PhD Proposal 2017

School: Ecole Centrale de Lille	
Laboratory: L2EP	**Web site: http://l2ep.univ-lille1.fr/**
Team: Equipe Réseaux	**Head of the team: Benoit ROBYNS**
Supervisor: Bruno FRANCOIS	**Email: bruno.francois@ec-lille.fr**
Collaboration with other partner during this PhD: 0	**In China:**

Title: Data analysis of sampled consumption profiles and clustering techniques applied to distribution electrical topologies
Scientific field: Electrical Engineering
Key words: Database management, power systems, electrical networks, load demand response, big data

Details for the subject:

Electrical generation based on the use of renewables (wind, photovoltaics...) in distribution systems is continuously increasing. However, those networks have not been initially designed in order to face such penetration. Therefore, in order to help the secure development of renewables while limiting the required investments, ensuring sufficient load flexibility is a solution currently considered by distribution system operators.

In that way, in addition to storage systems, the roll-out of smart meters is imagined and should allow consumers to modify their consumption according to a price signal. The installation of those meters will lead to increased data flows to be analyzed and clustered (in time, geographically...) via statistical procedures (incremental and/or iterative approach, principal component analysis...). The goal of this proposal is thus to provide to the distribution system operator a set of mathematical tools that will help him to reduce the quantity of data (definition of representative networks, customers or, even, periods of time...) to be considered in the context of an optimized flexibility market.

Research subject, work plan:

The aim is to extract statistical information representing the set of data retrieved following a large-scale deployment of smart meters in the electrical distribution network. More particularly, the low voltage network is focused here, and the idea is to limit the amount of information to board the optimization problem underlying this installation of metering devices. The objective function involves the continued maximizing technical and economic advantage of the DSO via a reduction at limited cost (price of deployment awareness campaign ...) while enabling a good system operation (no overvoltage, no imbalances, ...).

To extract useful information of all data from the so-called smart meters, data aggregation techniques will be implemented. Clustering will be both temporal (definition of a limited number of representative quarters of an hour but also a limited number of days of deviation over the year) as topological (extraction BT type networks, representative topologies) even geographical (based on correlations). To do this, the use of iterative algorithms ('k-means', …) and / or will be used. To reduce the size of the problem (size of data vectors to group), the principal component analysis can also be very useful.

The implementation of temporal and geographical grouping of technical data from smart meters consists of the first major part of the project.

The second part aims to use obtained representative data (in time and space) for the development of a decision support tool. This tool will define market models enabling the maximization, at least cost, technical and economic benefits of various load demand strategies. This is an optimization problem with astatistical layer (needed to consider the probability of belonging to a cluster and needed to convolute the temporal and geographical aspects).

Skills

- Knowledge in electrical engineering and in operating of power systems (electrical networks),
- Knowledge in clustering (principal component analysis, k-means, incremental methods, …)
- Knowledge in stochastic optimization
- Knowledge in load demand strategies
- Motivation for practical works and applications,
- Simulation software: Matlab / Simulink,
- English or French language

References:

[1] I.H. Witten, E. Frank, M.A. Hall, "Data Mining Practical Machine Learning Tools and Techniques", *The Morgan Kaufmann Series in Data Management Systems*, 3rd edition, 2011.

[2] G. Guojun, M. Chaoqun, W. Jianhong, "Center-based Clustering Algorithms", *Data Clustering Theory, Algorithms & Applications, ASA-SIAM, Series on Statistics & Applied Probability*, 2007.

[3] P.N. Tan, M. Steinbach, V. Kumar, "Cluster Analysis: Basic Concepts and Algorithms", *Introduction to Data Mining, Pearson*, 2006.

[4] K. Turytsin, P. Sulc, S. Backhaus, M. Chertkov, "Options for Control of Reactive Power by Distributed Photovoltaic Generators", *Proceedings of the IEEE*, 99(6), pp.1063-1073, 2011.

[5] F. Pilo, G. Pisano, G.G Soma, "Optimal coordination of energy resources with a two-stage online active management", *IEEE Transactions on Industrial Electronics*, 58(10), pp.4526 – 4537, 2011.

[6] I.T. Papaioannou, A. Purvins, E. Tzimas, "Demand Shifting Analysis at High Penetration of Distributed Generation in Low Voltage Grids", *Electrical Power and Energy Systems*, 44, pp.540-546, 2013.

[7] CIGRE WG C6.19, "Planning and optimization methods for active distribution systems*", SC C6 Distribution Systems and Dispersed Generation*, Technical Report, 2015.

[8] S. Mahieu, D. Ernst, B. Cornélusse, "Macroscopic analysis of interaction models for the provision of flexibility in distribution systems", *International Conference on Electricity Distribution*, CIRED 2015, Lyon, June 2015.

[9] I.T. Jolliffe, "Principal Component Analysis", *Springer Series in Statistics*, 2nd Edition, Springer-Verlag Editions, chap. 2&3, pp. 10-59, 2002.