Introduction to the Special Issue on Software-Intensive Autonomous Systems: methods and applications

Nesrine Khabou, Ismael Bouassida Rodriguez, Khalil Drira

Introduction to the Special Issue on Software-Intensive Autonomous Systems: methods and applications

Nesrie Khabou[b,*], Ismael Bouassida Rodriguez[b], Khalil Drira[a]

[a] *University of Toulouse, CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France*
[b] *ReDCAD, ENIS, University of Sfax.*
nesrine.khabou@redcad.org, bouassida@redcad.org, drira@laas.fr

1. Introduction

The focal concerns are Software-Intensive Autonomous Systems (SIAS). A SIAS is, by definition, any system where software influences, to a large extent, the design, construction, deployment, and evolution of the system as a whole. Some examples include computer-based systems ranging from individual software applications, information systems, embedded systems for automotive applications, telecommunications, wireless ad hoc systems, business applications with an emphasis on web services, software product lines and product families, cyberphysical systems, and systems-of-systems.

The emerging software-intensive systems become more and more considered as autonomy enabling solutions in different ICT-related domains. However, their increasing complexity makes them difficult to design, develop and maintain, and rises many challenges for researchers, architects, and developers. On the one hand, they must meet very stringent guarantees of adaptiveness, flexibility, performance and reliability, both for business as well as for safety reasons. On the other hand, their development requires interaction between engineers from control system and software domains, whose differing backgrounds are often a source of confusion and misunderstanding.

To master complex aspects of software-intensive systems, it is important to combine efforts from foundational research and recent engineering techniques that are based on mathematically well founded theories and approaches. The new methods should support the system life cycle including requirements, design, implementation, maintenance, reconfiguration and adaptation. This ensures the required levels of quality and trust, putting change and adaptation at all levels of system development.

*Corresponding author

2. Overview of the Special Issue

The theme of this special issue is “Software-Intensive Autonomous Systems”. We solicited the submission of high-quality papers describing original and significant work in the SIAS domain as well as submissions of extended papers from the workshop of Adaptive and Reconfigurable Systems and Architectures (AROSA 2020). The call for papers attracted 10 submissions covering diverse relevant topics. Each submitted article was carefully evaluated by at least two experts in the field. After a rigorous peer review process, two high-quality research papers have been selected for the issue.

Paper 1 titled “Model-Based Safety Engineering for Autonomous Train Map” by Nadia Chouchani focuses on a model-based approach to match between safety concepts expressed as an ontology, a derived safety model and a safety-extended railway infrastructure map model for autonomous trains. The proposed approach is validated by railway safety case studies for autonomous train map. The integration of this model-based safety solution from the early stages of the map system design improves the safety decisions management process.

Paper 2 titled “Practical Hybrid Confidentiality-based Analytics Framework with Intel SGX” by Abdulatif Alabdulatif focuses on the development of a privacy-preserving data analytics framework for the adaption of confidentiality-based data analysis in various domains in the realm of IoT. The developed framework aims to build a hybrid privacy-preservation solution that combines both software- and hardwarebased techniques to maintain data confidentiality in volatile and untrusted cloud environments. The framework comprises techniques, including advanced encryption standard (AES) and Intel as software guard extensions (SGX). The proposed framework can be beneficial for end-to-end confidentiality-based data computations across IoT domains, such as health care and smart-grid applications.

3. Acknowledgments

We would like to thank all the authors for their high-quality contributions to the special issue. In addition, our appreciation is due to all the reviewers for their great effort and constructive comments. We are also grateful to the editors-in-chief (Paris Avgeriou and David C. Shepherd), the special issue managing guest editor (Wing-Kwong Chan and Raffaela Mirandola), and the journal manager of JSS for their support throughout the process of preparing the special issue.

- **Khalil Drira** is Research Director at the French National Center for Scientific Research(CNRS). He chaired the Program Committee of several international conferences including ICSOC, ECSA, and IEEE-WETICE. He has co-organized several workshops and tracks including AROSA, ASOCA, SISOS, and CASA. He served in the Steering Committee of the international conferences IEEE-WETICE and ECSA. He served in the Program Committee of over 100 international conferences, including, recently, ECSA, ICSOC, and COOPIS. He is member of the editorial board of several journals including IEEE IoT Journal, Internet Technology Letters, Future Internet, and Smart Science Journal. He has been guest editor of more than 10 Special Issues in international journals including recently: CCPE, JSS, and FGCS. He was editor of several SPRINGER volumes including LNCS 2236, 7957, 10380, and 11895. He was (co-)editor of over 20 international conferences and workshops proceedings.

- **Ismail Bouassida Rodriguez** received the Ph.D degree in Computer Science from the National School of Engineering of Sfax-Tunisia and National Institute of Applied Sciences of Toulouse-France, in 2011. He is since September 2012 an associate Professor at Higher Institute of Computer Science and Multimedia of Sfax-Tunisia. His research interests include graphs grammars and software engineering of distributed systems. He has co-organized the following tracks: AROSA IEEE-WETICE 2013-2022; ASOCA ICSOC 2017-2022. He has served in the Program Committee of international conferences, including CRiSIS, ICCCI. He has been guest editor of Special Issues in international journals including The Journal of Supercomputing, Future Generation Computer Systems, Journal of Systems and Software. He is also involved in different European and Tunisian projects (ENI CBC MED, Erasmus+, DAAD, PRF).
- **Nesrine Khabou** received the Ph.D degree in Computer Science from the National School of Engineering of Sfax-Tunisia, in 2017. She is since September 2018 an associate Professor at Higher Institute of Arts and Crafts of Sfax-Tunisia. Her research interests include software engineering of distributed systems and context aware systems. She has served in the Program Committee of international conferences, including recently: WETICE, ASOCA, ICTAC and CRiSIS.

Guest Editors

Nesrine Khabou
ReDCAD-ENIS, University of Sfax, Tunisia
nesrine.khabou@redcad.org

Ismael Bouassida Rodriguez
ReDCAD, University of Sfax, Tunisia
bouassida@redcad.org

Khalil Drira
LAAS-CNRS, Univ. Toulouse, France
khalil@laas.fr

Editors-in-chief

Paris Avgeriou
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen

paris@cs.rug.nl

David C. Shepherd
Richmond, VA, United States
shepherdd@vcu.edu

Special Issue Editors

Wing-Kwong Chan
Department of Computer Science City University of Hong Kong
wkchan@gapps.cityu.edu.hk

Raffaela Mirandola
Dipartimento di Elettronica, Informazione e Bioingegneria Politecnico di Milano
raffaela.mirandola@polimi.it

Model-Based Safety Engineering for Autonomous Train Map *

Nadia Chouchani*, Sana Debbech, Matthieu Perin

Institut de Recherche Technologique Railenium, 180 Rue Joseph-Louis Lagrange, Valenciennes 59300, France

Abstract

As a part of the digital revolution of railway systems, an autonomous driving train will use a complete and precise map of railway infrastructure to conduct operational actions. Nevertheless, the full autonomy of trains depends on the safety decisions management capacity both on-board and track-side. These decisions must be refined into safety requirements in order to continuously check the consistency between the perceived infrastructure and safety related properties. However, traditional practices of the safety analysis integration are based on human competences. This may be error-prone and in interference with the embedded aspect of the train map. In this paper, we propose a model-based approach to match between safety concepts expressed as an ontology, a derived safety model and a safety-extended railway infrastructure map model for autonomous trains. This approach is validated by railway safety case studies for autonomous train map. The integration of this model-based safety solution from the early stages of the map system design improves the safety decisions management process.

Keywords: Model-Based Safety Engineering, Safety Ontology, Model-Driven

*This research work contributes to the french collaborative project TFA (autonomous freight train), with SNCF, Alstom Transport, Hitachi Rail STS, Altran and Apsys. It was carried out in the framework of IRT Railenium, Valenciennes, France, and therefore was granted public funds within the scope of the French Program "Investissements d'Avenir".

*Corresponding author

Email addresses: nadia.chouchani@railenium.eu (Nadia Chouchani), sana.debbech@railenium.eu (Sana Debbech), matthieu.perin@railenium.eu (Matthieu Perin)

Engineering, Safety/assurance case, Railway Infrastructure model, Autonomous Train

1. Introduction

The context of this research is the autonomous train project launched in 2016 as part of *Tech4Rail*, an ambitious technological program initiated by the direction of railway systems at *SNCF*, in France. The future system that follows from this vision is based on automatic train control (*ATC*) system. The latter is organized on the basis of three functional layers, i.e, (*i*) Automatic Train Protection, (*ii*) Automatic Train Operation (*ATO*) and (*iii*) Automatic Train Supervision. The second level of the *ATC* architecture aims to automate the driving functions of the train. Thus, the *ATO* performs railway driving by executing all the operational functions without human intervention. It is structured around several transverse and functional on-board subsystems like the train positioning, signaling recognition and environment monitoring. These main subsystems require a precise description of the rail network infrastructure. In this work, we propose a model-based approach to develop an on-board map for the autonomous train referring to the topology of the tracks and signaling. Indeed, the proposed model provides a topological description of the railway infrastructure and the signaling objects geo-located by a positioning system. However, the autonomous railway transportation are complex systems that require high safety integrity level. Traditionally, regular human interventions rely on the skills of human agents to ensure the integration of safety analysis. Nevertheless, these practices make the system verification difficult and challenging for safety assurance. Thus, to make the train become autonomous and safe, we identify the following research question (RQ) for this study :

RQ: How could the development of on-board map be enhanced by the integration of safety-related information for assisting the overall autonomous train subsystems ?

To avoid potential hazards, we provide a general framework for design and verification of the mapping system of the autonomous train. In order to have a consistent design process, domain ontologies are used to consider safety rules into system's components and to clarify safety management concepts. The main contribution is the proposal of a novel model-based safety approach which takes into account railway infrastructure information for autonomous train driving.

The outline of this paper is as follows. The next section 2 introduces an overview and the motivations of our work. Section 3 details the proposed model-based approach. Section 4 is devoted to describe railway case studies for the autonomous train map. In Section 5, we present the related work. Finally, the paper concludes and introduces the future work.

2. Overview and motivations

2.1. Safety ontologies

In order to deal with the complexity of safety management process, safety analysis results must be considered from the first design stages of critical systems [1]. This practice is widely recommended by safety standards, e.g., *EN50129* [2] for railway systems and *ISO/DIS 26262-1* [3] for the automotive domain. With the aim to provide a conceptualization of dysfunctional analysis, a reference domain ontology called *DAO* (Dysfunctional Analysis Ontology) was previously developed [4]. *DAO* is grounded on Unified Foundational Ontology (*UFO*) which is an upper-level ontology [5]. It establishes a common vocabulary for the knowledge sharing between safety engineers and system designers. *DAO* integrates both human errors and technical failures from both system and environment perspectives. It has been used on the safety analysis of railway systems. Based on the clarification of the ambiguous use of the failure concept, its causes, effects and related hazards, a set of safety measures may be identified in order to mitigate hazards. Otherwise, *DAO* is developed with the purpose of allowing a well-established formalization of a "Failure" and its surrounding concepts, which is used for the development of new safety critical systems, such as autonomous trains. In order to have an interoperable view of safety analysis methods, *DAO* is compliant with safety standards definitions of concepts. In other words, the proposed conceptual clarification aims to approximate the ideal conceptualization and to have an unambiguous interpretation of dysfunctional analysis concepts. As an example, we may refer to the proposed definiton of the concept of Hazard from the standard EN50126 [6] as "a condition that may lead to accidents". In order to clarify the ambiguous use of these terms, we proposed to define a Hazard as a subtype of a situation (in regard to *UFO*), which is inherent to an exposure (it is activated by a hazardous state) and is prevented by safety measures. Furthermore, *DAO* has been formlized in Web Ontology Language (*OWL*) and evaluated using logic reasoning in order to have a knowledge basis.

Indeed, the development of safety measures requires a control organization which is integrated in adaptive socio-technical systems, such as railway systems. From this point of view, *GOSMO*-a Goal-Oriented Safety Management Ontology- was developed with the aim of matching the safety knowledge and the Goal Oriented Requirements Engineering (*GORE*) concepts [7]. The safety measures development process is proposed based on the Organization-Based Control Access (*Or-BAC*) model, which is traditionally used to ensure the information systems security [8]. This contribution is motivated by the reinterpretation of *Or-BAC* concepts from a safety perspective and their alignment with safety and *GORE* concepts. Thus, *GOSMO* incorporates 3 main modules:

- *Or-BAC* concepts for the safety management process representation ;
- *GORE* concepts for the semantic bridge between safety and requirements engineering phases ;
- A set of *DAO* concepts for the matching between safety measures and safety goals and their management ;

Furthermore, *GOSMO* is grounded on *UFO* in order to help the semantic matching with *DAO*. Otherwise, *UFO* provides a complete set of concepts and relations which is able to cope with the semantic heterogeneity induced by knowledge domains combination. Then, *GOSMO* is built using the Systematic Approach for Building Ontologies (*SABiO*) [9]. *SABiO* methododlogy incorporates best practices of ontology engineering and ontological distinctions of foundational ontologies. In order to provide a high level of semantic expressivity and to have a reasoning support, *GOSMO* is formalized in *OWL* and evaluated using logic reasoning. Finally, the integrated railway knowledge is validated by the application of *GOSMO* to two real critical accidents and a remotely-operated task of autonomous trains [10]. This ontological approach is used from the first design stages in order to integrate dysfunctional analysis and to support the safety decisions making process. The integrated safety measures are adaptive to contexts and they are defined to satisfy safety goals. The formalisation of this semantic link between safety measures and safety goals is crucial since it improves the safety assurance and hazards mitigation. Further details about *DAO* and *GOSMO* development process may be found, respectively in [4] and [7]. In the present study, *DAO* and *GOSMO* are used and combined with other models to have a structured safety model-based process. In order to fulfill autonomous system's needs, a specific fragment of *DAO* is extracted and used in this approach. The reused *DAO* and *GOSMO* concepts are defined in Section 3.

2.2. Railway Infrastructure modelling

Upcoming autonomous transportation systems such as driver-less trains, need a dense, coherent and high-definition representation of their surroundings in order to accomplish their mission safely and efficiently. Thus, digital maps are a key challenge for the railway industry, mainly because this topic has not been known as a core competence of manufacturers nor researchers until now. Especially, the autonomous train on-board mapping subsystem must be capable of gathering a wide variety of data and providing them to a set of different users, i.e the other subsystems, with a strong variation in the nature of needed information. In order to overcome these challenges and since traditional digital maps may not be optimal nor capable, our proposal is to design the autonomous train map following a Model-Based Engineering approach. The proposed solution is associated with state-of-the-art results from international initiatives on digital twin representation for railway infrastructures. In this paper, we propose the Autonomous Train Map Ontology (*ATMO*) which is a Conceptual Independent Model (see next section) representing all the infrastructure objects needed by the future train in order to provide safe and accurate service, based on users requirements. Some modelling research proposed to model the railway infrastructure but they are limited to a domain or a single use case. They are presented briefly in section 5. Such limitations are incompatible in our opinion with the multiple map users such as perception, navigation, positioning, environment monitoring, and safety automation subsystems. *ATMO* is also aligned with existing standards like *RailTopoModel*[1] [11] for abstract and topological representation, *Eulynx*[2] for the physical and functional modelling of the signaling system and *IFC Rail*[3] for civil engineering-related elements such as track structures. Platform Independent and Platform Specific models can then be derived from *ATMO* through automatic processes to generate an implementation that will hold all the needed objects data.

2.3. Model-Based Engineering

In an attempt to ensure consistency between safety analysis and autonomous train map design, we propose to follow a model-based approach. In this multidisciplinary context, we opted for conceptual modelling with the aim to tackle the complexity of the system [12]. This modelling is a key element to generalize the use of *Model-Based Engineering* (*MBE*) and to clarify the semantic interpretation of domain concepts. But which architecture is suitable to build conceptual models for safety critical domains?

[1] From UIC: International union of railways, https://uic.org/

[2] https://www.eulynx.eu/

[3] Industry Foundation Class, Rail part: https://www.buildingsmart.org/ifc-rail-candidate-standard-is-available-for-review-and-comment/

According to *OMG* [13], the *MBE* consists in using a set of complementary models, each corresponding to a specific aspect of the system. A model, being an abstraction of reality, makes it easier to understand the system to be developed. However, it does not represent all of reality but at best the aspect that we want to exploit. Therefore, a view is a representation of the model in a projection of an hyper space to simplify it. In this work, the representation is based on *UML* (Unified Modeling Language) [14], a semi-formal, enrichable and structured language. The modelling task is structured around the expertise knowledge and competency questions, and based on semantic formalisms, transformation rules and frameworks for transition from one model to another [4]. Indeed, the *MBE* can ensure the traceability of business and safety requirements as described in our proposal. These requirements are modelled from the early stages of the development process, hence the minimisation of the downstream design effort. Three main types of models are defined :

CIM (Computational Independent Model) : represents the business model which is independent of any computer system. At this level, we used a safety and railway infrastructure ontologies.

PIM (Plaform Independent Model): independent of the technical platform, this model is a partial view of a *CIM*. It represents the business functional logic and describes the system, using classes and *OCL* constraints (Object Constraint Language). At this level, two *PIM*s are derived from the ontologies.

PSM (Platform Specific Model): depending on the technical platform, it is used as a basis for code generation [15].

The transition from one model type to another is done by tools for model transformation according to user designed rules. A transformation is defined as an operation on a model that produces another one, and which conforms to formal syntax and semantics [16].

MBE is a valuable methodology to conceive system assurance cases argumentation. The assurance cases are claims, arguments and evidence concepts that justify and assess confidence in the system critical properties, such as safety and security [17]. For instance, assurance case reports can be generated by model-to-text transformation [18]. Recently, model-based system assurance has attracted considerable research attention. In this context, the Structured Assurance Case Metamodel (*SACM*) [17] was specified by the Object Management Group for representing structured assurance cases. This metamodel was intended to improve standardisation and interoperability. Its specification evolved from experts collective safety/security knowledge and the associated experiences in the domain.

3. The safety model-based approach

The general architecture of our approach is given in Figure 1. It is composed of three components : (*i*) safety analysis, (*ii*) model extension, and (*iii*) safety management. The subsections below provide details on these components.

3.1. Safety Analysis

The first step is the extraction of relevant concepts from *DAO* in order to perform safety analysis for autonomous systems. Figure 2 shows the *DAO* fragment which represents the required concepts and relations between them in *OntoUML* [5]. The latter is a *UML* profile for conceptual modeling and it incorporates foundational distinctions defined in *UFO*. The interpretation of *failure* and its related concepts in real-world semantics may be found in [4]. The semantic interpretation of the main *DAO* concepts are detailed in Table 1, based on the knowledge acquisition step from safety engineering standards.

Once the autonomous system's structure is known, this *DAO* fragment may be applied in order to identify failures and their effects for each system's component. The obtained DAO instantiation is considered to be the *Safety Model* which depends on a specific dangerous event. This safety model is deduced from *DAO* and includes individuals of *DAO* concepts and relations between them.

Table 1: The semantic interpretation of *DAO* concepts

Concepts	**Definitions**	**Source**
Failure	A **Failure** is a *subtype of* **UFO::Event**. It *brings about* a **Failure State** and *is triggered by* a **Hazardous State**. A **Failure** *causes* another one (cascading failure) and *is manifestation* of an **Exposure**.	IEC 61508 [19]
Exposure	An **Exposure** is a *subtype of* **UFO::Disposition** (a special type of **Moment**). It denotes the **Exposure Moment** which *inheres in* **UFO::Object** and is *activated by* the **Hazardous State** (a *subtype of* **UFO::Situation**).	EN50126 [6]
Defect & Fault	A **Defect** is a *subtype of* **Exposure**. A **Defect** denotes a **Fault** when it is *manifested by* a **Fault emergence Failure**. A **Fault** *subsumes* an **Environment Object Fault** and a **System Equipment Fault**.	IEC 61508 [19]
Fault emergence Failure	A **Fault emergence Failure** is a *subtype of* a **Failure**. It represents any **Failure** caused by an **Object Fault**.	IEC 61508 [19]
Hazard	**Hazard** is a subtype of a **UFO::Situation**, which is inherent to an **Exposure** (it is activated by a **Hazardous State**) and *is prevented by* **Safety Measures**.	EN50126 [6]
Safety measure	**Safety Measure** is an **UFO::Action** which *prevents* a **Hazard** and *satisfies* a **Safety Goal**.	EN50126 [6]

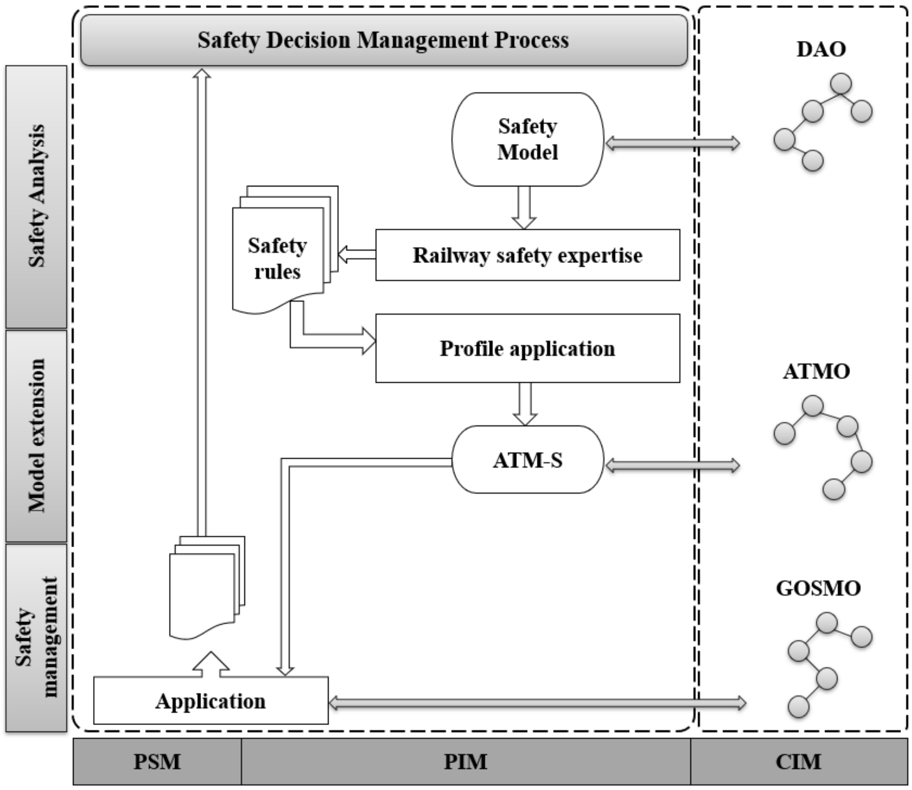


Figure 1: General architecture.

Individuals of DAO concepts represent the safety analysis elements of the considered system. According to the performed safety analysis, a set of safety measures are defined in order to mitigate the perceived hazard. Safety rules are defined as a set of actions or safety measures to be realized within a task in order to achieve the required safety integrity level. Furthermore, safety rules are assumed to be available in a specific context which may be composed of sub-contexts. They are defined based on the railway expertise acquired from domain experts and referential. Thus, safety rules are considered as an aggregation of 3 concepts: safety measures, a specific context and conditions that validate the rules application. These safety rules are expressed from a high level of abstraction in order to prevent perceived hazards, such as collisions. Fur-

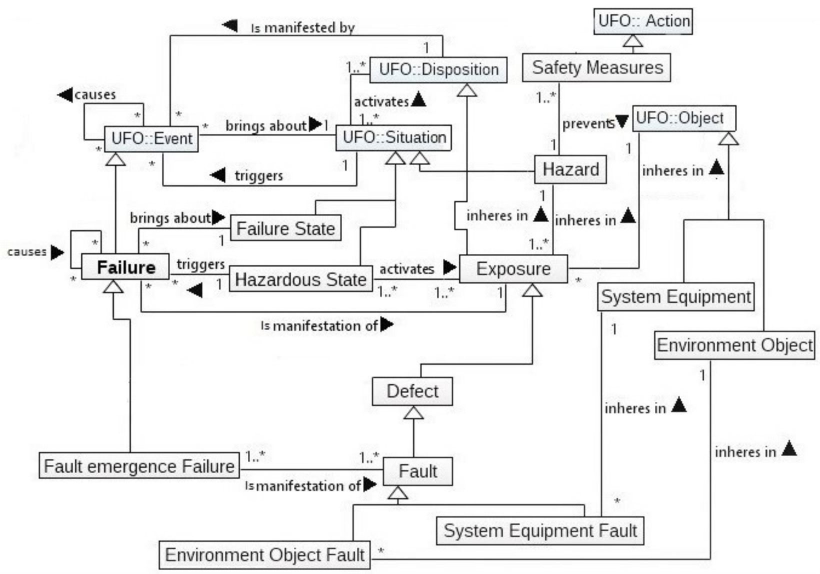


Figure 2: A fragment of *DAO* conceptual model.

thermore, they are integrated from the first design stages in order to prevent, as soon as possible, safety properties violation.

3.2. Model extension

As part of this approach, we are working on the modelling of a high definition on-board map or cartography, represented by *ATMO* and based on different standards including all information on infrastructure, signaling and even constructions such as tunnels and bridges, eventually in *3D* representation. Data integration and interoperability are complex challenges due to the heterogeneous nature of data and standards. To overcome this problem, we propose to apply semantic data modelling techniques to allow integration of heterogeneous information and make coherencies of cartographic elements in addition to the safety rules obtained from the previous step. The adopted methodology is structured around the following main steps :

3.2.1. Specification

The specification of the data model is defined by a set of functional and non-functional requirements derived from the established needs of the implementation of the autonomous train in the context of the project.

3.2.2. Knowledge acquisition

Several areas of knowledge are at the heart of this work. This step was carried out by defining Ontology Design Patterns (*ODP*s). It involves defining all the concepts to be used in the ontology, the relationships between them and also a documentation corresponding to the different concepts. In order to extract the domain knowledge of the ontology *ATMO*, we used three sources for explicit and implicit acquisitions. First bibliographic research of articles and books was necessary to form a background on the whole field and questions on more specific use cases. Then we collaborated with experts, especially in the signaling field. We had discussions around *EULYNX UML* model to which we had a read access. Finally, the reuse and re-engineering of non-ontological resources were applied to the model construction. The analysis of the various cited resources allowed to define data dictionary that meets the needs to be covered by *ATMO*.

3.2.3. Conceptualization

The vocabulary and the *ATMO* model are mainly based on the elements of *RTM*, *IFC Rail* and *EULYNX*, relying on both their *UML* models and natural language documentation. The designed model contains four packages, each one references one module of *ATMO*. An excerpt from the *UML* package of "Track" is shown in Figure 3 and described in Table 2. Due to confidentiality restrictions linked to the project, not all packages can be detailed here.

The methodology of *ATMO* design follows a compositional approach. The different modules, each corresponding to a dimension of the railway map, are constructed and subsequently composed to constitute the global model.

Table 2: *UML* "Track" package description.

Entity	Description
LocatedNetEntity	From *RTM*, it represents a functional object in the rail network located on the topology.
EntityLocation	From *RTM, it is the location of a network entity.*
LinearLocation	From *RTM*, a linear location consists on an ordered list of network elements.
AreaLocation	From *RTM*, it is an area located in the network.
Panel	It is a homogeneous section in configuration inheriting from "LocatedNetEntity" allowing a tiling of the infrastructure.
PanelArea	It is an area preempted by the functional object represented by the "Panel" which carries the topological objects. It is a geographic area ("EntityLocation")
TrackPanel	A simple, homogeneous track section, inheriting from "Panel"
CrossingPanel	A section representing a crossing of tracks inheriting from "Panel" linked to a geographical area "AreaLocation".
TurnoutPanel	A section of track representing a switch inheriting from "Panel" linked to a geographical area "AreaLocation".
Frog	Frog of turnout inheriting from "LocatedNetEntity".
Track	Functional and organizational object representing a channel inheriting from "LocatedNetEntity" and references "Panel type objects.
TrackSegment	Functional cut-out of the train guidance which carries the *RTM* "LinearElement" topological object.
LinearElement	From *RTM*, a linear segment representing a network element.
StoppingPoint	Fouling point to stop the train.
Ballast	Track ballast.
Sleeper	Track sleepers.

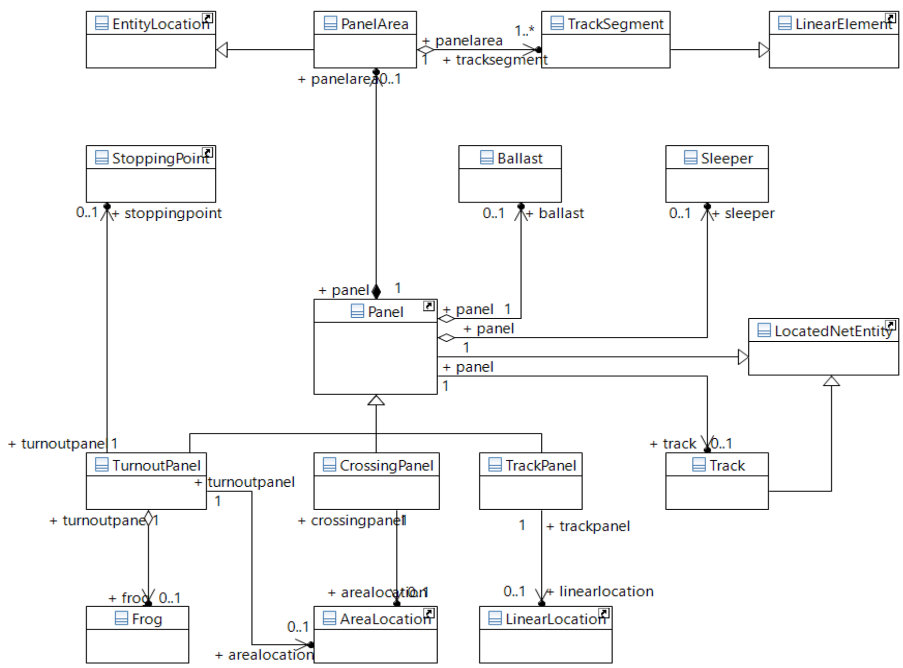


Figure 3: An excerpt of the *UML* “Track” package of the map *PIM*.

3.2.4. Integration

The purpose of this step is to integrate the safety rules into the conceptual map model (*PIM*). The aim is to get a view of the rail infrastructure system coupled with safety measures in order to be able to take on-board safety decision actions in an autonomous way. The extracted safety rules from the previous component, are expressed in natural language. In order to have a safety decision-making framework, safety rules are transformed from natural language to a machine-readable language. In this work, the *SWRL* (Semantic Web Rule Language) [20] is chosen thanks to its formal syntax and semantics and to its capabilities to express and integrate rules into ontologies.

For the safety decision management process, detailed in the following subsection, we relied on the safety actions (“DAO::Safety Measures”) associated to each context.

In order to integrate these safety measures into the map conceptual model,

we defined and apply a *UML* profile, derived from *DAO*, to the *PIM* obtained from *ATMO*. The resulting *PIM* is *ATM-S*, the autonomous train map model integrating safety assurance aspect. The main aim of the profile is to capture the different situations related to the infrastructure objects and make the correspondence with the integrated safety rules.
For example, "Exposure", "Hazard" and "Hazardous State" are stereotypes applied to the "TurnoutPanel" entity of the "Track" *UML* package.

3.3. Safety Management

Safety management is a crucial process for autonomous systems safety assessment since it is based on both perception and decision steps. In order to provide a structured safety management, safety measures derived from safety analysis must be linked to safety goals. This knowledge merging allows a shared view between safety and system objects with the aim of goals satisfaction. This is the subject of the third step of the proposed approach using *GOSMO* in order to orchestrate safety decisions management process. Figure 4 shows the *GOSMO* fragment which includes pertinent concepts for autonomous systems safety management.

The organization-based control model allows the assignment of roles using the concept **Stakeholder Role** to *ATMO* components. Then, a **Permission** is assigned to perform a **Task** that realises **Safety Measures** in a specific application **context**. Safety rules expressed in SWRL allow an allocation of safety measures to specific ATMO objects in a specific operational context.

This *Or-BAC* reinterpretation from a safety-perspective is suitable for the adaptive safety management of autonomous systems, such as railway systems. *GOSMO* conceptual model may be used to annotate the ATMO model as a profile in order to have a semantic link between them. This semantic annotation avoids ambiguities and allows consistency with system models. The considered goal-oriented perspective is useful for the requirements analysis process in a later stage of system development.

Table 3 shows *GOSMO* concepts definitions in order to improve readability.

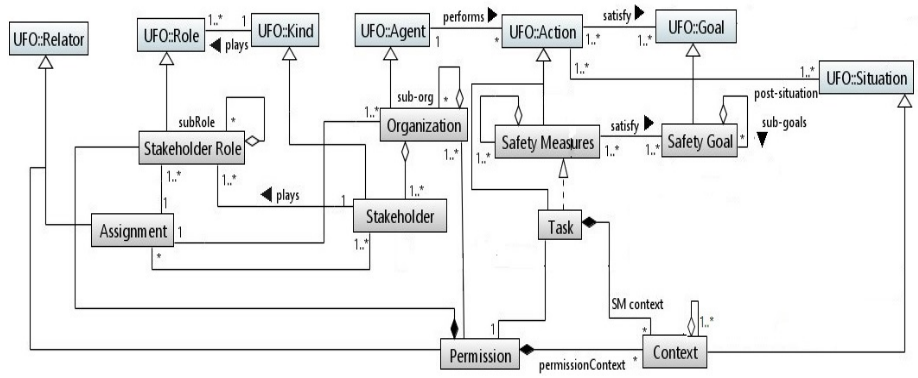


Figure 4: A fragment of *GOSMO* conceptual model for autonomous systems safety management

This ontology has been formalized in Ontology Web Language-Description Logic (OWL DL)[4], with the aim to reach a high level of semantic expressivity and to have a reasoning framework for safety decisions management. A set of DL axioms has been defined to constrain the proposed terminology and to help the data retrieval process [7]. Otherwise, the proposed framework aims to have an automatic safety decisions making process thanks to the predefined SWRL rules. It may be used from the first design stages of safety critical systems design

4. Safety cases : Application for autonomous train map

The proposed map system performs critical functions thus requires safety justifications. In the following, we detail assurance cases, in particular, safety cases. As specified by [21], a safety case should communicate a clear, comprehensible and defensible argument that a system is acceptably safe to operate in a particular context.

Safety cases can be represented either textually, in natural language, or graphically. In this section, we refer to goal structuring notation in order to analyze ans validate the satisfaction of safety goals by the integration of safety

[4]https://www.w3.org/2007/OWL/wiki/images/9/9a/Pfps-f2f1.pdf

Table 3: *GOSMO* concepts definitions

Concepts	Definitions
SafetyMeasure	A **SafetyMeasure** is a *subtypeOf* **Action**. It *hasPart* **SubSafetyMeasures**. It *satisfies* a **SafetyGoal** that *hasPart* **SubSafetygoals**. A **SafetyGoal** is *refinedIn* **SafetyRequirement** *gotFrom* a **Stakeholder**. When the **Task** is performed, a **postSituation** occurs and *satisfies* a **Proposition** (**Goal**).
Task	A **Task** is accomplished by a **Permission** assigned to **StakeholderRole** by an **Organization** according to a specific **Context**.
StakeholderRole	A **StakeholderRole** is a *subtypeOf* **Role**. It *is played by* a **Stakeholder** (a *subtypeOf* **Kind**).
Context	A **Context** is a *subtypeOf* **Situation**. It denotes the specific **Situation** (circumstances) in which the **Permission** is assigned to a **StakeholderRole** to perform the **Task**. It *hasPart* **SubContexts**. It *extends* a **SafetyRequirement** and a **FunctionalRequirement**.
Organization	An **Organization** is a *subtype of* **Agent** and it *hasPart* **suborganizations**. An **Organization** *hasPart* one or many **Stakeholders** that are a *subtypeOf* **Kind**.
Assignment	An **Assignment** is a *subtypeOf* **Relator** and it denotes the **StakeholderRole** assignment to a **Stakeholder** by an **Organization**.
Permission	A **Permission** is a *subtypeOf* **Relator** and it denotes the **Stakeholder Role** authorization to accomplish the **Task** according to a **Context**, which is a specific *subtypeOf* **Situation**.

rules. Then, the proposed approach is illustrated by two railway case studies.

4.1. Goal Structuring Notation

The Goal Structuring Notation (*GSN*) [21], widely adopted in the literature, is a graphical notation used to express system properties argumentations in a clear and well-structured way. Thanks to its powerful notation, *GSN* enables to represent structural system safety arguments. In order to produce a robust safety case, we followed the *GSN* metamodel which is compliant to *SACM* and represents the most popular approach for system assurance [18]. An excerpt of the resulted goal structure is shown in Figure 5.

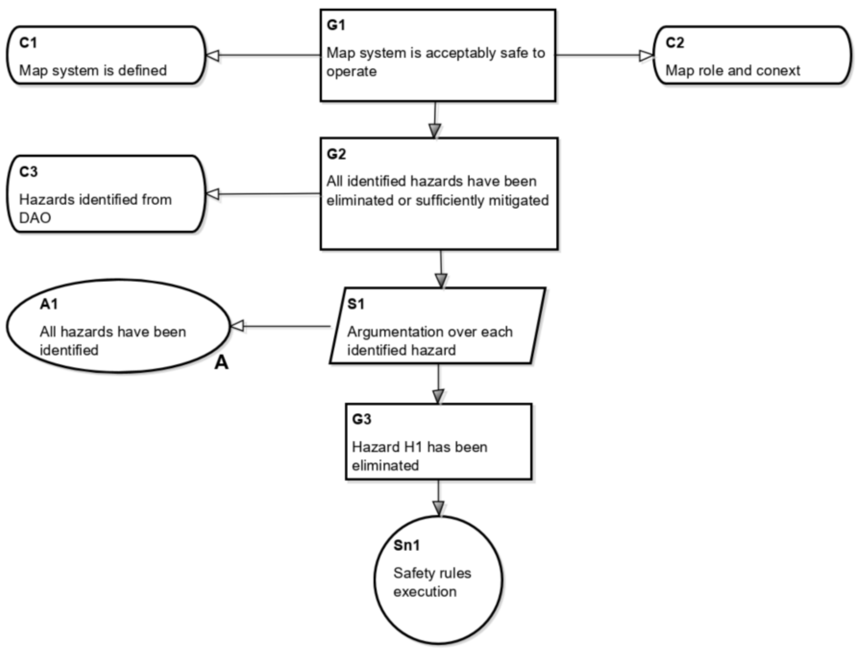


Figure 5: An excerpt of the goal structure using *GSN*.

The main goal (*G1*) of this structured safety case, is to operate the autonomous train map system safely with compliance to safety requirements. A sufficient mitigation and the avoidance of hazards are the key features to attend this goal. The latter is decomposed and sub-goals (*G2* and *G3*) are then identified. The demonstration of safety depends on contexts (*C1*, *C2* and *C3*) and is based on assumptions or justifications (*A1*). The solution (*Sn1*) guarantees to avoid hazards.

With the aim to show the attainability of the identified goal *G1* and therefore the safety of the proposed *ATM* system, the following sections present two case studies detailing the different hazrads from *DAO* (*C3*) and safety rules (*Sn1*) application for each case.

4.2. Case study 1: Side collision

In order to validate the proposed approach, we refer to a railway case study which illustrates its three phases. As a potential risk related to infrastructure or rolling stock failures, the side collision occurs when a train hurts another one at a track section which connects two tracks with different provenances. Figure 6 represents the side collision between two trains that intend to join the same track and direction.

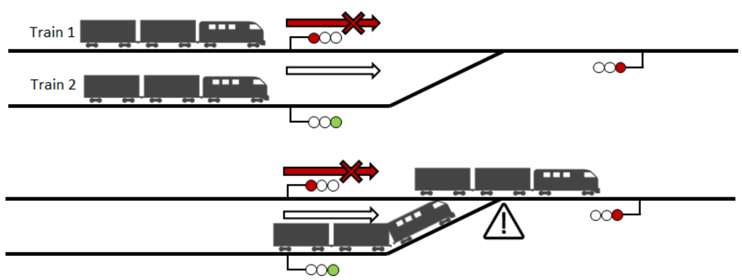


Figure 6: Presentation of the side collision risk in railway operation.

Indeed, train 1 crosses the first closed signal (red light) and keeps immobile at the merging track. Train 2 crosses the open signal (green light) and longitudinally hurts train 1.

The application of the proposed approach to this case-study allows representation of several zones to the infrastructure description in order to perform safety analysis. Side collision represents the “Hazard” concept in the *DAO* conceptual model. As depicted in Figure 7, the extracted candidate concepts after matching with *DAO* are the following :

- **Exposure Zone** represents the zone which activates the hazard occurrence.
- **Danger Zone** represents the zone which inheres in the hazard (Side collision).

- **System Equipment** represents infrastructure components such as signal and tracks.
- **Hazardous Zone** represents the danger zone.
- **No train zone** represents the failure state.
- **Perception Zone** The perception of context to manage safety decisions.

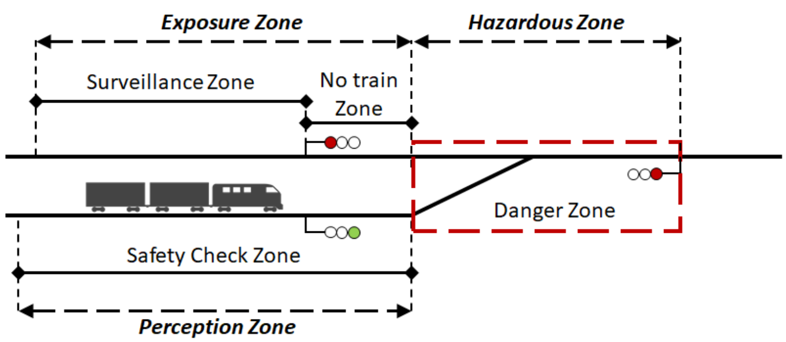


Figure 7: Added Safety-related zones for a turnout related to the side collision hazard.

The topological elements corresponding to this section of the infrastructure are :

- **Turnout** represented by “TurnoutPanel”
- **Signal** represented by a “LocatedNetEntity”
- **Area** represented by “AreaLocation”

This infrastructure decomposition allows the development of tailored safety rules. In order to avoid side collision, a set of safety rules are defined in natural language as follows :

1. The train must be in 30km\h as a maximal speed in the surveillance zone in order to perceive the context.
2. In the case of crossing of a closed signal, a deployment of technical device of train protection system, such as crocodile must be performed in order to trigger the emergency stop before the danger zone.

In order to automatize the safety decisions management process, these safety rules are transformed in SWRL as shown in Figures 8 and 9.

```
Safety Rule 1:

<swrlx:classAtom>
   <owlx:Class owlx:name="SystemEquipment" />
   <ruleml:var>x1</ruleml:var>
</swrlx:classAtom>
<swrlx:classAtom>
   <owlx:Class owlx:name="Train" />
   <owlx:SubclassOf>
   <owlx:Class owlx:name="SystemEquipment">
   </owlx:SubclassOf>
</swrlx:classAtom>
<owlx:Class owlx:name="Train" />
        <owlx:ObjectRestriction owlx:property="hasSpeed">
        <swrlx:datarangeAtom>
        <owlx:DataValue owlx:datatype="&xsd;int">30</owlx:DataValue>
        <ruleml:var>x1</ruleml:var>
        </swrlx:datarangeAtom>
</swrlx:classAtom>
<swrlx:classAtom>
        <owlx:Class owlx:name="Task" />
        <ruleml:var>x1</ruleml:var>
        <swrlx:individualPropertyAtom  swrlx:property="hasContext">
        <ruleml:var>task</ruleml:var>
        <ruleml:var>theSurveillanceZone</ruleml:var>
        </swrlx:individualPropertyAtom>
</swrlx:classAtom>
```

Figure 8: The first SWRL safety rule for case study 1

```
Safety Rule 2:

<swrlx:classAtom>
           <owlx:Class owlx:name="Task" />
                 <ruleml:var>x1</ruleml:var>
         <swrlx:individualPropertyAtom  swrlx:property="realizes">
         <ruleml:var>task</ruleml:var>
         <ruleml:var>deploymentofTechnicalDeviceofTrainProtectionSystem</ruleml:var>
         </swrlx:individualPropertyAtom>
         <owlx:IntersectionOf>
       <swrlx:individualPropertyAtom  swrlx:property="hasContext">
       <ruleml:var>task</ruleml:var>
       <ruleml:var>crossingOfaClosedSignal</ruleml:var>
       </swrlx:individualPropertyAtom>
         </owlx:IntersectionOf>
</swrlx:classAtom>
<swrlx:classAtom>
           <owlx:Class owlx:name="SafetyMeasure" />
    <ruleml:var>x1</ruleml:var>
         <swrlx:individualPropertyAtom  swrlx:property="satisfy">
         <ruleml:var> deploymentofTechnical deviceofTrainProtectionSystem</ruleml:var>
         <ruleml:var> triggerTheEmergencyStopBeforeTheDangerZone </ruleml:var>
                  </swrlx:individualPropertyAtom>
</swrlx:classAtom>
```

Figure 9: The second SWRL safety rule for case study 1

These safety decisions management is performed according to *GOSMO* conceptual model. Figure 10 represents the safety management related to this case study. The permission is assigned to the technical device to trigger emergency stop if the speed curve profile is in state * or *KO* and the train position is close to the closed signal. These elements represent the perceived context related to this task.

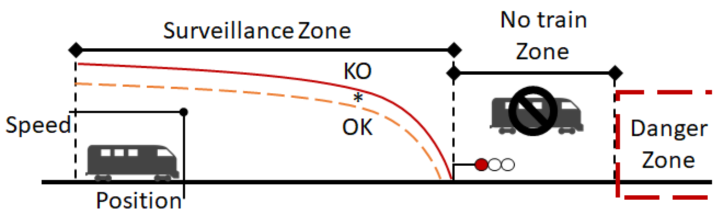


Figure 10: Detail of the safety zone related to the presence and displacement of a train on the merging track.

The proposed case study illustrates the rigorous choice of *DAO* and *GOSMO* concepts for autonomous systems and their matching with ATMO. The proposed approach may be applied to other case studies in order to validate the flexibility to cover several critical situations.

4.3. Case study 2: Real railway accident of Saint-Romain-En-Gier

In order to validate the capability of the proposed solution to represent real critical scenario, we illustrate it by a railway accident of Saint-Romain-En-Gier [22]. This accident denotes a frontal collision and occured on April 5th, 2004 between an empty high speed train and a works train on the french line Lyon/Saint-Etienne. The accident was due to track works between the cities of Rive-de-Giers and Givors, in a railway section equipped with reverse signalling. The works carried out on the night of the 4th to 5th of April took longer than expected, and consequently the works trains were behind schedule on their return journey. The ballast works train return journey conflicted with the first commercial morning run between Lyon and Saint-Etienne. Due to series of human errors, these two trains were running in opposite directions but moving towards each other on the same track and a head-on collision could not be avoided. Consequently, both train drivers were injured and considerable damage impact rolling stock. Figure 11 represents the infrastructure of the line Lyon/Saint-Etienne in which the accident occured.

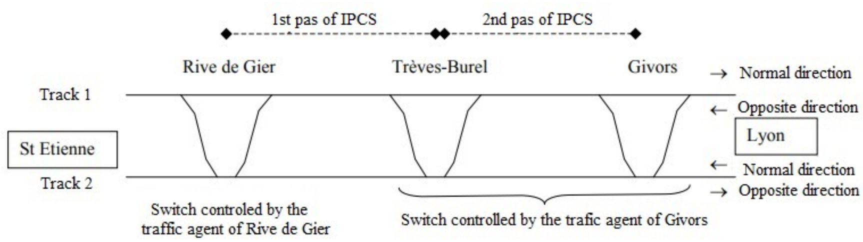


Figure 11: The line infrastructure of Lyon/Saint-Etienne [22]

The first human error comes from the safety agent who did not protect this area. Furthermore, the traffic agent emitted an erroneous authorisation to the works train due to a false interpretation of the situation. This works train crossed two closed signals which are out of its operating institution. More details about the accident factors and effects may be found in [22].

The proposed approach aims to analyse and anticipate critical situations in order to improve safety from the first design stages. Indeed, the application of DAO to this accident scenario allows a thorough safety analysis which prevents the occurrence of this collision. In order to mitigate the frontal collision as Hazard, DAO concepts are instantiated and represent safety-related information of this accident. Figure 12 depicts safety integrated concepts into the infrastructure section representation. Zones decomposition facilitates the safety decisions management process in order to ensure a safe system operation.

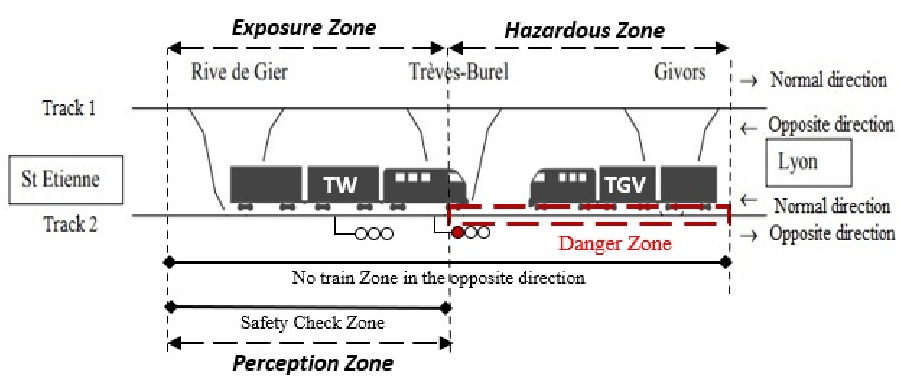


Figure 12: Added safety-related zones for the turnout of the frontal collision

The alignment between topological concepts derived from *ATMO* and the presented infrastructure section, is performed as follows:

- **Turnout** represented by "TurnoutPanel"
- **Signal** represented by a "LocatedNetEntity"
- **Area** represented by "AreaLocation"
- **Rive de Giers/Trèves-Bruel segment** represented by "TrackSegment"
- **Trèves-Bruel/Givors segment** represented by "TrackSegment"

Once the safety analysis performed, a set of safety rules may be integrated in order to avoid frontal collision between commercial and works trains. These organisational rules are defined in order to mainly enforce the following railway procedures:

1. When the works train is running outside of its operating area, the verification of signalling instructions must be integrated in the on-board signalling detection subsystem.
2. In the presence of switches for both running directions and tracks interception devices, the running direction must be indicated on-board.

The first safety rule allows the capture of signalling data for the overall area in order to avoid the crossing of closed signals (**SafetyGoal1**). The second safety rule is proposed with the aim to prevent the traffic on the opposite direction (**SafetyGoal2**). Figures 13 and 14 show the SWRL transformation of these safety rules in order to automate the safety decisions management process.

The illustration of the proposed approach by the accident of Saint-Romain-En-Gier shows that the integration of safety rules as soon as possible in the system development process could have avoided this collision. The matching between safety concepts and real data validates the powerful capabilities of semantics to represent, analyse and anticipate several critical scenarios.

```
Safety Rule 1:

<swrlx:classAtom>
  <owlx:Class owlx:name="SystemEquipment" />
  <ruleml:var>x1</ruleml:var>
</swrlx:classAtom>
<swrlx:classAtom>
  <owlx:Class owlx:name="OnboardSignallingDetectionSubsystem" />
  <owlx:SubclassOf>
  <owlx:Class owlx:name="SystemEquipment">
  </owlx:SubclassOf>
</swrlx:classAtom>
<swrlx:classAtom>
<owlx:Class owlx:name=" OnboardSignallingDetectionSubsystem " />
      <swrlx:individualPropertyAtom  swrlx:property="verifies">
      <ruleml:var> onboardsignallingdetectionsubsystem </ruleml:var>
      <ruleml:var>SignallingInstructions</ruleml:var>
      </swrlx:individualPropertyAtom>
</swrlx:classAtom>
<swrlx:classAtom>
      <owlx:Class owlx:name="Task" />
      <ruleml:var>x1</ruleml:var>
      <swrlx:individualPropertyAtom  swrlx:property="hasContext">
      <ruleml:var>task</ruleml:var>
      <ruleml:var>areaoutsideofitsoperatinginstitution</ruleml:var>
      </swrlx:individualPropertyAtom>
</swrlx:classAtom>
<swrlx:classAtom>
      <owlx:Class owlx:name="SafetyMeasure" />

      <ruleml:var>x1</ruleml:var>
      <swrlx:individualPropertyAtom  swrlx:property="satisfy">
      <ruleml:var> verificationofsignallinginstructions</ruleml:var>
      <ruleml:var>avoidthecrossingofclosedsignals </ruleml:var>
     </swrlx:individualPropertyAtom>

</swrlx:classAtom>
```

Figure 13: The first SWRL safety rule for case study 2

5. Related work

This section represents existing approaches and studies which tackle the different perspectives of the proposed methodology, such as safety ontologies for safety-critical systems, railway infrastructure models and MBSE approaches. Then, a comparative discussion is presented in order to highlight the original contributions of this paper.

Safety Rule 2:

```
<swrlx:classAtom>
      <owlx:Class owlx:name="Task" />

     <ruleml:var>x1</ruleml:var>

       <swrlx:individualPropertyAtom  swrlx:property="realizes">
       <ruleml:var>task</ruleml:var>
       <ruleml:var>IntegrationOfrunningdirectionOnboard</ruleml:var>
       </swrlx:individualPropertyAtom>
       <owlx:IntersectionOf>
       <swrlx:individualPropertyAtom  swrlx:property="hasContext">
       <ruleml:var>task</ruleml:var>
       <ruleml:var>Presenceofswitchesforbothrunningdirectionandtracksinterceptiondevices</ruleml:var>
       </swrlx:individualPropertyAtom>
       </owlx:IntersectionOf>
</swrlx:classAtom>
<swrlx:classAtom>
      <owlx:Class owlx:name="SafetyMeasure" />

       <ruleml:var>x1</ruleml:var>
       <swrlx:individualPropertyAtom  swrlx:property="satisfy">
       <ruleml:var> IntegrationOfrunningdirectionOnboard</ruleml:var>
       <ruleml:var>avoidthetrafficontheoppositedirection</ruleml:var>
      </swrlx:individualPropertyAtom>

</swrlx:classAtom>
```

Figure 14: The second SWRL safety rule for case study 2

5.1. Safety analysis for critical systems

Developing automated driving systems faces safety challenges since verifying such critical systems represents a difficult task. [23] raises discussion on safety challenges in terms of normative requirements. However, the absence of autonomous trains in mainline railway results in technological and fundamental risk assessment challenges. These same challenges were also raised in the automotive field [24, 25]. Indeed, the *Safety Of The Intended Functionality* (*SOTIF*) standard, shorthand for *ISO/PAS 21448* [26], testifies to the progress of the standardization of the autonomous vehicle safety. It provides design, verification and validation measures to achieve safety when identifying hazardous events. Unlike *ISO 26262* [3], it is concerned with mitigating risks without a system failure. In order to disambiguate safety analysis concepts and clarify their semantic from the first phases of the development cycle, knowledge representation is a key activity which facilitates this task. Indeed, ontologies have been widely used in the safety analysis of critical systems and their design. Authors of [27] proposed a safety ontology to formalize the safety knowledge and its link with information models. This ontology allows the automated safety planning for job hazard analysis using Building Information Modeling (*BIM*). In [28], an ontology was proposed to represent and manage Failure Modes and Effects Analysis (*FMEA*) knowledge in the automotive domain. Furthermore, it defines actions to mitigate the anticipated risk and allows the extraction of safety information using its operational version in OWL. On the other hand, a conceptualization of hazard-related knowledge (Hazard Ontology) [29] was proposed. This ontology aims to identify hazards from the early design stages of safety critical systems and elicit safety requirements that mitigate them. From the same context, [30] proposed an approach to increase the validation of hazard-mitigating requirements based on an Ontology for Hazard Relation Diagrams. It allows to generate the Hazard Relations Diagram which satisfies a specific safety goal. This solution is built based on the same motivations and the identified research goal of our proposed approach. Nevertheless, authors did not use a specific ontology, such as *GOSMO* to establish and maintain the semantic link between safety concepts and goal-oriented requirements concepts.

In their study, [31] developed a domain ontology to capitalize safety risk knowledge in metro construction. The built ontology is evaluated using case-studies and provides a decision-making support for safety risk identification. In order to provide a conceptualization of Functional Resonance Analysis Method (FRAM), [32] proposed a foundational ontology-based model using *UFO*. The conceptualization focused on the function concept and its surrounding aspects. The *FRAM* model is applied to a case study from the aviation domain in order to validate the integration of complex socio-technical system's features into this ontological analysis.

Most of these safety ontologies allow only the safety analysis by representing concepts of a specific method or based on a safety principle. However, none of them explored the overall dysfunctional analysis conceptualization which is independent of classic safety methods like *DAO*. Furthermore, their objectives are limited to safety analysis without a focus on how to exploit safety results and link them to the safety management process. This research goal is satisfied differently by other approaches [33] to align safety and systems models without conceptual clarification of semantic links. An approach to validate safety of perception software and system in autonomous driving systems has been proposed based on fault injection but it did not consider the safety management [34]. Finally, to the best of our knowledge, there is lack of an approach which integrates safety concerns with railway infrastructure ontologies. In this paper, we fill this gap and we propose a new approach which is able to deal with innovative industrial locks of future systems.

5.2. Infrastructure modelling

Previous works like [35] proposed modeling of railway infrastructure using *UML* and *UML* profiles. The aim was to obtain control-command models for signaling in tramway, but unlike *ATMO* only one usage for the infrastructure data is provided and no addition of safety-related information is present. Our approach differs because all the users of on-board mapping will benefit from the safety concepts added into *ATM-S*. The work presented in [36] focuses on the instance-level description of a railway infrastructure using *RailML*[5]. This study may be used by extending the scope of *RailML* to hold the safety information needed in order to instantiate *ATM-S* in a static file-based format. In [37], a component-based topology is used to model the infrastructure, as performed in *RailTopoModel* and subsequently *ATMO*. Therefore, the work presented in this paper may be seen as a follow-up of the proposed principle. Finally, [38] presented a full method from *UML* model of the infrastructure down to *SCADA* implementation for railway interlocking, aside the limitation to a sole user. In [39], an Ontlogies-based approach was proposed to support the integration of domain-specific models in the development process of critical systems. In a future work, the result of [39] may be extended to link the system behavior with an ontological level.

“Ontorail”[6] is an ongoing project to support the scientific initiatives for im-

[5] https://www.railml.org/en/

[6] *https://ontorail.org/ontorail/index.php?title=Main_Page*

plementing a shared railway dictionary using terminology adopted in several national and international standards, and technical specifications for interoperability. Their work is based on "MediaWiki"[7] and it semantic extension "Semantic MediaWiki"[8] . It attempts to use the power of its semantics and extension tool-set to develop a *CIM* for railway field represented by an ontology.

Recent works from domains such as autonomous road vehicles are tackling infrastructure modeling, generally focusing on on-board mapping service, with interesting development in semantic layer [40] to help manage dynamic information and graph-based layer [41] to help autonomous control on road lane driving. These works show interesting ideas close to railway infrastructure modeling topics but are not taking into consideration safety-related properties.

Now, to the best of our knowledge, there is no scientific research work that has proposed a general framework for modeling the railway infrastructure and joint safety requirements for autonomous trains.

5.3. Model-based system assurance

The model management operations and its consequent automation capabilities provided by *MDE* have proven that the consistency and efficiency are improved significantly. Several assurance cases tools have then adopted *MDE*, such as *CertWare* [42], *AdvoCATE* [43] and *D-Case Editor* [44].

Historically, the safety cases expressed safety arguments in free texts using natural language. The main problem is that these texts are unstructured and can be unclear. To guarantee the production of clear and well-structured cases and avoid the problems issued by expressing safety arguments in natural language, graphical argumentation notations were proposed. *GSN* and Claims-Arguments-Evidence (*CAE*) [45] are examples of these notations. *CAE* presents assurance cases as a set of claims which are supported by safety arguments. However, *GSN* provides a more detailed decomposition of arguments. Further-

[7]https://www.mediawiki.org/wiki/MediaWiki

[8]https://www.semantic-mediawiki.org

more, it supports additional features like modularity, controlled vocabulary and automated assurance case instantiation. These features are also adopted by *SACM*.

The use of *GSN* proved that the quality of argument approaches was improved, in addition to time development reduction [18]. A major problem with the tools based on *GSN* is that they define their own metamodel. In [18] a methodology was proposed to resolve interoperability problems by proposing a *GSN* metamodel compliant with *SACM*.

6. Conclusion

In order to make the trains become fully automated driver-less, high precision embedded map of the railway infrastructure is required. Our proposal is being sought to consider safety engineering to design the autonomous train map. This paper proposes a solution allowing the safety requirements to be integrated inside a map conceptual model in order to be embedded on-board. Our work is based on a modelling approach using *MDE* and safety engineering. Two safety cases were presented and allowed to validate our solution. The first is expressed textually in natural language to describe a side collision case study. The second provided a structural assurance case using *GSN* with compliance to *SACM* metamodel.

Safety rules are integrated to the map conceptual model and this allows are to automate their incorporation on-board and safety decisions management. Our solution offers an on-board safety-extended model for the railway infrastructure. The conceptual clarification and matching of different perspectives, namely safety analysis, railway infrastructure modeling and safety management allow a structured safety integration based on an ontological framework.

In future works, we intend to extend the proposed approach by integrating the requirement engineering concepts and to provide an operational solution for requirements traceability. This aspect is important in the system development process especially with dynamic aspect of safety requirements. Furthermore, we aim to reuse this approach for other components of future railway systems and validate the on-board application of the autonomous train map. Finally, we will investigate the formal verification aspect in order to check the safety rules consistency and the safety justification.

References

[1] S. Debbech, P. Bon, S. Collart-Dutilleul, Improving safety by integrating dysfunctional analysis into the design of railway systems, WIT Transactions on The Built Environment 181 (2018) 399–411.

[2] CENELEC, NF EN 50129, Applications ferroviaires - Systèmes de signalisation, de télécommunication et de traitement -Systèmes électroniques de sécurité pour la signalisation (2003).

[3] ISO/DIS 26262-1, Road vehicles - Functional safety - Part 1 Glossary (2009).

[4] S. Debbech, S. Collart-Dutilleul, P. Bon, An ontological approach to support dysfunctional analysis for railway systems design, Journal of Universal Computer Science (J.UCS) 26 (5).

[5] G. Guizzardi, Ontological foundations for structural conceptual models, Ph.D. thesis, University of Twente, Enschede, The Netherlands (2005).

[6] CENELEC, NF EN 50126-1, Applications ferroviaires : Spécification et démonstration de la fiabilité, de la disponibilité, de la maintenabilité et de la sécurité (FMDS)-partie 1 (2017).

[7] S. Debbech, P. Bon, S. Collart-Dutilleul, Conceptual Modelling of the Dynamic Goal-Oriented Safety Management for Safety Critical Systems, in: ICSOFT 2019-14th International Conference on Software Technologies-Volume 1, 2019, pp. 287–297.

[8] A. A. El Kalam, S. Benferhat, R. El Baida, C. Saurel, P. Balbiani, Y. Deswarte, G. Trouessin, et al., Organization based access control, in:

The IEEE 4th International Workshop on Policies for Distributed Systems and Networks (POLICY 2003), Lake Come, Italy, June, IEEE, 2003, p. 120.

[9] R. de Almeida Falbo, Sabio: Systematic approach for building ontologies., in: 1^{st} Joint Workshop ONTO.COM / ODISE on Ontologies in Conceptual Modeling and Information Systems Engineering. FOIS, Rio de Janeiro, 2014.

[10] S. Debbech, S. Collart-Dutilleul, P. Bon, Cas d'étude de mission ferroviaire télé-opérée, Rapport de recherche, IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (Nov 2018).
URL https://hal.archives-ouvertes.fr/hal-02020997/1

[11] RailTopoModel, UIC International Railway Standard IRS 30100, UIC, The Worldwide Railway Organisation (2016).

[12] A. Rodrigues da Silva, Model-driven engineering: A survey supported by the unified conceptual model, Computer Languages, Systems & Structures 43 (2015) 139 – 155.

[13] MDA Guide revision 2, OMG Document ormsc/14-06-01, Object Managment Group (june 2014).

[14] Unified Modeling Language v2.5, OMG Norm, Object Management Group (Mar. 2015).

[15] N. Chouchani, M. Abed, Automatic generation of personalized applications based on social media, Procedia Computer Science 170 (2020) 825 – 830, the 11th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops.

[16] K. Lano, S. Kolahdouz-Rahimi, S. Yassipour-Tehrani, M. Sharbaf, A survey of model transformation design patterns in practice, Journal of Systems and Software 140 (2018) 48 – 73.

[17] omg, Structured assurance case metamodel v2.1.
URL https://www.omg.org/spec/SACM/2.1

[18] R. Wei, T. P. Kelly, X. Dai, S. Zhao, R. Hawkins, Model based system assurance using the structured assurance case metamodel, CoRR abs/1905.02427.

[19] IEC 61508, Norme Internationale, Sécurité fonctionnelle des systèmes électriques électroniques programmables relatifs à la sécurité (2000).

[20] M. O'connor, H. Knublauch, S. Tu, B. Grosof, M. Dean, W. Grosso, M. Musen, Supporting rule system interoperability on the semantic web with swrl, in: International semantic web conference, Springer, 2005, pp. 974–986.

[21] T. Kelly, R. Weaver, The goal structuring notation–a safety argument notation, in: Proceedings of the dependable systems and networks 2004 workshop on assurance cases, 2004.

[22] Bureau d'Enquêtes sur les Accidents de Transport Terrestre, (BEA-TT), Rapport d'enquête technique sur l'accident ferroviaire survenu à Saint-Romain-En-Gier le 5 Avril 2004 (Nov 2004).
URL http://www.bea-tt.developpement-durable.gouv.fr/saint-romain-en-gier-english-summary-a15.html

[23] S. Rangra, M. Sallak, W. Schön, F. Belmonte, Risk and safety analysis of main line autonomous train operation: context, challenges and solutions, in: Congres Lambda Mu 21 de Maitrise des Risques et de Surete de Fonctionnement, 2018.

[24] K. Philip, W. Mickael, Autonomous vehicle safety: An interdisciplinary challenge, IEEE Intelligent Transportation Systems Magazine 9 (1) (2017) 90–96.

[25] R. Kianfar, P. Falcone, J. Fredriksson, Safety verification of automated driving systems, IEEE Intelligent Transportation Systems Magazine 5 (4) (2013) 73–86.

[26] ISO/PAS 21448, Safety of the intended functionality (2019).

[27] S. Zhang, F. Boukamp, J. Teizer, Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (jha), Automation in Construction 52 (2015) 29–41.

[28] Z. Rehman, C. V. Kifor, An ontology to support semantic management of fmea knowledge., International Journal of Computers, Communications & Control 11 (4).

[29] J. Zhou, K. Hänninen, K. Lundqvist, L. Provenzano, An ontological interpretation of the hazard concept for safety-critical systems, in: The 27th European Safety and Reliability Conference ESREL'17, 18-22 Jun 2017, Portoroz, Slovenia, 2017, pp. 183–185.

[30] B. Tenbergen, T. Weyer, K. Pohl, Hazard relation diagrams: a diagrammatic representation to increase validation objectivity of requirements-based hazard mitigations, Requirements Engineering 23 (2) (2018) 291–329.

[31] X. Xing, B. Zhong, H. Luo, H. Li, H. Wu, Ontology for safety risk identification in metro construction, Computers in Industry 109 (2019) 14–30.

[32] A. Lališ, R. Patriarca, J. Ahmad, G. Di Gravio, B. Kostov, Functional modeling in safety by means of foundational ontologies, Transportation Research Procedia 43 (2019) 290–299.

[33] K. Clegg, M. Li, D. Stamp, A. Grigg, J. McDermid, Integrating existing safety analyses into sysml, in: IMBSA, 2019.

[34] D. Rao, P. Pathrose, F. Huening, J. Sid, An approach for validating safety of perception software in autonomous driving systems, in: Y. Papadopoulos, K. Aslansefat, P. Katsaros, M. Bozzano (Eds.), Model-Based Safety and Assessment, Springer International Publishing, Cham, 2019, pp. 303–316.

[35] K. Berkenkötter, U. Hannemann, Modeling the Railway Control Domain Rigorously with a UML 2.0 Profile, in: J. Górski (Ed.), Computer Safety, Reliability, and Security, no. 4166 in Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2006, pp. 398–411.

[36] M. Bosschaart, E. Quaglietta, B. Janssen, R. M. P. Goverde, Efficient formalization of railway interlocking data in RailML, Information Systems 49 (2015) 126–141.

[37] C. Xiangxian, H. Yulin, H. hai, A component-based topology model for railway interlocking systems, Mathematics and Computers in Simulation 81 (9) (2011) 1892 – 1900.

[38] F. Mecitoğlu, M. T. Söylemez, A UML Modelling Approach for a Railway Signalization System Simulator and SCADA System, IFAC Proceedings Volumes 46 (25) (2013) 77–82.

[39] M. Perin, L. Wouters, Using Ontologies for Solving Cross-Domain Collaboration Issues, in: IFAC Proceedings Volumes, 19th IFAC World Congress, Vol. 47, 2014, pp. 7837–7842.

[40] T. Eiter, H. Füreder, F. Kasslatter, J. X. Parreira, P. Schneider, Towards a Semantically Enriched Local Dynamic Map, International Journal of Intelligent Transportation Systems Research 17 (1) (2019) 32–48.

[41] S. Ulbrich, T. Nothdurft, M. Maurer, P. Hecker, Graph-based context representation, environment modeling and information aggregation for auto-

mated driving, in: 2014 IEEE Intelligent Vehicles Symposium Proceedings, 2014, pp. 541–547.

[42] M. R. Barry, Certware: A workbench for safety case production and analysis, in: 2011 Aerospace Conference, 2011, pp. 1–10.

[43] E. Denney, G. Pai, Tool support for assurance case development, Automated Software Engg. 25 (3) (2018) 435–499.

[44] Y. Matsuno, H. Takamura, Y. Ishikawa, A dependability case editor with pattern library, in: 2010 IEEE 12th International Symposium on High Assurance Systems Engineering, 2010, pp. 170–171.

[45] P. Bishop, R. Bloomfield, A methodology for safety case development, in: F. Redmill, T. Anderson (Eds.), Industrial Perspectives of Safety-critical Systems, Springer London, London, 1998, pp. 194–203.

Practical Hybrid Confidentiality-based Analytics Framework with Intel SGX

Abstract

Massive cloud infrastructure capabilities, including efficient, scalable, and elastic computing resources, have led to a widespread adoption of Internet of Things (IoT) cloud-enabled services. This involves giving complete control to cloud service providers (CSPs) of sensitive IoT data by moving data storage and processing in cloud. An efficient and lightweight advanced encryption standard (AES) cryptosystem can play a major role in protecting IoT data from exposure to CSPs by protecting the privacy of outsourced data. However, AES lacks computation capabilities, which is a critical factor that prevents individuals and organizations from taking full advantage of cloud computing services. When Intel software guard extensions (SGX) is used with AES cryptosystem, the developing framework can provide a practical solution to build a confidentiality-based data analytics framework for IoT-enabled applications in various domains. In this paper, a privacy-preserving data analytics framework is developed that relies on a hybrid-integrated approach, in which both software- and hardware-based solutions are applied to ensure confidentiality and process-sensitive outsourced data in the cloud environment.

Keywords: Cloud computing, Confidentiality, Data clustering, Intel SGX, Internet of Things

1. Introduction

The advent of Internet of Things (IoT) and edge computing has opened numerous dimensions in technology and prompted researchers to innovate at a rapid rate. IoT technology is developing quickly and has introduced serious concerns about data privacy and integrity. With IoT, the volume of data production and the sharing of data among worldwide networks is unparalleled. As more organizations, private and public, are acquiring IoT to provide solutions in health care, sustainability and other vital sectors, the need for cloud adoption is also increasing. They are bound to obtain the cloud services for storing, managing, and processing massive amounts of data. The cloud services shorten the delivery time for solutions, thereby increasing productivity. Another significant benefit is the analysis and visualization of data for timely and informed decisions, promoting efficiency.

With all these advantages of cloud ecosystem, there is an increasing number of attacks and risks associated with it that can lead to the exposition of highly sensitive data. This creates additional challenges to fundamental aspects of data confidentiality, availability, and integrity (Zissis and Lekkas, 2012). Further, immense dependence on third-party cloud providers presents a risk of corruption, illegal exposure, and misuse of organization-owned data (Sundareswaran et al., 2012, Ren et al., 2012). The extant literature confers different strategies and frameworks to eradicate the problem of data protection and preservation in an outsourced (public cloud-based) environment. The techniques include strict access-control rules, implementation of different anonymization methods and application of multi-party computation (MPC) (Atallah et al., 2001, Wang et al., 2010, Zhou et al., 2011, Chadwick and Fatema, 2012, Backes et al., 2013, Li et al., 2014). However, these techniques are limited to providing privacy-preservation solutions in a specific context, excluding the power of data computation. Even if they possess the computational capability, they are either not intelligent enough or too expensive to provide constructive data analysis for informed decisions.

The objective of this paper is to develop a practical and efficient framework for the adaption of confidentiality-based data analysis in various domains in the realm of IoT. The developed framework aims to build a hybrid privacy-preservation solution that combines both software- and hardware-based techniques to maintain data confidentiality in volatile and untrusted

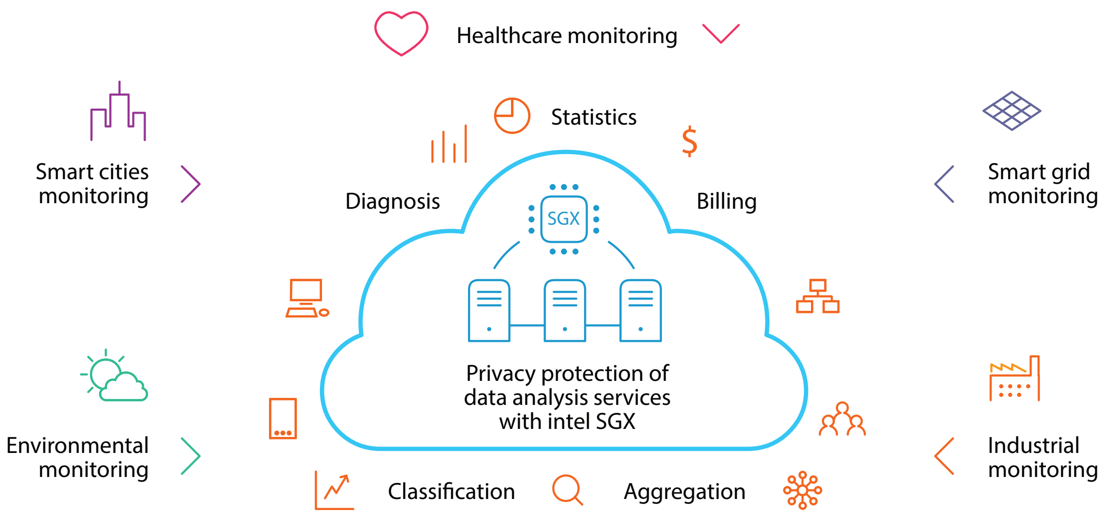


Figure 1: Overview of a secure data analytic approach for IoT cloud-enabled framework using Intel SGX.

cloud environments. The framework comprises techniques, including advanced encryption standard (AES) (Nechvatal et al., 2001) and Intel as software guard extensions (SGX) (McKeen et al., 2013). The practical implications of AES cipher are acknowledged worldwide with regard to protection of digital data, but it does not encompass analytical computation capabilities. An alternative is homomorphic cryptosystems. However, these are either impractical or cost heavy at a large scale. The latest versions of Intel processor generations—starting from $6th$ to the currently $10th$ generation—come with the Intel SGX component that has a security feature developed to ensure the confidentiality of outsourced data at the hardware level. To overcome these limitations, SGX provides the migration of processing and data storage to an isolated memory compartment to perform computations securely without compromising data confidentiality. This embedded framework can be beneficial for end-to-end confidentiality-based data computations across IoT domains, such as health care and smart-grid applications. Figure 1 represents a blueprint of the proposed secure data analytic framework. Applications that require processing sensitive data in various domains can benefit from the proposed framework, such as e-health diagnosis and assisted-living systems, through which patients' sensitive data can be processed efficiently while ensuring confidentiality. Further, industrial-scale applications (e.g., machine process and smart-grid monitoring systems) generate sensitive data from an industrial espionage perspective, in which disclosing this data can reveal sensitive customer data. These realistic scenarios of possible sensitive data disclosure can be eliminated when the data are stored and processed based on the proposed hybrid confidentiality-based analysis framework.

1.1. Motivation

According to Right Scale's cloud survey, (Flexera, 2019), 91% of enterprises outlined public cloud adoption in 2019 alone. According to Gartner report, the public cloud market investment is expected to increase by 17% in 2020 to reach 266.4 billion up from 227.8 billion in 2019. This shows the impact of rapid migration of cloud services, especially for small- and medium-sized enterprises as they equip them with essential resources for data storage and development within a small budget. While there is no doubt of the potential of cloud computing, offering cost-effective and reliable resources to organizations, several security and privacy concerns in the cloud ecosystem need to be addressed (Grobauer et al., 2010). With IoT in the frame, the need to develop privacy-preservation frameworks focused on processing and exchange of data to and from cloud resources has become of prime importance to ensure the protection of sensitive data. Ensuring the privacy of migrating data is critical to the realization of the full potential and advantages of cloud resources.

1.2. Contributions

The main contributions of this paper are as follows.

1. The development of a practical and hybrid confidentiality-based data analytics framework that combines the software AES cryptosystem and hardware Intel SGX-based security solutions to ensure end-to-end privacy protection at all phases of data communication, processing, and storage.
2. The evaluation of the developed framework in terms of analysis performance and accuracy. The experimental outcomes show that the proposed framework achieves a high level of accuracy of the overall analysis process similar to the insecure version of analysis tasks while ensuring full confidentiality protection for the data being processed in cloud computing.

The rest of the paper is further divided into the following. The literature review is presented in Section 2. The architecture of the developed framework is shown in Section 3. The threat model and applied machine-learning techniques are explained in Section 4 and 5. Section 6 presents the security discussion, while section 7 focuses on experimental evaluation. The concluding remarks are presented in Section 8.

2. Literature review

This section presents the prevailing research entailing secure data analytics techniques and Intel SGX implications.

Several approaches are adopted by researchers for preservation of privacy in data analytics models. The randomization- and cryptography-based approaches are widely utilized. Randomization-based approaches mask the data by adding random noise, thereby protecting data in processing phase (Agrawal and Srikant, 2000, Du and Zhan, 2003). However, to mask the data, these approaches also reduce the analytical accuracy by tampering the original data with noise Patel et al. (2015). The evidence of formal methods for security provisioning is also lacking. Conversely, the cryptography-based approaches lean on the MPC for data analysis (Goldreich, 2005). Though the discussed cryptography approaches can achieve a high level of privacy provisioning, the overhead costs and increased computation complexity are inevitable. The authors of (Inan et al., 2007, Doganay et al., 2008, Rivest et al., 1978) discussed three cryptography techniques: oblivious transfer, secret sharing, and homomorphic encryption. Oblivious transfer and secret sharing are not applicable for larger datasets because of high computation and communication costs (Duan and Canny, 2014). In contrast, homomorphic encryption techniques can perform complex computations on encrypted datasets and have two categories, as mentioned in (Gamal, 1985, Gentry and Halevi, 2011) (i.e., somewhat homomorphic encryption and fully homomorphic encryption). However, it is also deemed impractical at a large scale because of the increased cost and complexity. This paper focuses on developing a practical hybrid-analytical framework that will take advantage of both software- and hardware-based solutions. Advanced Encryption Standard (AES) (Daemen and Rijmen, 2020) is a well-known cryptosystem that has been proven and adopted world wide. AES cryptosystem can be used effectively to protect sensitive data, while it is at rest, or during transmission between different entities. Several approaches have been developed to enhance the efficiency of AES cryptosystems as in (Oukili and Bri, 2017, Rao Rupanagudi et al., 2019, Langenberg et al., 2020). AES cryptosystems have been applied in various domains, such as healthcare and smart grids, to ensure the confidentiality of sensitive data. Recently, there has been a shift toward developing hardware-based solutions for providing protection. The aim is to add another layer at the hardware level to enhance the secrecy of data processing. These solutions are termed trusted execution environments (TEE). The Intel SGX is leveling up as a competent TEE that can provide elite privacy with reduced costs associated with data analytic computations in cloud environment. The authors of (Schuster et al., 2015) explained how SGX has been applied in the Hadoop MapReduce framework for big data processing. The application of Intel SGX was also described by (Zheng et al., 2017) as building a distributed data analytics service with oblivious computing. In (Hunt et al., 2018), it was stated that Ryoan—a distributed sandbox specific to untrusted computations on sensitive data—has utilized SGX to improve its own effectiveness and security. As observed in previous research, there are several standalone solutions to overcome the problem of privacy-preserving analytic services. However, this paper has presented a practical hybrid approach that combines software- and hardware-based framework to provide end-to-end protection in the IoT outsourced data analytics environment. Unlike the existing solution, the developed framework aims to support the efficient implementation of various advanced analytics models, in a completely automated cloud-based platform, while taking full advantage of a cloud-computing environment, including storage and processing resources, that in turn will offer unlimited capabilities for adapting various analytical service applications, without compromising data privacy.

3. Hybrid confidentiality-based analytics framework

This section presents the proposed hybrid confidentiality-based analytics framework. This involves describing the entities, their roles, and how the entities interact to accomplish analysis tasks of sensitive IoT data in a privacy-preservation manner in the cloud.

The architecture of the proposed framework has three main entities:

- **Remote (edge) entity:** This is the data source. It can be either an

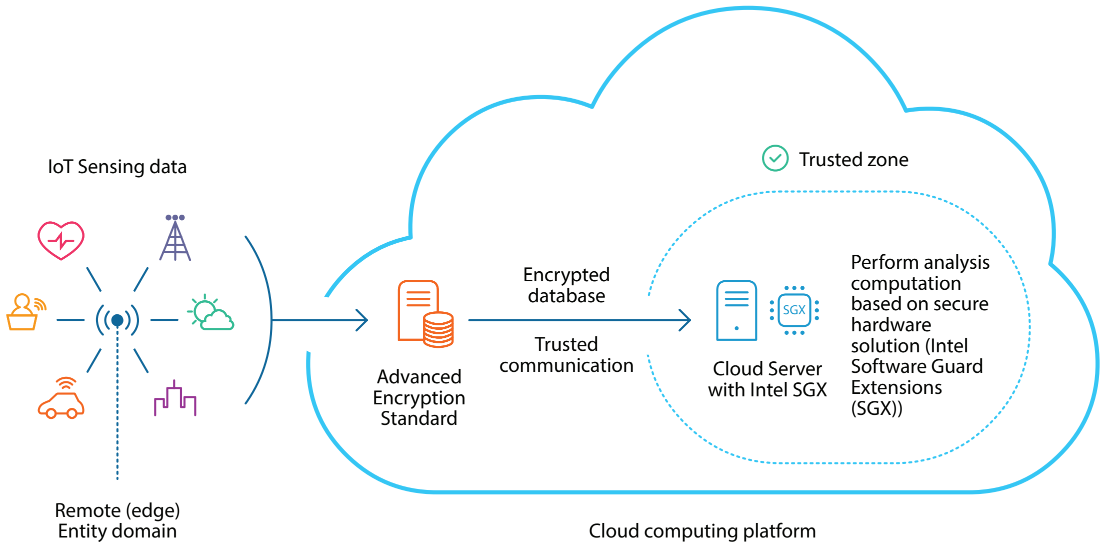


Figure 2: Overview of the proposed hybrid confidentiality-based analytics framework for IoT cloud-enabled framework using Intel SGX.

end-user or a sensor-enabled IoT device in which data are collected and later disseminated to cloud storage.

- **Cloud storage entity:** This is the storage place for the data coming from edge devices. The data are in encrypted form, using an AES cryptosystem.
- **Analytic engine entity:** This is the fundamental entity of the proposed framework. In this entity, the encrypted data in cloud storage are manipulated using data-clustering techniques.

The framework entities collaborate to aggregate, store, and perform data analysis tasks while providing end-to-end privacy. The developed framework comprises two main zones of the developed framework, including a trusted zone (trusted zone as shown in Figure 2). In the trusted zone, an isolated SGX is used to perform analysis tasks for applied analytic models including KMC and FCMC algorithms. For this, ECALL functions are used as a trusted component of SGX architecture to implement analytic models. The untrusted zone is assumed to be completely exposed to the adversary. Therefore, the AES cryptosystem (assuming the cryptosystem parameter initial-

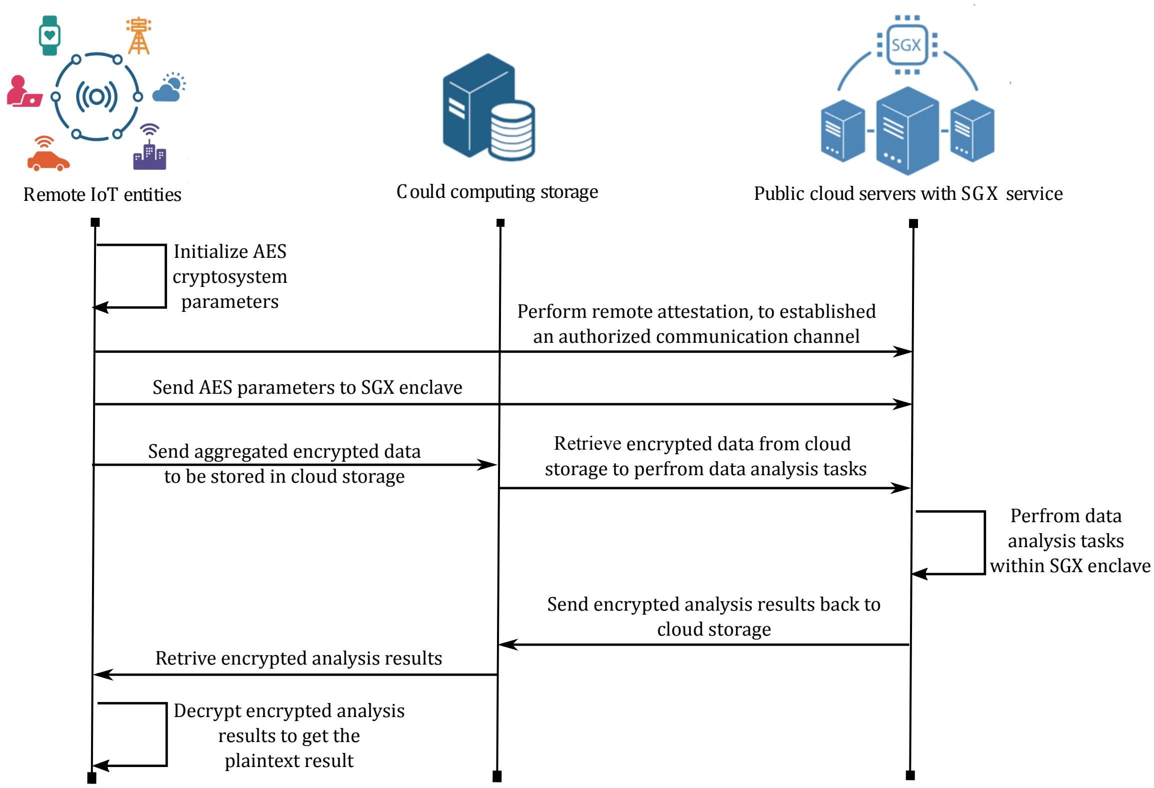


Figure 3: Workflow model of the proposed hybrid privacy-preserving analytics framework.

ization occurs in the secure remote edge entity) and the aggregated data from the remote edge entity that are transmitted for processing inside the SGX enclave. The remote (edge) entity can retrieve analysis results, for which OCALL functions are employed.

Regarding the communication channel between the trusted and untrusted zones, the remote attestation, an advanced feature of Intel SGX, plays a critical role to established an authorized communication channel between the SGX enclave and the remote (edge) entity to exchange encryption/decryption parameters and to facilitate any further data exchange, as shown in Figure 3. The remote attestation ensures a secure communication channel for sending sensitive collected to cloud storage and retrieving analysis results. The remote attestation includes three main services: verifying the identity of the analysis services within an SGX enclave, verifying their correctness (ensuring they have not been tampered with), and ensuring that analysis services run securely within an enclave on an Intel SGX-enabled platform. After the remote attestation process is completed, the encrypted data are sent to the cloud storage entity. The analytic engine entity can complete data processing independently. Data owners (individuals or enterprises) can retrieve the encrypted result through the cloud resource and present it to the beneficiaries through dedicated and secure sites. Later in Section 5 , the data analytic entity is discussed in detail, showing five clustering techniques as a proof of concept for the IoT cloud-enabled paradigm. The overall workflow model is shown in Figure 3.

4. Threat model

Before discussing the entities for the proposed framework, an assumption is made to shape the threat model—that the remote entity (i.e., end-user and edge devices) are secure to collect and receive the sensing IoT data. The rest of the model entities are vulnerable to internal and external threats. Therefore, identification of a security mechanism is essential to make the proposed framework resilient enough to withstand any compromise. This section will shed light on the way users' sensitive data and associated analytical operations will be protected through the complete lifecycle of end-to-end communication in IoT ecosystem.

4.1. Remote (edge) Entity and Communication Channel

It has already been stated that the communication channel to and from the remote entity is not secure, despite the remote devices being secure themselves. It is essential to transfer data between the devices and storage entity in an encrypted form. To achieve this, a privacy-protection mechanism must be devised to exchange the highly sensitive information between the remote entities and Intel SGX enclave. Remote attestation can establish a secure communication channel with the remote entity. This enables the remote secure entity to transfer AES cryptographic primitives to the SGX enclave securely. It is assumed that the adversary cannot compromise the secure enclaves and their relevant keys—in this case, seal, and attestation keys. Advanced side-channel attacks, as in (Chen et al., 2020, Murdock et al., 2020), can be prevented by applying current defense techniques, as in (Orenbach et al., 2020). However, this concern, along with physical and denial-of-service attacks on the remote entities, are beyond the scope of this article.

4.2. Data Analytic Entity

As discussed previously, the processing component of the proposed framework, the analytic engine entity, is used to perform the computational tasks. The primary feature of the proposed framework is that the computational tasks will be performed inside the Intel SGX architecture. We also assume that the computations are processed inside the SGX enclave environment. It is further assumed that the cloud service provider (CSP) is a semi-honest party that follows framework transactions but attempts to gain more information than is allowed. The SGX enclaves hosted by CSPs are assumed to be isolated completely from BIOS, I/O, and even power of cloud servers, which are considered potentially untrustworthy. Further, an adversary may control computing resources or software, such as operating systems or hypervisors, to attack the protected analysis processes. Therefore, it is assumed that the analysis functions that run inside the enclaves are the only trusted components. The analytic based clustering computations are only dependent on built-in C/C++ libraries within SGX enclave environments. Particularly, the only computations implemented are standard arithmetic operations supplemented with exponentiation and polynomial evaluations of the initial inputs, along with intermediate results through which SGX enclaves completely assist these operations. Therefore, assuming that the SGX internal state is secure implies that the analysis computations processing inside SGX enclave are also secure.

5. Analytic services-based data clustering

Data-clustering analysis is used to categorize objects (data points) that share similar properties into different groups called clusters. For initial exploration of input data, data clustering is deemed a popular technique. It is used in various fields, including image analysis, pattern recognition, information retrieval and bioinformatics. In this paper, two principal centroid-based clustering algorithms are applied as proof of concept for the proposed model, including K-means clustering (KMC) and fuzzy C-means (FCM) clustering algorithms. The procedural steps for both algorithms are illustrated next.

KMC can be accomplished as follows and is diagrammatically presented in Figure 3.

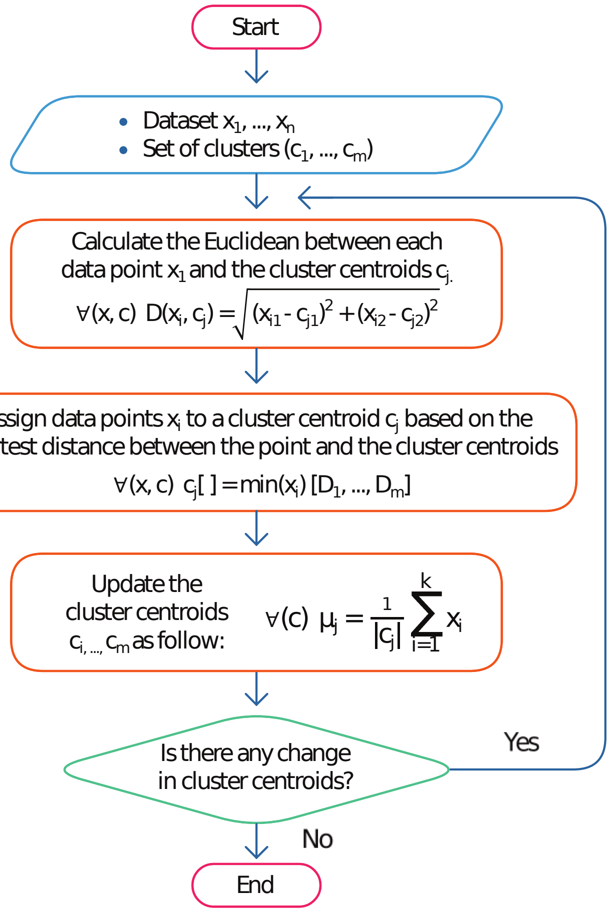


Figure 4: The procedural steps of K-means clustering algorithm.

1. Let $x_1, \ldots, x_n$ be a set of two-dimensional data points. The algorithm randomly selects a set of cluster centroids $c_1, \ldots, c_m$.
2. Calculate the Euclidean distance between each data point x_i and the cluster centroids c_j.

$$\forall(x,c) \quad D(x_i, c_j) = \sqrt{(x_{i1} - c_{j1})^2 + (x_{i2} - c_{j2})^2} \tag{1}$$

3. Assign data points x_i to a cluster centroid c_j based on the shortest distance between the point and the cluster centroids.

$$\forall(x,c) \quad c_j[\ ] = min(x_i)[D_1, \ldots, D_m] \tag{2}$$

4. Update the cluster centroids $c_i, \ldots, c_m$.

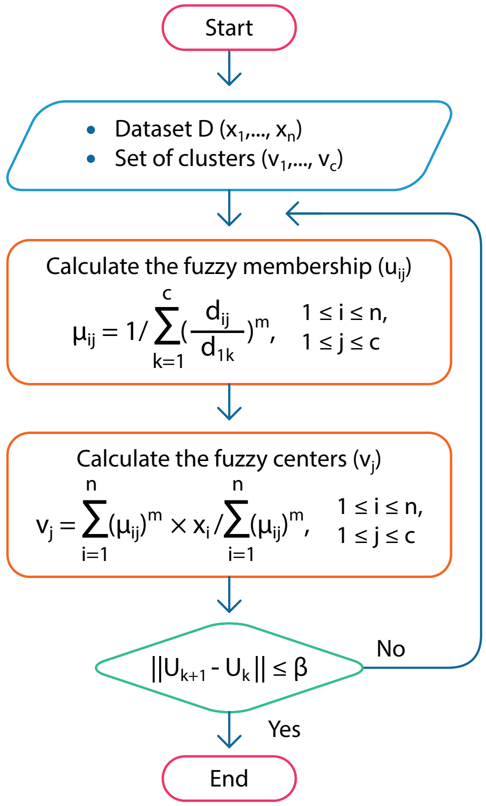


Figure 5: The procedural steps of fuzzy c-means clustering algorithm.

$$\forall(c) \quad \mu_j = \frac{1}{|c_j|}\sum_{i=1}^{k} x_i \tag{3}$$

Where k is the number of data points that are assigned to a cluster centroid c_j and μ_j is the updated mean of a cluster centroid c_j.

5. Repeat Steps 2,3 and 4 until there is no longer change in the updated cluster centroids.

The FCM clustering algorithm can be accomplished as follows and is diagrammatically presented in Figure 4.

1. Data objects are assigned to possible clusters based on calculated membership matrices.

$$\mu_{ij} = 1/\sum_{k=1}^{c}(\frac{d_{ij}}{d_{1k}})^m \quad (4)$$

Where μ_{ij} is a membership value between a data object i and a cluster centroid j. d_{ij} is an Euclidean distance between a data object i and a cluster centroid j as shown in Equation 1.

2. Cluster centroids are updated by calculating the new means of data objects in the current clusters through the following function:

$$\nu_j = \sum_{i=1}^{n}(u_{ij})^m x_i / \sum_{i=1}^{n}(u_{ij})^m \quad (5)$$

where ν_j is the j^{th} cluster.
The membership values of data points and cluster centroids are updated based on Equations 4 and 5 until the following condition is satisfied:

$$||U^{k+1} - U^k|| < \beta \quad (6)$$

where U is $(\mu)_{n*c}$ the fuzzy membership matrix and β is the termination criterion value that is pre-determined.

6. Security discussion

The developed hybrid privacy-preservation analysis framework aims to protect the privacy of aggregated IoT-based data and perform analysis tasks securely to prevent any malicious activities. Thus, the developed framework is secured against the threat model. In the event of an eavesdropping-based attack on the communication channel between remote entities and Intel SGX enclaves, a possible adversary could only intercept protected data through encryption, when an AES cryptosystem is applied on aggregated sensed data upon receipt to ensure its confidentiality. Further, the injection of illegitimate key material during communication can be another attack that also not possible for the attacker with Intel's SGX attestation process. The supporting defense layer effectively mitigates such vulnerabilities. This type of compromise is sometimes referred to as the Eve mechanism and was first observed as a vulnerability for naive Diffie–Hellman.

In the case of eavesdropping attacks targeting Intel SGX enclaves, the only known feasible methods to eavesdrop the sensitive data from protected the SGX enclave memory are the spectre techniques, such as an adversary being able to launch side-channel attacks. Developed schemes, SCONE (Arnautov et al., 2016) and Varys (Oleksenko et al., 2018) can be deployed to overcome such attacks. Moreover, the patterns of memory access can compromise the privacy of data during data exchange and inside enclave (Sasy et al., 2018). Therefore, analytic models, such as machine-learning algorithms, can be implemented based on oblivious techniques to eliminate and execute data-dependent patterns (Ohrimenko et al., 2016). After discussing the security of individual entities in the proposed framework, the research can conclude that the entire system is secure. There is no computationally feasible mechanism to extract either data or results from the system, except with negligible probability.

7. Experimental Evaluation

In this section, a set of varying experiments are conducted to assess the functionality and performance of the proposed framework. For these experiments, the primary data mining algorithms used are KMC and FCMC algorithms. The performance of adapted AES cryptosystem and communication overhead of exchanging encrypted data between IoT device (in this case, Raspberry Pi node) and Intel SGX enclave are evaluated in detail. Furthermore, clustering-based algorithms are implemented and used for plaintext and ciphertext versions comparison. The plaintext implementations are used as a baseline against the measurement of encrypted system. Two fundamental questions are asked:

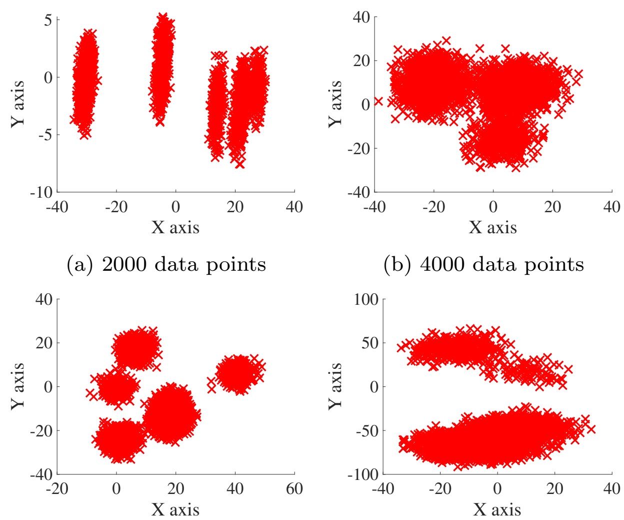
1. Do the developed privacy-preservation analytic models (KMC and FCMC algorithms) achieve high level of analytic accuracy compared with the plaintext versions of analytic models?

2. What are the relative performance overheads between the developed privacy-preservation analytic models and the plaintext versions of analytic models?

This section outlines the results obtained after series of experiments with observed comparisons between functionality and performance.

7.1. Datasets

The developed framework is evaluated using a public set of benchmark clustering datasets. These datasets are specifically designed for cluster analysis and consider varying characteristics (Franti and Virmajoki, 2006). They are represented in Figure 6. The datasets consist of 2000, 4000, 6000, and 8000 two-dimensional data points with corresponding class labels and numerous 12 centroid clusters with different degrees of overlap. To demonstrate various aspects of the proposed framework, the datasets are divided into subsets to examine the analytic accuracy and performance overheads with varying dataset sizes.



(a) 2000 data points (b) 4000 data points

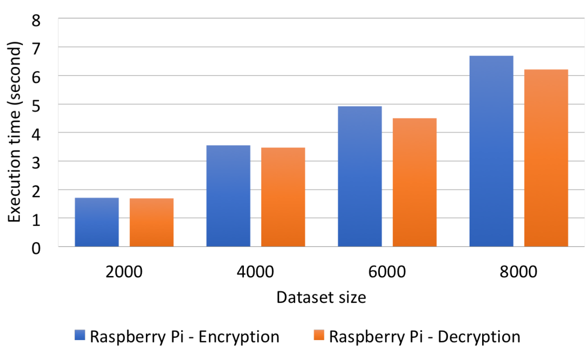
(c) 6000 data points (d) 8000 data points

Figure 6: The distribution of two-dimensional synthetic datasets. The datasets consist of 2000, 4000, 6000, and 8000 two-dimensional data points with corresponding class labels and varying number of cluster centroids with different degrees of cluster overlap.

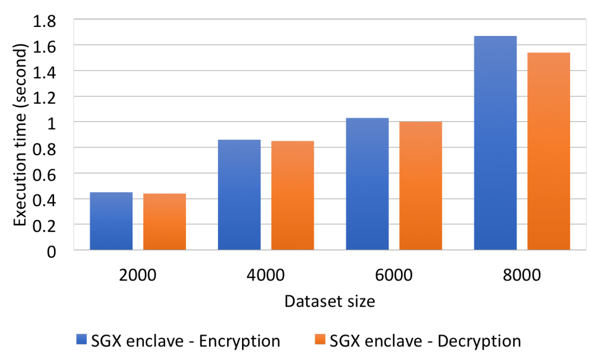
7.2. Experimental Setup

To demonstrate the experimental evaluation, we deployed a server on Microsoft Azure. We used the DCsv2 series machines, which offers SGX-enabled processors. Intel ®Xeon CPU ®E-2288G @ 3.70 GHz with 8 cores and 32 GiB RAM, running on Ubuntu 20.04 OS is used with a processor supports 256MB of enclave size (a total usable memory of 168MB). Moreover, Raspberry Pi 3 with 4 GB memory is used to collect and send aggregated data to the Intel SGX enclave. It is of interest to measure the performance and functionality of a complete developed encrypted-based data analytic framework and the corresponding plaintext version of analytic models.

The experiments comprise several phases. First, in the initialization phase, the AES cryptosystem encryption/decryption key material is generated. Second, during the key sharing phase, remote attestation is enabled to transfer key material. Third, during the encryption phase, the datasets are encrypted in the remote IoT entity. Fourth, in the transmission phase, the encrypted data are sent to the secure Intel SGX processing unit. Fifth, during the data analysis phase, the Intel SGX processing unit decrypts the data that are transferred in the second phase and performs the analysis tasks before encrypting the analysis results. Sixth, during the receiver phase, the encrypted results are transmitted back to the remote entity. In the final phase, the results are decrypted for any further processing tasks in the remote secure entity.



(a) Raspberry Pi



(b) Cloud-based SGX enclave

Figure 7: Execution time of AES encryption and decryption processes in both IoT-based Raspberry Pi and cloud-based SGX enclaves with varying dataset sizes.

7.3. Performance Metrics

The performance evaluation demonstrates two main criteria: analysis task accuracy and performance overheads. Figure 7 shows the extracted execution times for the developed privacy-preservation analytic framework for both encryption and decryption processes with varying dataset sizes for both IoT-based Raspberry Pi and cloud-based SGX enclaves. Overall, IoT-based Raspberry Pi takes longer to process compared with cloud-based Intel

SGX enclaves because of the limited resource capabilities of IoT-based devices. Further, it is observed that the developed privacy-preservation analytic framework and corresponding plaintext versions of KMC and FCMC algorithms produce identical analysis results regarding analytic accuracy. The result is as expected since the presence of encryption in each part of the data transmission and data receiver phases will not modify the values of the raw data. Moreover, the analysis processing of the developed framework is performed in plaintext version inside the SGX enclave, which results in similar analysis results.

From the performance perspective, the notable differences can be observed in KMC and FCM algorithms' execution time, including data encryption at remote entity, data transmission, decryption, and analysis tasks, and finally send the encrypted results back to secure remote entity. This is directly proportional to the dataset size and number of clusters. These differences are represented in Figure 8. For example, it has been observed that the KMC algorithm takes an average time of 193 milliseconds for $2,000$ data points while it takes 824 in FCMC algorithm for the same dataset size. Further, the KMC algorithm performs analysis tasks for $6,000$ data points in about 266 milliseconds, while it takes 1693 in FCMC algorithm for the same dataset size. The FCMC algorithm has a higher performance overhead for the analysis tasks compared with the KMC algorithm, which is related the computation complexity of the FCMC algorithm compared with the KMC algorithm.

Regarding Intel SGX enclave memory usage for storing encrypted data, a dataset of $2,000$ encrypted data points consumes around 608 kilobytes of memory while the memory size increases in linear relation to the size of input dataset, as shown in Figure 9. Finally, one of the main obstacles in building SGX-based solutions for analytic models is the communication overhead, which is an essential component of analytic processes in which data are sent inside the SGX enclave through a secure established communication channel with third parties. Figure 10 shows the approximate communication overhead between the remote IoT entity and the cloud server based on the size of the dataset, which provides a visible insight into the developed model's capabilities and limitations. For instance, it takes approximately 41 milliseconds to transmit $2,000$ data points and approximately 82 milliseconds to transmit $4,000$ data points. This shows a linear increase in the communication overhead with the size of input dataset, as shown in Figure 10.

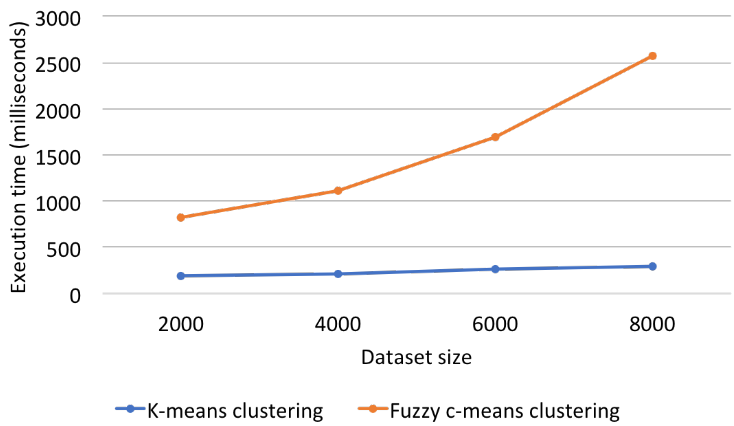


Figure 8: Execution time for processing privacy-preservation KMC and FCMC algorithms with varying dataset sizes.

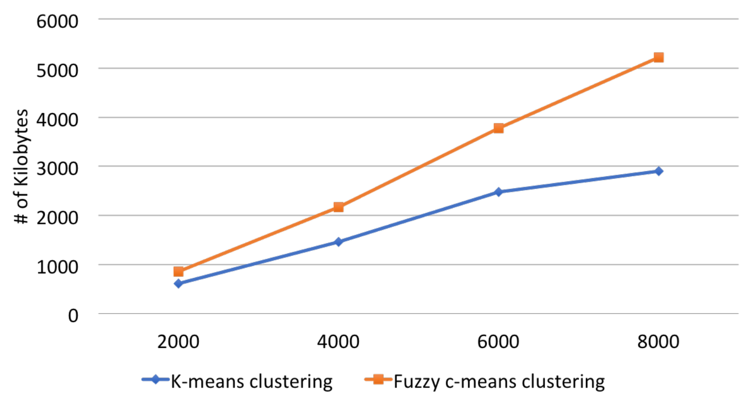


Figure 9: The memory usage of Intel SGX enclave with varying dataset sizes.

8. Conclusion

In this paper, a practical hybrid confidentiality-based analytic framework is based on Intel SGX. It relies on a hybrid-integrated model, including both software- and hardware-based solutions, to ensure the confidentiality and process sensitivity of outsourced data in the cloud environment. The developed framework aims to provide secure data-analytic services for IoT-enabled applications in various domains, such as smart grid and healthcare applications. The experimental evaluation shows a high level of analysis accuracy in a privacy-preserving manner, while indicating differences in execution times and processing overheads. The developed framework can be adapted efficiently for various analytical service applications, to take advantage of public

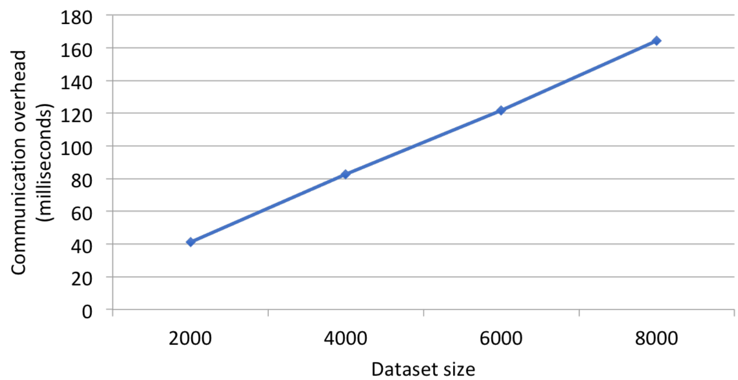


Figure 10: Communication overheads for exchanging encrypted data with Intel SGX encrypted datasets of varying sizes.

cloud computing without compromising data privacy. Future research will focus on building more advanced analytical models, in order to overcome challenges such as communication and storage limitations, because of their complexity in both computational and analytical structure.

References

Agrawal, R. and Srikant, R. (2000). Privacy-preserving data mining. In Chen, W., Naughton, J. F., and Bernstein, P. A., editors, *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 16-18, 2000, Dallas, Texas, USA*, pages 439–450. ACM.

Arnautov, S., Trach, B., Gregor, F., Knauth, T., Martin, A., Priebe, C., Lind, J., Muthukumaran, D., O'Keeffe, D., Stillwell, M., Goltzsche, D., Eyers, D. M., Kapitza, R., Pietzuch, P. R., and Fetzer, C. (2016). SCONE: secure linux containers with intel SGX. In Keeton, K. and Roscoe, T., editors, *12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016, Savannah, GA, USA, November 2-4, 2016*, pages 689–703. USENIX Association.

Atallah, M. J., Pantazopoulos, K. N., Rice, J. R., and Spafford, E. H. (2001). Secure outsourcing of scientific computations. *Advances in Computers*, 54:215–272.

Backes, M., Fiore, D., and Reischuk, R. M. (2013). Verifiable delegation of computation on outsourced data. In *2013 ACM SIGSAC Conference on Computer and Communications Security, CCS'13, Berlin, Germany, November 4-8, 2013*, pages 863–874.

Chadwick, D. W. and Fatema, K. (2012). A privacy preserving authorisation system for the cloud. *Journal of Computer and System Sciences*, 78(5):1359–1373.

Chen, G., Chen, S., Xiao, Y., Zhang, Y., Lin, Z., and Lai, T. (2020). Sgxpectre: Stealing intel secrets from SGX enclaves via speculative execution. *IEEE Security & Privacy*, 18(3):28–37.

Daemen, J. and Rijmen, V. (2020). *The Design of Rijndael - The Advanced Encryption Standard (AES), Second Edition*. Information Security and Cryptography. Springer.

Doganay, M. C., Pedersen, T. B., Saygin, Y., Savas, E., and Levi, A. (2008). Distributed privacy preserving k-means clustering with additive secret sharing. In *Proceedings of the 2008 International Workshop on Privacy and Anonymity in Information Society, PAIS 2008, Nantes, France, March 29, 2008*, pages 3–11.

Du, W. and Zhan, J. Z. (2003). Using randomized response techniques for privacy-preserving data mining. In Getoor, L., Senator, T. E., Domingos, P. M., and Faloutsos, C., editors, *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, pages 505–510. ACM.

Duan, Y. and Canny, J. F. (2014). Practical distributed privacy-preserving data analysis at large scale. In Gkoulalas-Divanis, A. and Labbi, A., editors, *Large-Scale Data Analytics*, pages 219–252. Springer.

Flexera (2019). Cloud computing trends: 2019 state of the cloud survey.

Franti, P. and Virmajoki, O. (2006). Iterative shrinking method for clustering problems. *Pattern Recognit.*, 39(5):761–775.

Gamal, T. E. (1985). A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472.

Gentry, C. and Halevi, S. (2011). Implementing gentry's fully-homomorphic encryption scheme. In Paterson, K. G., editor, *Advances in Cryptology - EUROCRYPT 2011 - 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tallinn, Estonia, May 15-19, 2011. Proceedings*, volume 6632 of *Lecture Notes in Computer Science*, pages 129–148. Springer.

Goldreich, O. (2005). Foundations of cryptography - A primer. *Foundations and Trends in Theoretical Computer Science*, 1(1).

Grobauer, B., Walloschek, T., and Stocker, E. (2010). Understanding cloud computing vulnerabilities. *IEEE Security & privacy*, 9(2):50–57.

Hunt, T., Zhu, Z., Xu, Y., Peter, S., and Witchel, E. (2018). Ryoan: A distributed sandbox for untrusted computation on secret data. *ACM Transactions on Computer Systems*, 35(4):13:1–13:32.

Inan, A., Kaya, S. V., Saygin, Y., Savas, E., Hintoglu, A. A., and Levi, A. (2007). Privacy preserving clustering on horizontally partitioned data. *Data & Knowledge Engineering*, 63(3):646–666.

Langenberg, B., Pham, H., and Steinwandt, R. (2020). Reducing the cost of implementing the advanced encryption standard as a quantum circuit. *IEEE Transactions on Quantum Engineering*, 1:1–12.

Li, J., Huang, X., Li, J., Chen, X., and Xiang, Y. (2014). Securely outsourcing attribute-based encryption with checkability. *IEEE Transactions on Parallel and Distributed Systems*, 25(8):2201–2210.

McKeen, F., Alexandrovich, I., Berenzon, A., Rozas, C. V., Shafi, H., Shanbhogue, V., and Savagaonkar, U. R. (2013). Innovative instructions and software model for isolated execution. In *HASP 2013, The Second Workshop on Hardware and Architectural Support for Security and Privacy, Tel-Aviv, Israel, June 23-24, 2013*, page 10.

Murdock, K., Oswald, D., Garcia, F. D., Bulck, J. V., Gruss, D., and Piessens, F. (2020). Plundervolt: Software-based fault injection attacks against intel SGX. In *2020 IEEE Symposium on Security and Privacy, SP 2020, San Francisco, CA, USA, May 18-21, 2020*, pages 1466–1482. IEEE.

Nechvatal, J., Barker, E., Bassham, L., Burr, W., Dworkin, M., Foti, J., and Roback, E. (2001). Report on the development of the advanced encryption standard (aes). *Journal of Research of the National Institute of Standards and Technology*, 106(3):511.

Ohrimenko, O., Schuster, F., Fournet, C., Mehta, A., Nowozin, S., Vaswani, K., and Costa, M. (2016). Oblivious multi-party machine learning on trusted processors. In Holz, T. and Savage, S., editors, *25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016*, pages 619–636. USENIX Association.

Oleksenko, O., Trach, B., Krahn, R., Silberstein, M., and Fetzer, C. (2018). Varys: Protecting SGX enclaves from practical side-channel attacks. In Gunawi, H. S. and Reed, B., editors, *2018 USENIX Annual Technical Conference, USENIX ATC 2018, Boston, MA, USA, July 11-13, 2018*, pages 227–240. USENIX Association.

Orenbach, M., Baumann, A., and Silberstein, M. (2020). Autarky: closing controlled channels with self-paging enclaves. In Bilas, A., Magoutis, K., Markatos, E. P., Kostic, D., and Seltzer, M. I., editors, *EuroSys '20: Fifteenth EuroSys Conference 2020, Heraklion, Greece, April 27-30, 2020*, pages 7:1–7:16. ACM.

Oukili, S. and Bri, S. (2017). High speed efficient advanced encryption standard implementation. In *2017 International Symposium on Networks, Computers and Communications, ISNCC 2017, Marrakech, Morocco, May 16-18, 2017*, pages 1–4. IEEE.

Patel, S. J., Punjani, D., and Jinwala, D. C. (2015). An efficient approach for privacy preserving distributed clustering in semi-honest model using elliptic curve cryptography. *International Journal of Network Security*, 17(3):328–339.

Rao Rupanagudi, S., Vidya J, V., Bhat, V. G., Padmavathi, P., Darshan, G., Gurikar, S. K., Darshan, S., and Sindhu, N. (2019). A further optimized mix column architecture design for the advanced encryption standard. In

2019 11th International Conference on Knowledge and Smart Technology (KST), pages 181–185.

Ren, K., Wang, C., and Wang, Q. (2012). Security challenges for the public cloud. *IEEE Internet Computing*, 16(1):69–73.

Rivest, R. L., Shamir, A., and Adleman, L. M. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126.

Sasy, S., Gorbunov, S., and Fletcher, C. W. (2018). Zerotrace : Oblivious memory primitives from intel SGX. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018*. The Internet Society.

Schuster, F., Costa, M., Fournet, C., Gkantsidis, C., Peinado, M., Mainar-Ruiz, G., and Russinovich, M. (2015). VC3: trustworthy data analytics in the cloud using SGX. In *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, pages 38–54. IEEE Computer Society.

Sundareswaran, S., Squicciarini, A., and Lin, D. (2012). Ensuring distributed accountability for data sharing in the cloud. *IEEE transactions on dependable and secure computing*, 9(4):556–568.

Wang, J., Zhao, Y., Jiang, S., and Le, J. (2010). Providing privacy preserving in cloud computing. In *3rd International Conference on Human System Interaction*, pages 472–475.

Zheng, W., Dave, A., Beekman, J. G., Popa, R. A., Gonzalez, J. E., and Stoica, I. (2017). Opaque: An oblivious and encrypted distributed analytics platform. In Akella, A. and Howell, J., editors, *14th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2017, Boston, MA, USA, March 27-29, 2017*, pages 283–298. USENIX Association.

Zhou, M., Mu, Y., Susilo, W., Au, M. H., and Yan, J. (2011). Privacy-preserved access control for cloud computing. In *IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2011, Changsha, China, 16-18 November, 2011*, pages 83–90.

Zissis, D. and Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation Computing Systems*, 28(3):583–592.