allows prevention and increasing the opportunity window for conducting a successful reaction from the operator. This is especially important if urgent expertise and/or administration activity is required. These anomalies often develop from performance problems, component and system failures, or security indignant and leave some fingerprints within the monitored data: logs, metrics or distributed traces.

Depending on the origin of the data, the observable system data, describing the state in distributed IT system, are grouped into three categories: metrics, application logs, and distributed traces [1], [2]. The *metrics* are time-series data representing the utilization of the available resources and the status of the infrastructure, typically regarding CPU, memory, disk, network throughput, and service call latency. *Application logs* record which actions were executed at runtime by the software. The metrics and log data sources are limited on a service or

Self-Supervised Anomaly Detection from Distributed Traces. Bogatinovski, J.; Nedelkoski, S.; Cardoso, J. and Kao, O. In IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC), 2020

HUAWEI

Log Recommendation

How can logs be improved?

BACKGROUND & MOTIVATION

Logs are one of the best source of information to troubleshoot systems

log.info("Neutron successfully connected to Nova")
[Subject] [adverb] [verb] [Object]

log.info("Neutron successfully connected")
[Subject] [adverb] [verb] ~~[Object]~~

log.info("Successfully connected")
~~[Subject]~~ [adverb] [verb] ~~[Object]~~

Problem

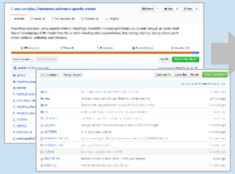
- Lack of log quality results in low efficiency when troubleshoot faults
- Many developers do not comply with log specifications

INNOVATION

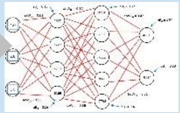
Exploit modern deep learning algorithms to provide log recommendations

GitHub

HUAWEI



Code projects



Deep Learning model

Objectives

- Level Evaluation. Message content must match the level
- Linguistic Evaluation. Correct English expressions
- Log Recommendation. How to improve log records

DESCRIPTION

APPROACH

Multi-head self-attention learning architecture

- Trained on log messages from 20k top ranked GitHub repositories
- Log instruction extraction from source code
- 4 classification scenarios: e.g. INFO-ERROR

HOW IT WORKS

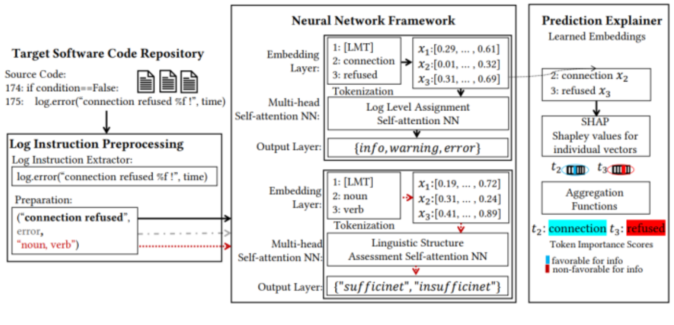


Figure 2: Internal architectural design of QuLog

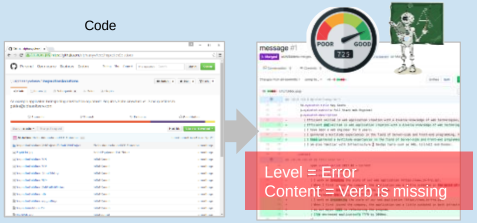
ASSUMPTIONS & LIMITATIONS

- AI model was parametrized for the English language
- It is not clear if a generic AI model can be reused for domain specific applications

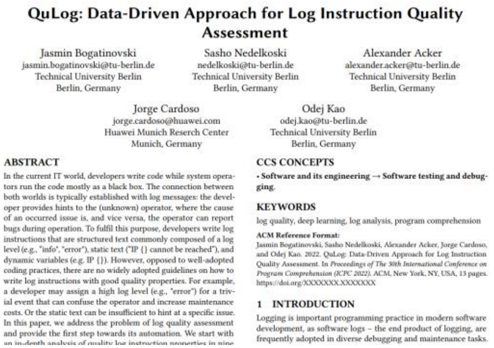
TRL 8. *Technology proven to work in under expected conditions. Include test and evaluation of the system in its intended context*

ANTICIPATED IMPACT

Better logs for better anomaly detection and root-cause analysis



Result. QuLog enables the generation of better log statements which is important for automated solutions for anomaly detection and root-cause analysis based on logs.



QuLog: Data-Driven Approach for Log Instruction Quality Assessment

Jasmin Bogatinovski, jasmin.bogatinovski@tu-berlin.de, Technical University Berlin, Berlin, Germany
Sasho Nedelkoski, nedelkoski@tu-berlin.de, Technical University Berlin, Berlin, Germany
Alexander Acker, alexander.acker@tu-berlin.de, Technical University Berlin, Berlin, Germany
Jorge Cardoso, jorge.cardoso@huawei.com, Huawei Munich Reserch Center, Munich, Germany
Odej Kao, odej.kao@tu-berlin.de, Technical University Berlin, Berlin, Germany

ABSTRACT

In the current IT world, developers write code while system operators run the code mostly as a black box. The connection between both worlds is typically established with log messages: the developer provides hints to the (unknown) operator, where the cause of an occurred issue is, and vice versa, the operator can report bugs during operation. To fulfil this purpose, developers write log instructions that are structured text commonly composed of a log level (e.g., "info", "error"), static text ("IP {} cannot be reached"), and dynamic variables (e.g. IP {}). However, opposed to well-adopted coding practices, there are no widely adopted guidelines on how to write log instructions with good quality properties. For example, a developer may assign a high log level (e.g., "error") for a trivial event that can confuse the operator and increase maintenance costs. Or the static text can be insufficient to hint at a specific issue. In this paper, we address the problem of log quality assessment and provide the first step towards its automation. We start with

CCS CONCEPTS

• **Software and its engineering → Software testing and debugging.**

KEYWORDS

log quality, deep learning, log analysis, program comprehension

ACM Reference Format:
Jasmin Bogatinovski, Sasho Nedelkoski, Alexander Acker, Jorge Cardoso, and Odej Kao. 2022. QuLog: Data-Driven Approach for Log Instruction Quality Assessment. In *Proceedings of The 30th International Conference on Program Comprehension (ICPC 2022)*. ACM, New York, NY, USA, 13 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Logging is important programming practice in modern software development, as software logs – the end product of logging, are frequently adopted in diverse debugging and maintenance tasks.

QuLog: Data-Driven Approach for Log Instruction Quality Assessment.
Bogatinovski, J.; Nedelkoski, S.; Acker, A.; Cardoso, J. and Kao, O. In 30th IEEE/ACM International Conference on Program Comprehension, 2022.

HUAWEI

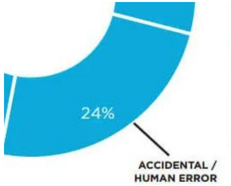
Secure Operations
Operation Risk Evaluation and Prevention

BACKGROUND & MOTIVATION

Gartner estimates that human error is a leading cause of costly IT outages

Fig. (Google) configuration errors are the 2nd major cause of service failures

Human errors are responsible for >24% of outages (Uptime Institute)



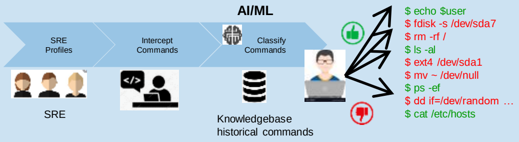
Problem

- Estimating operation risk and preventing service interruptions is difficult due to the large surface and lack of parameter specification of APIs

INNOVATION

Exploit modern deep learning algorithms to support secure operations

- Develop a way to address new security requirements of large-scale public cloud platforms without demanding significant changes to existing deployed



- Explore NER, Word2Vec, Doc2Vec, POS tagging, conditional random fields, recommendation algorithms, & collaborative filtering

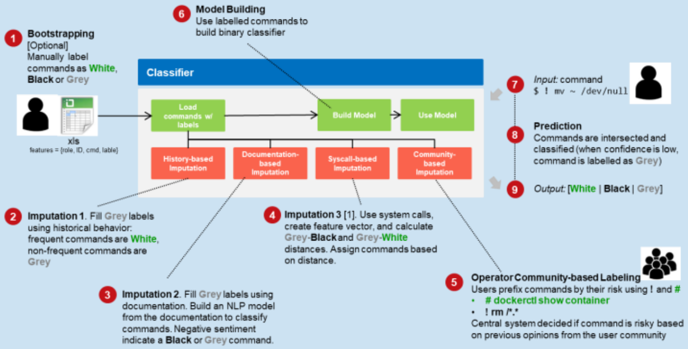
DESCRIPTION

APPROACH

Secure Operations (hybrid approach)

- Rules are matched according to regular expressions. Black, white, gray, and transparent lists are used
- ML/NLP is used for complex analysis (10% cases) for which operators are not able to provide rules

HOW IT WORKS



ASSUMPTIONS & LIMITATIONS

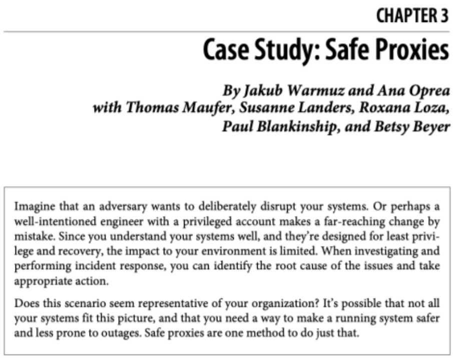
- Protecting several APIs with a large surface is complex (e.g., POSIX, MySQL)
- Rule management becomes a necessity

TRL 8. *Technology proven to work in under expected conditions. Include test and evaluation of the system in its intended context*

ANTICIPATED IMPACT

Higher security and reliability of HUAWEI CLOUD

Worldwide deployment of Secure Operations. (January 2022) the Secure Operations system was already deployed in 8 datacenters



CHAPTER 3

Case Study: Safe Proxies

By Jakub Warmuz and Ana Oprea with Thomas Maufer, Susanne Landers, Roxana Loza, Paul Blankinship, and Betsy Beyer

Imagine that an adversary wants to deliberately disrupt your systems. Or perhaps a well-intentioned engineer with a privileged account makes a far-reaching change by mistake. Since you understand your systems well, and they're designed for least privilege and recovery, the impact to your environment is limited. When investigating and performing incident response, you can identify the root cause of the issues and take appropriate action.

Does this scenario seem representative of your organization? It's possible that not all your systems fit this picture, and that you need a way to make a running system safer and less prone to outages. Safe proxies are one method to do just that.

Building Secure and Reliable Systems, Ana Oprea, Bets, et al. 2020. O'Reilly

HUAWEI

AIOps for Networks
Research on Network Verification

Requirements / Objectives

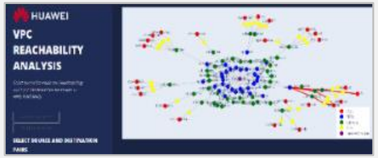
- Manage the increasing **complexity of Virtual Private Cloud networks (VPC)**
- Use **formal techniques** for **static network analysis**
- Use **active probing** to check the reachability of VM hosted by PC
- Analyze **misconfiguration** and **security violations** due to **human errors**

Research Domains

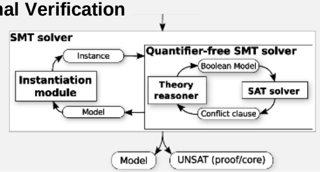
2x directions: Static Network Verification and Dynamic Network Verification

Static Network Verification
Infrastructure, monitoring, UI

User Interface of vReach



Formal Verification



Dynamic Network Verification
Orchestrate distributed AI models

User Interface of pReach

Static versus dynamic verification

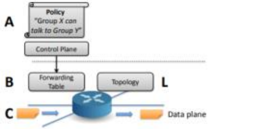


Figure 1: Static versus Dynamic Checking: A policy is compiled to forwarding state, which is then executed by the forwarding plane. Static checking (*e.g.* [14]) confirms that $A = B$. Dynamic checking (e.g. ATPG in this paper) confirms that the topology is meeting liveness properties (L) and that $B = C$.

Technology Trends

Technical maturity Fundamental research — Mass adoption

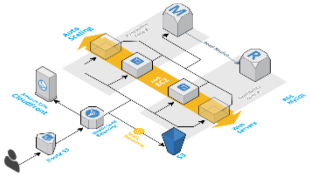
- Cloud networking market size was USD 3.32 billion in 2019 and is projected to reach **USD 14.61 billion by 2027**, a **CAGR of 21.1%**
- Market will grow at a CAGR of 6.36% up to 2026
- **AWS** (2020 [1]), **Google** (2019 [2]), **Azure** (2019 [3]) have rolled out network verification tools
- **Google Network Intelligence Center** [2] and **Microsoft Azure Network Watcher** enable a dynamic verification and connection troubleshooting
- **Network verification startups** are offering new solutions [5-6]

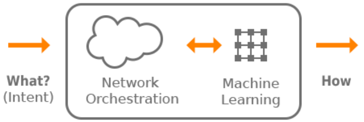
Industry Applicability

Industry applicability Niche — Industry cross-cutting

What it enables companies to do

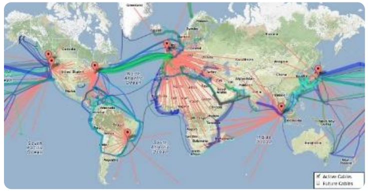
Virtual Private Cloud Networks





Intent-Based Network

Data centers



Cross-regions / AZ network analysis

[1] Backes, John, et al. "Reachability analysis for AWS-based networks." *Inter. Conf. Computer Aided Verification*. 2019.
[2] Network Intelligence Center | Google Cloud https://cloud.google.com/network-intelligence-center/
[3] Network watcher | Azure https://azure.microsoft.com/en-us/services/network-watcher/
[6] Best-in-class network modelling for mission-critical networks | Forward Networks
[7] Yang, Hongkun, et al.. "Real-time verification of network properties using atomic predicates." IEEE/ACM TON. 2015.

HUAWEI

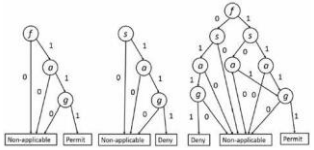
AIOps for Networks

Static and Dynamic Network Verification

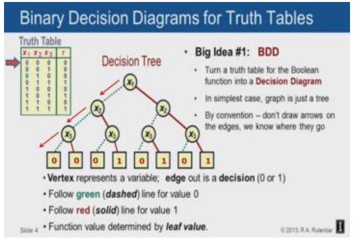
2021

AIOps for Networks Core Technologies

vReach Tech.: Multi-Terminal Binary Decision Diagrams



Multi-Terminal Binary Decision Diagrams (MTBDDs) [1]



Binary Decision Diagrams (BDDs) [2]

- **Network abstraction using packet filtering predicates** based on **Binary Decision Diagrams** (BDDs) [1] and Multi-Terminal BDD [2]

vReach Results

1 Pull VPC network models using Huawei Cloud SDK

Huawei/**cloud-sdk-python**

2 Fast and memory-efficient

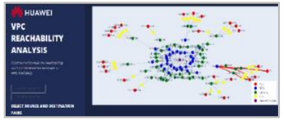
FAST

Reachability verification between 2 nodes **< 40ms**

Peak memory during inference **< 200 MB**

RCA analysis (worst case) of **1k ACL rules / node < 8ms**

3 UI and Integration



vReach's User interface

BatFish integration

pReach Tech. Active probing

.

1 Active Reachability Verification [3]

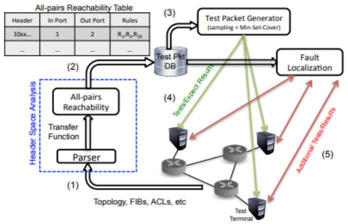


Figure 5: ATPG system block diagram.

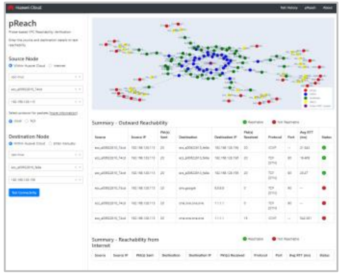
- **Packet trains are used as probes** between source & destination nodes
- Intelligence lying completely in the source node
- Scapy [4] is used to manipulate and send packets

2 Active Network Performance Monitoring

- Probes characterize **network quality in real-time** with metrics like **packet loss** and **latency**

pReach Results

.



pReach's User Interface (with test histories)

- **Dynamic reachability** verification between **VMs in VPCs**
- pReach cross-validates the results of symbolic reachability tool vReach

FAST

Reachability verification is completed **< 30 seconds**

Only 20 packets are used to verify reachability

[1] Yang, Hongkun, et al.. "Real-time verification of network properties using atomic predicates." IEEE/ACM TON. 2015.
[2] Fujita, et al. "Multi-terminal binary decision diagrams: An efficient data structure for matrix representation." Formal methods in system design. 1997.
[3] H. Zeng et. al, "Automatic test packet generation,". International conference on Emerging networking experiments and technologies (CoNEXT 2012)
[4] https://scapy.net/

Thank you.

Bring digital to every person, home and organization for a fully connected, intelligent world.

HUAWEI