NORIA-O

an Ontology for Anomaly Detection and Incident Management in ICT Systems

IETF119 AI4NET side meeting

Lionel Tailhardat, Orange, lionel.tailhardat@orange.com
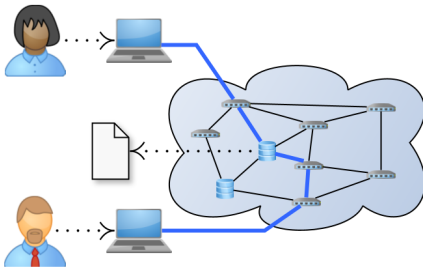Yoan Chabot, Orange, yoan.chabot@orange.com
Raphaël Troncy, EURECOM, raphael.troncy@eurecom.fr

Orange & EURECOM

March 18, 2024

EURECOM
Sophia Antipolis

orange™

Context & motivations: alarm spreading & heterogeneous networks

Scenario Networking / online collaboration	Diagnosis Situation understanding through causal models
Situation Impaired network service	Real world Alarm spreading phenomenon, heterogeneous networks
Observables Alarms and logs from multiple monitoring systems	(multi-technology, multi-vendor)

Scenario Networking / online collaboration
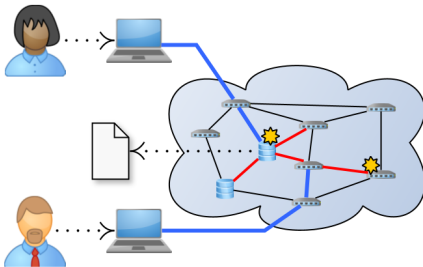
Situation Impaired network service

Observables Alarms and logs from multiple
monitoring systems

Diagnosis Situation understanding through
causal models

Real world Alarm spreading phenomenon,
heterogeneous networks
(multi-technology, multi-vendor)

Context & motivations: alarm spreading & heterogeneous networks

Scenario Networking / online collaboration
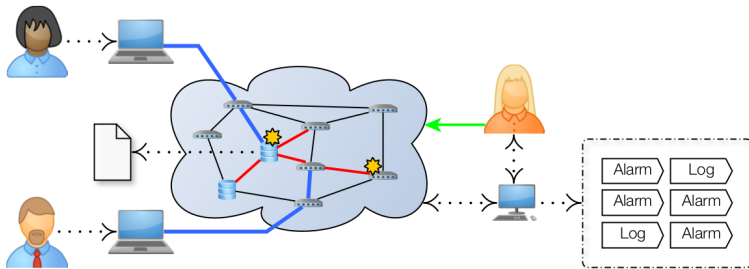
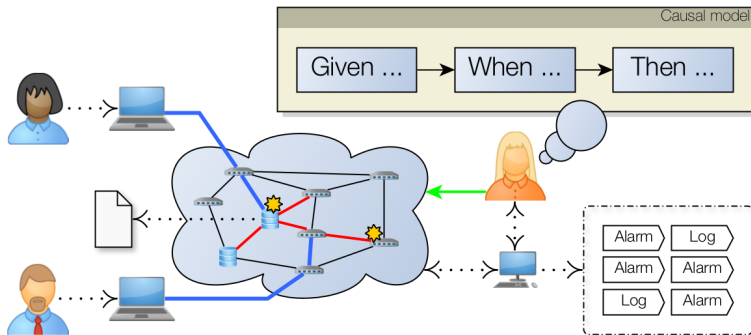**Situation** Impaired network service

Observables Alarms and logs from multiple monitoring systems

Diagnosis Situation understanding through causal models

Real world Alarm spreading phenomenon, heterogeneous networks (multi-technology, multi-vendor)

Context & motivations: alarm spreading & heterogeneous networks

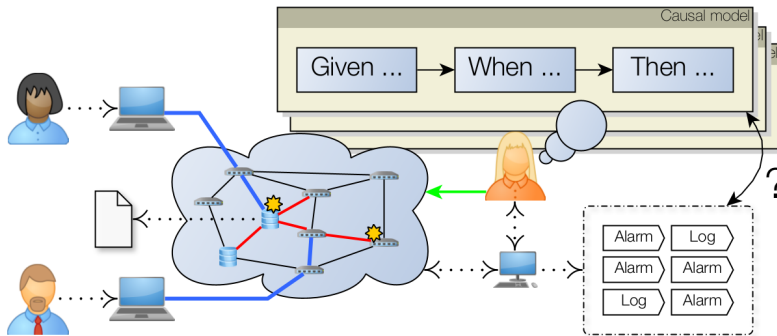| | | Scenario | Networking / online collaboration |
		Situation	Impaired network service
		Observables	Alarms and logs from multiple monitoring systems
		Diagnosis	Situation understanding through causal models
		Real world	Alarm spreading phenomenon, heterogeneous networks (multi-technology, multi-vendor)

Context & motivations: alarm spreading & heterogeneous networks

Scenario Networking / online collaboration	**Diagnosis** Situation understanding through causal models
Situation Impaired network service	**Real world** Alarm spreading phenomenon, heterogeneous networks
Observables Alarms and logs from multiple monitoring systems	(multi-technology, multi-vendor)

Having a comprehensive and integrated view of ICT systems for anomaly detection and decision support?

Challenges

- **Modeling a four-faceted domain of discourse with temporal evolution**
  - Structural
  - Functional
  - Dynamic
  - Procedural

- **Enabling logical & probabilistic reasoning**

- **Interoperability with third-party knowledge bases**
  - Vulnerability databases
  - Geographical information systems
  - Energy management
  - etc.

Approach

- Implementing a data model with Semantic Web technologies and reusing existing models/vocabularies.
- Experts panel interview, concepts and relations analysis, ontology requirements design.

Ren et al. "Towards Competency Question-Driven Ontology Authoring." in The Semantic Web: Trends and Challenges, Springer International Publishing, 2014.
Poveda-Villalón et al. "Linked Open Terms (LOT) Methodology", 2019.

Having a comprehensive and integrated view of ICT systems for anomaly detection and decision support?

Challenges

- Modeling a four-faceted domain of discourse with temporal evolution
  - Structural
  - Functional
  - Dynamic
  - Procedural

- Enabling logical & probabilistic reasoning
- Interoperability with third-party knowledge bases
  - Vulnerability databases
  - Geographical information systems
  - Energy management
  - etc.

Approach

- Implementing a data model with Semantic Web technologies and reusing existing models/vocabularies.
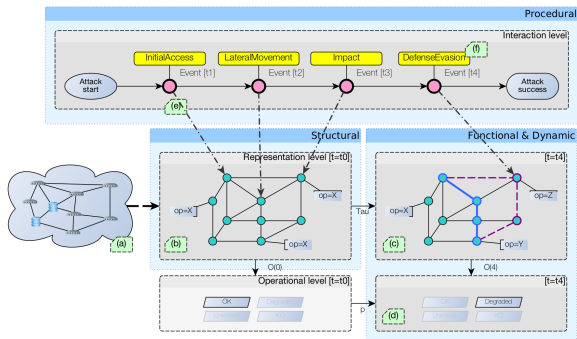- Experts panel interview, concepts and relations analysis, ontology requirements design.

(CQs) Ren et al. "Towards Competency Question-Driven Ontology Authoring." In The Semantic Web: Trends and Challenges, Springer International Publishing, 2014.

(LOT) Poveda-Villalón et al. "Linked Open Terms (LOT) Methodology", 2019.

Overview of the NORIA-O v0.3 data model

NORIA-O implementation: https://w3id.org/noria/ (open source release under BSD-4 license)
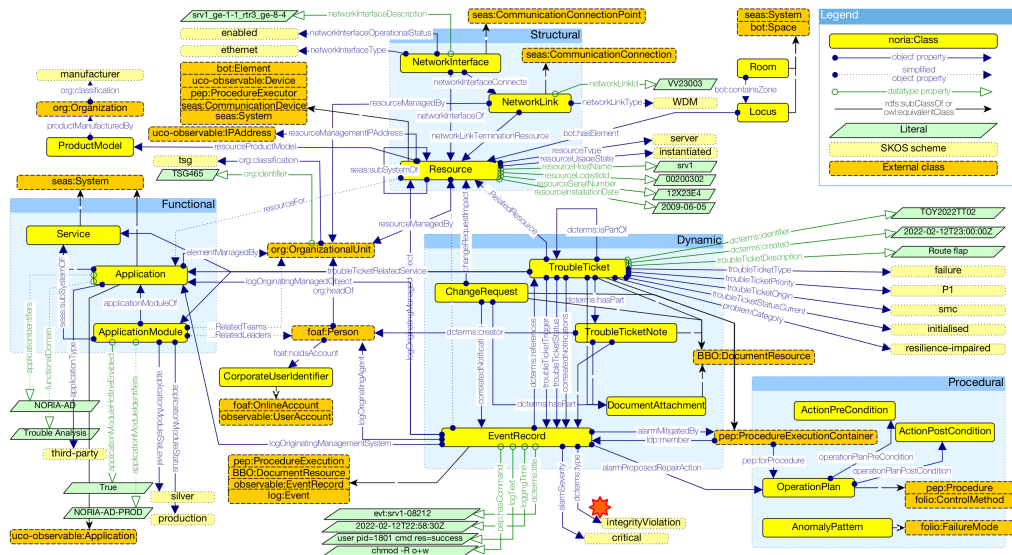
A toy example from the NORIA-O v0.3 project

NORIA-O dataset: https://w3id.org/noria/dataset/
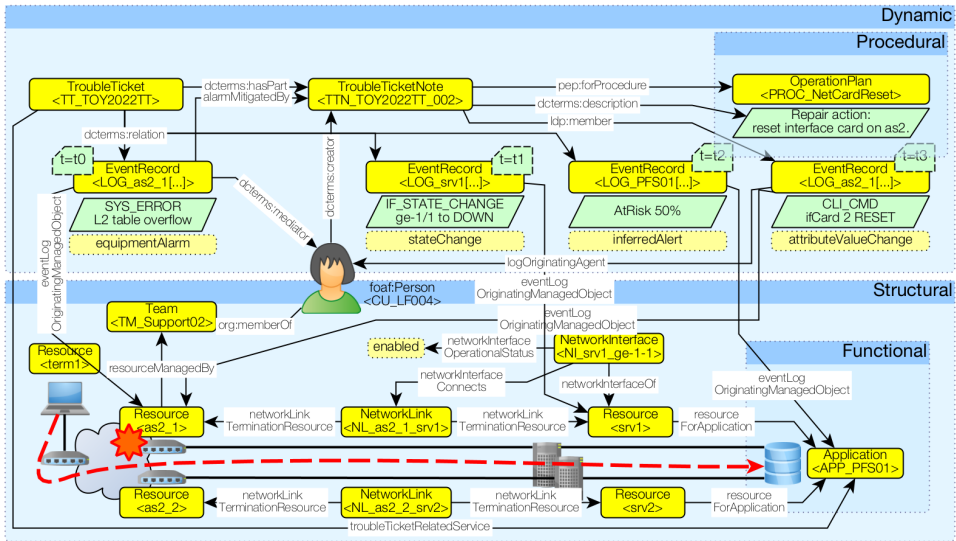
Evaluating NORIA-O with Authoring Tests

Evaluation set 26 Competency Questions (CQs), available at https://w3id.org/noria/cqs/, translated into 25 Authoring Tests (SPARQL queries).

Evaluation results	#CQs	Remarks
OK	16/26	Answered using a single or several simple SPARQL queries and the ontology.
AI	9/26	Require the implementation of more complex AI-based algorithms such as anomaly detection algorithms.
Extension	1/26	Require the introduction of new concepts or relations via an extension of the NORIA-O model.

Examples

OK "Which entity (resource/application/site) is concerned by a given incident?"

AI (1) "What was the root cause of the incident?",
→ the explicit representation of alarms and logs associated with a given incident is not enough and needs to be enhanced with root cause analysis algorithms.

AI (2) "What are the vulnerabilities and the associated risk levels of this infrastructure?",
→ can be answered only by looking for non-desirable network topology shapes or relations to third-party cybersecurity vulnerability entities based on structure and security scanners.

Extension "What is the financial cost of this incident if it occurs?",
→ involves information about the cost of an incident.

Evaluating NORIA-O for anomaly detection and situation understanding

Data integration Knowledge graph-based platform [1]

Model-Based Design Query the graph to retrieve anomalies and their context [2]
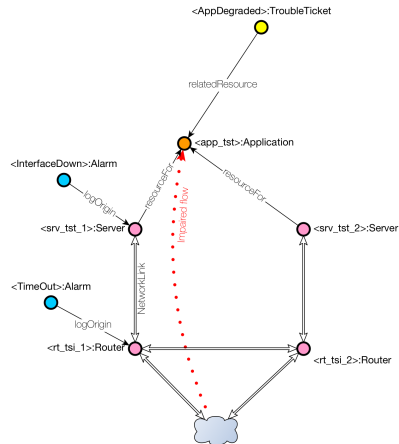
- k out-of n devices with faults
- User with unusual account rights
- Absence of traffic on an interface supposed to be active

Process mining Align a sequence of entities to activity models, then use this relatedness to guide the repair [3]

- (EnergyLoss)->(TimeoutAlert)->(LossOfSignal)
- (LoginFail)->(LoginFail)->(LoginFail)

Statistical Learning Relate entities based on context similarities, then use this relatedness to alert and guide the repair [2]

- The hidden cause of the trouble ticket on server 1 is a "data leak" attack that started on server 2

[1] Tailhardat, et al. 2023. "Designing NORIA: a Knowledge Graph-based Platform for Anomaly Detection and Incident Management in ICT Systems" (ESWC'2023)

[2] Tailhardat, et al. 2023. "Leveraging Knowledge Graphs For Classifying Incident Situations in ICT Systems" (ARES'2023)

[3] Tailhardat, et al. 2024. "Graphameleon: Relational Learning and Anomaly Detection on Web Navigation Traces Captured as Knowledge Graphs" (WWW'2024)

Evaluating NORIA-O for anomaly detection and situation understanding

Data integration Knowledge graph-based platform [1]

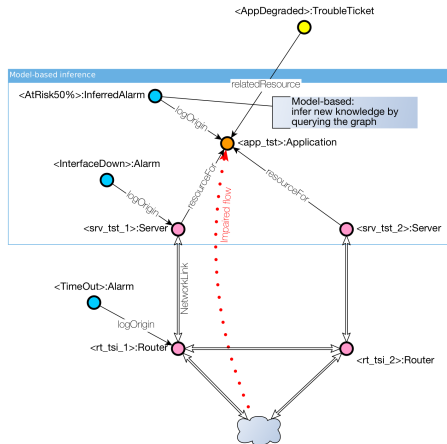**Model-Based Design** Query the graph to retrieve anomalies and their context [2]

- k out-of n devices with faults
- User with unusual account rights
- Absence of traffic on an interface supposed to be active

Process mining Align a sequence of entities to activity models, then use this relatedness to guide the repair [3]

(EnergyLoss)~>(TimeoutAlert)~>(LossOfSignal)
(LoginFail)~>(LoginFail)~>(LoginFail)

Statistical Learning Relate entities based on context similarities, then use this relatedness to alert and guide the repair [2]

The hidden cause of the trouble ticket on server 1 is a "data leak" attack that started on server 2

[1] Tailhardat, et al. 2023. "Designing NORIA: a Knowledge Graph-based Platform for Anomaly Detection and Incident Management in ICT Systems" (ESWC'2023)

[2] Tailhardat, et al. 2023. "Leveraging Knowledge Graphs For Classifying Incident Situations in ICT Systems" (ARES'2023)

[3] Tailhardat, et al. 2024. "Graphameleon: Relational Learning and Anomaly Detection on Web Navigation Traces Captured as Knowledge Graphs" (WWW'2024)

Evaluating NORIA-O for anomaly detection and situation understanding

Data integration Knowledge graph-based platform [1]

Model-Based Design Query the graph to retrieve anomalies and their context [2]
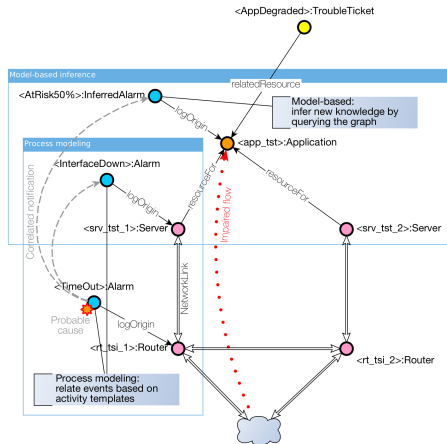
- k out-of n devices with faults
- User with unusual account rights
- Absence of traffic on an interface supposed to be active

Process mining Align a sequence of entities to activity models, then use this relatedness to guide the repair [3]

- (EnergyLoss)=>(TimeoutAlert)=>(LossOfSignal)
- (LoginFail)=>(LoginFail)=>(LoginFail)

Statistical Learning Relate entities based on context similarities, then use this relatedness to alert and guide the repair [2]

- The hidden cause of the trouble ticket on server 1 is a "data leak" attack that started on server 2

[1] Tailhardat, et al. 2023. "Designing NORIA: a Knowledge Graph-based Platform for Anomaly Detection and Incident Management in ICT Systems" (ESWC'2023)

[2] Tailhardat, et al. 2023. "Leveraging Knowledge Graphs For Classifying Incident Situations in ICT Systems" (ARES'2023)

[3] Tailhardat, et al. 2024. "Graphameleon: Relational Learning and Anomaly Detection on Web Navigation Traces Captured as Knowledge Graphs" (WWW'2024)

Evaluating NORIA-O for anomaly detection and situation understanding

Data integration Knowledge graph-based platform [1]

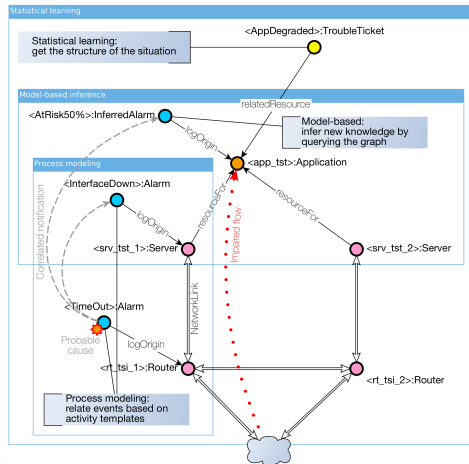**Model-Based Design** Query the graph to retrieve anomalies and their context [2]

- k out-of n devices with faults
- User with unusual account rights
- Absence of traffic on an interface supposed to be active

Process mining Align a sequence of entities to activity models, then use this relatedness to guide the repair [3]

- (EnergyLoss)=>(TimeoutAlert)=>(LossOfSignal)
- (LoginFail)=>(LoginFail)=>(LoginFail)

Statistical Learning Relate entities based on context similarities, then use this relatedness to alert and guide the repair [2]

- The hidden cause of the trouble ticket on server 1 is a "data leak" attack that started on server 2

[1] Tailhardat, et al. 2023. "Designing NORIA: a Knowledge Graph-based Platform for Anomaly Detection and Incident Management in ICT Systems" (ESWC'2023)

[2] Tailhardat, et al. 2023. "Leveraging Knowledge Graphs For Classifying Incident Situations in ICT Systems" (ARES'2023)

[3] Tailhardat, et al. 2024. "Graphameleon: Relational Learning and Anomaly Detection on Web Navigation Traces Captured as Knowledge Graphs" (WWW'2024)

Summary & future work

Problem Comprehensive and integrated view for anomaly detection and decision support in complex ICT systems.

Our approach Knowledge representation using SemWeb technologies, reusing and aligning with third-party vocabularies, and evaluating through authoring tests and real-world use cases.

Next Enriching/aligning the controlled vocabulary for specific technological domains, establishing a shared knowledge base of failure modes related to the nature of networks.

Paper

Lionel TAILHARDAT, Yoan CHABOT, and Raphaël TRONCY.
NORIA-O: an Ontology for Anomaly Detection and Incident Management in ICT Systems.
Semantic Web - 21st International Conference, ESWC 2024.

Code repository

NORIA-O – https://w3id.org/noria/