

Prédiction de la consommation par inférence grammaticale

Guillaume Guerard, Sonia Djebali

Pôle universitaire Léonard de Vinci, Research Center, La Défense, France
{prenom.nom}@devinci.fr

Mots-clés : *prédiction, inférence grammaticale, aide à la décision.*

1 Introduction

Le Smart Grid désigne un réseau énergétique intelligent, qui intègre le comportement des utilisateurs (producteurs, consommateurs, agrégateurs, etc.) afin de lisser la courbe de consommation et de limiter les rejets polluants et les fluctuations du coût de l'énergie [1, 5]. En effet, une courbe ne possédant pas de pics ou de creux important signifie que l'usage des énergies fossiles devient obsolètes.

Cependant, le Smart Grid est un système complexe. Comprendre et connaître l'émergence de comportements à l'échelle macro ou même méso est souvent impossible, le système doit être analysé et optimisé par étude systémique. Cette étude par l'échelle micro prend en compte les entités, les liens et les interactions entre les entités, ainsi que des régulateurs et des évolutions dans le temps [8]. Ces systèmes sont souvent décrits comme des systèmes multi-agents.

Le méthode présentée s'inscrit dans un objectif de planification de la consommation. La connaissance des scénarios possibles permet lors d'une décision d'appréhender les choix futurs. La méthode propose d'apprendre par inférence de chaque appareil consommateur afin d'établir un automate stochastique représentant le comportement énergétique de cet appareil. Ainsi, il est possible par l'algorithme de Viterbi [4] ou par marche aléatoire d'établir des scénarios de consommation avec leurs probabilités d'apparitions.

2 Méthodologie

La méthode comporte plusieurs étapes comme suivants :

1. Collecte des données : l'appareil produit une série temporelle décrivant sa consommation dans le temps (en terme d'évolution de consommation entre chaque point).
2. Création du dataset : à partir de la série temporelle, des séquences de consommation sont extraites. Chaque séquence décrit un cycle de consommation de l'appareil, et se termine par un arrêt ou mise en veille prolongée de ce dernier. Les données sont ensuite normées et le bruit réduit afin de lisser les données.
3. Analyse des séquences en deux étapes
 - (a) Un algorithme de clustering, ici le k-means, regroupe les évolutions de consommation possédant une valeur proche. L'étude de la silhouette détermine le nombre de clusters [6].
 - (b) Un algorithme de recherche de motifs sur des données contiguës repère les schémas de consommation les plus fréquents afin de rendre la sous-séquence indissociable [7]. Les motifs sont ensuite étendus dans les séquences concernées pour amoindrir le bruit restant.
4. Les données prétraitées sont ensuite stockées dans un trie (ou arbre de préfixes). Ce dernier représente les séquences sous la forme d'un automate à fréquence.

5. L'algorithme ALERGIA infère le trie afin d'établir un graphe (non plus un arbre), ce graphe est un automate stochastique [2, 3].

Une fois l'automate stochastique obtenu, ce dernier est exploité afin d'établir les schémas de consommation. A partir d'un préfixe connu, qui sera lu par l'automate, il est possible par marche aléatoire d'établir une liste de scénarios avec la probabilité d'apparition. Il est aussi possible, à partir d'un scénario donné, d'en connaître sa plus grande probabilité d'apparition par l'algorithme de Viterbi [4].

Ainsi, cette méthode de prédiction de la consommation est utile pour planifier les besoins énergétiques futures. Suite à la création de la liste de scénarios, chaque maison intelligente est capable de communiquer ses besoins à l'agrégateur et grâce à la théorie des jeux, de choisir le meilleur scénario que ce soit en terme de consommation, de lissage de courbe tout en tenant compte du confort de l'utilisateur.

Références

- [1] Michael Batty. *Cities and complexity : understanding cities with cellular automata, agent-based models, and fractals*. The MIT press, 2007.
- [2] Colin De la Higuera. *Grammatical inference : learning automata and grammars*. Cambridge University Press, 2010.
- [3] Rémi Eyraud, Colin De La Higuera, Makoto Kanazawa, and Ryo Yoshinaka. Introduction to the grammatical inference special issue of *fundamenta informaticae*, 2016.
- [4] G David Forney. The viterbi algorithm. *Proceedings of the IEEE*, 61(3) :268–278, 1973.
- [5] G. Guérard, S.B. Amor, and A. Bui. Survey on smart grid modelling. *International Journal of Systems, Control and Communications*, 4(4) :262–279, 2012.
- [6] Greg Hamerly and Charles Elkan. Learning the k in k-means. In *Advances in neural information processing systems*, pages 281–288, 2004.
- [7] Yu Hirate and Hayato Yamana. Generalized sequential pattern mining with item intervals. *JCP*, 1(3) :51–60, 2006.
- [8] John H Holland. Complex adaptive systems. *Daedalus*, 121(1) :17–30, 1992.