



Energy Research and Development Division

FINAL PROJECT REPORT

A Semantically Integrated Operational Dashboard for Management of a Smart Grid

Gavin Newsom, Governor November 2019 | CEC-500-2019-059

PREPARED BY:

Primary Authors:

Kimberly Garcia Jack Hodges Florian Michahelles

Siemens Corporation 1936 University Avenue Suite 320 Berkeley, California <u>Siemens Corporation website</u> (https://www.siemens.com/us/en/home.html)

Contract Number: EPC-15-046

PREPARED FOR:

California Energy Commission

Qing Tian, PhD, P.E. **Project Manager**

Fernando Piña Office Manager ENERGY SYSTEMS RESEARCH OFFICE

Laurie ten Hope Deputy Director ENERGY RESEARCH AND DEVELOPMENT DIVISION

Drew Bohan Executive Director

DISCLAIMER

This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees, or the State of California. The Energy Commission, the State of California, its employees, contractors and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission passed upon the accuracy or adequacy of the information in this report.

ACKNOWLEDGEMENTS

The Artificial and Human Intelligence Research Group at Siemen Corporate Technology acknowledges the sponsors of this project: the California Energy Commission and its Electric Program Investment Charge. Thanks to Qing Tian for his project management support through this research project.

Moreover, those responsible for this project acknowledge the following colleagues for their domain expertise, advisory, and support at different stages of this project:

- Dr. Ralf Mosshammer at Siemens Corporate Technology in Vienna, Austria
- Dr. Alfred Einfalt at Siemens Corporate Technology in Vienna, Austria
- Dr. Konrad Diwold at Pro2Future in Graz, Austria

Finally, special thanks to Dr. Simon Mayer at University of St. Gallen, Switzerland, for his hard work on this project and setting it up for success.

PREFACE

The California Energy Commission's Energy Research and Development Division supports energy research and development programs to spur innovation in energy efficiency, renewable energy and advanced clean generation, energy-related environmental protection, energy transmission and distribution and transportation.

In 2012, the Electric Program Investment Charge (EPIC) was established by the California Public Utilities Commission to fund public investments in research to create and advance new energy solutions, foster regional innovation and bring ideas from the lab to the marketplace. The California Energy Commission and the state's three largest investor-owned utilities—Pacific Gas and Electric Company, San Diego Gas & Electric Company and Southern California Edison Company—were selected to administer the EPIC funds and advance novel technologies, tools, and strategies that provide benefits to their electric ratepayers.

The Energy Commission is committed to ensuring public participation in its research and development programs that promote greater reliability, lower costs, and increase safety for the California electric ratepayer and include:

- Providing societal benefits.
- Reducing greenhouse gas emission in the electricity sector at the lowest possible cost.
- Supporting California's loading order to meet energy needs first with energy efficiency and demand response, next with renewable energy (distributed generation and utility scale), and finally with clean, conventional electricity supply.
- Supporting low-emission vehicles and transportation.
- Providing economic development.
- Using ratepayer funds efficiently.

A Semantically Integrated Operational Dashboard for the Management of a Smart Grid is the final report for the Developing a Distribution Substation Management System project (Grant Number EPC-15-046) conducted by Siemens Corporation. The information from this project contributes to the Energy Research and Development Division's EPIC program.

For more information about the Energy Research and Development Division, please visit the <u>Energy Commission's website</u> (www.energy.ca.gov/research/) or contact the Energy Commission at 916-327-1551.

ABSTRACT

As the electric distribution system becomes increasingly complex with integrating more energy resources, existing distribution automation systems must be enhanced with functions to manage more renewable energy connected at the distribution level and provide greater control over the operation of these energy resources. Distribution management systems must automate more in monitoring and controlling operations at substations to respond quickly to problems and reduce outage times. In this project, Siemens Corporation has developed a semantically integrated management system for future secondary substations. This system provides an unprecedented level of operational automation with more reliable services, which can benefit utilities and ratepayers. This research project created an operational dashboard for future electrical distribution substations capable of downloading and installing software from a marketplace. The conceived innovative system provides grid operators with an intuitive and real-time visualization of the current state of the grid and (potential) problems. When the system detects a problem, the system suggests the grid operator download, install, and configure a specific piece of software (application) capable of fixing the problem detected on the faulty substation. Moreover, a grid operator is provided with a global view of the cluster being managed, which allows the operator to react faster to detected anomalies and prevent severe problems such as outages.

Keywords: Smart grid, operational dashboard, smart grid operators, grid reliability, knowledge models, future secondary substations, augmented reality

Please use the following citation for this report:

Garcia, Kimberly, Jack Hodges, and Florian Michahelles. 2019. *A Semantically Integrated Operational Dashboard for the Management of a Smart Grid*. California Energy Commission. Publication. Number: CEC-500-2019-059.

TABLE OF CONTENTS

ACKNOWLEDGEMENTSi
PREFACEii
ABSTRACTiii
TABLE OF CONTENTSiv
LIST OF FIGURES
LIST OF TABLES vii
EXECUTIVE SUMMARY
Introduction
Project Purpose
Project Approach
Project Results
Advancing the Research to Market4
Benefits to California5
CHAPTER 1: Introduction7
Smart Grid of the Future7
The Future Secondary Substation7
Semantic Technologies8
User Interfaces
CHAPTER 2: Project Approach10
Semantically Integrated Operational Dashboard Architecture
Grid Knowledge Models11
Future Secondary Substation Knowledge Models12
FSSN Applications13
The Open Semantic Framework15
The Data Stream Classifier16
The Data Stream Classifier Backend16
The Data Stream Classifier Frontend17
The Application Placement Engine19
Model Formula20

Objective Function Components: Performance Metrics	21
App Placement Program	23
Global FSSN Manager: Approach	23
The Semantically Integrated Operational Dashboard	24
The 2D Dashboard	25
The Augmented Reality App	25
CHAPTER 3: Project Results	28
Data Streams Classifier	28
Testing Method	28
Test Results	29
Classifier Performance	30
Application Placement Engine	30
Testing Scenarios	31
Correctness	33
Parameter Effectiveness	36
Resource Conservation vs Replication	36
Load Balancing versus Minimal App Movement	37
Time versus Quality Tradeoff	39
The Semantically Integrated Operational Dashboard	43
CHAPTER 4: Technology/Knowledge Transfer Activities	49
CHAPTER 5: Conclusions	51
CHAPTER 6: Benefits to Ratepayers	53
LIST OF ACRONYMS	55
REFERENCES	56

LIST OF FIGURES

	Page
Figure 1: Architecture Overview	10
Figure 2: Extensions to the Operational Smart Grid Knowledge Model	13
Figure 3: The Classification Frontend Displays Available Gridline Data Points (left), Known Data Point Classes (right), and Observed Data Points (center)	17

Figure 4: The 2D Dashboard25
Figure 5: Storyboard Example for Installing Apps in Substation to Alleviate Anomaly26
Figure 6. Prototyping the Augmented Reality App27
Figure 7: Augmented Reality App Appearance During the User Tests
Figure 8: Classification Errors for Different Settings of the SAMPLE_LENGTH Parameter
Figure 9: Placement for Scenario 2
Figure 10: Placement for Scenario 2.1
Figure 11: Placement for Scenario 2.2
Figure 12: Placement for Scenario 3
Figure 13: Placement Emphasizing Resource Conservation
Figure 14: Placement Emphasizing Replication of Frequently Used Apps
Figure 15: Initial Placement for Load Balancing versus Minimal Movement Experiment 38
Figure 16: Placement Emphasizing Load Balancing
Figure 17: Placement Emphasizing Minimal App Movement
Figure 18: Placements for Scenario 2 After 1, 5, 10 and 15 Seconds (from top left to bottom right)41
Figure 19: Placements for Scenario 3 After 1, 5, 10, 30 Seconds, 30 Minutes and 14 Hours (from top left to bottom right)42
Figure 20. Installing an App Under Normal Conditions43
Figure 21: Review and Confirmation of the Installation of an App44
Figure 22: Installation of the Adaptative Assignment Module App44
Figure 23: Installation of the Gridlink XMPP Stack Onto Available Substations45
Figure 24: Voltage Violation in North Berkeley45
Figure 25: Dashboard Shows a Problem in a Substation46
Figure 26. Apps Suggested to Fix a Voltage Violation46
Figure 27: Using the App Placement Engine to Compute a Global Installation Plan47
Figure 28: Using the App Placement Engine to Compute a Global Installation Plan47
Figure 29: Installation of an App Completed48

LIST OF TABLES

Table 1: Applications Considered in the Knowledge Models	14
Table 2: Scenario 2, App Description	31
Table 3: Scenario 3 App Descriptions	33
Table 4: Execution Times	40
Table 5. Estimate of Reduction in Costs in PG&E Distribution Grid	54

EXECUTIVE SUMMARY

Introduction

Currently large quantities of data are produced everywhere, such as humans using applications, devices capable of sensing different environmental conditions, and machines able to record their operation and performance. However, the data produced is only collected and used in a restricted way to provide a very specific service. Some of the most popular trends in artificial intelligence are mainly concerned in processing and analyzing these large amounts of data to give meaning to it. Such methods give a degree of accuracy of the possible meaning of each piece of data, this understood data is then correlated with other probably relevant data that might help create more complex services.

Semantic technologies are a set of methods and tools providing advanced ways to process data and can automatically understand the data produced. These technologies use a bottom-up approach, instead of trying to make sense of all the data and coping with errors that probabilistic approaches have by nature. Semantic technologies use formal languages to create knowledge models that give meaning to different raw data and are built based on such knowledge models, which describe the system in a way that the data produced has a meaning and a context within the system. Moreover, by promoting and adopting standardized knowledge models, applications and systems of the same domain can interact. Semantic technologies create a common understanding among users, applications, and platforms, helping users collaborate better with systems, and machines, hence those systems and machines communicate, and act autonomously without human help that interprets data.

A "smart grid" combines "electrical" and "intelligent" infrastructures to the electrical systems to improve resiliency, reliability, flexibility, security, and efficiency. Substations are controlled and operated by supervisory control and data acquisition systems, which are enhanced by distribution management systems. On other end of the distribution grid, Internet of things (IoT)-enabled devices (smart meters, smart breakers) have been used to optimize the power system and guarantee a reliable cost-effective and carbon dioxide (CO₂)-neutral operation.

However, between substations and the other end of the grid, there is an intermediate point that can be automated to increase the resilience of the grid. This intermediate point corresponds to secondary substations that convert voltage from medium to low so that it can be used by customers. This project considered future secondary substations: smart hardware that can perform sophisticated functions, such as voltage and reactive power control (voltage control by the operator), and optimizing distributed generation over the low-voltage grid using specialized software that can be downloaded from a secured marketplace. The project developed an operational dashboard to manage future secondary substations effectively. The dashboard provides a useful tool for grid operators to react faster to detected anomalies and prevent severe problems such as outages in the low-voltage grid.

Project Purpose

This research project developed an operational dashboard display for smart grid distribution substations based on semantic technologies. This dashboard can display the current grid state and warnings for detected problems, and it can automatically suggest solutions to the found problems. This project is to achieve an unprecedented level of automation for the management and operation of smart grids. The specific objectives of this research project were to:

- Create an operational smart grid knowledge model to streamline the operation of the grid.
- Demonstrate the value of the developed model within an operational management interface that promotes the local operational control of individual distribution substations.
- Extend the semantically integrated operational dashboard with the capability to manage future secondary substations at a global level, across individual substations.

Project Approach

This research takes a semantics-based approach to provide the flexibility that the rapidly changing smart grid requires. It consists of a set of descriptions, expressed in a machine understandable way, of the future secondary substations, the software that controls them, and the software that can be installed in such substations. These descriptions provide meaning and give context to data that are produced in the grid. These descriptions make it possible to identify anomalies and problems and to recommend actions to resolve them.

To create the semantically integrated operational dashboard, Siemen Corporate Technology formed a team at its Artificial and Human Intelligence Research Group. The team had a common interest and background in knowledge modeling, knowledge representation, and user interfaces, but with different finer focuses. Moreover, the team conducted this research project under the advice of another Siemens research group in Vienna working on a smart cities project in Austria. Such a team is in charge of creating the software (applications) to help manage future secondary substations. Guided by its expertise and domain knowledge, the team was able to overcome the technical challenge regarding the relevant information that should be modeled in a machine in an understandable way. Moreover, the strategic alliance with a team located in Princeton, New Jersey, complemented the knowledge representation skills and expertise required for developing a global view for a set of managed substations. Finally, given that the objective of the project is to reduce the cognitive overload of grid operators at managing substations, an innovative augmented reality user interface was explored. Augmented reality is a technology that allows users to superimpose computergenerated images on the user's view of the real world through the usage of devices such as augmented reality glasses, smartphones or tablets. Thus, grid operators are able to use their physical space as their workspace without being limited to a monitor.

This project was divided in six technical tasks, with four main products:

- Knowledge models for managing the smart grid. These knowledge models are the foundation of this research project.
- Semiautomatic data classifier. This software component allows grid operators to create and train statistical classifiers that will be used to automatically classify the type of data stream coming to the system (such as voltage and current).
- Application placement engine. Given a set of constraints specified by the smart gird operator, this software component is in charge of computing a feasible installation of specialized software within the set of future secondary substations that the grid operator manages.
- Semantically integrated dashboard (2D and augmented reality user interface). This dashboard constitutes the final product of this research project, in which all the other software components come together to provide the grid operator with a user interface capable of making him or her aware of detected anomalies, suggest possible solutions, and let him or her install applications in a future secondary substation.

Since this project is basic research, all the developed components were tested individually to ensure they provided the required functionality, the performance of these components was adequate, and they were easy to install and use. For the classifier, Siemens used data from a Siemens SICAM P855 multifunctional device, as well as randomly generated data for switching decisions, and a set from the Lawrence Berkeley National Lab was used.

Project Results

This research project developed a semantically integrated operational dashboard for electrical substations. The tool provides grid operators with an intuitive and real-time visualization of the current grid state and (potential) grid problems. Moreover, the dashboard can suggest solutions to encountered problems and provide a global view of the set of substations that a grid operator is managing. Thus, the increasingly complex smart grid becomes more accessible to operators.

Although the results of the proposed project could be transferable to other types of grid nodes, the focus of this research was on distribution substations (secondary substations) for which main tasks are voltage transformation, voltage regulation, and the isolation of faults in the transmission or distribution systems. Those future secondary substations are "smart" hardware capable of downloading and installing apps from a secured market place in the cloud to help monitor and operate the grid (voltage band violations).

The main outcome of this project is a system that can suggest solutions to grid operators to solve anomalies detected in the grid. The team successfully validated each software component developed within this project in terms of its ease of installation and usability. Moreover, the software components' functionality was validated by building a simulated infrastructure that is constantly producing *real-time* data. This data is categorized by a statistical classifier, which feeds the knowledge models in charge of monitoring the data, raises alarms, and suggests solutions in case operators detect anomalies in the grid. The grid operator interacts with the system through either the two-dimensional or the augmented reality dashboard created in this project. These dashboards allow grid operators to go through the full managing process, from monitoring incoming grid data, to deciding if the suggested application to fix a potential problem in the grid is valid for the current problem, and whether the system computed installation plan for a set of future secondary substations is adequate.

The team built two versions of the dashboard in this project with the objective of decreasing the cognitive overload that grid operators experience in analyzing and assessing the data that traditional operational dashboards provide. Thus, the reaction time to detected anomalies can be reduced resulting in a higher probability of preventing outages. This project helps build the path to a self-healing grid, given that at the current implementation, the project aims to support gird operators. However, operator input is only needed to confirm the plan computed by the developed system. In future research, this step could be skipped, making the execution of the plan fully automatic.

Advancing the Research to Market

To raise awareness and demonstrate the benefits that adopting cutting-edge technologies such as semantic technologies could bring in smart grid settings, the team held several meetings with Siemens stakeholders from the Energy Management business unit and received positive feedback. Different versions of the dashboard were also presented in internal Siemens conferences, such as Siemens ConneCTs 2017, and the Data Analytics and Artificial Intelligence Conference 2018. Outside of Siemens, the team presented the developed work to members of the California Institute for Energy and Environment and a research group from the Lawrence Livermore National Laboratory. Following are some of the outreach activities done during the project:

- Met with Omnetric-Siemens looking for partners to apply for pilot project.
- Met with members of the California Institute for Energy and Environment to introduce them to this research.
- Met with Siemens stakeholders from Energy Management business unit.

- Met with the Lawrence Livermore National Laboratory research group looking for pilot partners.
- Presented the project at Siemens ConneCTs 2017 in Princeton, New Jersey.
- Distributed brochures at the EPIC symposium 2018 in Sacramento, California.
- Presented HoloLens prototype at Siemens DAAI Conference (Data Analytics & Artificial Intelligence) in Nuremberg, Germany.

The technical advisory committee for the project includes academics from the Lawrence Livermore National Laboratory and stakeholders from the Siemens Energy Management business unit. The committee held its first meeting in September 2017, and the project results presented included the knowledge models, data classifier, and mechanism for automatically finding applications to solve detected problems. The team received positive comments on the innovation of the approach, along with feedback about alternative hardware that could be used in a smart grid setting to acquire grid data and improve the accuracy of the data classifier.

During the latest steps of this project, the team presented and published several technical papers. The Siemens Corporate Technology research team is looking for partners at Siemens and at California utilities to conduct a pilot project and user studies of the developed two-dimensional and augmented reality dashboards with grid operators.

Benefits to California

By implementing the proposed automated approach in an actual smart grid setting, ratepayers would benefit from better electricity reliability and lower costs. Using this automated system could help utilities improve their operations, enhance their outage management, and reduce electrical losses through faster control decisions. Moreover, it can allow utilities to take better advantage of asset use derived from reducing manual grid operation efforts and could improve grid monitoring capabilities for better planning.

Adopting these technologies could have the potential to yield a reduction of the distribution grid System Average Interruption Duration Index value in the Pacific Gas and Electric Company (PG&E) distribution grid (excluding major events) of about 10 percent, which would translate to cost savings of more than \$1 billion per year to PG&E customers, including more than \$134 million per year to PG&E residential customers. Even using a more conservative estimate of 5 percent, the potential yearly economic benefit of the developed project would be upward of \$67 million, considering only California residential customers. Moreover, more automation in grid control rooms would also allow available staff to focus on services and locations of the grid, in which the highly automated proposed system cannot recover on its own.

These benefits can directly affect private, commercial, and industrial consumers who can profit from more reliable services and reduced business losses due to improved grid resiliency as well as potential bill savings thanks to a higher degree of automation,

which also implies a lower grid cost operation. More automation within the smart grid also implies benefits for society, since it could allow more renewable energy resources, reduce environmental emissions, improve the security of electricity delivery, and reduce the import of crude oil through transportation electrification.

CHAPTER 1: Introduction

Smart Grid of the Future

The current power distribution grid is controlled and operated by supervisory control and data acquisition systems (SCADA). Such systems have been enhanced by distribution management systems (DMS) that provide useful tools such as network, and fault location and control. Even though the degree of automation of the distribution grid is high up to substation level, the low voltage grid remains poorly equipped, providing hardly any measurements that can help managing the low voltage grid efficiently. IoTenabled power grid devices, such as smart meters, smart breakers, and smart storage systems have been highly adopted to optimize the power system and guarantee a reliable cost-effective and carbon dioxide (CO2)-neutral operation. However, these IoT components are at the end of the distribution chain. A complementary approach, and perhaps of higher ecological impact can be obtained from turning the currently passively operated secondary substations to active ones, capable of providing functions that impact the lower voltage grid, and provide services that consider external information, such as forecasting and weather information. To this end, Siemens Corporate Technology in Europe, specifically in Austria, in conjunction with the Austrian Technology Institute have proposed the Intelligent Secondary Substation (Faschang, et al. 2017), which is a smart secondary substation capable of performing sophisticated functions, such as voltage and re-active power control, and distributed generation optimization over the low voltage grid. To perform these functions, a Smart Grid Application (a piece of software) is installed in an Intelligent Secondary Substation (such as Voltage Guard Application).

This research project builds on top of the current efforts being done by the Siemens research groups in Europe. Thus, Smart hardware called Future Secondary Substations (FSSNs) are considered as the equivalent American smart hardware of the Intelligent secondary substations from Europe. In this research, a semantically integrated operational dashboard for Future Secondary Substations is proposed. This dashboard provides managing and control tools for grid operators to react quickly to the different conditions in the low voltage grid.

The Future Secondary Substation

A FSSN is a smart hardware on the low distribution grid. Applications (Apps) are the main functional elements of the FSSN system. These apps are software downloaded and installed in a FSSN from a market place (Appstore) to remedy detected grid problems as well as for performing routine monitoring tasks. Conception and implementation of these apps is beyond the scope of this project. These Apps are created in a joint effort by the Siemens groups in Vienna, Austria and the Austrian

Institute of Technology. The link between the semantically integrated operational dashboard developed in this project and the actuals Apps that perform an action over the grid, is the Operational Smart Grid Knowledge Model, a component of the dashboard developed within this project that contains "cognitive twins" of the Apps. Such "cognitive twins" are human-readable description as well as machine-interpretable information about the Apps.

Semantic Technologies

In today's world, there are large quantities of data been produced everywhere, such as humans using applications, devices capable of sensing different environmental conditions, and machines able to record their operation and performance. However, the data being produced is only collected and exploited in a very restricted manner to provide a very specific service. Some of the most popular trends in Artificial Intelligence are mainly concerned in processing and analyzing these large amounts of data to give meaning to it. Such methods give a degree of accuracy of the possible meaning of each piece of data, this understood data is then correlated with other probably relevant data that might help create more complex services. Thus, the question arises, would not it be possible to create systems that know what the data they produce means? And how can it be related to other relevant data? This would lead to even more powerful systems capable of communicating and interoperating with each other seamlessly to create more sophisticated services.

Semantic technologies originated by the very need of automatically understanding the data that is produced. These technologies use a bottom-up approach, instead of trying to make sense of all the data and coping with errors that probabilistic approaches have by nature. Semantic technologies use formal languages to create knowledge models that give meaning to heterogeneous raw data. Semantic Applications are built based on such knowledge models, which describe the system in a way that the data that is produced has a meaning and a context within the system. Moreover, by promoting and adopting standardized knowledge models, applications and systems of the same domain can interact. Thus, Semantic Technologies create a common understanding among users, applications, and platforms, helping users collaborate better with systems, and machines, hence those systems and machines communicate, and act autonomously without human help that interprets data.

User Interfaces

For more than forty years, the classic 2D Graphic User Interfaces (GUI) have been dominant in most software. In the early days, programmers realized that the success of their software relayed heavily on its ease of interaction, and the appeal of the graphical user interfaces. Thus, it was popularized the constant use of windows, icons, menus, and pointers for desktop computers. Methods to create these 2D user interfaces have evolved with time to simplify users' efforts. Particularly, the introduction of mobile devices had required to come up with different ways to display simplified content in a smaller display, and incorporate other means of interaction, such as taps on touchscreens. Nowadays all users are familiarized with a traditional 2D desktop user interface.

Innovative user interfaces have been on development and are becoming more relevant as the hardware that supports them becomes more mainstream. Virtual Reality (VR) (Whyte, et al. 2000) was proposed as a paradigm for making the access to information much more tangible and interactive. VR systems allow users to experience a virtual world, in which space is unlimited, laws of physics can be excluded to provide an immersive experimentation field to explore properties of objects yet to be build.

On its side, AR allows to superimpose virtual 2D or 3D objects, graphics or sounds in real-world environments, providing users with an immersive experience, and allowing them to use a much broader working space than a standard monitor can offer. The visual and sound resources available for AR app design can reduce the user's cognitive effort. Moreover, the freedom of movement around a room provides a more intuitive user experience. AR provides a 360° immersed into the data and virtual surrounding experience, whilst letting these operators to be part of the real-world environment. Thus, grid operators are still able to interact with other coworkers, decreasing the risk of motion sickness considerably.

CHAPTER 2: Project Approach

Semantically Integrated Operational Dashboard Architecture

The semantically integrated operational dashboard developed within this project can be seen as the control panel for managing a grid of future secondary substations (FSSNs), a FSSN is a secondary distribution substation capable of downloading and installing software (FSSN Apps) from a marketplace located in the cloud. The installed software is designed to have an effect in the gird, such as fixing an anomaly, monitoring, or retrieving information. Since a grid operator is in charge of more than one substation, the semantically integrated operational dashboard provides support for managing a set of FSSNs in a global way. Hence, the status of the grid is considered as a full, rather than individual substations.

Figure 1: Architecture OverviewFigure 1 shows the software architecture for the semantically integrated operational dashboard. Marker A shows the components that were created by other Siemens research groups prior this project. Those components include the: a) gridlink messaging bus (Faschang, et al. 2017), which is a topic-based messaging bus responsible for distributing grid events and data among its participants, b) a software container that runs in each FSSN capable of hosting FSSN Apps, and c) the cloud-based FSSN Appstore that manages the deployment of Apps to FSSNs and their lifecycle for example, configuration and deletion.



Figure 1: Architecture Overview

Source: Siemens CT

Within the scope of this project, a Classification Interface Backend (marker B) and a Classification Interface Frontend (marker G) were developed. The former gets data from the Gridlink bus and statistically classifies Gridlink data points (for example, it can differentiate a data point that contains load profiles from data points that contain voltage data). The Classification Interface Backend relies on statistical classifiers, which are trained with information provided by the grid operator through the Classification Interface Frontend. The classified values are stored in a central operational knowledge model via the Open Semantic Framework (OSF) (marker C) (Mayer, et al. 2017), which is a Web interface to meaningful data. Additionally, the OSF is in charge of processing, and monitoring the incoming data in order to detect grid problems such as voltage violations or overload events. Given a problem detected by the OSF, the FSSN Manager (marker E) determines the FSSN Apps that should be installed, and their configuration parameters. In turn, the Global FSSN Manager (marker F) or App Placement Engine computes optimized app placements across multiple FSSNs given installation parameters defined by the grid operator (for example, resource consumption optimization, or load balancing across the FSSN grid). Finally, the semantically integrated operational dashboard (marker D) is the user interface that brings together all the backend functionalities in order to support grid operators in the process of managing the FSSN grid.

Grid Knowledge Models

The fundamental component of this research project is the machine understandable representations of the elements involved in the management of a network of Future Secondary Substations (FSSNs). Such representations correspond to knowledge models (ontologies) that were customized from existing standardized models (such as SSN¹ ontology and QUDT²) and extended during this research work. As mention in Chapter 1, the objective of a knowledge model is to describe in a deeper level a specific domain (for example, the smart grid) in order to give meaning to the elements involved by stating the relationships and logic constraints among such elements. Knowledge modelling is ideal for achieving the level of adaptivity required for modeling complex, evolving systems with long lifetimes of individual components, such as the smart grid. Using this technology, descriptions of new elements that are not currently into play in the smart grid (such as new distributed energy resources) can be incorporated seamlessly.

As stated before, Semantic technologies are a powerful tool to describe a domain of the world in a machine understandable way. However, to allow interoperability among

^{1 &}lt;u>Semantic Sensor Network Ontology</u> (https://www.w3.org/2005/Incubator/ssn/ssnx/ssn).

² QUTD (http://www.qudt.org/).

different systems based on knowledge models, the creation of ad-hoc models should be kept to a minimum (Hodges, García and Ray 2017). Thus, this research work combined elements from existing ontologies and extended them (when strictly necessary) with specialized knowledge models that cover relevant concepts for the FSSNs use cases, including Gridlink-specific concepts, possible grid problems (for example, Voltage Band Violations), and their corresponding solutions (that is, FSSN apps). As the knowledge foundation, the FSSN ontologies build upon concepts from the QUDT ontology (which describes quantities, units, dimensions, and types) and from the SSN ontology (that describes common concepts in sensor networks). The integration with QUDT is particularly powerful since this ontology describes concepts such as voltages and power values in detail, allowing the identification and labeling of values with proper units. It also describes the properties of values themselves and captures their relationship to other values and data types. On its side, the SSN ontology describes most basic concepts of sensors and data observation. Thus, concepts from SSN that apply to Smart Grids were reused in this work.

Future Secondary Substation Knowledge Models

To understand the value of having a rich description of a concept, let us describe the DataPoint class. This class allows live data to be processed and stored directly in the semantic model. The incoming data is processed by means of SPIN³ rules. SPIN is an industry standard to represent logic constraints on Semantic Web models. Specifically, the DataPoint class is a subclass of SensingDevices (see), it has an identifier and it is responsible for observing time-stamped data values (TimeSeriesNumericOutput).

SPIN rules defined within the DataPoint class associate the observed values with their corresponding DataPoint instance in the knowledge model. In addition, instances of the DataPoint class are associated with a QuantityRange that represents upper and lower acceptable bounds for the values of a specific DataPoint instance. For example, it can be state that data points of the type AmericanLineVoltage may fluctuate in the range of 110V +/-5% not to trigger a violation.

Each DataPoint instance can have one of two states: a state in which all its values stay within the defined constraints and a state in which at least one of its value violates at least one constraint. The goal of the grid operator is to keep DataPoint instances in the non-violating state. The states as well as the goal are represented as instances in the model for each data point. Data point values that violate any of the constraints associated with their DataPoint instance are considered violations. In the knowledge model, this is implemented via associations of DataPoint instances with an (ideal, non-violating) State class (for example, "DataPoint values are in the acceptable range"); States in turn are associated with Goals (for example, "DataPoint values must be within

³ SPARQL Inference Notation (SPIN) (http://spinrdf.org/).

their acceptable range"). Thus, any violation constitutes a violating state of that goal and is a sign of an operating problem in the grid that could escalate into a bigger problem such as a blackout. Such a violation should be taken care of, for instance by installing an FSSN app that is designed to alleviate operational problems of the type that causes the given violation.



Figure 2: Extensions to the Operational Smart Grid Knowledge Model

Source: Siemens CT

SPIN rules are the entities verifying the meaning of the data. They make sure that incoming values are: a) properly categorized by associating them with correct DataPoint instances, b) correctly pre-processed (for example, by averaging over a given time window), and c) validated against the specific constraints associated with their DataPoint instances to detect violations and clean up old values.

FSSN Applications

FSSN Applications are the main functional elements of the FSSN system and are downloaded from a market place or Appstore, configured, and installed into a FSSN. These apps can provide information about the grid, monitor, detect, or remedy grid problems, among other functionalities. However, the conception and implementation of these apps is beyond the scope of this project; they are created in a joint effort by a Siemens Research Group in Corporate Technology in Vienna, Austria, and the Austrian Institute of Technology. The relationship between this research work and the FSSN Applications and Appstore is established by the extended knowledge models, which are "cognitive twins" of deployable applications. Thus, for each application, a subgraph of the knowledge model contains a human-readable description as well as machineinterpretable information about the purpose of the application, its dependencies, and suggested configuration. These "cognitive twins" are then linked to grid problems as means to alleviate them. Thus, when the system detects a problem, it decides on a proper reaction from the applications related to such a problem.

Table 1 shows the four categories of applications that the current knowledge models consider, namely: a) basic functions, b) voltage monitoring and switching decisions, c) topology monitoring, validation and verification, and d) interconnection with building management systems.

Category	Description
Basic Functions	Java Stack: a multi-purpose Java client stack
	Storage/Archiver: it takes streamed data as input and produces a
	time series archive that is stored locally and made available to other
	apps GridLink XMPP Stack: enables other apps to communicate via the
	XMPP protocol
Voltage	Voltage Guard I (VG1): it requires, as input, the prevalent bus bar
monitoring and	voltage at the FSSN and produces, as output, tap switching decisions
decisions	for the FSSN it is deployed on
decisions	Voltage Guard II (VG2): a more sophisticated Voltage Guard, it
	noint to make a switching decision
	Voltage Guard III (VG3): it considers the streamed voltage data from
	the monitored data point and the bus bar, and time-series data about
	the prevalent bus bar voltage at the FSSN to make tap switching
	decisions
Topology	Grid Representation Module (GRM): it keeps a representation of the
Monitoring,	current state of the low-voltage grid, particularly of its topology and
Varidation, and	supplies that information to other apps that depend on this
, or moution	relationships between ESSNs and connected nodes) and the
	"dynamic" grid topology that contains information about the current
	state of switches.
	Topology Verification Module (TOV): it verifies automatically the grid
	topology that is published by the GRM and marked as the currently
	valid dynamic topology.
Interconnection	GridLink Flexop Building Energy Agent (BEA): it enables other apps
With Building	to communicate with buildings
Systems	Building Energy Agent (BEA): It loads data from smart meters in buildings that are not natively Cridlink canable, and stores it using
-,	the Storage app
	Building Representation Module (BRM): it creates digital twins of all
	buildings that are supplied through a specific FSSN.
	Adaptive Assignment Module (AAM): it supplies information about the
	dynamic position of buildings in the low-voltage grid of an FSSN

 Table 1: Applications Considered in the Knowledge Models

Source: Siemens CT

The core function of each app is captured in the ontologies to enable the automatic resolution of smart grid problems, as well as the app dependencies, which correspond to other applications that need to run either locally in the same FSSN (for example, a JavaStack), or remotely in a neighbor FSSN, and the suggested app configuration parameters that could potentially automate the installation of apps in a FSSN. It should be noted that thanks to the evolving characteristics of Semantic Technologies, when new applications are created, their description, and relationship with potential grid problems can be seamlessly integrated to the current knowledge models.

The Open Semantic Framework

The software component in charge of the collection, curation, and access to ontologies that encapsulate knowledge and experience in a machine-understandable way, is the Open Semantic Framework (OSF), a tool created within the Siemens research team in charge of this research project. The OSF forms the basis for enabling automated reasoning and decision making on top of knowledge models and lets semantic applications use domain-specific and general knowledge models. The OSF furthermore tackles several major obstacles to the widespread use of integrated semantic models by supporting individuals who are not versed with ontologies in understanding and extending them, and by making these models more tangible with the help of advanced human–interface technologies.

The following functions are provided by the OSF to facilitate the operational management of FSSNs:

- Knowledge Management: the OSF contains core knowledge models that are relevant across domains (for example, models that capture data types, units, dimensions, etc.) as well as Knowledge Packs (KPs) that encode more specialized (for example, domain-specific) information for usage by specific clients. KPs thus enable vertical interoperability within a domain (for example, within FSSNs and smart meters) whereas their integration with core ontologies ensures horizontal interoperability across domains (for example, between an FSSN and a building management system).
- Knowledge Access: the OSF provides controlled access to stored knowledge both in the core ontologies and in KPs through a querying interface inside a REST API. This interface is based on prefabricated SPARQL query templates that are shipped with a specific KP. Thus, KPs not only determine the knowledge applications can access but also exactly how they access it. The purpose of this mechanism is to prevent unwanted modifications to the knowledge models and to forbid clients to extract all knowledge from the OSF; both aspects are of paramount importance for the commercial viability of any semantic framework. The OSF's Semantic API can also be used to insert new information using construct or update query templates. Within this project, this is particularly useful for integrating semantic models with real-time data streams.

 Knowledge Visualization: visualizing knowledge in an easily accessible and tangible way is beneficial for several purposes, including the semantic validation of ontologies by subject-matter experts and for supporting non-ontologists with the extension of vocabularies. The OSF supports knowledge visualization applications via a specialized, visualization specific KP with queries that enable client applications to explore knowledge models. Existing approaches to visualize semantic models often merely display the underlying schemas, and users can easily get lost because the complexity of conceptual relationships makes it difficult to visualize ontologies on 2D computer screens. Therefore, the usage of interfaces that support 3D interaction to produce practically usable visualizations of knowledge models is encouraged.

For this project, the smart grid knowledge models are packaged as a KP together with predefined queries (for example, for the insertion and loading of data points, and the identification of an app that resolves a given problem) and deployed within OSF. Clients of the OSF, for instance the data stream classifier and the Application Placement Engine use these queries as their interface to the knowledge models.

From an operational point of view, the semantic knowledge models of the system that are provisioned via OSF are the storage and processing unit of the system, since they contain all required information and are able to evaluate and classify data values to assess the current state of the grid in real time.

The Data Stream Classifier

The automatic classification of smart grid data points represents an essential step to enable the application of the smart grid knowledge model to incoming data. The goal of this component is to provide an interface that can be used by grid operators to create statistical classifiers for data points which origin and purpose are known to them. Once created, these classifiers are applied to new or unclassified data points.

This Data Stream Classifier consists of a back and a frontend. The backend is responsible for training a statistical classifier given training examples that are sent via the frontend. Moreover, the backend remembers one-to-one associations between data points and data point classes and stores trained associations persistently to be used across several runs of the system.

The Data Stream Classifier Backend

For classifying data points statistically, the backend uses a 3-dimensional KD-tree structure that stores a statistical fingerprint of each data point with a known association that is based on the ten most recent observations of a data point. The fingerprint consists of a 3D vector containing the mean of a data point, its standard deviation, and its skewness. To classify a new data point given its Gridlink identifier, the backend loads the most recent observations of such data point from the Gridlink and then runs a

nearest neighbor search inside the KD-tree. In addition, it computes a confidence that is based on the Euclidean distance between the new data point sample and its nearest neighbor and provides the confidence value to the client.

The REST interface of the backend provides its clients with the options of a) asking for a data point class given its Gridlink identifier, b) registering one-to-one mappings given a Gridlink identifier and the URI of the target class within the knowledge model, c) uploading new associations between Gridlink data points and a data point class as training examples for the statistical classifier, and d) deleting both types of mappings.

The Data Stream Classifier Frontend

The Classification Frontend is a graphical user interface (Figure 3) that represents the linking pin between the Classifier Backend, the OSF, and the Gridlink message bus.

a frank (and the contract of the
Absolution Standa Springer Standa Springer Standa Springer Standa	Construction of the Advance of the A

Figure 3: The Classification Frontend Display

Available Gridline Data Points (left), Known Data Point Classes (right), and Observed Data Points (center)

Source: Siemens CT

The Classification Frontend displays all data points that are available on the Gridlink at any given time together with their most recent values (left column in Figure 3), the frontend interacts with the Classification Backend to apply already trained classifiers to the data points and allows operators to categorize new points. To this end, it loads available data point categories from the knowledge model via the OSF and visualizes them (right column in Figure 3). In case a data point has already been successfully classified by the Classification Backend, the frontend displays the data point's class and the confidence of the classification in the left column. An operator can also switch data points into "Observation" mode (central panel in Figure 3) which triggers the insertion of newly recorded data points into the knowledge model and their processing via SPIN rules inside the OSF.

Operators interact with the frontend mainly by drag-and-dropping the "cards" that represent individual data points. The frontend provides three main functions:

- 1. Teaching a new association between a data point and a data point class to the backend. To this end, an operator drags the representation of the data point and drops it on the class that should be associated with it. The operator is then given two options on how to proceed:
 - The association can be sent to the backend as a training example. In this case, the frontend instructs the backend to fetch from the Gridlink the ten most recent values of such data point. These values are then used to train a statistical classifier. This classifier is subsequently applied to all other available data points. If any data point can now be successfully classified, the frontend is updated with this information.
 - The association can be recorded and stored as a one-to-one mapping between the data point and the class. The classifier remains oblivious of this association and it is thus not used to classify any other data points.
 - An operator would choose the first option whenever data points of the given class are similar enough that can be generalized (for example, a line voltage). The second option is applicable whenever the given data point has unique statistical properties that should not be used for generalization (for example, a specific building's power consumption).
- 2. To remove an association between a data point and a class, an operator drags a data point card on the "Remove Classification" icon. This triggers a REST request to the backend that removes the training example or one-to-one mapping that is associated with the given data point.
- 3. To place a classified data point in "Observation" mode, an operator drags the data point card to the central panel of the frontend. This action triggers several REST requests that reconfigure the system to observe that data point: a new DataPoint instance is created in the knowledge model to hold this point's data and new values are added to this instance as they become available on the Gridlink. Since the knowledge model now has data values that are associated to a specific data point class, OSF inferencing subsystem will trigger the SPIN functions knowledge model to observe the data point given its class and report any violations. If a violation is detected, this is recorded in the knowledge model and the frontend visualizes the violation (Figure 3). The operator can access details about the violation by clicking on the data point card in the observation panel.

Application Placement Engine

As stated before, the knowledge models help describing the problems that could be found on the grid as well as their potential solutions in terms of applications to install. On its side, the classifier is responsible for categorizing and validating incoming data through the usage of knowledge models. However, grid operators oversee not only one, but a cluster of substations.

The Application Placement Engine was created to help operators manage those FSSNs globally, this component considers the current state of the FSSN and performance metrics defined by the grid operator, which express his interest in different aspects of the installation, such as resource consumption or load balancing. Given this information, the placement engine computes a feasible installation plan.

To compute an installation plan, the problem of placing M apps on N nodes was formulated as a combinatorial optimization problem. Each node corresponds to a FSSN in a peer to peer communication network through the Gridlink bus. Thus, it is assumed that any app placed on any node $i \in \{1, ..., N\}$ can communicate with any other app placed on node j for any $j \in \{1, ..., N\}$; forming a complete graph.

Each node $i \in \{1, ..., N\}$ is equipped with three resource types: $r_{i,s}$ the amount of available disk storage, $r_{i,c}$ the amount of available CPU cycles and $r_{i,r}$ the amount of available RAM.

Given that the maximum amount of each resource required by each app during runtime is known, app $i \in \{1, ..., M\}$ consumes at most $C_{i,s}$ units of storage $C_{i,c}$ CPU cycles and $C_{i,r}$ units of RAM at runtime.

Given that apps may have dependencies; it can be said that app *j* is a dependency of app *i* if app *i* requires access to services provided by app *j* at runtime. Dependencies may themselves have dependencies, but it is assumed that a directed graph representing all app dependency relationships is acyclic (i.e., it cannot happen that app A has a dependency on app B, which in turn has a dependency on app A).

All app dependency relationships are described in the knowledge models. Such dependencies can be local or non-local. Local dependencies must be placed at the same node as the app depending on it, whilst non-local dependencies can be placed anywhere in the network.

Let D^l be the local dependency relation

$$d_{i,j}^{l} = \begin{cases} 1, & \text{if app i has a local dependency on app j} \\ 0, & \text{otherwise} \end{cases}$$
(1)

Let $D^{nl} \in \{0,1\}^{M \times M}$ be the matrix that encodes all non-local app dependency relationships such that its (i, j)th entry

$$d_{i,j}^{nl} = \begin{cases} 1, & \text{if app i has a non-local dependency on app j} \\ 0, & \text{otherwise} \end{cases}$$
(2)

The grid operator specifies a set of mandatory installation requirements (for example, install Voltage Guard I app on Node 1 to solve voltage violation).

Let $P \in \{0,1\}^{M \times N}$ be the matrix that encodes all node-specific app placements (i.e., mandatory installation requirements and local dependencies) such that its (i,j)th entry

$$p_{i,j} = \begin{cases} 1, & if app \ i \ must \ be \ placed \ on \ node \ j \\ 0, & otherwise \end{cases}$$
(3)

Some applications may provide crucial services to others, or there may be multiple applications that have the same dependency. The use of "*hot standby*" replicas of crucial dependencies is a common technique used to mitigate downtime associated with the failure of nodes hosting crucial applications. It is assumed that a grid operator provides a vector $H \in N^M$ such that its *jth* element h_i indicates the number of hot standby replicas of app *i* that are to be deployed within the network. In addition, a grid operator might want to install an app on any node, which is a special case of a "hot standby" with one replica.

Model Formula

The search space is defined as the set X of all binary matrices $x \in \{0,1\}^{M \times N}$ whose (i, j)th entry

$$x_{i,j} = \begin{cases} 1, & if app is placed on node j \\ 0, & otherwise \end{cases}$$
(4)

Note that such definition precludes the possibility of placing the same app twice on a given node. However, it is possible to place the same app on multiple nodes.

Given $D, P, H, Qr_{i,c}, r_{i,s}, r_{i,r}$ for $i \in \{1, ..., N\}$ and $c_{i,c}, c_{i,s}, c_{i,r}$ for $i \in \{1, ..., M\}$ the goal of the Application Placement Engine is to find a feasible $x^* \in X$ Feasibility is defined by the following constraints:

a) Capacity Constraint: for each resource type, the total resource usage by all apps hosted on a node must not exceed the total amount of resource available on such a node (i.e., storage, ram, and cpu):

$$\sum_{i=1}^{M} x_{i,j} c_{i,s} \le r_{j,s}, \forall j \in \{1, \dots, N\},$$
(5)

$$\sum_{i=1}^{M} x_{i,j} c_{i,r} \le r_{j,r}, \forall j \in \{1, \dots, N\},$$
(6)

$$\sum_{i=1}^{M} x_{i,j} c_{i,c} \le r_{j,c}, \forall j \in \{1, \dots, N\}$$
(7)

b) Installation Requirements Constraint: local app placements encoded by *P* must be enforced

$$x_{i,j} = 1, \ \forall \ (i,j) \ such \ that \ p_{i,j} = 1$$
 (8)

c) Local Dependency Placement Constraint: if there is an app *i* placed on node *j* that has a local dependency *d*, then that dependency must be placed on the same node. This contraint is expressed as follows:

$$x_{i,j} = 1 \to x_{d,j} = 1, \ \forall \ (i,d) \ such \ that \ d_{i,d}^l = 1, \ \forall \ j \in \{1, \dots, N\}$$
(9)

d) Non-local Dependency Placement Constraint: if there is a placed app *i* that has a non-local dependency *d* then that dependency must be placed somewhere on the network. This constraint is expressed as follows:

$$\sum_{j=1}^{N} x_{i,j} \ge 1 \to \sum_{j=1}^{N} x_{d,j} \ge 1, \forall (i,d) \text{ such that } d_{i,d}^{nl} = 1, \forall j \in \{1, \dots, N\} (10)$$

e) Hot Standby Constraint: for each $i \in \{1, ..., M\}$ there must be at least h_i replicas of app *i* within the network:

$$\sum_{i=1}^{N} x_{i,i} \ge h_i \,\forall \, i \in \{1, \dots, M\}.$$
(11)

Objective Function Components: Performance Metrics

The constraints in the prequel do not preclude placements that consume most, if not all the available resources in the network. For example, it is possible that a primary app *i* has dependency *j* and *j* is placed on every node in the network even if no other app requires its services. Such a placement, though feasible for some input data, wastes resources and undermines resiliency since under node failure conditions, these unnecessarily replicated apps consume resources that could otherwise be allocated to apps requiring migration from failed nodes.

For this reason, a performance model that promotes minimal network-wide resource consumption and load imbalance is proposed:

a) Network-wide Resource Consumption: the total network-wide resources consumed by a placement *x* is given by

$$R(x) = w_{R_s} R_s(x) + w_{R_r} R_r(x) + w_{R_c} R_c(x),$$
(12)

where w_{R_s} , w_{R_r} , and w_{R_c} are tunable weighting parameters, and

$$R_s(x) = \sum_{j=1}^N \sum_{i=1}^M x_{i,j} c_{j,s}$$
(13)

$$R_r(x) = \sum_{j=1}^N \sum_{i=1}^M x_{i,j} c_{j,r}$$
(14)

$$R_c(x) = \sum_{j=1}^{N} \sum_{i=1}^{M} x_{i,j} c_{j,c}$$
(15)

b) Load Balancing: in addition to minimizing overall resource consumption R(x), for maximization of computational resiliency it is desirable to minimize load imbalance. If all resource consumption was perfectly balanced within the network, each node would supply exactly

$$\phi_{s}(x) = \frac{R_{s}(x)}{\sum_{i=1}^{N} r_{i,s}}$$
(16)

$$\phi_r(x) = \frac{R_r(x)}{\sum_{i=1}^N r_{i,r}}$$
(17)

$$\phi_c(x) = \frac{R_c(x)}{\sum_{i=1}^N r_{i,c}}$$
(18)

amount of resource to all apps it hosts, i.e., the total amount of resources consumed, divided by the total amount of resources available in the network.

Perfect balancing may not be possible since the placement problem is binary, and apps require different amounts of each resource type. Thus, the network imbalance imbalance is minimized, which can be encoded as

$$B(x) = w_{B_s}B_s(x) + w_{B_r}B_r(x) + w_{B_c}B_c(x), \qquad (19)$$

where $w_{B_s}(x)$, $w_{B_r}(x)$ and $w_{B_c}(x)$ are tunable weighting parameters,

$$B_{s}(x) = \sum_{j=1}^{N} \left\| \frac{\sum_{i=1}^{M} x_{i,j} c_{i,s}}{r_{j,s}} - \phi_{s}(x) \right\|$$
(20)

$$B_r(x) = \sum_{j=1}^N \left\| \frac{\sum_{i=1}^M x_{i,j} c_{i,r}}{r_{j,r}} - \phi_r(x) \right\|$$
(21)

$$B_{c}(x) = \sum_{j=1}^{N} \left\| \frac{\sum_{i=1}^{M} x_{i,j} c_{i,c}}{r_{j,c}} - \phi_{c}(x) \right\|,$$
(22)

and $\|\cdot\|$ may stand for either $(\cdot)^2$ or $|\cdot|$, for example. The $\|\cdot\|$ terms in $B_s(x), B_r(x)$ and $B_c(x)$ measure the difference between node *j*'s actual resource consumption, and the ideally balanced resource consumption.

c) Minimal Movement: the App Placement Engine can be triggered by either a human operator, or when a node failure is detected. In either case, it is desirable to minimize the number of configuration changes since movement of apps from one node to another is costly in terms of time and bandwidth. This objective can be expressed as follows:

$$M(x) = \sum_{i=1}^{M} \sum_{i=1}^{N} |\tilde{x}_{i,j} - x_{i,j}|, \qquad (23)$$

where $\tilde{x}_{i,j}$ indicates whether app *i* is currently placed on node *j* or not.

- d) Replicated Placement of Frequently Used Dependencies: to improve the resiliency of the system, installing the most frequently used apps across the network multiple times is considered.
 - The set *i* in equation *S* dentifies all apps that are dependencies of at least one of the *M* apps to be placed. For each dependency app $i \in S$, let n_i be the number of apps (either primary or dependency apps) that require interaction with *i*. Note that n_i is the i'th column sum of the dependency matrix *D*.

 Maximization of the following term promotes the replicated placement of the most frequently used dependencies, in proportion to how many deployed apps use them:

$$U(x) = \sum_{i \in S} \sum_{j=1}^{N} f_i x_{i,j},$$
(24)

where

$$f_i = \frac{n_i}{\sum_{k \in S} n_k}.$$
 (25)

App Placement Program

Based on the model formulas above, the app placement program is given by

$$\min_{x \in X} J(x)$$

s.t. $x \in C$, (26)

where

$$J(x) = w_R R(x) + w_B B(x) + w_M M(x) - w_U U(x)$$
(27)

 $w_R, w_B, w_M, w_U \in R_+$ are appropriately chosen normalization parameters, R(x), B(x) and M(x) are as in (12), (19), and (23), and *C* is the subset of *X* whose elements *x* satisfy (4), (6), (7), (8), (9), (10), and (11). The term U(x) is substracted since the objective is to maximize it.

This model is structurally flexible; setting $w_B = 0$, for example results in a program that does not take balancing into account.

Global FSSN Manager: Approach

Following the approach in (Tucker, et al. 2007), also used in (Pradhan, et al. 2017), the SMT solver Z3 (Urgaonkar, Arnold L and Prashant 2007) is iteratively applied to problem (26). SMT solvers are designed to accept a set of logical or algebraic constraints and to either produce a feasible point that satisfies all the constraints, or indicate that the constraints are not satisfiable.

In addition to the algebraic constraints (5), (6), (7), (8), (9), (10) and (11), at the kth iterated application of the SMT solver, it is supplied a constraint that imposes a strict upper bound b_k on the value of the cost function J(x):

$$J(x) < b_k \tag{28}$$

If the problem is feasible, the SMT solver returns either a feasible app placement configuration

$$x_k \in C \cap \{x \in X \mid J(x) \le b_k\},$$
 (29)

or it indicates that the problem is infeasible, i.e. the set in (29) is empty. This process is iterated a maximum of *T* times. Given the feedback from the previous iteration, a sequence generator produces a sequence of upper bounds b_k with the aim of finding b^* , the smallest upper bound for which the set $C \cap \{x \in X \mid J(x) \le b_k\}$ is nonempty. The corresponding placement configuration x^* is accepted as the optimal app placement configuration.

The Semantically Integrated Operational Dashboard

The semantically integrated operational dashboard is the comprenhensive product of the knowledge models, and software components that have been presented in the previous sections. The objective of this dashboard is to provide smart grid operators with a user interface that allows them managing a set of Future Secondary Substations. The managing activities include:

Showing the status of the FSSNs

Showing a detected anomaly in the grid

Computing and showing suggested apps to alleviate the anomaly

Showing the preferred configuration parameters values for the installation of the app

Allowing the grid operator to compute a global installation plan

Resolving app dependencies and informing about their automatic installation, and

Reflecting the effect that the installation of the app might have caused in the FSSN

The dashboard is a user interface that relies directly on the semantic backend, which is in charge of detecting problems in the grid by monitoring the received data from the Data Stream Classifier, suggesting apps to remedy problems in the grid, and resolving app dependencies. Moreover, the dashboard is the user interface for the App Placement Engine to compute a global installation plan.

Traditional grid management tools, such as SCADA systems provide grid operators sitting in control rooms with data that needs to be correlated and interpreted. Such data is displayed in traditional two-dimensional user interfaces such as computer screens. Thus, operators are sitting in front of a computer continuously monitoring the state of the grid, analyzing the incoming data, and reacting when necessary.

In this research project two dashboard approaches were developed. On the one hand, a more traditional 2D dashboard was created as a Web application. On the other hand, an Augmented Reality (AR) (Mizutani, et al. 2017) (Funk, Kritzler and Michahelles 2017) app was concieved to shows how cutting-edge technologies could enhance the experience of managing FSSNs.

The 2D Dashboard

The 2D dashboard is a Node.js⁴ Web application that is connected to the backend components through their REST APIs. This dashboard allows a grid operator to visualize the FSSNs being managed; they are represented as green nodes placed in the map of the corresponding area (Figure 4). When selecting one of the FSSN, the detailed status of the FSSN is presented in the left of the screen (Figure 4 left side). Moreover, the grid operator has access to the application store by clicking on the three lines button. Then, he can drag and drop the app to be installed (Figure 4 right side).



Figure 4: The 2D Dashboard

Source: Siemens CT

The Augmented Reality App

The design and development of the AR app for the Smart Management of Smart Grid project followed six phases:

- 1. User Research and Benchmarking. A questionnaire was sent to colleagues in the Siemens Energy Management division in France in order to obtain information about how current grid operators work. From this questionnaire, it was learned that grid operators are situated in monitoring rooms with open spaces, dedicated tables, and work stations. Such rooms are not very noisy, since the amount of people in them is regulated. Moreover, grid operators are constantly reading manuals with instruction on how to react when events appeared in the grid.
 - Based on the spacious grid operator work conditions and the low level of noise, it was decided to develop an AR app for a head mounted display

⁴ Node.js® (https://nodejs.org/en).

(for example a Microsoft HoloLens⁵). The head mounted display provides users with the freedom of movement, has voice command capabilities, and allows users to use their hands freely and to consult external manuals or talk to coworkers in case it is necessary without needing to leave the user interface.

2. Ideation and Storyboarding. Given that a 2D dashboard for a desktop application was also built, the user story was clear up to the point in which a grid operator wants to manage not just one individual substation, but the whole cluster of substations he is in charge of. This corresponds to the functionality provided by the Application Placement Engine. Incorporating this functionality did not represent a huge effort on the traditional 2D dashboard. Only some widgets and a few buttons were added to the interface. However, for the AR app, scenes needed to be constructed to guide grid operators through the grid management process. Thus, sketching and story boarding were essential to decide which elements needed to be presented in the 3D space (Figure 5). These methods also helped to consider the possible workflows when managing substations.

Figure 5: Storyboard Example for Apps in Substation to Alleviate Anomaly



Source: Siemens CT

3. Initial Prototyping. After sketching, and talking to the team members, the scenes sequence, were decided, and started to build a prototype in Unity⁶. which is commonly used for the implementation for HoloLens applications. The objective

⁵ Hololens 2 (https://www.microsoft.com/en-us/hololens).

⁶ https://unity3d.com/

was to create a quick prototype to get a preview of how the designs will look in 3D. Figure 6 shows two early prototypes of the app, in which substations are depicted as 3D cubes, and the app information panels are all copies of the same app. The early prototyping helped us to identify the limits of the HoloLens and sparked more ideas on how to display the components in more user-friendly manner.







Figure 7: Augmented Reality App Appearance During the User Tests



Source: Siemens CT

- 4. User Testing. To polish, confirm, or rework the design assumptions made during the implementation of the AR app, six short user tests were conducted with Siemens colleagues. The feedback obtained from the user tests guided the last phase of development of the app. Figure 7 shows the general appearance of the AR app at the time the user tests were conducted.
- 5. Prototype Polishing and Futureproofing. Given the feedback obtained from the user testing, more time on the development of the AR app was spent to address some of the shortcomings that had been over looked during the first development period.
- 6. Providing Functionality to the Prototype. The following natural step was to connect all the backend components (i.e., OSF, and App Placement Engine) to the AR application.

CHAPTER 3: Project Results

Data Streams Classifier

To verify that the classification backend can successfully classify data points that are relevant to the context of the operational smart grid dashboard, a sequence of tests was performed using data from several field measurements. The data sources used are:

- 1. A data set from a Siemens SICAM P855 Multifunctional Device that is deployed in the low voltage grid in central Europe:
 - a. Data Points: RMS Voltage (3 phases); RMS Current (3 phases); System Frequency; Total Active Power; Active Power (3 phases); Total Reactive Power; Reactive Power (3 phases); Active Power Factor (3 phases)
 - b. Time Frame: The data point values were obtained over a timeframe of ~16.5h, from midnight until 16:25pm; 1433596 individual data point values were recorded across all data points. The time interval between two data point values varies between ~4s (for example, Total Active Power) and ~40s (for example, Current on 3rd phase)
- 2. A data set of 1500 random Binary Switching Decisions was generated as this kind of data was not available from the in-field deployment and can easily be generated.
- 3. Voltage Total Harmonic Distortion (THD) and Instantaneous Flicker data from a deployment at Lawrence Berkeley National Lab⁷ was obtained. Data recorded throughout one day in 1s-intervals was used for the tests presented in the following.

Testing Method

One thousand iterations of the following test setup were performed to verify the accuracy of the classifier. ASAMPLE_LENGTH parameter, corresponding to the sample size, was set between 3 and 103 (in increments of five):

- Training: for each data point, select, at random, an index in the time series and train the classifier using the data point values at that index and the following SAMPLE_LENGTH indices.
- Classification: for each data point, select, at random, an index in the time series and classify the data point using the data point value at that index and the

⁷ maps.pqube.com

following SAMPLE_LENGTH indices. Record an error every time the category given by the classifier is different from the true category

Test Results

The method previously described was implemented for all data points. However, for some data one-to-one mappings are more appropriate (for example, power measurements should be considered unique to a specific deployment and should not be used for training a generic "Power" data point classifier). The developed classifier delivered very good results (Figure 8) and behaves as expected when the training and or classification sample size is increased. At a sample size of three data point values (i.e., the minimum sample size for the setup, since the classifier needs to compute the skewness of the data sample), the recorded error rate was of 8.78 percent; when the sample size was increased to 103 samples, the error rate decreased to 1.66 percent.



Figure 8: Classification Errors for Settings of SAMPLE_LENGTH Parameter

When the data points that should be mapped using a one-to-one classifier were excluded, the error rates dropped to 0 percent for all sample lengths. This is due to the very nature of the implemented classifier, since it should not have trouble distinguishing between defined values, such as Voltages, that hover around a nominal value for example, 230V or 110V and Frequencies (that do also remain in close to a defined value). All errors in the tests indeed stem from pairs that the classifier is not expected to distinguish successfully, such as Active Power being classified as Reactive Power and vice versa.

It was also verified that the Classification Backend can successfully de-classify sample data points, i.e. that the data point - class associations can be removed by the operator, by creating a test case that checks whether its classification result stays the same after adding classifiers and removing them again.

The results obtained from this test prove that the developed classifier can detect and flag Under and Overvoltage, and Unbalance-related conditions.

Source: Source: Siemens CT

Classifier Performance

The classifier was exposed to simulated data, and its response to automated queries was timed. The acceptable response time was <10s which is value based on commonly accepted time limits for human perceptual abilities.⁸ The following test cases were verified:

Classification Response Time: the time required for training and classifying 19 random data point samples (over 1000 iterations) using the setup described in the Testing Method, and with a SAMPLE_LENGTH parameter of 10 is 5ms per sample.

Server Response Time: the time required for classification is dominated by the communication overhead between the backend and the frontend. The overhead was tested by simulating 20 interacting users over a 1min period and a 1s period between two interactions by the same user. All users interacted with a server on the same machine. On average, the Classification Interface responded after 14ms.

Network Latency: In a distributed real-world deployment, both response times would be dominated by the network latency with a remote server. This latency would be of 10ms to 50ms.⁹ In addition, adding TLS-based security in a real-world deployment could add up to 500ms per request (Naylor, et al. 2014).

Summary: Overall, the response time of the Classification Backend is expected to stay well below 1s even for secured, real-world deployments.

Application Placement Engine

To validate the Application Placement Engine, three types of experiments were conducted:

- 1. Correctness: to verify that the Placement Engine produces correct placements, several examples were used, and the fulfillment of the set requirements was verified
- 2. Parameter Effectiveness: to test the effectiveness of the weight parameters, they were varied, and the resulting placements were compared
- 3. Time vs. Quality tradeoff: given the complexity of the problem, the t_{max} parameter was introduced, which enables a tradeoff between quality of the placement and time needed to compute it. This tradeoff is illustrated by varying the t_{max} parameter for the same scenarios.

The Placement Engine was implemented as a python module that can be either run on the command line or used in a larger project as a library. It uses Z3 (De Moura and

^{8 &}lt;u>Nielsen Norman Group</u> (https://www.nngroup.com/articles/response-times-3-important-limits/).

^{9 &}lt;u>3Scale</u> (http://tech.3scale.net/2013/08/19/performance-variance- between-aws-zones).

Nikolaj 2008), a SMT solver, to compute the placements. All tests were conducted on a Microsoft Surface Book running Windows 10 1803. The machine features an Intel Core i7-6600U CPU and 8 GB of RAM. The Python version used was 3.6.5 and the version of Z3 used was 4.7.1.

Testing Scenarios

Three scenarios were considered to validate the Application Placement Engine. Scenario 1 corresponds to a toy example consisting of three FSSN nodes, and four applications. However, due to its simplicity, the results are not significant, hence they are not presented in this report. Table 2: Scenario 2 shows the app descriptions for Scenario 2, and its variances. Such descriptions include the apps non-local and local dependencies, and resource consumption. This table makes us aware that when choosing to install an app, the dependencies need to be installed as well.

Арр	Local Dependencies	Non-Local Dependencies	Resource Consumption
App1	-	Арр3, Арр8, Арр9	CPU: 2 RAM: 1 Storage: 1
App2	-	Арр3, Арр8	CPU: 1 RAM: 1 Storage: 1
Арр3	-	App8	CPU: 5 RAM: 3 Storage: 1
App4	-	App8	CPU: 1 RAM: 2 Storage: 2
App5	-	Арр4, Арр8	CPU: 1 RAM: 1 Storage: 2
App6	-	Арр4, Арр8	CPU: 1 RAM: 1 Storage: 1
App7	-	App8	CPU: 3 RAM: 1 Storage: 1
App8	-	-	CPU: 1 RAM: 1 Storage: 2
App9	-	App8	CPU: 2 RAM: 1 Storage: 2

Table 2: Scenario 2, App Description

Source: Source: Siemens CT

Following, the variations of Scenario 2 are listed, including required apps to install, and the initial app placement.

- Scenario 2
 - Required to install:
 - App6 on Node2
 - App7 on Node4
 - App7 on Node5
 - Network wide installations:
 - App2 with 1 replica
 - Initial Placement:
 - Node1: None

- Node2: None
- Node3: None
- Node4: None
- Node5: None
- Scenario 2.1
 - Required to install:
 - App9 on Node2
 - Network wide installations:
 - App1 with 1 replica
 - App6 with 1 replica
 - Initial Placement:
 - Node1: App3
 - Node2: None
 - Node3: App1, App4, App6, App8, App9
 - Node4: None
 - Node5: None
- Scenario 2.2
 - Required to install:
 - App6 on Node2
 - App7 on Node4
 - App7 on Node5
 - Network wide installations:
 - App2 with 1 replica
 - App4 with 3 replicas
 - Initial Placement:
 - Node1: None
 - Node2: None
 - Node3: None
 - Node4: None
 - Node5: None

Table 3 shows the app descriptions for Scenario 3. Such descriptions include the apps non-local and local dependencies, and resource consumption.

Арр	Local dependencies	Non-local dependencies	Resource Consumption
App1	-	Арр3	CPU: 2 RAM: 1 Storage: 1
App2	-	Арр3, Арр8	CPU: 1 RAM: 1 Storage: 1
Арр3	-	Арр7, Арр8	CPU: 5 RAM: 3 Storage: 1
App4	App10	Арр1, Арр3	CPU: 1 RAM: 2 Storage: 2
App5	-	Арр9	CPU: 1 RAM: 1 Storage: 2
Арр6	App10	Арр9	CPU: 1 RAM: 1 Storage: 1
App7	App10	App8	CPU: 3 RAM: 1 Storage: 1
App8	App10	-	CPU: 1 RAM: 1 Storage: 2
App9	App10	-	CPU: 2 RAM: 1 Storage: 2
App10	-	-	CPU: 1 RAM: 1 Storage: 4
App11	-	Арр3	CPU: 3 RAM: 2 Storage: 3

Table 3: Scenario 3 App Descriptions

Source: Source: Siemens CT

- Scenario 3
 - Required to install:
 - App1 on Node5
 - App2 on Node3
 - App4 on Node7
 - Network wide installations:
 - App11 with 1 replica

Correctness

A placement is correct if the following requirements are fulfilled:

- All apps for which an installation requirement exists are placed on the respective FSSN
- All apps for which a network wide installation requirement exists are installed on (at least) as many FSSNs as needed
- All local dependencies for a given app are installed on the same FSSN

- All non-local dependencies for a given app are installed somewhere in the network
- The available resources are not exceeded for any of the FSSNs

The parameters were set to $1 w_R = w_B = w_M = w_U = 1$ and optimal placements were computed for the scenarios presented before. However, it was not possible to get the optimal placement for Scenario 3, since after more than 14 hours of computation it was decided to abort.

The placement for Scenario 2 is shown in Figure 9: Placement for Scenario 2Figure 9. Looking at the placement, it is possible to verify its correctness, since:

- App6 is installed on Node2 and App7 is installed on Node4 and Node5. Thus, the installation requirements are fulfilled
- There is one copy of App2 installed on Node4, so the network wide installation requirements are fulfilled
- None of the apps have local dependencies so the requirement is trivially fulfilled.
- The non-local dependencies (i.e., App3, App4 and App8) are all placed, fulfilling the non-local dependency requirement
- The figure also shows that none of the FSSNs resources are exceeded

The placement for Scenario 2.1 is shown in Figure 10. It was verified whether all requirements are fulfilled:

- App9 is installed on Node2
- There is one copy of App1and one of App6.
- There are not non-local dependencies
- The non-local dependencies (i.e., App3, App4, App8, and App9) are placed at least once as required.
- The available resources are not exceeded



Figure 9: Placement for Scenario 2

Source: Siemens CT



Figure 10: Placement for Scenario 2.1

Source: Siemens CT

The difference between Scenario 2 and Scenario 2.2 is that the latter requires at least 3 copies of App4. Figure 11, shows that the requirements for Scenario 2 are fulfilled and that App4 is placed on Node1, Node4 and Node5. Thus, the replication requirement is also met.

Figure 12 shows Scenario 3 placements, in which all the requirements are also fulfilled:

- App1 is installed on Node5, App2 is installed on Node3 and App4 is installed on Node7.
- App11 is placed once on Node2.
- App4, App7 and App8 have a local dependency on App10. These apps are installed on Node5 and Node7 which both also have App10 installed.
- The non-local dependencies (i.e., App3, App7 and App8) are placed at least once as required.
- No resources are exceeded.



Figure 11: Placement for Scenario 2.2





Source: Siemens CT

Parameter Effectiveness

The proposed objective function offers two tradeoffs:

- Network wide resource conservation vs. replication of frequently used apps
- Load balancing between FSSNs vs. minimal movement of apps during reconfiguration

For both tradeoffs two experiments were conducted to show the effect of the model parameters.

Resource Conservation vs Replication

To explore the tradeoff between resource conservation vs replication of apps, Scenario 2 was used, which requires installing App6 on Node 2, App7 on Node 4, App7 on Node 5, and App2 should be placed at least once in the network.

To clearly see this tradeoff, the balancing and movement parameters were set to 0, i.e., $w_B = w_M = 0$, and two optimal placements were computed. For the placement shown in Figure 12 $w_R = 10$ and $w_U = 1$ were set. For the placement shown in Figure 14, the replication was prioritized instead of resource consumption hence $w_R = 1$ and $w_U = 10$.

As Figure 13 shows, when emphasizing low resource consumption, the Placement Engine only places the apps that are strictly required, keeping the amount of available resources as high as possible. In case the weight for the replication term in the objective function is higher, the engine replicates the two most frequently used apps (App4 and App8) across all FSSNs, as shown in Figure 14.



Figure 13: Placement Emphasizing Resource Conservation

Source: Siemens CT





Source: Siemens CT

Load Balancing versus Minimal App Movement

To explore the tradeoff between load balancing and minimal app movement, Scenario 2.1 was used, which requires installing App9 in Node 2 and App1 and App6 should be

placed at least once in the network. Figure 15 shows the initial placement of this scenario.



Figure 15: Initial Placement for Load Balancing vs Minimal Movement

Similar to the resource conservation vs resource replication tradeoff, two optimal placements were computed, one with $w_B = 20$, $w_M = 1$ and another one with $w_B = 1$, $w_M = 20$. The other two weights were set to $w_R = 50$ and $w_U = 0$ for both runs. The high value for the resource conservation term is there to ensure that the engine does not place additional apps that help the load balancing, instead of distributing the necessary apps fairly across FSSNs.

Figure 16 shows that when load balancing is emphasized, the Placement Engine distributes the use of resources fairly across all the FSSNs moving previously installed apps to different nodes if necessary. Figure 17 shows that when minimal movement is emphasized, the engine tries to keep all apps where they are. Thus, the only movement made is the one needed to fulfill the requirements of installing App9 on Node2 instead of Node3.

Source: Siemens CT



Figure 16: Placement Emphasizing Load Balancing



Figure 17: Placement Emphasizing Minimal App Movement





Time versus Quality Tradeoff

An important requirement when managing the smart grid is to have short execution times, as the Placement Engine should be able to compute a new network configuration in emergency situations, such as complete failure of a FSSN. However, using non-linear terms for the load balancing and minimal app movement objectives makes the optimization problem computationally hard, leading to high execution times. Table 4 lists the time it took to compute each placement described in the Testing Scenarios.

Scenario	Execution Time	Note
Scenario 2	40.571s	Optimal
Scenario 2.1	796.089s	Optimal
Scenario 2.2	95.711s	Optimal
Scenario 3	51964.510s	Aborted

Table 4: Execution Times

Source: Siemens CT

Clearly, some of these times are not practical. A grid operator might have just seconds to react to a problem. Thus, in an emergency, a suboptimal placement might be good enough to resolve the issue; the network could still be reconfigured later to increase efficiency. As mentioned before, finding an optimal placement for Scenario 3 takes more than 14 hours. However, a first placement is found after just a fraction of a second. By adding the t_{max} parameter it was possible to get best effort placements in the available time. To demonstrate how the engine computes better and better placements over time, Scenarios 2 and 3 from were executed with different t_{max} values. The evolution of the placements can be seen in Figure 18.

Figure 18 shows the placements obtained for Scenario 2 when stopped after 1, 5, 10, and 15 seconds. While the placement after 15 seconds is already optimal, the engine needs another 25 seconds to verify that indeed there is no better placement.

Scenario 3 is significantly larger than Scenario 2, since it comprises 8 nodes and 11 apps. When performing the testing, its execution had to be aborted after 14h, but interesting results were obtained. Figure 19 shows the placements obtained after 1, 5, 10, 30 seconds, 30min and 14h, it can be seen that, even after one second, the Placement Engine already computes a solution that fulfills the installation requirements. Assuming that a grid operator might be able to wait for a placement up to 30s in case of an emergency. However, in case the operator decides to make changes after, the placement can be re-computed for longer to obtain a better result.



Figure 18: Placements for Scenario 2

After 1, 5, 10 and 15 Seconds (from top left to bottom right)

Source: Siemens CT



Figure 19: Placements for Scenario 3

After 1, 5, 10, 30 Seconds, 30 Minutes and 14 Hours (from top left to bottom right) Source: Siemens CT

The Semantically Integrated Operational Dashboard

It was verified that the dashboard, on its AR version, as well as on the FSSN 2D dashboard is capable of supporting a grid operator on the management of FSSNs. To this end, five core functionalities were tested:

1. Allowing a grid operator to install apps when the status of the grid is normal: As shown in Figure 20, a grid operator is able to install an application on a FSSN under normal conditions. Figure 20a shows a view of the map in which four FSSNs are located. Those correspond to the cluster the grid operator is in charge of managing. When getting closer to a substation, gazing, and air tapping into the "Manage apps" button (Figure 20b), a menu with cards describing the apps available is shown (Figure 20c). Each card shows the name, and resource consumption of the app. A user can choose to see more information, or to install the app. When the installation of the app is selected, the user is presented with the configuration parameters, he is able to enter the values he decides, or he can choose to use the system suggested values (Figure 20d).



Figure 20. Installing an App Under Normal Conditions

Source: Siemens CT

After confirming the parameters, the AR app allows the user to review the changes on the FSSN (Figure 21a). When the user confirms the installation, they can gaze at the substation and thus confirm that the app has been installed (Figure 21b).



Figure 21: Review and Confirmation of the Installation of an App

Source: Siemens CT

A grid operator using the 2D dashboard is able to install apps on the selected FSSN by dragging the corresponding app into the selected FSSN. Figure 22 shows the installation process of the Adaptative Assignment Module app. The system suggests the parameters values (Figure 23a). A user can confirm the installation, and the app dependencies are also computed, and installed as shown in Figure 22 (left side).





Source: Siemens CT

Moreover, the dashboard allows grid operators to install an app in all FSSNs at the same time. Figure 23 shows the installation of the Gridlink XMPP Stack onto the four FSSNs available in the Alameda cluster.



Figure 23: Installation of Gridlink XMPP Stack On Available Substations

Source: Siemens CT

2. Making the grid operator aware of potential anomalies in the grid: When an anomaly is detected the user is made aware of it by two visual indicators: a) the main panel that shows the FSSNs status highlights the substation that presents a problem (Figure 24). In Figure 24a, a voltage band violation is presented in the North Berkeley FSSN; and b) a visual element draws the user's attention to the FSSN in the map as shown in Figure 24b.





Source: Siemens CT

On the 2D dashboard, a problem in a FSSN is shown with an animation over the circle representing the FSSN presenting a problem as shown in Figure 25.



Figure 25: Dashboard Shows a Problem in a Substation

Source: Siemens CT

3. Showing suggested solutions to an encountered anomaly: Figure 26 shows the suggested apps to solve the voltage violation. Figure 26a shows three versions of a Voltage guard. However, once the Voltage Guard I app has been installed in the North Berkeley FSSN, if a new violation occurs the system only suggests installing either the Voltage Guard II or the Voltage Guard III app as shown in Figure 26b.

Figure 26. Apps Suggested to Fix a Voltage Violation



Source: Siemens CT

4. Allowing a grid operator to use the app placement engine to compute a global installation plan: Figure 27 shows the installation of the Gridlink XMPP stack app in the Berkeley Hills FSSN. In the step shown in Figure 27a, the grid operator selects the values of the RBMU parameters (i.e., resource consumption, load balancing, minimal movement or apps, and replication of apps) that the App Placement Engine uses to compute a global installation plan. In Figure 27b, the user reviews the changes that the installation will cause in the FSSNs. In this case, the stack is installed not only in the Berkeley Hills FSSN, but also in the Berkeley Cal one. This is due to the higher value on the maximize parameter corresponding to the replication of frequently used apps.



Figure 27: Using App Placement Engine to Compute Global Installation Plan

Source: Siemens CT

In the 2D Dashboard, the app placement engine is activated by ticking the box "Use AI" when installing an application as shown in Figure 28.



Figure 28: Using App Placement Engine to Compute a Installation Plan

Source: Siemens CT

5. Showing that the installation of apps has been completed: Figure 29 shows the AR and 2D dashboards once the installation of an app has been completed. The operator can gaze at the FSSN to see its details (AR app) or click on the FSSN (2D dashboard).



Figure 29: Installation of an App Completed

Source: Siemens CT

CHAPTER 4: Technology/Knowledge Transfer Activities

In order to raise awareness and demonstrate the benefits that adopting cutting-edge technologies such as Semantic Technologies could bring in real Smart Grid settings, several meetings with Siemens stakeholders from the Energy Management Business Unit were held. Moreover, different versions of the developed dashboard were presented in internal Siemens conferences. External to Siemens, the developed work was presented to members of the California Institute for Energy and Environment (CIEE), and a research group from the Lawrence Livermore National Laboratory. Following is a non-exhaustive list of activities done along the duration of this project:

- Held meeting with Omnetric-Siemens looking for partners to apply for pilot project (5 people).
- Held meeting with California Institute for Energy and Environment (CIEE) to introduce them to this research work (3 people).
- Held meetings with Siemens stakeholders from Energy Management Business Unit (15 people).
- Held meeting with Lawrence Livermore National Laboratory (LLNL) research group, looking for pilot partners (6 people).
- Presented the project at Siemens ConneCTs 2017 Siemens CT in Princeton, New Jersey (50 people).
- Distributed brochures at the EPIC symposium 2018 in Sacramento, California (30 people).
- Presented HoloLens prototype at Siemens DAAI (Data Analytics & Artificial Intelligence) in Nuremberg, Germany (100 people).

The feedback collected from the different events where this research project was presented was positive. Specifically, the AR dashboard was well received in the DAAI conference, which is an internal Siemens conference on Data Analytics and Artificial Intelligence. In this event colleagues from various Siemens locations were able to try themselves the AR dashboard. Such presentation led us to fruitful discussions with stake holders from other research groups, and from the Siemens Energy Management Business Unit.

The visit made to the Lawrence Livermore National Lab as well as the meetings the team had with the CIEE made them aware of other projects that both institutions have regarding smart grids and open channels for collaboration in future projects.

The technical advisory meeting was formed by academics from the Lawrence Berkeley National Laboratory, and stakeholders from the Siemens Energy Management Business Unit. The first meeting was conducted in September 2017, the results presented in this meeting included the knowledge models, the data classifier, and the mechanism for automatically finding applications to solve detected problems. The comments received were positive on the innovation of the approach, and feedback was received about alternative hardware that can be used in an actual smart grid setting to acquire grid data, which could improve the accuracy of the data classifier.

As part of the technology transfer activities, the following scientific papers were published:

- Hodges, Jack, Kimberly García, and Steven Ray. 2017. *Semantic Development and Integration of Standards for Adoption and Interoperability*. Computer 50, no. 11 (2017): 26-36.
 - 271 full text views from IEEE Xplore, and 3 citations source Google Scholar
- Mayer, Simon, Jack Hodges, Dan Yu, Mareike Kritzler, and Florian Michahelles. 2017. *An open semantic framework for the industrial Internet of Things*. IEEE Intelligent Systems 32, no. 1 (2017): 96-101.
 - $\circ~$ 1166 full text views from IEEE Xplore, and 24 citations source Google Scholar
- Diwold, Konrad, Simon Mayer, Jack Hodges, et. al. 2018, *Grid Watch Dog: A Stream Reasoning Approach for Lightweight SCADA Functionality in Low-Voltage Grids*. Proceedings of the 8th International Conference on the Internet of Things, IoT 2018, CA, USA.
 - 17 downloads from ACM and 1 citation source Google Scholar

During the latest steps of this research project, several efforts were made to show the developed demonstrator to critical stakeholders within Siemens to bring this research project forward. However, to set this project in a real scenario, it is necessary to get access to modern hardware which is not highly available in the current United States grid. Thus, the Siemens Corporate Technology team in charge of this research work is actively looking for partners within Siemens and the utilities in California to run a pilot project that involves modern hardware such as the Future Secondary Substations, and that allows to conduct user studies of the developed 2D and AR dashboards with grid operators.

CHAPTER 5: Conclusions

The semantically integrated operational dashboard supports grid operators in managing a set of FSSNs. This can allow faster reaction to anomalies detected within the grid, which in consequence provides greater resiliency within the grid. The semantically integrated operational dashboard corresponds to the comprehensive user interface that brings together the software components developed in this research project. Technically, this Operational Dashboard can be divided into four main components, namely: Knowledge Models, Data Classifier, Application Placement Engine, and the semantically integrated operational dashboard in its 2D and AR versions. Each of those components were tested separately at the time they were implemented, and final tests were conducted to the semantically integrated operational dashboard.

The Knowledge Models that were extended within this research project are the foundation for increasing the degree of local automation of FSSNs, which has advantages with respect to the cost of operation of the smart grid and its resiliency, since decisions can be made at a much higher pace. To be able to evaluate the state of the grid, all data point values need to be evaluated, irregularities need to be flagged, and the grid operator needs to be informed to act accordingly to prevent escalation. The Smart Grid Knowledge Models within the developed system allow the validation and monitoring of arbitrary data points, including, for instance, voltages of sensors on distribution substations. Labeled data streams are monitored by feeding the data into the Knowledge Models. On every inferencing iteration, the semantic backend checks the values by comparing them to their associated range boundaries and generates violations. Possible solutions for violations, such as the installation of FSSN apps are also described in the Knowledge Models, as well as possible dependencies that such FSSN apps have.

The Data Classifier is a system component that allows a grid operator to classify single data points and to generalize classifications to new data points using a statistical classifier that is based on a 3-dimensional KD-tree. Such a classifier enables the accurate classification and labeling of data points within the knowledge model, which is a prerequisite for the semantic validation and monitoring of data points. This component increases the degree of local automation of FSSNs, which has advantages with respect to the cost of operation of the smart grid and its resiliency, since decisions can be made at a much higher pace and, potentially, even automatically.

The app placement engine provides a global view of the grid, since it considers the general state of the grid, instead of just one FSSN. The Global FSSN Manager takes as input the current state of each stations (for example, load, current resource consumption), as well as, the app description (for example, required resources,

dependencies), and grid operator preferences that reflects the interest in the four proposed performance metrics (i.e., network wide resource consumption, load balancing, minimal movement, and replicated placement of dependencies). The objective of the Application Placement Engine is to produce a feasible app placement that maximizes fault tolerance and satisfies all app dependencies. The approach presented in this research project were tested with different scenarios, producing satisfactory results. However, when testing with a larger scenario, the system took too long to compute the optimal placement. To cope with this limitation, and still offer a feasible placement, it was introduced a time stopping criterion, a grid operator could stop the placement engine even after just 1 second, or he can decide to run the system for longer to improve the response.

The semantically integrated dashboard corresponds to the user interface that allows the management of FSSs. Two versions of the dashboard were developed: the traditional 2D dashboard, which corresponds to a user interface running in a browser in a computer and an AR app, which not only demonstrates the capabilities of the backend components, it also provides an immersive user experience to grid operators. Grid operators can perform the same operations as in the traditional 2D dashboard, but their interaction with the system is much richer, dynamic and user friendly.

As mentioned, the ideal path of this research project is to find the right partners within Siemens, and a California utility to run a pilot project that allows to see the system working in an actual environment, and that permits quantifying the effectiveness of the developed dashboards in reducing the cognitive overload of grid operators.

CHAPTER 6: Benefits to Ratepayers

The implementation of this project could help utilities improve their operations, enhance their outages management, and reduce electrical losses through faster control decisions. Moreover, it can allow utilities to take better advantages of asset utilization derived from reduction in manual grid operation efforts and could improve grid monitoring capabilities to enable better planning. These benefits can directly affect private, commercial, and industrial consumers who can profit from more reliable services and reduced business losses due to improved grid resiliency as well as potential bill savings thanks to a higher degree of automation which also implies a lower grid cost operation. A higher degree of automation within the Smart Grid also implies benefits for society as a whole, since this can enable a higher percentage of renewable energy resources, to reduce environmental emissions, improve the security of electricity delivery, and reduce the import of crude oil through transportation electrification.

The sample calculation shows an estimation of a possible reduction in the System Average Interruption Duration Index (SAIDI) value, as a result of the deployment of the prototypes developed within the project in the Pacific Gas & Electric Company (PG&E) distribution grid (Table 5). According to PG&E's 2017 annual reliability report, the SAIDI for PG&E in 2017 was of 357.7 hours, of which 90 hours are due to distribution system outages excluding major events (Pacific Gas & Electric Company 2017). Such an index is higher than the 2016 one, which was of 83 hours. Considering only PG&E's customers (4.6 million residential and 700,000 commercial, industrial and others), a SAIDI of 90h translates to a total load not served of about 1138.7 GWh. Assuming a value of service of \$2.5/\$10/\$25 per kWh for residential/commercial/industrial customers (numbers from (Vicenzo and Steven 2012)), this implies an economic cost to PG&E customers due to sustained outages to the distribution system (excluding major events) in 2017 of an estimated \$10.3 billion (Pacific Gas & Electric Company 2017), (Pacific Gas & Electric Company_b 2017), and (Vicenzo and Steven 2012).

	Number of PG&E accounts in 2017	Load not served due to PG&E distribution grid outages in 2017, excluding major events	Estimated Value of Service. Estimates from (Vicenzo and Steven 2012)	Estimated economic cost of sustained outages, PG&E customers and distribution grid outages only, excluding major events
Residential	4.6M	538.2 GWh	\$2.5 / kWh	\$1,345,500,000
Commercial	0.5M	400.5 GWh	\$10 / kWh	\$4,005,000,000
Industrial	0.2M	200 GWh	\$25 / kWh	\$5,000,000,000
Total		1138.7 GWh	-	\$10,350,500,000

Table 5. Estimate of Reduction in Costs in PG&E Distribution Grid

Source: Siemens CT

The adoption of the technologies developed in this project could have the potential to yield a reduction of the distribution grid SAIDI (excluding major events) of about 10 percent (nine hours per year), which would translate to cost savings of more than \$1 billion per year to PG&E customers. This would imply more than \$134 million per year to PG&E residential customers. However, even using a more conservative estimate of 5 percent (a SAIDI reduction of 4.5 hours per year), the potential yearly economic benefit of the developed project would be upwards of \$67 million, considering only California residential customers. Moreover, a higher degree of automation in grid control rooms would also allow available staff to focus on services and locations of the grid, in which the highly automated proposed system cannot recover on its own.

LIST OF ACRONYMS

Term/Acronym	Definition	
AC	Alternating Current	
AR	Augmented Reality	
ADR	Automatic Demand Response	
API	Application Programming Interface	
Арр	Application	
FSSN	Future Secondary Substation	
КР	Knowledge Pack	
OSF	Open Semantic Framework	
QUDT	Quantities, Units, Dimensions, Types	
RDF	Resource Description Framework	
REST	Representational State Transfer	
SPARQL	SPARQL Protocol and RDF Query Language	
SPIN	SPARQL Inference Notation	
SSN	Semantic Sensor Network	
URI	Uniform Resource Identifier	
VR	Virtual Reality	
2D	Two-dimensional	
3D	Three-dimensional	

REFERENCES

- De Moura, Leonardo, and Bjørner Nikolaj. 2008. "Z3: An efficient SMT solver." *In International conference on Tools and Algorithms for the Construction and Analysis of Systems.* Springer, Berlin Heidelberg. 337-340.
- Energy, United States Department of. 2009. *Smart Grid System Report.* Accessed 01 20, 2019. http://energy.gov/sites/prod/files/2009%20Smart%20Grid%20System%20Report.pdf. DOE.
- Faschang, Mario, Stephan Cejka, Mark Stefan, Albin Frischenschlager, Alfred Einfalt, Konrad Diwold, Filip Pröstl Andrén, Thomas Strasser, and Friederich Kupzog.
 2017. "Provisioning, deployment, and operation of smart grid applications on substation level." *Computer Science-Research and Development* 32(1-2):117-30.
- Funk, Markus, Mareike Kritzler, and Florian Michahelles. 2017. "HoloLens is more than air Tap: natural and intuitive interaction with holograms." *In Proceedings of the Seventh International Conference on the Internet of Things.* ACM.
- Hodges, Jack, Kimberly García, and Steven Ray. 2017. "Semantic Development and Integration of Standards for Adoption and Interoperability." *Computer* 26-36.
- Mayer, Simon, Hodges Jack, Yu Dan, Kritzler Mareike, and Michahelles Florian. 2017. "An open semantic framework for the industrial Internet of Things." *IEEE Intelligent Systems* 96-101.
- Mizutani, Iori, Kritzler Mareike, García Kimberly, and Michahelles Florian. 2017. "Intuitive interaction with semantics using augmented reality: a case study of workforce management in an industrial setting." *In Proceedings of the Seventh International Conference on the Internet of Things.* ACM. In Proceedings of the Seventh International Conference on the Internet of Things.
- Naylor, David, Finamore Alessandro, Leontiadis Ilias, Grunenberger Yan, Mellia Marco, Munafò Maurizio, Papagiannaki Konstantina, and Steenkiste Peter. 2014. "The cost of the S in HTTPS." *In Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies.* ACM. 133-140.
- Needham, H, D Brown, and L Carter. 2012. *Impacts and Adaptation Options in the Gulf Coast.* Louisiana State University. Center for Climate and Energy Solutions. http://www.c2es.org/docUploads/gulf-coast-impacts-adaptation.pdf.
- Pacific Gas & Electric Company. 2017. *2017 Annual Electric Reliability Report.* Accessed 01 20, 2019. https://www.pge.com/pge_global/common/pdfs/outages/planning-and-preparedness/safety-and-preparedness/grid-reliability/electric-reliability-reports/AnnualElectricDistributionReliabilityReport2017.pdf.

- Pacific Gas & Electric Company_b. 2017. *PG&E Overview.* Accessed January 20, 2019. http://www.pgecorp.com/corp_responsibility/reports/2016/bu01_pge_overview.j sp.
- Pradhan, Subhav, Dubey Abhishek, Khare Shweta, Nannapaneni Saideep, Gokhale Aniruddha, Mahadevan Sankaran, Schmidt Douglas, and Lehofer Martin. 2017.
 "Chariot: A holistic, goal driven orchestration solution for resilient iot applications." *ACM Transactions on Cyber-Physical Systems.*

Sichon, Connie. 2014. My Story. Sacramento: Energy Press.

- Tucker, Chris, Shuffelton David, Jhala Ranjit, and Lerner Sorin. 2007. "Opium: Optimal package install/uninstall manager." *In Proceedings of the 29th international conference on Software Engineering.* IEEE Computer Society. 178-188.
- Urgaonkar, Bhuvan, Rosenberg Arnold L, and Shenoy Prashant. 2007. "Application placement on a cluster of servers." *International Journal of Foundations of Computer Science 18* 1023-1041.
- Vicenzo, Giordano, and Bossart Steven. 2012. *Assessing Smart Grid Benefits and Impacts: EU and US initiatives.* Join Report of EC JRC and US DOE, JRC.
- Whyte, Jennifer, Bouchlaghem N., Thorpe A., and McCaffer R. 2000. "From CAD to virtual reality: modelling approaches, data exchange and interactive 3D building design tools." *Automation in Construction 10* 43-55.