

Simulation prédictive pour la gestion des consommations électriques dans un quartier

Eric Ferreri*¹, Jean-Marc Salotti¹, Pierre-Alexandre Favier¹

¹ 109, avenue Roul 33400 Talence
33400 Bordeaux
France,
*eric.ferreri@ensc.fr

RESUME. La simulation de consommations électriques permet d'obtenir rapidement une estimation des besoins présents et futurs en énergie électrique. Les simulations purement statistiques ne parviennent pas à résoudre divers problèmes. Les paramètres de la plupart des modèles proposés dans la littérature n'ont pas de lien concret avec des comportements d'utilisateurs. De plus, lorsque les systèmes et les usages évoluent, la base de données devient inadaptée. Une approche intéressante consiste à exploiter un modèle adaptatif centré utilisateur et usages permettant une extrapolation fine des prédictions de consommation en dehors des profils et contextes répertoriés. Dans cet article, nous proposons une approche multi-agents centrée sur un nouveau modèle d'agent bio-inspiré. Celui-ci s'appuie sur des à priori de comportements d'utilisateurs obtenus par analyse multivariée de bases de données de consommations. Notre approche a été évaluée quantitativement, les résultats obtenus sont précis et les paramètres de notre modèle sont directement interprétables en termes comportementaux.

MOTS-CLÉS : Système multiagent, comportement humain, extrapolation

ABSTRACT. The simulation of electrical power consumption allows quick estimations of future needs in this domain to be obtained. However, purely statistic methods suffer from diverse limitations. Very often, a model's parameters are not directly related to inhabitant's behavior. Furthermore, when habits or equipment evolve, the database may become unsuited or even obsolete. An interesting approach to overcome these limitations is to use an adaptive model, centered on the behavior of inhabitants thus allowing robust extrapolations, even among user's profiles that are not part of the database. In this paper, we propose a multi-agent system based on a new model of bio-inspired agent. The agent relies on knowledge extracted from consumption databases with multivariate analysis. The parameters of our model are mainly behavior-related and the results obtained are precise.

KEYWORDS : Multiagent system, human behavior, extrapolation

INTRODUCTION

Le domaine de modélisation des consommations électriques, connaît actuellement un essor. En effet, la demande sans cesse croissante d'énergie électrique, n'est pas toujours suivie d'une augmentation de la production de celle-ci. Afin de répondre adéquatement à cette demande en augmentation, il est important de pouvoir connaître à l'avance les besoins des habitants. Pour ce faire, plusieurs approches ont été explorées. La plupart des méthodes statistiques décrivent les comportements des usagers en termes de probabilités d'actions. Widen et al. (Widen et al. 2010) proposent l'utilisation de chaînes de Markov non-homogènes. Dans leur méthode, des probabilités de transitions entre activités (dormir, cuisiner, allumer la télévision...) sont calculées pour chaque habitant d'un logement. Les résultats obtenus correspondent bien à la réalité des consommations.

Cependant, leur méthode requiert des bases de données de comportements et de consommations très précises : une centaine d'habitants devait noter, à la minute près, chaque activité réalisée. De plus, ce modèle est statique, il est impossible de connaître l'effet qu'aura la variation d'un paramètre comportemental sur la matrice de probabilités.

Par exemple, en cas de modification importante du contexte, il n'y a aucun moyen de prédire son impact car le modèle n'intègre aucune information comportementale. En réalité cette méthode, comme la plupart des modélisations statistiques, résume le comportement humain à une suite d'activités où la probabilité de réaliser une certaine activité à l'instant t , dépend au mieux de l'activité réalisée à l'instant $t-1$. D'autres méthodes parviennent à obtenir de bons résultats en requérant des bases de données moins exhaustives. Par exemple l'application 'Courbotree' proposée par Stephan et al. (Stephan 1989) se base sur une technique d'arbre de régression multivariée. Une méthode d'apprentissage automatique permet de regrouper les consommateurs en classes, grâce à un ensemble de variables explicatives (par exemple la proportion d'énergie électrique consommée pendant les heures creuses, le type de logement...). Il devient ainsi possible d'associer une courbe espérée de consommation à un individu inconnu à partir d'un minimum d'information. Mais une fois de plus, il est difficile d'établir un lien causal entre le comportement des usagers et les valeurs des paramètres du modèle.

Offrant des solutions aux problèmes cités précédemment, les modélisations multi-agents suscitent un grand intérêt. En s'appuyant sur une structure qui cherche à prédire les processus comportementaux de chaque habitant en le représentant par un agent, les paramètres du modèle deviennent interprétables en termes d'intentions et d'habitudes. D'autre part, il devient possible d'étudier l'influence de certaines habitudes sur la consommation électrique d'un logement. En effet, les bases de données REMODECE et IRISE dont nous disposons, permettent d'associer aux occupants des actions portant sur l'utilisation d'appareils électroménagers. Dans cet article, nous proposons un nouveau schéma d'agent basé sur un modèle cognitif ayant fait ses preuves. Notre démarche se déroule en deux étapes. Une première étape vise à utiliser des données comportementales afin de paramétrer le processus de formation des décisions de l'agent, et une seconde étape cherche à utiliser ce processus de formation des décisions pour générer des comportements et des habitudes plus variés que ceux de la base de paramétrage. Nos contributions portent donc sur le nouveau schéma de l'agent cognitif, et ses capacités à prédire dans une certaine mesure, un comportement motivé. Dans la suite de cet article, nous présentons dans la section 2, le principe de la modélisation multiagent appliqué au domaine des activités domestiques ainsi que notre méthode de modélisation basée sur un modèle bio-inspiré. Nous présentons dans la section 3, les résultats de cette nouvelle approche ainsi qu'une conclusion concernant l'efficacité et les perspectives de celle-ci.

2. SIMULATION DE CONSOMMATION ELECTRIQUES PAR METHODES MULTIAGENT

2.1 PRINCIPE ET CHOIX DE L'AGENT

Les Méthodes de simulation multi-agents offrent un cadre privilégié lorsque l'on souhaite connaître l'influence du comportement et des habitudes sur la consommation électrique. En plus d'utiliser une

description statistique des usages, on utilise un schéma cognitif pour établir un lien entre les consommations électriques et des variables comportementales.

Dans leur article Le et al. (Le, et al 2010) utilisent une architecture d'agent belief-desire-intention (BDI) qui leur permet de reconstruire des données à partir de règles comportementales et environnementales issues de bases de données adéquates. Cependant, le schéma d'agent BDI, bien que très utile, n'est pas totalement adapté pour simuler des comportements d'habitants. En effet, dans le cadre qui nous intéresse, le comportement ne peut être réduit à une liste de moyens à utiliser de manière optimale pour satisfaire des besoins. Un comportement est l'aboutissement d'un processus de réflexion de niveau variable qui n'a pas nécessairement d'utilité associée.

C'est ce qui a motivé notre utilisation d'un schéma de raisonnement basé sur le modèle développé par Fiske et Neberg (Fiske et Neberg, 1990) en psychologie sociale. Leur modèle définit d'une part des réponses schématiques, basées sur les habitudes et d'autre part des réponses spécifiques basées sur une réflexion plus approfondie et propre à la situation. Les habitudes permettent de savoir rapidement et sans effort cognitif quel type d'action peut être adapté à une situation courante. Dans le cas où les habitudes ne suffisent pas à décider, des informations plus spécifiques à la situation et un raisonnement approfondi permettent de choisir une action appropriée.

Dans le cadre qui nous intéresse, cette vision hiérarchique du raisonnement est très adaptée. En effet, elle permet d'expliquer que l'on trouve une certaine régularité horaire dans les activités quotidiennes, mais elle permet aussi d'expliquer pourquoi parfois ces habitudes ne sont pas respectées : d'autres envies peuvent prendre le dessus.

Toutefois, il s'agit d'un modèle issu de la psychologie, certaines adéquations au formalisme informatique seront nécessaires afin de le rendre opérationnel au sein d'un système multi-agents.

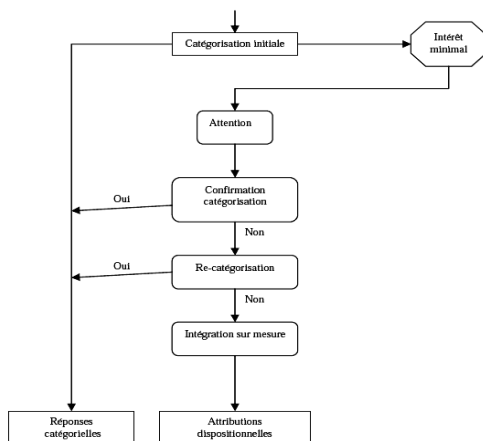


Figure 1 : modèle de Fiske et Neberg

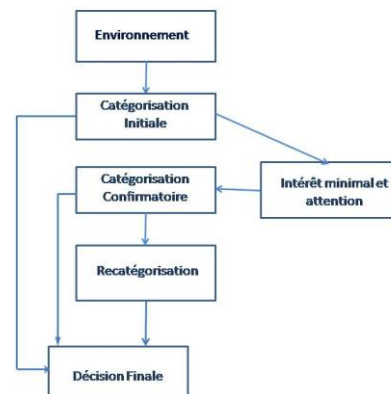


Figure 2 : Modèle proposé

2.2 Implémentation

Nous exposons ici notre approche permettant d'utiliser dans un système informatique un schéma d'agent proche de celui présenté dans la figure 1. Le modèle a dû être modifié, l'étape d'intégration sur mesure proposée dans le modèle de Fiske et Neberg, qui permet de faire face à des situations complètement nouvelles, n'est pas de notre ressort. Le modèle proposé dans cet article, présenté dans

la figure 2, permet de simuler le comportement humain grâce à des raisonnements de base. Dans les prochaines sections, nous expliquerons comment nous avons traduit une partie des processus de pensée et d'action en termes mathématiques.

2.2.1 Catégorisation initiale

L'objectif de cette étape est de proposer un type d'action à un instant t de la journée en cours. Si nous nous en tenons au modèle psychologique, la première étape de traitement, doit être faite grâce à un descripteur global, c'est-à-dire portant sur des classes de comportement vastes. La catégorisation devra différencier globalement deux classes d'actions bien différentes : celles qui concernent les besoins physiques (cuisiner, laver...) et celles qui concernent les besoins récréatifs (allumer le four, la TV...). Dans notre base de données, nous disposons d'informations concernant les usages de la télévision, la chaîne hifi, le four, le lave linge, le lave vaisselle et la friteuse. Ces informations sont fournies par périodes de 24 heures. La méthode présentée permet de formuler des profils de consommation pour une nouvelle période de 24 heures, fixant de cette manière l'horizon de prédiction.

Afin de regrouper adéquatement les profils de consommation concernant des appareils différents, nous avons utilisé une méthode de classification non supervisée : la méthode des mélanges de gaussiennes. L'idée consiste à modéliser un historique de 24 heures pour chaque activité par un mélange de gaussiennes. Ainsi à chaque appareil électrique, est associée une somme pondérée de N_g gaussiennes ($N_g = 24$), qui décrit son usage typique au long de la journée. La probabilité d'observer une consommation de I wattheures à l'instant t , pour l'appareil p est estimée par :

$$p(I^t(p) | \mathbf{I}_p) = \sum_{i=1}^{N_g} w_i^t \cdot \eta(I^t(p); \bar{u}_i^t, \Sigma_i^t) \quad (1)$$

où $\mathbf{I}_p = \{ I^1(p), I^2(p), \dots, I^{24}(p) \}$ est l'historique des consommations électriques de l'appareil p ; w_i^t est le poids de la gaussienne i à l'heure t ; \bar{u}_i^t et Σ_i^t sont la moyenne et la matrice de covariance de la gaussienne. La valeur de la puissance consommée, de l'appareil p à l'heure t , notée $I^t(p)$ appartient à la gaussienne i s'il vérifie :

$$\|I^t(p) - \bar{u}_i^t\| < K \sigma_i^t$$

où K est un paramètre de la méthode. Dans le cas qui nous intéresse, il faut classifier les usages des 6 appareils précédemment cités. La classification par mélanges gaussiennes, permet d'obtenir deux prototypes, chacun d'eux représentant fidèlement sa classe d'appartenance. Il est possible de classifier les activités de la base de données en fonction de leur distance à ces prototypes par la méthode du plus proche voisin. Voici les résultats que nous avons obtenus:

Appareil électrique	Groupe attribué
Télévision	2
Lave vaisselle	1
Lave Linge	1
Chaîne Hi-Fi	2
Friteuse	1
Four	1

Tableau 1 : Résultats de la classification non supervisée

La méthode de classification non supervisée réussit à regrouper correctement les activités qui appartiennent à une même classe. Par la suite, si l'on rajoute de nouvelles activités, leur groupe d'appartenance pourra être connu en calculant sa distance au plus proche prototype.

Pour décider quel type d'action doit être entreprise, on calcule la probabilité de réaliser une activité quelconque à l'instant t notée $P_t(A)$. Dans ce premier niveau d'analyse, on considère que :

$$P_t(A) = P(A_1|H) \cup P(A_2|H) \quad (3)$$

$P(A_i|H)$ est connu grâce à la base de données. A présent, il faut définir une condition permettant de savoir s'il est nécessaire de passer au niveau délibératif suivant. Soit r un nombre aléatoire généré tel que $r \sim \mathcal{U}(0,1)$, si la condition $r < P(A)$ est remplie, l'agent passe au niveau délibératif suivant. L'équation 3 permet de modéliser la formation d'une intention en lien avec les habitudes, et la règle de décision qui utilise r permet de savoir si le processus de réflexion ira plus loin ou s'il s'agira uniquement d'une intention restée sans suite. L'utilisation des prototypes de classification, ainsi que la règle de décision constituent le bloc 'Catégorisation Initiale' de la figure 1.

2.2.2. Catégorisation confirmatoire

L'étape de traitement qui suit la catégorisation initiale, se caractérise par une attention plus soutenue et une évaluation plus précise de la situation que celle effectuée précédemment. Ce qui nous intéresse à ce stade, n'est pas de reproduire des patterns d'utilisation d'appareils ou d'habitudes observés, mais d'en générer de nouveaux. En effet, même en absence d'imprévu, notre comportement s'éloigne parfois de nos habitudes. Par exemple l'instant où nous prenons un repas n'est pas seulement conditionné par l'habitude de dîner à heure fixe, mais aussi par l'heure du repas précédent. Une loi de distribution de probabilité peut adéquatement rendre compte de ce phénomène. En raison de régularités horaires liées aux besoins physiques et aux habitudes, une loi de Poisson décrit adéquatement la probabilité d'utiliser k fois dans un intervalle de temps T , un même appareil électrique. Lorsque le nombre moyen λ d'utilisations dans un intervalle de temps T de cet appareil A est connu, la fonction de distribution de probabilité correspond à:

$$P(X_A = k_A) = \frac{\lambda_A^{k_A}}{k_A!} e^{-\lambda_A} \quad (5)$$

Le paramètre λ_A de cette distribution peut être estimé à partir du profil moyen de consommation de l'appareil A . En plus de $P(A_i|H)$ qui est obtenue directement à partir de la lecture du profil journalier de consommation, nous disposons de (5). Une technique appelée Méthode de la Fonction Inverse (MFI) permet de générer des suites aléatoires et non répétitives qui suivent la loi de distribution de l'équation (5). Cette méthode a déjà été appliquée dans un contexte semblable et expliquée de manière très détaillée (Page et al. 2008). En utilisant la MFI, on génère une décision concernant l'appareil A_i d'après la fréquence observée d'utilisation de celui-ci.

En utilisant cette nouvelle information, nous pouvons donner une définition plus précise de $P_t(A_i)$:

$$P_t(A_i) = P(A_i | (H, n_{A_i})) \quad (6)$$

L'utilisation de l'équation (6) pour chaque appareil électrique donne un ensemble de probabilités d'utilisation de chaque appareil A_i .

Pour chacun des appareils, on génère un nombre aléatoire $s \sim \mathcal{U}(0,1)$, si $s < P(A_i | (H, n_{A_i}))$, alors ce premier stade de décision propose de réaliser l'action A_i .

Dans ce module de décision, nous avons pris en compte la période entre réalisations d'une même activité. Dans ce qui suit, nous allons présenter un autre ensemble de règles qui prendra en compte encore plus finement l'information contenue dans les profils journaliers d'utilisation de notre base de données.

2.2.3. Recatégorisation

Jusqu'ici la plupart des travaux réalisés dans le domaine des consommations électriques domestiques ont très peu traité du lien existant entre utilisations de divers appareils électriques (Wilke 2013). Pourtant, l'information contenue dans la corrélation entre activités possède un double intérêt que nous allons expliquer par le biais d'un exemple. Le repassage du linge est normalement effectué une fois que celui-ci a été lavé et séché, il est donc naturel de prévoir qu'un certain temps après la mise en marche de la machine à laver, suivra une utilisation du fer à repasser. Le délai habituellement écoulé entre ces deux activités peut être calculé par corrélation croisée normalisée des profils de consommation de la machine à laver et du fer à repasser. Ce délai révèle deux liens : un premier lien statistique entre deux variables aléatoires de répartition connue, et un lien plus profond qui correspond à une planification d'activités. Les processus de décision, tels que décrits en psychologie sociale, confirment que lors d'évaluations attentionnelles, des interactions de haut niveau sont prises en compte dans la formulation de décisions (Turner et al. 1986).

Lors de l'analyse par corrélation croisée de notre base de données, nous avons constaté, un délai significatif séparant l'utilisation du lave linge et celui du lave vaisselle, dans ce cas plus que d'une planification il s'agit de la prise en compte de restrictions matérielles. Ainsi, il n'est souvent pas possible d'allumer le lave-vaisselle et le lave-linge en même temps, car le débit d'eau ou de courant requis peut dépasser les capacités de l'installation domestique. Afin de prendre en compte ces liens corrélatifs entre activités, nous allons à nouveau utiliser la méthode de la fonction inverse sur une distribution de poisson, mais cette fois le calcul du paramètre λ est quelque peu différent :

$$\lambda_{i,j} = \frac{24}{\text{argmax}(A_i * A_j)}, i \neq j \quad (6)$$

Le symbole * est utilisé pour représenter l'opération d'autocorrélation croisée normalisée, ainsi $\text{argmax}(A_i * A_j)$ correspond à l'intervalle de temps modal entre l'utilisation des appareils A_i et A_j .

En procédant de cette manière pour chaque appareil électrique de notre base de données, nous disposons d'une distribution décrivant sa périodicité, et de 5 distributions de probabilité décrivant

l'intervalle de temps entre utilisations d'appareils pour chaque paire i, j d'appareils différents. Cette nouvelle information permet de proposer un nouveau moyen de calcul pour $P_t(A_i)$:

$$P_t(A_i) = P\left(A_i \mid (H, n_{A_i}, \lambda_{i,j})\right), i \neq j \quad (7)$$

La règle de décision utilise les probabilités $P_t(A_i)$ de la même manière que lors de la catégorisation confirmatoire : Pour chaque $P_t(A_i)$, on génère un nombre aléatoire $s \sim \mathcal{U}(0,1)$, si $s < P(A_i | (H, n_{A_i}))$, alors l'appareil A_i sera utilisé.

3. RESULTATS ET CONCLUSIONS

Précédemment, nous avons présenté un schéma d'agent qui permet de reproduire une part de variance des comportements plus importante que si l'on avait uniquement recours aux méthodes d'analyse statistique. Dans ce qui suit, nous allons montrer comment nous avons généré des profils de consommation différents de ceux présents initialement dans la base de données.

Afin de valider la méthode proposée, nous avons utilisé une procédure de type leave k-out. Parmi les 200 profils d'utilisation d'appareils électriques répartis par habitation, nous utilisons 20 profils pour générer automatiquement un nouveau groupe correspondant à la consommation de 5 habitations. Par la suite, nous évaluons la pertinence et l'originalité des profils générés en les comparant à l'ensemble des profils de la base de données. Ainsi, il est possible de savoir si les nouveaux profils de consommation horaire générés correspondent à des comportements humains observés, mais aussi de vérifier que les profils générés sont bien différents des profils de départ.

Un exemplaire de profil initial est montré dans la figure 3, la figure 4 compare le profil d'utilisation généré à un profil présent dans la base de données mais qui n'a pas été utilisé par notre programme informatique. Le tableau récapitulatif 2, présente les résultats obtenus. Pour comparer la ressemblance entre deux profils d'utilisation journalière, nous avons considéré chaque profil comme étant un point dans un espace à 24 dimensions puis avons utilisé le carré de la distance entre ces deux points pour métrique d'évaluation.

Le tableau montre d'une part l'espérance du carré des distances (noté D) entre profils générés (notés NP) et profils de consommation réels (notés PBD), ainsi que le carré des distances entre les profils utilisés par le programme (notés AP) et les nouveaux profils générés.

	Télévision	Chaîne Hi-Fi	Lave-linge	Lave-vaisselle	Four	Friteuse
$D(NP, PBD)$	2.9	2.4	2.3	4.1	6.3	5.3
$D(NP, AP)$	4.1	4.8	5.93	9.9	10	8.9

Tableau 2 : récapitulatif des résultats

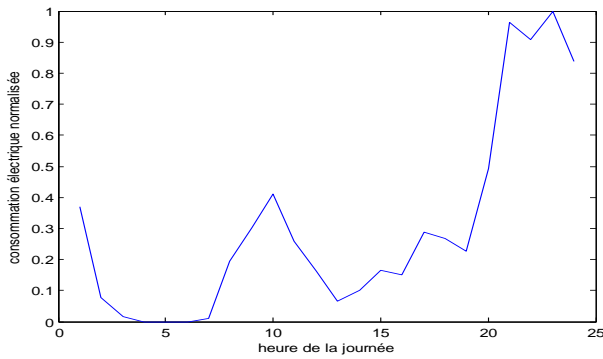


Figure 3 : Profil d'utilisation de la télévision.

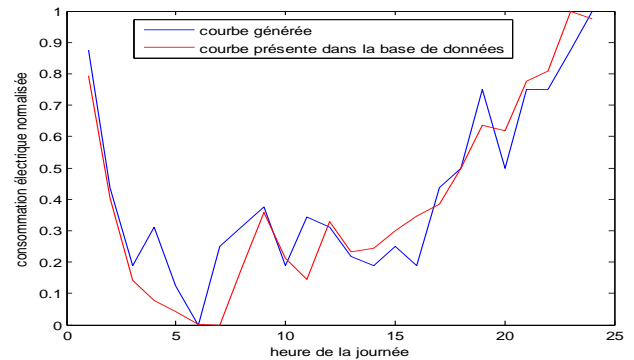


Figure 4 : profil généré comparé à un profil réel.

Le tableau 2 ainsi que les figures 3 et 4, montrent que pour l'échelle de prédiction ciblée, celle d'un petit quartier de 5 domiciles, les courbes générées sont bien différentes des courbes originelles, mais aussi que les profils générés correspondent dans une bonne mesure à des profils réels.

Ce travail a proposé une architecture d'analyse et de génération de données de consommations électriques permettant de modéliser et d'anticiper correctement les consommations dans les bâtiments ou quartiers résidentiels. Cette architecture est centrée sur un modèle comportemental bio-inspiré et éprouvé qui permet de donner un sens comportemental aux données statistiques. De plus, les résultats des simulations permettent de valider les algorithmes concernant la création de nouveaux profils d'utilisateurs.

BIBLIOGRAPHIE

- Fiske, Susan T. et Neuberg, Steven L. 1990. « A continuum of impression formation, from category—based to individuating processes: Influences of information and motivation on attention and interpretation. » *Advances in experimental social psychology*, vol. 23, p. 1-74.
- Le, X. H. B., Kashif, A., Ploix, S., Dugdale, J., Di Mascolo, M., & Abras, S. 2010. « Simulating inhabitant behaviour to manage energy at home. » In *International Building Performance Simulation Association Conference*, Moret-sur-Loing, France.
- Page, J., Robinson, D., Morel, N., Scartezzini, J.-L., 2008 A generalised stochastic model for the prediction of occupant presence, *Energy and Buildings*, 40(2) p83-98
- Stéphan, V., Cogordan, F. « Courbotree: Application Des Arbres De Regression Multivaries Pour La Classification De Courbes. » <http://www.agro-montpellier.fr/sfds/CD/textes/stephan1.pdf>
- Turner, John; Oakes, Penny (1986). « The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. » *British Journal of Social Psychology* 25 (3): 237–252.
- Widén, J., Wäckelgård, E. 2010. « A high-resolution stochastic model of domestic activity patterns and electricity demand. » *Applied Energy*, 87(6), 1880-1892.
- Wilke, U., Scartezzini, J. L., et Haldi, F. 2013. « Probabilistic Bottom-up Modelling of Occupancy and Activities to Predict Electricity Demand in Residential Buildings » (Doctoral dissertation, EPFL)