SOGNO

D3.1 v1.0

Description of new SOGNO techniques for autonomous and self-healing power systems

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Program, under Grant Agreement no 774613.

Project Name	SOGNO
Contractual Delivery Date:	31.10.2018
Actual Delivery Date:	31.10.2018
Author(s):	RWTH TI
Work package:	WP3– Advanced data analytics techniques for autonomous and self-healing power systems
Security:	PU
Nature:	R
Version:	V1.0
Total number of pages:	21

Abstract
This document describes the progress of task T3.1 and T3.2 in WP3. The main focus of those tasks is to characterize and develop the required techniques for a self-healing power system. Within WP3, such techniques can be represented in the form of two SOGNO services: first, fault location isolation and service restoration (FLISR), and the load and power generation prediction (LP and GP), respectively.

Keyword list
Autonomous and self-healing power systems, fault location isolation and service restoration (FLISR), load prediction (LP), (power) generation prediction (GP), machine learning, data-driven methods.

Disclaimer
All information provided reflects the status of the SOGNO project at the time of writing and may be subject to change.

Executive Summary

The reliable operation of grids and the continuity of power supply to final customers are two of the main goals of electric grid operators. With the evolution of the power system due to the increasing penetration of renewable sources and other distributed energy resources, grid management becomes more complex and requires additional intelligence to achieve these targets. Traditional approaches are not sufficient anymore to deal with the challenges brought by the emerging technologies connected to the grid, above all at the distribution level of the electric system. For this reason, new solutions have to be conceived to guarantee the reliable and efficient operation of the network and intelligent automation needs to be deployed to ensure the prompt reaction to possible anomalies in the power system operation.

While several efforts have been made in last years by the scientific community to propose new concepts and solutions for the automation of the distribution grids, the goal of the SOGNO project is to perform a further step towards the smart automation of the electric system. This can be achieved by introducing innovative machine learning approaches to deal with the complexity of today's power systems. The objective from this standpoint is to exploit the benefits associated to data driven intelligence for improving the conventional approaches for power system automation. To this purpose, two services are designed in WP3 to provide autonomous self-healing features to the distribution grid: A Fault Location Isolation and Service Restoration (FLISR) service and a Load and Generation Prediction (LP and GP, respectively) service.

Therefore, this document provides the definition of an autonomous and self-healing power system. Additionally, it presents the progress made in tasks T3.1 and T3.2 of the SOGNO project for the design of the two services mentioned previously until the submission date of this report. In T3.1, the concepts behind the machine learning based services are developed. In Task 3.2, the methods, procedures and approaches to implement the two services are defined.

Authors

Partner	Name	e-mail
RWTH TI		
	Arash Behboodi	arash.behboodi@ti.rwth-aachen.de
	Rana Alhalaseh	rana.alhalaseh@ti.rwth-aachen.de
GH		
	Artur Löwen	aloewen@gridhound.de

Table of Contents

1. Introduction

The project *Service Oriented Grid for the Network of the Future* (SOGNO) is funded by the Work Program H2020-LCE-2017-SGS. It has officially started in January 2018.

1.1 Related Project Work

The report is based on the work done in task T3.1 and T3.2 of work package WP3. These tasks deal with the definition and design of the services aimed at providing the power system with autonomous self-healing features. The methods, procedures and approaches to design a Fault Location Isolation and Service Restoration (FLISR) service and a Load and Generation Prediction (LP and GP, respectively) service based on innovative machine learning techniques are developed here. As shown in Figure 1, the services designed in WP3 are then provided to WP4 to integrate them within the cloud platform. This step provides virtualization of the distribution grid intelligence, before the final deployment into the SOGNO laboratory and field trials.

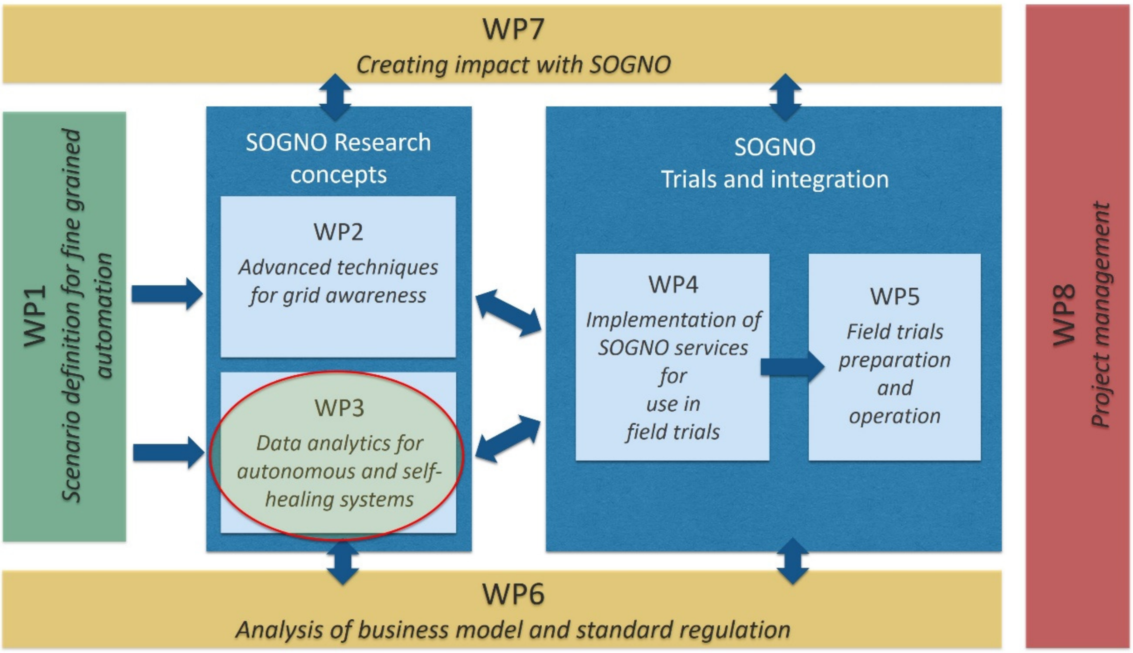

Figure 1: Overview of SOGNO activities.

1.2 Objectives of the Report

The objective of this document is to provide the main concepts behind autonomous and self-healing power systems. In particular, the report presents the main services implemented in WP3 to this purpose, namely Fault Location Isolation and Service Restoration (FLISR) and Load and Generation Prediction (LP and GP), which lead to a self-healing power system. Machine learning based methods are used as basis for the development of these services and therefore an insight on the deep learning techniques considered for the implementation of these algorithms is also presented.

The FLISR service aims at detecting faults in the power grid upon occurring. When a fault occurs, the faulty segment or section of the grid is detected (fault location) and the affected area is isolated (fault isolation). If the grid has multiple options to be reconfigured, the accurate location of the fault allows minimizing the portion of the system that has to be isolated and gives the possibility to restore the power supply for the rest of network. Thus, minimizing the number of customers affected by the outage. The FLISR algorithm allows performing these tasks automatically, promptly reacting to the fault occurrence, hence limiting the disservice for all the customers located in a non-faulty area. Moreover, the accurate location of the faulty section provides useful information to the operators for quickly detecting the fault and fixing it, thus determining an overall reduced disconnection time also for the customers within the faulty area of the grid.

On the other hand, the LP and GP services provide beneficial amount of information to the power grid to plan its operation upon power demands from one side and renewable energy generation from the other side. In a scenario where both load and generation are volatile and difficult to predict (such as in the electric distribution grid) and the grid is operated much closer to its operating limits, the availability of these services is essential for grid operators to forecast possible issues in the system and to take adequate countermeasures.

1.3 Outline of the Report

The first part of this report (Chapter 2) presents the general concept behind autonomous and self-healing power system. Chapter 3 provides information on the concepts and the initial ideas for the design of the techniques for FLISR, LP and GP. At the date of submission of this report, the details of the implementation are still under discussion, due to the necessity of customizing the services for different distribution grid scenarios and according to the specific availability of hardware infrastructure on the field. As a consequence, a more in-depth presentation of the algorithm details will be provided in D3.3. Finally, Chapter 4 gives an insight on the approaches considered for the application of machine deep learning techniques to the problems at hand.

1.4 How to Read this Document

The Deliverable can be read as a standalone document, but other Deliverables can be read to get a more thorough view of the concepts advanced in the SOGNO project. In particular, other Deliverables closely related to this one, are the following:

- D1.1 – Scenario & architectures for stable & secure grid (M12): it gives the description of power system scenarios investigated in the project and a high-level description of the services with their mapping to the different SGAM layers.
- D3.4 - Description of initial interfaces & services for autonomous self-healing power systems (M12): it provides details on the interfaces needed for the integration of the services described in this report into the SOGNO virtualized platform.
- D4.1 – SOGNO system architecture definition (M10): it provides the view of the overall SOGNO architecture used for the integration of the designed services.
- D3.3 – Validation and description of the techniques, interfaces and services for autonomous & self-healing power systems (M22): it will provide the detailed description of the designed algorithms together with the results obtained during the evaluation and validation phase.

2. Definition of autonomous and self-healing power systems

One of the main tasks of a grid operator is the provision of a continuously available power system infrastructure and its reliable operation for a continuous source of power to the consumers and a continuous sink for prosumers and generators. The grid operators are always keen to improve the resilience of the grid, since it allows them to reduce penalties as well as economic losses due to loss of productivity and customer inconvenience from power supply disruptions.

Every now and then unforeseen events happen, leading to a fault in the electric grid. A fault in the grid leads to the intervention of circuit breakers, which in turn will cause power supply disruptions. A multitude of events can determine a fault in the power system. As an example, during a storm a tree can fall into overhead lines (~25% of the times this is the source of an outage), equipment, like transformers, can fail (~18%), underground cables or joints can break (~10%), accidents can happen (~9%) to name just four of the most likely events. An autonomous and self-healing power system must be capable of promptly reacting to any abrupt event in the best possible way. In case of fault, an important target is to limit the number of customers affected by the power supply disruption. For all those users affected by the outage, instead, the main objective is to limit as much as possible the duration of the supply interruption.

The main feature for enabling such an autonomous and self-healing systems is the availability of a fault location, isolation, and service restoration (FLISR) intelligence. Such systems are also known as FLISR systems. The task of a FLISR module, as the name suggests, is first to determine the exact zone or section of the grid where the fault occurred. Based on this information, it is then possible to reconfigure the electric grid so that only the faulted area is disconnected (isolated), while the rest of the grid is supplied back through the service restoration procedure. For the practical implementation of the FLISR functionalities, an autonomous self-healing grid system requires sensors and actuators (switches) installed in the power system and an advanced software to translate the sensor inputs into actuation commands, when needed, without the action of a human operator. The sensors provide grid measurements that allow the identification of a fault as well as its location. The switches allow isolating the fault and reconfiguring the grid to restore the power supply in all the areas not directly affected by the fault.

While FLISR is the most important component in an autonomous and self-healing power system in order to react to unforeseen events, other components or software modules can also be important in the daily operation of the power system to provide the intelligence necessary for enabling “preventive” self-healing functionalities. This is the case for example of the load prediction (LP) and generation prediction (GP) services. The primary goal of the LP and GP algorithms is to provide a forecast about the future power consumption and generation, respectively, of consumers or generators in the grid. In the distribution grid scenario, this function is becoming more and more crucial to foresee the future operation patterns, since both load (demands) and generation are volatile at distribution level and, with the increasing levels of power consumption and installed renewable generation, many grids are approaching their operational limits. The accurate predictions provided by ad hoc designed LP and GP services can thus serve as an essential element for grid operators to foresee the possible occurrence of contingencies in their network and to promptly take adequate countermeasures. If the LP and GP service are interconnected to other automatic functionalities, such as grid topology reconfiguration algorithms, the manual intervention of grid operators is not needed anymore (in case, operators will only get system awareness and will take care of supervising tasks) and the LP and GP will be a fundamental part of the automation loop aimed at providing “preventive” self-healing functionalities.

Due to their characteristics, the FLISR, LP and GP services implemented in SOGNO are deemed as the most important services to enable an autonomous self-healing power system.

3. Services and techniques to achieve autonomous and self-healing power systems

3.1 Fault location isolation and service restoration (FLISR) service

The main task of the advanced software for autonomous and self-healing grids is to perform data analysis in real-time, and to act based on the FLISR concept. The FLISR software needs first to identify that a fault has occurred, pin-point its location, and then to isolate the faulty section and restore the power supply to all the non-faulty sections that were automatically disconnected by the protection devices. Through this procedure, the FLISR service allows automatically reacting to a fault occurrence in the grid in order to minimize the number of customers affected by the power disruption, which usually results in an outage.

Ideally, the main steps triggered by the occurrence of a fault in a self-healing grid equipped with automated FLISR functionalities are the following:

- Fault occurrence: when a fault occurs, the protective devices installed in the grid act to disconnect the power supply to avoid any safety problem for the final customers and to prevent the damage of the electrical equipment connected to the grid. Different types of fault can occur, which can be mainly classified in single, double or three phase-to-ground faults, or line-to-line faults.
- Recloser operation: most of the faults happening in the grid are temporary faults, for which a short disconnection of the power supply is sufficient to extinguish the fault itself. For this reason, shortly after the intervention of the circuit breakers, reclosers try to restore immediately the power supply to verify if the fault was a temporary one and, in such a case, to minimize the duration of the power disruption.
- Outage: if the fault is however not temporary, then it is classified as permanent. It means that a crew needs to intervene for fixing the problem. In this moment, all the sections of the grid subtended by the circuit breaker(s) that reacted to the fault are disconnected from the power supply (it is worth noting that currently, in most of the distribution grids, and particularly at the low voltage grid level, the detection of an outage from the grid operators usually happens only thanks to the calls from disconnected customers and this is one of the reasons for large outage durations).
- Fault location: if the fault is permanent and determines an outage, the FLISR has to react for reducing the number of customers affected by the supply interruption and to limit the duration of the disservice. Regardless of the specific technique implemented to carry out the fault location, the associated algorithm receives as input some measurement data related to the period in which the fault had occurred. Through the processing of these data it identifies the location, namely the specific section or line of the grid, where the fault happened.
- Fault isolation: once the position of the faulty line is detected, the FLISR service opens the switches immediately upstream and downstream (if present) the faulty section to guarantee the isolation of the faulty area.
- Service restoration: since the faulty part of the grid has been already isolated, the rest of the grid can be supplied back to guarantee that only a minimum number of customers suffers the power disruption. This can be performed for example by acting on normally open switches that, during the emergency conditions, can be closed to reconfigure the grid topology and to allow the restoration of the power supply in the non-faulty areas of the grid.
- Fault fixing: thanks to the information provided by the FLISR algorithm, the grid operators immediately know the location of the grid where the fault has happened and can directly go there for fixing the fault. This guarantees a reduction of the outage duration, since the time dedicated to “find” the fault can be significantly reduced.
- Return to normal operation: after the fault has been fixed by the crew in charge, the control centre can be notified about the successful intervention and grid operators can restore the power supply into the previously faulty area (by acting on the switches opened during the fault isolation step) and can reconfigure the grid to work with the default topology used during normal operating conditions.

Several techniques have been investigated in the scientific literature to apply the FLISR concept and mainly to obtain the accurate location of a fault in the grid. Conventional solutions range from simple approaches, mainly tailored to passive grids, based on the collection of the logical information associated to the switches and breakers that detected a faulty current, to much more complex solutions based on advanced mathematical techniques, such as impedance-based methods, traveling-wave approaches or Wavelet transforms. These techniques, while generally able to deliver high accuracy results in specific scenarios, also present some drawbacks. One of the main disadvantages is the requirement to have a very high number of measurement devices deployed in the grid, often aimed at a full coverage of measurements at the different nodes of the electric network. Such a requirement is hardly achievable in most of the cases by the Distribution System Operators (DSOs), at least in short times. Furthermore, some techniques also pose specific requirements on the type of measurement data to be received (e.g. analog samples of the voltage and/or current waveforms), or on the synchronization of these measurements (very accurate time synchronization usually can be only achieved via devices equipped with a GPS receiver, such as Phasor Measurement Units), or for the communication with the software intelligence, which can be an additional constraint for the application of the method.

In the SOGNO project, one of the main goals is to enable the automation of distribution grids by using only limited number of devices on the field, thus reducing the requirements for hardware deployment on the field. For this matter, one of the most promising solutions to deliver power system automation services, based on the information provided by only few devices on the field, is the use of data driven approaches based on last generation machine learning techniques. The main idea is that faults can be identified by analyzing the patterns from the measurements. Machine learning based methods are also widely found in the literature to implement FLISR. In general, they are the most appropriate approach to utilize mixed sources of information read from the field to identify specific patterns associated to power system events. As a consequence, such techniques have been successfully applied also to detect the faulty sections in a grid [1]. Since this approach is based on the processing of measurement data, measurement devices need to be strategically placed for a good coverage of the grid. Otherwise the accuracy of the results can be affected. Measurement information can come from measurement devices specifically installed for this purpose or from instrumentation already available in the field, such as the measurement equipment associated to the same breakers and switches in the grid.

3.2 Load and power generation prediction service

In the last decades, electrical load prediction as well as power generation prediction have been widely subject to investigation, as they present an enormous amount of useful information regarding network operation, which contributes to reliable and sustainable power supply [1] [2]. The business case behind LP and GP comes from the importance of having a stable power network, which meets the costumers' electricity demands, while maximizing the utilization of renewable energy resources. In the context of autonomous, self-healing power systems, load and generation prediction provide forecast information about the operating conditions of the grid in future instants of time. Such information can be used by grid operators to verify if the expected power profiles allow a reliable operation of the grid or, in case, to identify possible contingencies that could occur. In a smart grid scenario, the forecast information can be coupled with services specifically devised to automatically process the forecast information so that the detection of possible contingencies in the system is also performed automatically. Based on this information, grid operators can be alerted and promptly take countermeasures if needed or, again, automatic routines can be triggered to reconfigure the grid for guaranteeing its reliable operation.

The load prediction (LP) and generation prediction (GP) service consists basically of two different parts: the load forecast component and the power generation forecast component, where the former provides an idea about the electrical power demand at future time instants, and the latter provides the information about the electrical power supplied by renewable energy resources, also at future time instants. Within the SOGNO project, different cases are taken into account for the LP service, such as the possibility to forecast the power consumption of secondary substations or the footprint of power consumption for individual customers. It is worth noting that, as known from the literature, when going down to the prediction of power consumption for single customers, random events and random behavior of the users can play a relevant role on the accuracy and reliability of the forecasting results. For the GP service, the renewable energy resources of interest at distribution grid level are mainly solar and wind power [1]. Within the SOGNO project, the focus is manly to predict the power generated by photovoltaic (PV) systems, mostly relying upon weather information such as solar irradiance, temperature, and cloud coverage.

Machine learning based methods are considered as an appropriate approach to implement and develop the LP and GP services by utilizing appropriate historical information. In case of the LP service, the historical power consumption values are used as an input in order to understand the pattern of the power consumption in a certain day of the week, or a certain month of the year. The historical power consumption values are in the form of active and reactive power, which are sampled at a certain uniform rate, for example every 5, 10 or 15 minutes. Hence, the machine learning based prediction scheme is trained using such historical information and is used to predict the power consumption of the next hours ahead, always at the same rate as the starting data.

On the other hand, historical weather conditions and the corresponding power generated in a certain PV system are utilized to predict the future power generation. Such historical information includes the past weather conditions and the corresponding power both measured at a certain rate. As an intermediate step, the most dominant and relevant weather conditions are chosen to take part in the prediction process. Similar to LP service, the historical information is used to understand the pattern of the generated power based on certain weather conditions in order to train the machine learning based prediction scheme. Upon obtaining the weather forecast of the next day, the corresponding power generation is hence predicted. In other words, for example, the pattern of the generated power given at cloudy days is understood in order to predict the corresponding power for similar future conditions.

In terms of the prediction horizon, i.e., the future prediction duration or future number of predicted samples, both the LP and GP services can be divided into the following three categories:

- Short term prediction: the prediction horizon can be for several hours ahead, for example three to four hours ahead. Mainly, this will be the prediction horizon considered within SOGNO project.
- Medium term prediction: the prediction horizon can be up to one day.
- Long term prediction: the prediction horizon can be for the following day.

Given the two services under consideration, both of them can be understood as a direct regression problem which utilizes historical sets of data. Irrespective of the prediction horizon, many statistical and mathematical approaches have been considered in literature for such a problem. The major and widely applied approach employs time series methods such as autoregressive moving average (ARMA) and Box-Jenkins autoregressive integrated moving average (ARIMA) [3] [4]. Such statistical and parametric methods are capable of achieving acceptable prediction results in the case of stationary time series data and without missing data. However, the main drawback is their incapability and limitation in case of abrupt changes in the environment [5].

Machine learning and artificial intelligent based techniques can be utilized for the problem under consideration, where the previously mentioned limitation can be overcome. Such techniques have drawn the attention lately, where different work can be found in literature. Support vector machines, deep neural networks, decision trees and fuzzy logic are among the utilized machine learning based techniques [6] [4] [7] [8] [9].

4. Machine learning and data analytics methods to achieve autonomous and self-healing power systems

The field of machine learning and big data analytics has expanded in the last decades. These methods have proved their reliability to solve some problems automatically, with minimum human interaction. Furthermore, they are capable of extracting useful information to analyze and predict behaviors and patterns. For example, artificial neural networks, support vector machines, decision trees, and other machine learning based methods have been widely employed to solve classification, prediction, and pattern recognition or detection problems. Given the complexity of the power systems, particularly in future scenarios, the previously explained services in WP3 can gain great benefit if recent machine learning and data analytics-based methods are adopted as basis for their implementation.

4.1 A machine learning based FLISR approach

The design of a FLISR approach is particularly challenging due to multitude of aspects, also it dependent on the specific grid characteristics to be taken into account. In the power system, different types of faults can happen, which lead to a different evolution of the main electrical quantities (current and voltage) also according to the impedance characteristics of the grid and the details on its grounding system. Moreover, faults can be characterized by different features, for example, a different fault impedance, which again significantly affect the transient voltages and currents arising after the fault occurrence. Last but not least, the topology of the grid under consideration and the presence of possible energy sources in the network able to “supply” the fault has a relevant impact on the post-fault grid conditions. As a matter of fact, based on these considerations, the implementation details of the FLISR algorithm need to be tailored for the specific grid where they will be implemented. The FLISR service developed during the SOGNO project will thus follow a general methodology based on the use of machine learning techniques, which is introduced in the following. However, the implementation details will be specific for each one of grids or the test cases taken into account. The developed FLISR algorithms will be tested in laboratory environment and then delivered to the field trials during phase 2 (which starts at the beginning of 2019) of the SOGNO project.

4.1.1 Data aggregation in a preliminary work

The FLISR service can be divided in two steps: first, identifying the faulty parts (lines) in the grid; second, isolating the faulty portion of the grid and reconnecting the power supply for all those loads that are not connected to this part of the network. The first part of the problem has been initially considered. In order to clearly understand the fault location part of the problem, and the requirements to apply it on grids with very large number of nodes and lines, several benchmark grids have been considered. As a starting point, the IEEE 13 and 37 bus feeders, which are shown in Figure 2 and Figure 3, respectively, are considered.

The GridLabD power flow simulation tool has been used for data aggregation, i.e. in order to extract data from the grid in the form of voltages and currents at different nodes and lines of the grid, respectively. All possible scenarios of faults are simulated, such that each scenario represents a fault occurring at a certain line in the grid and the no fault scenario. Hence, the total number of scenarios reflects the total number of lines in the grid, plus the normal operation case. In simulation, in order to obtain multiple examples of grid measurements under each fault scenario, the loads connected to the grid have been varied randomly between 0 and 150% of their nominal values. Single-phase grids are simulated at this stage of the work.

The measurements obtained from the simulation tool, along with its corresponding fault locations are used as a dataset. The dataset is divided into two parts: training and testing dataset, where the former is used to train and learn a machine learning based model to classify the measurements according to their fault location. The latter is however used to test the classification

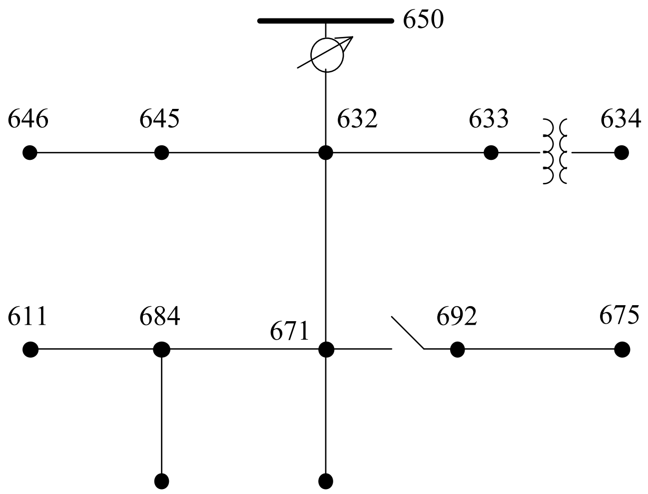

Figure 2: IEEE 13 bus feeder [11]

accuracy of the implemented network [11].

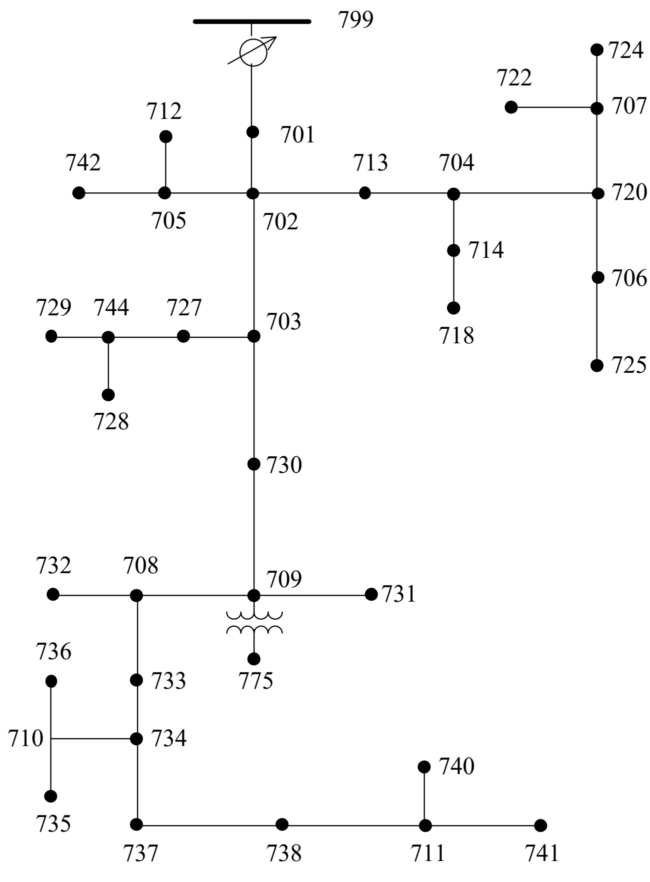

Figure 3 - IEEE 37 bus feeder [11].

In IEEE 13 grid model, there are 12 number of classes (included the normal operation case) which the aggregated measurements belong to. On the other hand, IEEE 37 grid model contains a total of 36 classes. At this point, it is worth mentioning that extra nodes have been included into the previously mentioned grid models for simulation purposes.

4.1.2 Machine learning based feature selection methods

Given the obtained measurements along with the corresponding fault location (line segment) in the form of a dataset, feature selection methods are used in our preliminary work in order to rank the measurements, i.e., features, based on their importance and relevance to the classification problem. The output of the feature selection methods is useful to provide the following information:

- Reduce the dimensionality of the feature vector, i.e. the measurement values per fault scenario, in order to reduce the computational cost and memory usage.
- Remove the features which provide redundant and irrelevant information and keep those relevant and important ones. This process helps improving the learning process and provides a classifier with superior classification accuracy.
- By ranking the features (measured quantities from a grid) based on their relevance to the problem in hand, the feature selection method provides an optimal measurement unit placement for an optimal FLISR performance.

Different feature selection methods have been tested in this preliminary work, which are variance thresholding, mutual information and correlation coefficient. The main idea behind variance thresholding is to remove features which has low variance [12], as they contribute less to the problem in hand. As one approach could be to define specific thresholds, the threshold has been kept to zero in the preliminary analysis and the output of this feature selection method is the features being ranked based on their relevance. Furthermore, the mutual information measure is based on the same concept from information theory background, which provides the mutual dependency between two random variables [13]. Within the problem in hand, it measures it between each individual feature with the class or label. Finally, the correlation coefficient measure is the covariance between a certain feature vector and the corresponding class vector divided by the multiplication of individual standard deviations. The correlation coefficient does not exceed 1, which indicates highly relevant and important features, whereas 0 indicates irrelevant features, which are to be neglected.

4.1.3 Machine learning based classification methods

An updated dataset is formed based on the relevant features, which are chosen based on the explained feature selection methods. The dataset contains several examples of fault scenarios, for all possible fault locations, along with the normal operation. The dataset is divided into two parts: training and testing dataset, where the earlier trains the machine learning based classifier, and the latter tests it to determine its detection accuracy.

For the detection issue, several classification methods are going to be tested; such as deep neural network (DNN), decision tree (DT), and random forest (RF). The DNN, or in specific, the feed forward DNN (FFDNN), have captured the attention due to their outstanding performance in different applications, such as classification, pattern recognition, and prediction. The DNN consists of multiple hidden layers, each layer contains a certain number of weighted neurons. The input is summed over all layers and transformed through the activation function at the final layer [14]. During the training process, the cross-entropy loss function is minimized.

A DT classifier is suitable for classifying data with common attributes [15], which have a set of nodes, edges and leaves, which results in a structure similar to a tree. Whenever the data is to be classified into two classes, a node is generated. The resulting number of edges at a node indicates the possible categories that a certain feature has, whereas the leaves are the resulting classification output, which is the estimated output class of the input vector.

Finally, a RF is a bootstrap ensemble of multiple DTs, combined in order to construct and come up with a more powerful, reliable and robust classifier [16]. Different from DT, the main characteristics behind RF is randomness, in order to decorrelate the underlying trees, so that the resulting ensemble scheme has a low variance [17].

4.1.4 Service restoration methods

The service restoration part of the FLISR service is a very complex and challenging problem faces the grid operation. One way to approach this task is by simply considering it as a multi-objective, multi-constraint optimization problem. The effective plan for restoration must take care of the following [18]:

- Minimize the number of healthy out-of-service sectors.
- Minimize the number of switching operations for the reduction of interruption time.
- Prioritize switching operations in remotely controlled switches (RCSs) for the reduction of interruption duration and cost.

Such requirements can be addressed as the objectives of the service restoration problem, which are conditioned on [18]:

- Non-violation of the pre-defined limits of node voltage, network loading and substation loading.
- Quickly implementing the service restoration algorithm in order to ensure customer's satisfaction and avoid any penalization.

The optimization problem of the service restoration part of FLISR has been differently addressed in literature. For example, Multi-objectives Evolutionary Algorithm (MOEA) based methodologies can be a straight forward choice, as they treat each objective function separately, thus better exploration of the search space to obtain better results [18]. As explained in [19] [20], table based MOEA methodologies provide reliable and acceptable performance for large scale networks.

The mathematical expression of the restoration problem can be formulated by minimizing the following objective functions [18]:

$$\psi_{RCS}(G, G^i), \psi_{MCS}(G, G^i), X(G), B(G), V(G) \qquad (1)$$

where G in (1) denotes a configuration of the grid under consideration, which is represented by a graph forest, whereas G^i is the grid configuration after a certain section has been identified and isolated. As explained in [18], the functions in (1).**Fehler! Verweisquelle konnte nicht gefunden werden.** are defined as following:

- $\psi_{RCS}(G, G^i)$ is the number of switches operations in RCS. On the other hand, $\psi_{MCS}(G, G^i)$ denotes the number of switches necessary to obtain G from G^i
- $X(G)$ is the maximum network loading in G, which is defined as $X(G) = max \frac{x_j}{\bar{x}_j}$, where $\bar{x}_j$ is an upper bound for electric current for each current magnitude x_j in the line j.
- $B(G)$ is the maximum loading in G, which is defined by $B(G) = max \frac{b_s}{\bar{b}_s}$. $\bar{b}_s$ is an upper bound for each magnitude of current injection b_s provided by a substation s.
- $V(G)$ is the maximum relative voltage drop in $V(G)$ defined as $V(G) = max \frac{|v_s - v_k|}{\delta}$, where v_s and v_k is the node voltage magnitude at a substation bus s and at a network bus k, respectively, whereas δ is the maximum permissible voltage drop.

The objective function can be formulated as a weighted sum of each functions presented in (1) and solved using MOEA algorithms [21].

4.2 Load and power generation prediction

The application of machine learning algorithms perfectly fits the problem of load and generation prediction, since both rely upon the analysis of historical data to determine the forecast. Whether the purpose is to implement the future load demand (LP service) or the future generated power (GP service), the considered prediction approach is similar. Historical data are required to learn and train the predictor scheme under consideration. Afterwards, the process of predicting the future values takes place using the trained scheme. Similar to the FLISR service, the design of both the LP and GP services is started in the first phase of the project, but their implementation on the pilots is expected for phase 2 (which starts at the beginning of 2019).

4.2.1 Load prediction

Irrespective of the fact that single or multiple customers are considered as target for the prediction service, historical load consumption in the form of active and reactive consumed power is required, along with the associated time stamp. The historical power values are sampled at a certain rate, for example every 5, 10 or 15 minutes. As a starting point for the design of the algorithm, the online UCI machine learning repository in [21] was used, since it provides the active and reactive power consumption for 47 months (from December 2006 until November 2010). The power consumption data available in this database is sampled at a one-minute rate.

The load prediction process is explained in Figure 4. The concept behind it is to use several examples of the load consumption at a certain month (or day of the week), like January (or Monday), in order to predict the consumption at January the next year (or Monday the next week).

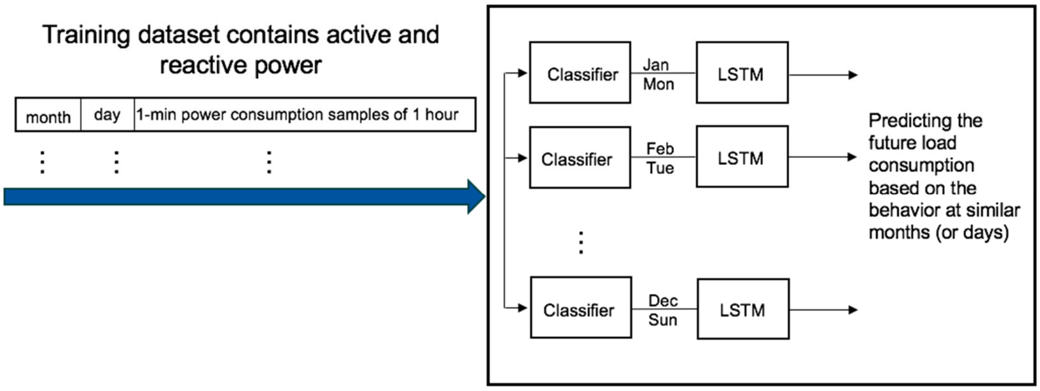

Figure 4: Block diagram of the load prediction service.

The classifier makes sure to train each deep neural network with examples of the same month (or the same day), in order to understand and recognize the pattern at each month (or day) individually. Afterwards, it is able to predict the behavior in the future time slots, based on the required duration.

The long-short-term memory (LSTM) recurrent neural network is used at this point as a suitable prediction scheme. It is able to explore the long dependencies in the electric load and its time series characteristics, in order to accurately predict the future load consumption [11]. The LSTM belongs to the family of recurrent neural network (RNN). Unlike the popular feed forward deep neural networks (FFDNN), RNNs are the most suitable neural networks for time series data [11]. They allow the input data to travel forward and backward as they introduce loops and allow internal connections among the hidden units [11]. Figure 5 represents the block diagram of the LSTM network.

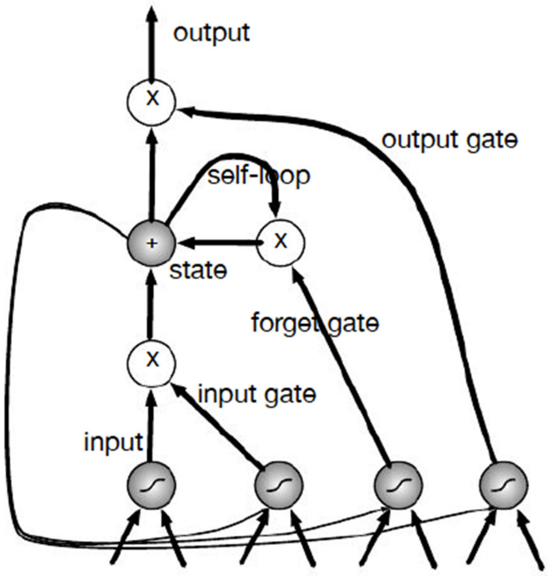

Figure 5: Block diagram of the LSTM network [12].

4.2.2 Power generation prediction

As the power generated from a certain PV system is considered, the historical data in this case consists of the weather conditions along with the corresponding generated power measured at a certain rate. Solar irradiance is an important weather information to be included as input information, along with other information such as cloud coverage, and temperature. Further features, i.e., other specific weather conditions can be investigated in the future to achieve more accurate prediction results. Figure 6 presents the block diagram of the GP service.

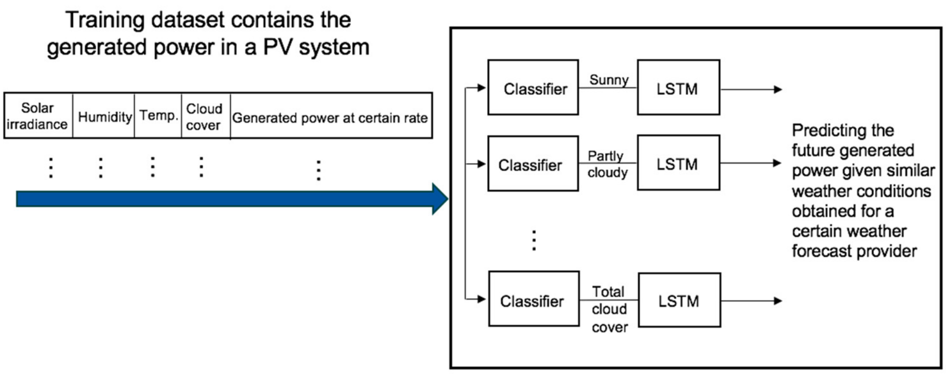

Figure 6: The block diagram of the power generation prediction service.

Referring to Figure 4, it can be shown that the concept behind both LP and GP services are pretty similar. Even in this case, the LSTM is utilized to predict the future generated power. The dataset used for the starting design of the algorithm has been explained in [22].

5. Conclusion

This document provides an overview of the concepts and initial work and implementation of the services implemented in WP3, namely the fault location isolation and service restoration (FLISR) as well as the load and generation prediction services. Furthermore, the report gives an insight on the possible application of different machine learning techniques, with the associated strengths and drawbacks, for the design of the desired services. The procedures followed to implement those two services are explained in this document based on the exchanged information among the different partners.

The report gives an initial overview of the design of the algorithms. However, these algorithms can be subject to modifications within the duration of the project, and according to the feedback received from the pilot owners. A final version of the designed algorithm delivered to the pilots will be presented in D3.3.

6. List of Figures

7. References

[1] R. A. G. A. R. M. H. A. Tokel, „ A New Approach for Machine Learning-Based Fault Detection and Classification in Power Systems," in *2018 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Washington DC, USA, 2018.

[2] N. S. a. P. S. a. D. I. a. P. Shenoy, „Predicting solar generation from weather forecasts using machine learning," in *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011.

[3] R. T. a. A. M. a. M. M. M. a. A. S. A. a. M. R. S. a. F. Behdad, „Medium term load forecasting in distribution systems based on multi linear regression amp; principal component analysis: A novel approach," in *2014 19th Conference on Electrical Power Distribution Networks (EPDC)*, 2014.

[4] T. H. a. P. P. a. S. F. a. H. Z. a. A. T. a. R. J. Hyndman, „Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting,* pp. 896 - 913, 2016.

[5] S. P. a. L. E. a. P. K. a. D. K. a. S. K. a. G. C. a. P. Skafidas, „Electricity demand load forecasting of the Hellenic power system using an ARMA model," *Electric Power Systems Research,* pp. 256 - 264, 2010.

[6] V. Ş. E. a. S. Akar, „ARIMA forecasting of primary energy demand by fuel in Turkey," *Energy Policy,* pp. 1701 - 1708, 2007.

[7] A. B. a. K. G. a. S. K. Mishra, „A Review of Short Term Load Forecasting using Artificial Neural Network Models," *Procedia Computer Science,* pp. 121 - 125, 2015.

[8] D. N. a. Y. W. a. D. D. Wu, „Power load forecasting using support vector machine and ant colony optimization," *Expert Systems with Applications,* pp. 2531 - 2539, 2010.

[9] J. Z. a. C. X. a. Z. Z. a. X. Li, „Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017.

[10] K. P. W. a. M. S. a. N. Soni, „Short-term electrical load forecasting using predictive machine learning models," in *2016 IEEE Annual India Conference (INDICON)*, 2016.

[11] K.-B. S. a. Y.-S. B. a. D. H. H. a. G. Jang, „Short-term load forecasting for the holidays using fuzzy linear regression method," in *IEEE Transactions on Power Systems*, 2005.

[12] I. P. A. D. T. F. W. Group, „IEEE PES Power & Engergy Society," 2018. [Online]. Available: http://sites.ieee.org/pes-testfeeders/. [Zugriff am 20 07 2018].

[13] Y. H. a. L. Yu, „A Variance Reduction Framework for Stable Feature Selection," in *2010 IEEE International Conference on Data Mining*, 2010.

[14] S. K. a. T. K. a. S. Nasreen, „A survey of feature selection and feature extraction techniques in machine learning," in *2014 Science and Information Conference*, 2014.

[15] I. J. G. a. Y. B. a. A. Courville, Deep Learning (Adaptive Computation and Machine Learning), The MIT Press, 2016.

[16] Z.-S. P. a. S.-C. C. a. G.-B. H. a. D.-Q. Zhang, „Hybrid neural network and C4.5 for misuse detection," in *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*, 2003.

[17] L. Breiman, „Random Forests," *Machine Learning,* 2001.

[18] T. H. a. R. T. a. J. Friedman, The elements of statistical learning: Data mining, inference, and prediction, Springer Series in Statistics, 2001.

[19] L. T. M. a. A. C. B. D. a. J. B. A. London, „Towards the improvement of multi-objective evolutionary algorithms for service restoration," in *2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2017.

[20] L. T. M. a. A. C. B. D. a. J. B. A. L. a. M. H. M. Camillo, „Service restoration in large-scale distribution systems considering three levels of priority customers," in *2015 IEEE Eindhoven PowerTech*, 2015.

[21] A. C. S. a. A. C. B. D. a. J. B. A. L. a. N. G. Bretas, „Node-Depth Encoding and Multiobjective Evolutionary Algorithm Applied to Large-Scale Distribution System Reconfiguration," *IEEE Transactions on Power Systems,* pp. 1254-1265, August 2010.

[22] Q. Z. a. H. Li, „MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition," *IEEE Transactions on Evolutionary Computation,* pp. 712-731, 2007.

[23] G. Hebrail, „UCI Repository," [Online]. Available: https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption # . [Zugriff am 2018].

[24] J. Z. a. C. X. a. Z. Z. a. X. Li, „Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, 2017.

[25] T. H. a. P. P. a. S. F. a. H. Z. a. A. T. a. R. J. Hyndman, „Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting,* pp. 896 - 913, 2016.

[26] M. L. McHugh, „The Chi-square test of independence," *Biochemia Medica,* 2013.

8. List of Abbreviations

WP3	Work package 3 in SOGNO project
FLISR	Fault location isolation and service restoration
LP	Load prediction
GP	Generation prediction
SGAM	Smart grid architecture model
PV	Photovoltaic
LSTM	Long short-term memory
RNN	Recurrent neural network
DT	Decision tree classifier
RF	Random forest classifier
DNN	Deep neural network
FFDNN	Feed-forward DNN
MOEAs	Multi-objectives evolutionary algorithms