
Simulation de l'activité humaine

Une étude sur le réalisme multi-niveau

**Quentin Reynaud¹, Nicolas Sabouret², Yvon Haradji³,
François Sempé⁴**

1. QRCI - 63000 Clermont-Ferrand, France
quentin.reynaud.pro@gmail.com

2. LIMSI, CNRS, Univ. Paris-Sud, Université Paris-Saclay
Bât 508, Campus Universitaire, 91405 Orsay, France
nicolas.sabouret@limsi.fr

3. EDF R&D, EDF Lab Paris-Saclay
7 Boulevard Gaspard Monge, 91120 Palaiseau, France
yvon.haradji@edf.fr

4. François Sempé AE, Paris, France
sempe.francois@gmail.com

RÉSUMÉ. Cet article présente une méthode de simulation de l'activité humaine qui s'appuie à la fois sur un système multi-agent (SMA) et sur des études statistiques. Il est possible de considérer le réalisme d'une simulation (i.e. sa proximité avec les comportements observés sur le terrain), à différents niveaux. La prise en compte de ces niveaux dans la littérature se fait généralement de manière séparée. Nous proposons un modèle qui considère deux approches complémentaires : la simulation multi-agent, visant traditionnellement un réalisme local, et les études statistiques, visant un réalisme agrégé. Nous présentons notre architecture, sa mise en œuvre sur des données issues d'une enquête « emploi-du-temps ». Nous étudions le réalisme des simulations produites dans un contexte local et agrégé.

ABSTRACT. In this paper, we present a methodology of human activity simulation, based both on a Multi-Agent System (MAS) and statistical data. We show that the notion of simulation realism (i.e. the proximity with behaviors observed in the field) can be studied at various levels. Traditionally, these levels are never considered simultaneously in a single architecture, hence we propose a methodology allowing the combination of multi-agent simulation (aiming at a “microscopic” realism) and statistical data (aiming at a “macroscopic” realism). We present our architecture within the use of the last “time use survey”. We study the simulation's realism within both local and aggregated context.

MOTS-CLÉS : simulation multi-agent de l'activité humaine, réalisme multi-niveau.

KEYWORDS: multi-agent based simulation of human activity, multi-level realism.

DOI:10.3166/RIA.32.197-221 © 2018 Lavoisier

1. Introduction

La simulation de l'activité humaine définit un champ de recherche pluridisciplinaire. On y trouve des travaux en ergonomie (Haradji *et al.*, 2012), en informatique (Barot *et al.* 2013 ; Drogoul *et al.* 2013), en physique (Mohammadi et Saiac, 2003), en psychologie (Marsella et Gratch 2009), etc. Le contexte applicatif de nos travaux est celui de la simulation de la consommation électrique en milieu résidentiel, dans une optique de réduction de la consommation énergétique. Dans ce cadre, la simulation de l'activité humaine permet de produire des « courbes de charge » décrivant la consommation électrique des ménages. La modélisation de l'activité humaine ne vise ici aucune réalité cognitive ou émotionnelle, mais seulement une imitation de surface de l'activité humaine réelle (Haradji et Poizat, 2017).

L'activité humaine est, par nature, un phénomène complexe, difficile à simuler. La question de la validité d'une simulation de l'activité humaine est donc cruciale. Dans cet article nous proposons d'étudier cette question sous l'angle du réalisme, c'est-à-dire de la *proximité* avec l'activité humaine réelle. Nous montrons que l'activité humaine peut être étudiée à différents niveaux et que la notion de réalisme diffère en fonction des niveaux considérés. Nous présentons une approche permettant de combiner différents niveaux de réalisme, ainsi que de produire des simulations avec différentes échelles de temps (une semaine, un an, etc.) et de populations (un ménage, une ville, un pays, etc.).

2. Niveaux de simulation et de réalisme : 2 axes dans l'étude de la simulation de l'activité humaine

Toute simulation informatique pose la question de sa validité (Sargent 2013). En effet, simuler consiste à construire un modèle de l'objet d'étude, donc à effectuer des simplifications, puis à produire des traces de simulation qui sont les observations sur lesquelles se basent l'étude. Lorsqu'on s'intéresse à la simulation de systèmes complexes avec un système multi-agent, la conformité de ces traces avec des observations terrains est une question difficile de par la nature du phénomène étudié. En effet, le comportement d'un système complexe simulé émerge des interactions locales entre les différents composants de la simulation : il ne peut pas être décrit simplement en termes d'équations s'appliquant à l'ensemble du système. Il est donc difficile de définir quelles observations caractérisent le mieux son fonctionnement.

Pour étudier la simulation de l'activité humaine, nous proposons de considérer deux axes : le niveau de simulation et le niveau de réalisme.

2.1. Niveau de simulation

Dans le domaine de la simulation de l'activité humaine, les observations sur le comportement du système peuvent être faites à différents niveaux : au niveau d'un individu (on parle alors de niveau microscopique ou individuel), d'un groupe

d'individus ou de l'ensemble de la population simulée (on parle alors de niveau macroscopique ou agrégé). À chaque niveau, des observations peuvent être définies. Par exemple, on peut s'intéresser, au niveau individuel, à l'agenda de la journée. Au niveau du groupe, on peut s'intéresser à l'évolution de la dynamique de composition des groupes et au niveau de la population, au temps moyen passé chaque semaine devant la télévision. Chacune de ces observations est liée au niveau de l'étude. Dans la suite, nous parlons d'*indicateurs* pour désigner les valeurs observées.

Le réalisme de la simulation est généralement vu comme une mesure de la distance entre les indicateurs issus de la simulation et ceux issus de données terrain. Il s'agit donc de comparer les traces de la simulation (c'est-à-dire les valeurs des indicateurs observés sur des données de simulation) avec le relevé de ces mêmes indicateurs dans la réalité. Le choix du niveau de simulation est souvent contraint par les données terrain qu'il est possible de collecter, donc des indicateurs qu'il sera possible de recueillir et de comparer. Or, plus un indicateur est proche du niveau microscopique (on parle alors d'indicateurs « bas niveau »), plus il est précis mais plus les données réelles sont difficiles à obtenir. Par opposition, un indicateur de « haut-niveau » donne des informations moins précises, mais les données réelles sont plus faciles à obtenir. Par exemple, mesurer précisément l'usage de tous les appareils de la maison est difficile à obtenir dans la réalité (alors que cela s'obtient facilement en simulation). Au contraire, la consommation totale du foyer est relativement facile à obtenir, mais décrit moins finement ce qu'il se passe dans le foyer.

2.2. Niveau de réalisme

Dans le domaine de l'activité humaine, la notion de réalisme de la simulation peut aussi être considérée à plusieurs niveaux :

- au niveau **individuel**, c'est-à-dire au niveau des activités d'un individu. Par exemple, si nous simulons les activités de Mme Dupont et si nous prenons comme indicateur bas niveau l'activité effectuée à un instant t par cette personne, nous pouvons comparer la trace d'activité avec l'activité réelle (à condition de pouvoir l'observer) ;

- au niveau **agrégé**, c'est-à-dire au niveau des activités d'une population (ou d'un ensemble) d'individus. Par exemple, si nous considérons l'ensemble des habitants de l'immeuble de Mme Dupont et si nous sommes capables de produire une population synthétique ayant des caractéristiques similaires à ces habitants, nous pouvons étudier la similarité entre les activités agrégées des habitants de l'immeuble et de ceux de la simulation (par exemple, quelle proportion des habitants est en train de faire la cuisine entre 19h et 20h).

Mais s'agissant de l'activité des individus dans un groupe, il existe un troisième niveau à considérer :

- Au niveau **local**, le réalisme de l'activité est celui d'un individu au sein d'une simulation de population, c'est-à-dire un individu ne correspondant pas à un

individu réel. Par exemple, est-ce qu'il est réaliste que tel agent de la simulation ne fasse jamais la cuisine en semaine ?

La nécessité de définir ce niveau local de réalisme provient de l'impossibilité de simuler à l'identique la population considérée. Tout d'abord, il est illusoire de recréer une population synthétique qui aurait exactement les mêmes caractéristiques que la population considérée. Toute modélisation nécessite des simplifications qui font perdre de l'information. En quand bien même il serait possible de simuler chaque individu de manière parfaitement réaliste, nous ne disposerions pas nécessairement d'indicateurs de bas niveaux pour pouvoir valider ce réalisme individuel. Ce réalisme est plus délicat à définir, puisqu'il n'est pas possible de définir directement une mesure de distance avec une cible réelle. Il est par contre relativement facile de repérer des comportements aberrants, comme par exemple partir travailler sans s'être habillé, ou s'habiller pour prendre une douche.

Ces niveaux de réalisme sont complémentaires, le réalisme d'un niveau ne dit rien sur le réalisme d'un autre niveau. Par exemple, il est possible de produire une simulation dans laquelle chacune des traces d'activité des individus simulés est réaliste : pour chaque trace t_i , il existe un individu dans la population réelle dont l'activité r_j est proche de t_i . Pour autant, l'activité agrégée peut ne pas être réaliste (par exemple, tous nos individus font la cuisine entre 19h et 20h, ce qui est réaliste au niveau individuel, mais pas au niveau agrégé). Réciproquement, l'activité agrégée peut être réaliste (la proportion d'individus cuisinant à l'heure dite est correcte) sans que les activités individuelles ne le soient (il existe des traces t_i ne correspondant à aucune activité réelle r_j).

2.3. Cadre applicatif et objectifs

Le contexte applicatif de nos travaux est celui de la simulation de la consommation électrique des ménages, dans une optique d'anticipation et de réduction de la consommation énergétique dans l'habitat (Amouroux *et al*, 2013). Dans ce cadre, il est impératif que nos simulations visent un réalisme des activités au niveau agrégé afin de simuler des consommations électriques ayant une validité statistique et que la population simulée soit statistiquement réaliste au regard de la population cible. De plus, la question de l'efficacité énergétique nous oblige à considérer des simulations dont le réalisme est capable de supporter de longues durées (jusqu'à un an). Par exemple, une étude typique est l'étude comparative de la consommation de 10 000 ménages représentatifs de la population de France métropolitaine, pendant 1 an, avec différents jeux de données météo, afin de tester l'impact du climat sur la consommation agrégée à l'échelle du pays.

Cependant, comme l'indiquent Haradji *et al* (2012), la simulation de la consommation nécessite d'intégrer une relation structurelle à l'activité humaine, car la consommation dans l'habitat résulte pour beaucoup de la dynamique de vie quotidienne (préparer le repas, faire le bain des enfants, laver le linge etc.). Le réalisme des activités au niveau local est donc obligatoire afin de produire des consommations pertinentes à un niveau local et à un niveau agrégé.

De plus, pour répondre à notre défi de maîtrise de l'efficacité énergétique, nous souhaitons pouvoir simuler l'impact d'un tarif sur la consommation individuelle ou collective, tester les conséquences d'un nouvel équipement sur la facture du client, évaluer l'effet d'une politique énergétique à l'échelle locale, régionale ou nationale, etc. Il doit donc être possible de simuler des activités sur de longues périodes temporelles (semaines, mois), et de les rendre adaptables à des modifications de l'environnement. Ainsi, une étude envisagée est l'impact sur la courbe de charge nationale d'une généralisation des voitures électriques (et de leur longue charge généralement nocturne).

2.4. Plan de l'article

Dans la section 3 nous présentons un état de l'art de l'évaluation du réalisme d'une simulation, ainsi que de la simulation multi-agent de l'activité humaine. Dans ces deux domaines, le besoin de données réelles nous incite à nous intéresser également aux enquêtes statistiques portant sur l'activité humaine. Dans la section 4 nous présentons notre modèle : la plate-forme SMACH (simulation multi-agent du comportement humain), qui utilise conjointement des méthodes appartenant au monde multi-agent et des méthodes appartenant au monde de la statistique dans le but de créer des simulations réalistes au niveau microscopique et macroscopique. Nous exposons également les différentes validations réalisées concernant le réalisme individuel. Dans la section 5 nous montrons comment les données statistiques des enquêtes « emplois du temps » peuvent être utilisées pour générer des emplois du temps réalistes à un niveau agrégé. La 6^e section est consacrée à la nécessité d'intégrer de la variabilité dans l'emploi du temps généré. Dans la section suivante, nous présentons une implémentation de cet algorithme avec les données de la dernière enquête emploi du temps de l'INSEE, ainsi que les diverses expérimentations ayant été réalisées afin de valider le réalisme des simulations au niveau agrégé. Enfin, en section 8 nous discutons du réalisme local de nos simulations, ainsi que des détails spécifiques portant sur la simulation de consommation électrique.

3. État de l'art sur le réalisme des simulations de l'activité humaine

Ces dernières années, une importante activité de recherche s'est développée autour de l'évaluation du réalisme en simulation multi-agent. Nous allons montrer comment ces travaux se situent dans les différents niveaux de réalisme que nous avons identifié. Nous présenterons ensuite les différentes approches existantes en simulation multi-agent de l'activité humaine, et leurs intérêts respectifs.

3.1. Évaluation des différents niveaux de réalisme

Le réalisme d'une simulation de l'activité humaine peut être considéré sous différents angles. Par exemple, comme Marsella et Gratch (2009), on peut chercher à modéliser avec le plus précision possible l'intervention du niveau psychologique

dans le processus décisionnel, ou comme Thalmann et Musse (2007), on peut chercher à reproduire le plus fidèlement possible les mouvements d'une foule. Quel que soit le type de réalisme visé, son évaluation peut se faire de deux manières différentes : par une méthode quantitative ou par une méthode qualitative.

3.1.1. Méthode quantitative

La méthode quantitative consiste à calculer une mesure de distance entre des indicateurs issus de la simulation et ceux issus de données réelles. Cette méthode est donc applicable à n'importe quel niveau de simulation, tant que les données réelles sont disponibles.

Traditionnellement, les simulations de l'activité humaine ne s'intéressent qu'à un seul niveau de réalisme à la fois. Ainsi, au niveau individuel, Taillandier *et al.* (2010) s'intéressent à la gestion de l'espace de l'environnement des SMA, alors que Six *et al.* (2013) simulent le comportement de conducteurs de camions en se basant sur des données réelles portant sur la vitesse et la distance entre véhicules. Au niveau agrégé, le réalisme des activités produites est mesuré en termes de proximité statistique avec les comportements observés à l'échelle de la population étudiée, ainsi Caillou (2010) valide-t-il des simulations du marché du travail en comparant les données de la simulation et des données réelles. Au niveau local, (Darty *et al.*, 2014) proposent l'utilisation d'une méthode de classification afin de valider le réalisme des comportements de conduite en utilisant à la fois des données terrain et des données simulées. L'approche défend l'idée que si les clusters obtenus sont mixtes, c'est-à-dire contiennent des données réelles et simulées, alors les activités simulées sont indiscernables des activités réelles par rapport aux critères considérés.

Dans certaines approches, par exemple celles portant sur la variation dynamique du niveau de détail (Navarro *et al.*, 2011), le niveau de réalisme visé change en cours de simulation afin d'économiser les ressources de calcul. Ainsi, des mesures de distance entre agents regroupant profil psychologique et localisation sont calculées, dans le but de regrouper les agents les plus « proches ».

3.1.2. Méthode qualitative

La validation quantitative est limitée par les données réelles disponibles. La méthode qualitative permet de s'affranchir de cette limite, ce qui est indispensable pour des situations où les données réelles ne sont pas disponibles ou n'existent pas (par exemple pour les simulations de prospective). Comme l'indique Sargent (2013), une des méthodes les plus classiquement utilisées consiste à demander à des experts de valider le réalisme d'une simulation.

Au niveau individuel, il est possible d'utiliser un individu comme juge de son propre comportement, en le confrontant à ses actions simulées, par exemple grâce à de la simulation participative (Drogoul *et al.*, 2003 ; Haradji *et al.*, 2012). Dans ces approches, un expert évalue la proximité ressentie par l'humain entre son expérience vécue et la simulation. Au niveau agrégé, on peut par exemple demander à un expert énergéticien d'analyser une courbe de charge de consommation simulée pour dire si elle est réaliste. Au niveau local, Bosse *et al.* (2013) travaillent sur l'évaluation de

décisions prises collectivement, à partir des états mentaux individuels et des interactions entre individus.

3.2. Simulations multi-agents de l'activité humaine

Dans le domaine de la simulation de l'activité humaine, une grande partie des travaux utilise le paradigme SMA. Dans ces approches, on peut identifier trois catégories : 1) les approches utilisant des comportements « scriptés » ; 2) les approches orientées vers des comportements autonomes et/ou réactifs; 3) les approches hybrides visant à combiner les deux précédentes.

3.2.1. Approches scriptés

Dans les approches utilisant des comportements « scriptés », les agents suivent un comportement prédéfini (Ulicny et Thalmann, 2001). Ce type d'approches est typiquement utilisé lorsque les comportements attendus sont connus et suffisamment bien formatés. Un des intérêts majeurs de l'approche est la certitude que le scénario se déroulera selon un plan défini à l'avance, ce qui est utile par exemple dans les jeux, via le domaine de la scénarisation interactive (Paul *et al.*, 2010), ou dans les simulations à visée éducative ou d'entraînement (Sharma et Otunba, 2012), puisqu'il assure que l'apprenant sera exposé aux situations désirées.

Dans ce genre d'approches, le principe général est de limiter l'autonomie des agents afin d'augmenter le contrôle sur le déroulement de la simulation.

3.2.2. Approches réactives

Lorsque l'environnement est instable, que son évolution est difficile à prédire, ou que sa taille est importante, l'utilisation de scripts devient difficile à mettre en œuvre à cause du nombre de cas à traiter. Dans ces conditions, l'utilisation d'agents capables de réagir à leur environnement et de s'y adapter permet de simplifier le travail du modélisateur (Shendarkar *et al.*, 2008). Par exemple lorsque de nombreux agents coopèrent ou négocient, la capacité à réagir et à s'adapter aux comportements des autres prend de l'importance (Traum *et al.*, 2003). Un modèle largement utilisé est le modèle BDI (*Belief-Desire-Intention*) proposé par Rao et Georgeff (1991). Grâce à la modélisation des objectifs, intentions et croyances de chaque agent, les comportements individuels sont la résultante des tentatives des agents d'atteindre leurs objectifs.

Dans ce type d'approches, la contrepartie est la diminution de la maîtrise du déroulement des simulations.

3.2.3. Approches hybrides

L'idée générale de ces approches est de définir des « scénarios d'activité » qui servent de cadre comportemental de « haut-niveau » aux agents, et de les coupler avec un mécanisme autonome de sélection de l'action qui gère les comportements à un niveau plus élémentaire. Ces approches sont utilisées pour différents types

d'application. Par exemple Lanquepin *et al.* (2013) utilisent des modèles de coordination prescrits, combinés avec des mécanismes de raisonnement, afin de modéliser des comportements collectifs. Il existe plusieurs manières de combiner les comportements prescrits et les comportements autonomes. Ainsi, Hubner et Sichman (2007) proposent de construire un système de contraintes organisationnelles que chaque agent doit respecter. D'un autre côté, Tambe (1997) propose d'équiper les agents de modèles de coordination leur permettant de travailler en équipe de manière adaptative et flexible. Grosz et Kraus (1996) permettent aux agents de se coordonner pour planifier collectivement leurs actions.

La difficulté de ce type de modèles est d'arriver à calibrer les scénarios d'activités. En effet, les comportements sont autonomes et les éléments du scénario définissent des « bornes » à ne pas dépasser. Il est donc nécessaire de calibrer ces bornes.

C'est dans cette dernière catégorie que notre approche se situe. Par ailleurs, dans un système complexe, il n'est possible de produire des simulations réalistes que si la population d'entités qui représentent les composants du système est elle-même réaliste. En simulation multi-agent, ce problème est classiquement résolu grâce à l'utilisation de générateurs de population (domaine de recherche de la *population synthesis*) qui permettent de générer une population « virtuelle » similaire à une population cible (Beckman *et al.*, 1996 ; Gargiulo *et al.*, 2010 ; Müller et Axhausen, 2011). Cependant, la prise en compte de l'activité au niveau de cette génération est un problème majeur, non traité par les méthodes existantes. Afin de dépasser cette limitation, nous allons nous intéresser aux travaux issus des statistiques sur l'activité humaine.

3.3. Les approches statistiques basées sur l'enquête emploi du temps

Des enquêtes statistiques nationales de type « enquête emploi du temps » (*Time-Use Survey*) (Stinson, 1999) ont été menées dans différents pays afin de caractériser l'emploi du temps des gens. Dans ces enquêtes, les participants remplissent des questionnaires (ou « carnets ») portant sur l'emploi de leur temps au cours de la journée. Ils sont constitués d'un ensemble de plages horaires (typiquement, toutes les 10 minutes pendant 24h). Les participants indiquent quelle activité ils ont réalisée lors de chaque plage horaire. Ensuite des experts regroupent les activités déclarées en catégories. Un intérêt majeur de ces enquêtes est qu'elles respectent un formalisme mondialement partagé. Il est ainsi possible d'obtenir le même type de données pour des individus de différents pays dans le monde.

Ces enquêtes sont couramment utilisées par des approches statistiques pour reproduire l'activité humaine quotidienne. On distingue deux tendances dans l'utilisation de ces enquêtes pour la reproduction d'activités humaines : les approches *top-down* et *bottom-up*.

3.3.1. *Approches « top-down »*

Les données provenant des enquêtes sont utilisées pour calculer une matrice qui détermine, à chaque moment de la journée, la probabilité (pour un type d'individu donné) de passer d'une activité à une autre. La génération des activités est ensuite effectuée à chaque pas de temps, en sélectionnant l'activité suivante selon cette table de transition (Chiou, 2009 ; Richardson *et al.*, 2010 ; Widén *et al.*, 2012)).

D'après Yamaguchi et Shimoda (2015), ces approches sont limitées par trois facteurs. Elles nécessitent l'accès aux données brutes des enquêtes, qui ne sont pas toujours disponibles librement selon les pays ; elles manquent de précision quant à la durée des différentes activités (ces approches se concentrent sur les transitions entre les activités, et non sur leurs durées effectives) ; il n'est pas possible de gérer la coordination entre les membres d'un même ménage : les matrices de changement d'activités ne prennent pas l'environnement en compte.

3.3.2. *Approches « bottom-up »*

Dans ce type d'approche, les données des enquêtes sont utilisées pour calculer la durée moyenne et la répartition des activités pendant la journée. A partir de ces informations, les approches « bottom-up » construisent des emplois du temps en sélectionnant itérativement l'activité suivante grâce à une distribution probabiliste des durées. Tanimoto *et al.* (2008) montrent que cette approche ne nécessite pas un accès aux données brutes des enquêtes. Les seules informations nécessaires sont : la durée moyenne (et les écarts types associés) de chaque activité, ainsi que le pourcentage d'individus qui adoptent une activité spécifique à un moment donné.

Yamaguchi et Shimoda (2015) améliorent cette méthode en ajoutant le concept de « comportement de routine ». Un emploi du temps partiel est généré, uniquement constitué des activités « dormir », « travailler », « aller à l'école », ainsi que celles liées aux repas et à l'hygiène. Les autres comportements sont ajoutés à cet emploi du temps de manière à combler les « vides ». Le traitement différent des activités de routine (les premières à être placées), permet une coordination basique des individus (mais limitée à ces activités).

3.4. *Conclusion de l'état de l'art*

Ces méthodes statistiques permettent d'obtenir un réalisme des activités au niveau agrégé. Mais aucune d'entre elles ne s'intéresse au réalisme local. Le réalisme n'est que statistique : les individus simulés sont des « individus moyens ». C'est pourquoi nous proposons de coupler approches statistiques et approches multi-agents, afin de coupler réalisme agrégé et individuel. Ainsi, dans notre approche, nous utiliserons une méthode qualitative pour valider le réalisme individuel grâce à des simulations participatives, et des méthodes quantitatives pour valider le réalisme agrégé et local.

De plus, même si certains travaux en simulation multi-agent (Navarro *et al.*, 2011), permettent à une simulation de changer dynamiquement de niveau de détails,

aucune méthode ne permet à ce jour de coupler simultanément un réalisme local et un réalisme agrégé. Nous proposons de combler ce manque par notre approche : SMACH (simulation multi-agent de comportement humain).

4. SMACH : un modèle multi-agent de l'activité humaine

Notre proposition peut être résumée par les points suivants :

1. utiliser un système multi-agent capable de produire des simulations réalistes à un niveau individuel, grâce à des agents adaptatifs,
2. utiliser un générateur de population virtuelle capable de produire une population virtuelle réaliste par rapport à une population cible,
3. intégrer à ce générateur un mécanisme permettant de doter les individus générés d'un emploi du temps réaliste à un niveau agrégé (grâce à des données provenant d'une enquête emploi du temps),
4. simuler les individus et leur emploi du temps via le système multi-agent du point 1)

L'originalité de notre approche repose sur l'utilisation d'un système multi-agent adaptatif axé sur le réalisme individuel, couplé avec des données statistiques réalistes à un niveau agrégé. Cette articulation entre SMA et statistiques présente plusieurs avantages : elle résout le problème récurrent en multi-agent de la phase d'initialisation ; elle améliore le réalisme des simulations basées sur les statistiques en prenant en compte le niveau local ; et elle permet de passer de simulations visant uniquement une reproduction d'activités connues, à des simulations de prospective. Notre hypothèse est que le réalisme local naîtra de ce couplage.

4.1. Généralités

Le modèle d'agent utilisé est celui de la plate-forme SMACH (Amouroux *et al.*, 2014). À l'origine, cette plate-forme de simulation multi-agent a été développée pour simuler de manière réaliste des individus au sein d'un ménage.

Dans cette section nous allons commencer par présenter le modèle d'agent utilisé, puis le modèle d'activité. Pour finir, nous présenterons le travail effectué pour la validation du réalisme au niveau individuel.

4.2. Le modèle d'agents

Dans cette plate-forme, chaque individu est modélisé comme un agent avec des objectifs (activités à réaliser), des connaissances (sur les autres individus et sur l'environnement) et des préférences (en termes de confort, de comportement et d'emploi du temps). Les agents sont capables d'échanger des informations, de coordonner leurs activités, de planifier leurs journées et leurs semaines.

Chaque jour de simulation, les agents reçoivent une liste d'activités à effectuer : leur emploi du temps. A chaque pas de temps, ils sélectionnent leur activité courante dans cette liste, grâce à un mécanisme de sélection de l'action basé sur un niveau de priorité qui varie dynamiquement en fonction des connaissances de l'agent, de ses préférences, et de la situation. C'est ce processus qui permet de rendre les individus simulés adaptatifs. De nombreux paramètres sont pris en compte, certains sont détaillés dans la section suivante, mais Amouroux *et al.* (2013) en fait une présentation plus systématique.

4.3. Le modèle d'activités

Les activités du modèle SMACH sont dotées des paramètres suivants :

- **Durée.** Chaque activité est définie par une durée minimum et une durée maximum. Lorsqu'un agent réalise une activité, son niveau de priorité augmente par rapport à celui des autres activités. Ce bonus représente le bonus d'engagement dans l'activité courante. Une fois que l'activité a été réalisée pendant une durée au moins égale à sa durée minimum, ce bonus d'engagement diminue, jusqu'à atteindre zéro lorsque l'activité a été réalisée pendant une durée égale à sa durée maximum.

- **Rythme.** À chaque activité est attribué un nombre de répétitions par jour ou par semaine. Par exemple l'activité de sommeil est généralement répétée une fois par jour.

- **Période préférentielle (PP).** A chaque activité est associée une période préférentielle (notée PP) indiquant les périodes de la journée (ou de la semaine) qui sont préférées pour la réalisation de l'activité. A l'intérieur de la PP, le niveau de priorité de l'activité augmente. En dehors de la PP, il diminue.

Exemple : pour un individu donné, l'activité « douche » peut être caractérisée par une durée comprise entre 5 et 10 minutes, 1 fois par jour, entre 6h et 8h.

4.4. Validation du réalisme individuel

Le réalisme individuel des simulations SMACH a été validé grâce à une série de 10 simulations participatives (Haradji et Poizat, 2017). Plusieurs entretiens individuels ont été réalisés avec des familles, afin de déterminer leur emploi du temps hebdomadaire. Finalement, un humain de chaque foyer a été confronté à la simulation de sa vie quotidienne (c'est-à-dire de son propre emploi du temps, prenant en compte les interactions avec les autres membres de la famille). Il devait évaluer la dynamique de vie quotidienne de « son » agent, mais aussi la dynamique des interactions avec les autres agents du foyer. Il commentait et modifiait cette activité « artificielle » en lien avec son activité réelle.

Par ailleurs, plusieurs simulations « à l'aveugle » ont été effectuées. Une simulation à l'aveugle est une simulation d'un ménage (ici une colocation) dans laquelle plusieurs individus du ménage sont contrôlés de manière automatique, et d'autres sont pilotés par des humains. L'objectif était pour les humains de découvrir

qui parmi les membres de la colocation sont pilotés par la machine, et lesquels sont pilotés par d'autres humains. Les résultats ont montré qu'il était très difficile pour les humains de distinguer les avatars automatiques des avatars pilotés par les humains.

4.5. Limites de la plate-forme

En l'état, la plate-forme de simulation SMACH permet de simuler des activités réalistes au niveau individuel. Afin de viser également un réalisme agrégé, il est nécessaire de générer une population réaliste (l'étape est prise en charge par un générateur de population virtuelle) et d'intégrer des données statistiques sur les activités humaines. C'est le rôle du module de génération d'emploi du temps.

5. La génération d'emplois du temps

Afin de générer des emplois du temps ayant un réalisme au niveau agrégé, nous utilisons les données issues des enquêtes emploi du temps. Étant donné que l'ensemble de ces enquêtes partagent un formalisme commun, les données de n'importe quelle enquête peuvent être utilisées indifféremment.

5.1. Formalisme des données

Les données statistiques d'une enquête emploi du temps sont composées d'un ensemble C de « carnets » (les questionnaires de l'enquête) $c : C = \{c_1, \dots, c_n\}$, d'un ensemble ACT d'activités $A : ACT = \{A_1, \dots, A_p\}$, et d'un ensemble IND d'individus ind (les personnes interrogées) : $IND = \{ind_1, \dots, ind_k\}$. Les individus sont dotés d'un sexe, d'un âge, et d'un statut de travail (actif, retraité, chômeur, etc.). Chaque carnet c est lié à un individu ind , et à un jour j . Il contient un ensemble d'instances d'activités a , telles que : $c_{ind,j} = \{a_1, \dots, a_m\}$.

Nous définissons le « type d'individu » i (défini par les caractéristiques de sexe, âge et statut) et j le « type de jour » (typiquement jour de semaine/week-end). Ces catégories permettent de créer des groupes de carnets correspondant à des situations similaires, c'est-à-dire remplis par des individus d'un même type, pour un même type de journée. Grâce à ces groupes de carnets, nous pouvons établir des hypothèses statistiques sur la façon dont les individus d'un certain type emploient leur temps pendant des jours d'un certain type. Soit J l'ensemble des types de jour $j : J = \{j_1, \dots, j_q\}$ et I l'ensemble des types d'individus $i : I = \{i_1, \dots, i_y\}$.

5.2. Modélisation des données

À partir de ces données brutes, nous calculons pour toutes les activités $A \in ACT$, pour tous les types d'individu $i \in I$, pour tous les types de jour $j \in J$, et pour tous les carnets $c \in C$, les indicateurs suivants (tableau 1).

Tableau 1. Données calculées à partir des enquêtes emploi du temps

$N(A,c)$	le nombre d'instances de l'activité A , dans le carnet c
$C_{ij}=\{c_1,\dots,c_n\}$	l'ensemble des carnets liés à un individu de type i et à un jour de type j
$P_N(A,i,j)$	la distribution de probabilité des valeurs de $N(A,c)$ pour tout $c \in C_{ij}$
$C_N(A,i,j)=\{c_1,\dots,c_m\}$	l'ensemble des carnets $c \in C_{ij}$ dans lesquels il y a N instances de l'activité A
$D_N(A,i,j) = \{a_1,\dots,a_p\}$	l'ensemble des instances a de l'activité A dans les carnets $c \in C_N(A,i,j)$
$moy(D_N(A,i,j))$	la durée moyenne des instances $a \in D_N(A,i,j)$
$et(D_N(A,i,j))$	l'écart-type des durées des instances $a \in D_N(A,i,j)$
$PST(D_N(A,i,j))$	la distribution probabiliste des heures de début des instances $a \in D_N(A,i,j)$
$PFT(D_N(A,i,j))$	la distribution probabiliste des heures de fin des activités $a \in D_N(A,i,j)$

5.3. Définition des « emplois du temps bruts »

Ces données permettent de calibrer les paramètres du modèle d'activité (voir section 4.3). Ainsi, pour tout agent simulé (de type i), et pour toute journée de simulation (de type j), noté $ag_{i,j}$, on calcule, pour toute activité $A \in ACT$:

– $Rythme(A,ag_{i,j}) = N^*$: le nombre de fois que l'agent ag répètera l'activité A lors de cette journée de simulation, grâce à un tirage aléatoire dans $P_N(A,i,j)$.

– $Dmin(A,ag_{i,j}) = moy(D_{N^*}(A,i,j)) - 2 \cdot et(D_{N^*}(A,i,j))$: la durée minimum de chaque répétition de l'activité A lors de cette journée de simulation.

– $Dmax(A,ag_{i,j}) = moy(D_{N^*}(A,i,j)) + 2 \cdot et(D_{N^*}(A,i,j))$: la durée maximum de chaque répétition de l'activité A lors de cette journée de simulation.

– $PP(A,ag_{i,j})$: la période préférentielle pour réaliser ces répétitions est calculée grâce à des tirages aléatoires dans $PST(D_{N^*}(A,i,j))$ et $PFT(D_{N^*}(A,i,j))$

Notons qu'en raison des tirages aléatoires effectués, on ne manipule à aucun moment un « emploi du temps moyen », chaque emploi du temps généré est unique (2 agents ayant le même type auront des emplois du temps différents, mais générés à partir des mêmes données).

Ainsi, grâce aux données d'une enquête emploi du temps, il est possible de paramétrer les activités du modèle de telle sorte que leurs durée, rythme, et PP soient statistiquement réalistes pour tous les types d'individu et de journée. Cependant, ce réalisme n'est valable qu'à l'échelle d'une journée puisque les données des enquêtes emplois du temps ont été compilées à partir d'une, voire de 2 journées (non consécutives) pour chaque répondant. Or, comme le montrent Feldman et Pentland

(2003) et Haradji *et al.* (2012), l'activité humaine est caractérisée par des activités de routine journalière, mais aussi par des variations autour de ces routines. La question qui se pose alors est donc : à partir des données des enquêtes emploi du temps, qui ne portent que sur des journées individuelles, comment construire un emploi du temps sur une période plus longue (par exemple sur une semaine) et qui articule routines et variabilité dans les activités simulées ?

6. Rythme hebdomadaire et variabilité de l'activité

Générer des copies d'une même journée pendant plusieurs jours ne permet pas d'obtenir une variabilité dans les comportements. D'un autre côté, générer des journées sans lien les unes avec les autres ne permet pas d'obtenir de comportements de routine. Afin de résoudre ce problème, nous proposons la création d'emplois du temps « de routine ».

6.1. Les emplois du temps de routine

Notre solution pour dépasser cette limitation des données brutes est de poser l'hypothèse **H** suivante : « les comportements humains les plus atypiques à un niveau macroscopique (i.e. les comportements les moins représentés dans l'enquête) sont également des comportements inhabituels à un niveau individuel (c'est-à-dire qu'ils ne peuvent pas être des comportements de routine d'un individu) ».

Par exemple, si pour une enquête donnée, très peu de carnets recensent une activité de sommeil d'une durée inférieure à 2 heures par nuit, alors, d'après **H**, nous considérerons qu'aucun individu n'a pour habitude de dormir moins de deux heures par nuit. Les individus simulés peuvent toujours dormir moins de 2 heures, mais uniquement de manière exceptionnelle, en raison de circonstances particulières.

Cette hypothèse est une simplification de la réalité, puisqu'il existe des individus qui dorment en moyenne 2 heures par nuit. Cependant, la marge d'erreur de cette hypothèse se limite à des cas particuliers rares. Grâce à **H**, il est possible de « recréer » les informations liées à la variabilité des comportements qui manquent dans les données statistiques. Ainsi, pour chaque activité, on peut séparer les instances qui sont « habituelles », des instances « inhabituelles » en fonction de leur fréquence dans l'enquête. Les instances habituelles sont les comportements « de routine », alors que les instances inhabituelles représentent les variations individuelles autour de ces routines. Pour chaque individu, il devient possible de créer un emploi du temps « de routine », uniquement basé sur les instances de routine de chaque activité. En pratique, on considère qu'une instance est habituelle si sa durée est comprise dans l'intervalle :

$$[moy(D_N(A,i,j)) - et(D_N(A,i,j)) ; moy(D_N(A,i,j)) + et(D_N(A,i,j))]$$

c'est-à-dire si sa durée est située à moins d'un écart-type de la moyenne. D'après la définition de l'écart type, 68,2 % des instances de chaque activité sont dans cet

intervalle, donc les emplois du temps de routine se basent sur les 68,2% des instances les plus « habituelles ».

6.2. Génération des emplois du temps de routine

La génération des emplois du temps de routine est similaire à celle des emplois du temps brut, à l'exception des durées (dont l'intervalle est réduit) :

– $Rythme_R(A, ag_{i,j}) = NR^*(A)$: le nombre de fois que l'agent ag répètera l'activité A lors de cette journée de routine, grâce à un tirage aléatoire dans $P_{NR^*(A,i,j)}$.

– $Dmin_R(A, ag_{i,j}) = moy(D_{NR^*(A,i,j)}) - et(D_{NR^*(A,i,j)})$: la durée minimum de chaque répétition de l'activité A lors de cette journée de routine.

– $Dmax_R(A, ag_{i,j}) = moy(D_{NR^*(A,i,j)}) + et(D_{NR^*(A,i,j)})$: la durée maximum de chaque répétition de l'activité A lors de cette journée de routine.

– $PP_R(A, ag_{i,j})$: la période préférentielle pour réaliser ces répétitions de routine est calculée grâce à des tirages aléatoires dans $PST(D_{NR^*(A,i,j)})$ et $PFT(D_{NR^*(A,i,j)})$.

6.3. Génération de l'emploi du temps définitif

Les emplois du temps de routine sont uniques (grâce aux tirages aléatoires dans $P_{NR^*(A,i,j)}$). Chaque individu peut répéter son emploi du temps de routine sur plusieurs jours : c'est sa routine personnelle. Cependant, son comportement doit également varier autour de cette routine. Pour représenter cela, chaque nouveau jour de simulation, l'emploi du temps qui sera réellement utilisé sera légèrement différent de l'emploi de routine : chaque instance a présente dans l'emploi du temps de routine a une probabilité de 31,8 % de devenir une instance inhabituelle. Cette probabilité vient du fait que l'emploi du temps de routine ne représente que 68,2% des instances. Si a devient inhabituelle, cela se traduit par une modification de ses durées (qui sont éloignées de la moyenne afin de retrouver la bonne distribution des durées) :

$$Dmin(a) = moy(D_{NR^*(A,i,j)}) - 2.et(D_{NR^*(A,i,j)})$$

$$Dmax(a) = moy(D_{NR^*(A,i,j)}) - et(D_{NR^*(A,i,j)})$$

OU (avec une probabilité de 50%) :

$$Dmin(a) = moy(D_{NR^*(A,i,j)}) + et(D_{NR^*(A,i,j)})$$

$$Dmax(a) = moy(D_{NR^*(A,i,j)}) + 2.et(D_{NR^*(A,i,j)})$$

REMARQUE. — Les durées de 4,5 % des instances sont situées à plus de 2 écarts types de la moyenne (par définition de l'écart type). Nous choisissons de ne pas les modéliser, en raison du manque de représentativité des données statistiques associées.

6.4. Résumé du processus de génération d'emplois du temps

La génération des emplois du temps peut être résumée par les étapes suivantes :

- Le générateur de population construit une population
- Pour chaque individu de cette population, on détermine son type, et on calcule son emploi du temps de routine pour chaque type de journée
- Lors de chaque nouvelle journée de simulation, pour chaque individu de la simulation, nous calculons l'emploi du temps définitif du jour
- L'ensemble des activités de l'emploi du temps définitif est envoyé au mécanisme de sélection de l'action, qui calculera à chaque pas de temps la priorité de chacune de ces actions, et qui décidera de l'action courante

6.5. Conclusion sur le problème de la variabilité des comportements

Grâce à l'hypothèse *H*, en contrepartie d'une faible marge d'erreur, les activités des individus simulés exhibent à la fois des routines et des variations autour de celle-ci, tout en respectant les distributions statistiques des durées, rythmes, et périodes préférentielles. Dans la section suivante, nous allons présenter les différentes expérimentations ayant été accomplies afin de valider le réalisme agrégé et local.

7. Expérimentations

Ces expérimentations ont été réalisées en utilisant les données d'une enquête emploi du temps précise : la dernière enquête emploi du temps de l'INSEE, pour la population de France métropolitaine.¹

7.1. Présentation de l'enquête

Cette enquête emploi du temps a été réalisée par l'INSEE en 2009-2010. Elle porte sur 12 000 ménages, parmi lesquels une ou deux personnes ont été interrogées et ont rempli un ou deux carnets. Un carnet est composé de 162 fenêtres temporelles de 10 minutes, entre 21h et minuit le jour suivant. Cette enquête comporte environ 27 000 carnets, remplis par plus de 18 500 personnes. Les experts ont réparti les activités réalisées par les répondants dans 140 activités. Ces activités sont très précises, par exemple : « lire un journal », « lire un livre », « regarder un film », « regarder la TV », etc. Par rapport à nos objectifs de simulation (la simulation de la consommation électrique des ménages), nous préférons regrouper les activités similaires afin de simplifier le modèle et d'augmenter la représentativité de chaque activité. Nous créons donc 30 activités plus cohérentes par rapport à nos objectifs de simulation : « lire », « regarder la TV », « dîner », « travailler », etc.

1. <https://www.insee.fr/fr/metadonnees/source/s1224>

L'enquête emplois du temps recense un certain nombre d'informations concernant les individus, et leur ménage. Nous choisissons une typologie d'individus d'après les critères sociologiques les plus classiques : sexe, âge (tranches d'âges de 10 ans), et activité (actif, chômeur, retraité, étudiant). Toutes les combinaisons de ces critères ne sont pas présentes dans l'enquête (par exemple l'enquête ne recense aucune femme retraitée de moins de 20 ans). Nous obtenons au total 27 types d'individus. De plus, chaque journée enregistrée dans l'enquête est caractérisée par un type (jour de semaine/jour travaillé, week-end/jour de repos). Les journées d'enquête ont été choisies de telle sorte que les jours travaillés et les jours de repos sont bien équilibrés (51,4 % de journées travaillées).

7.2. Expérimentations macroscopiques : validation du réalisme agrégé

Les agents SMACH sont dotés de capacité d'adaptation grâce au processus autonome de sélection de l'action. Ce mécanisme leur permet, entre autres, de réorganiser les différentes activités de leur emploi du temps en fonction de la situation, voire de planifier la réalisation de certaines activités en avance si leur emploi du temps est fortement contraint. Il est donc nécessaire de vérifier que les activités effectivement simulées sont toujours conformes aux données statistiques, c'est-à-dire que l'autonomie des agents ne remet pas en cause la validité statistique des activités au niveau agrégé.

Dans un but de clarté, nous allons donner un exemple concret en comparant les activités réelles et simulées pour un type donné d'individu i_I : « homme actif entre quarante et cinquante ans », et pour le type de jour j_I : « jour de semaine ». Pour cela, on génère 100 emplois du temps (finaux) correspondant à un individu de type i_I et un jour de type j_I . Ces emplois du temps sont ensuite simulés dans la plateforme, puis comparés avec les activités réelles recensées dans l'enquête emploi du temps. Les figures 1 et 2 illustrent la répartition de différentes activités pendant les 100 jours simulés.

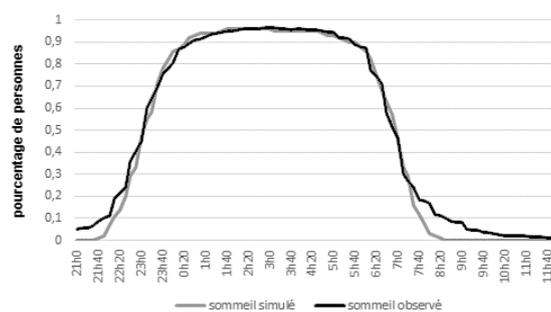


Figure 1. Comparaison des activités de sommeil simulées et observées

La seule différence notable sur la figure 1 est un écart vertical le soir et le matin. Il est principalement causé par les 4,5 % d'activités les plus rares, non simulées par notre modèle. Sur la figure 2, les écarts sont encore plus faibles : jamais plus de 3 points (écart vertical) et 20 minutes (écart horizontal) de différence.

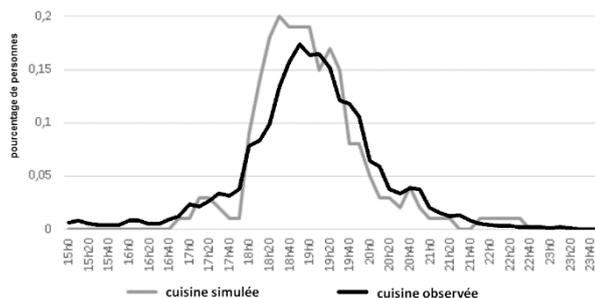


Figure 2. Comparaison des activités de cuisine simulées et observées

En ce qui concerne le nombre de répétitions des activités au cours de la journée (i.e. le rythme), regardons la figure 3, qui indique le nombre de répétitions des activités « préparer le dîner » et « faire le ménage » au cours d'une journée.

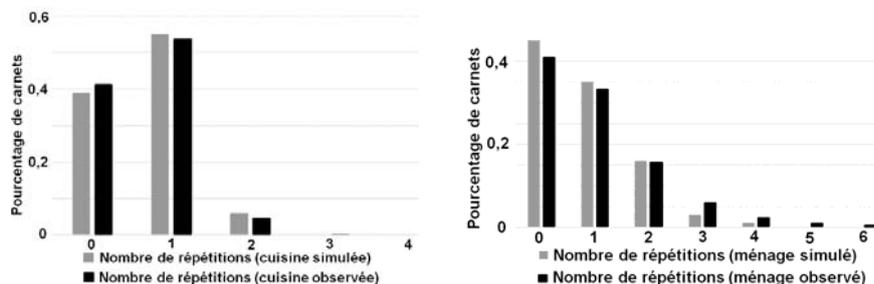


Figure 3. Comparaison des activités de cuisine et de ménage simulées et observées

Les écarts sont très limités pour l'activité de cuisine (< 2 points) et un peu plus importants (< 4 points) pour le ménage. En ce qui concerne le ménage, la raison principale est la difficulté pour nos agents simulés de trouver le temps durant leur journée de faire plus de 2 fois le ménage.

En conclusion, les activités réelles et simulées restent statistiquement très proches, tant sur le plan de la durée, du rythme, que des périodes préférentielles. L'impact des capacités d'autonomie de nos agents sur le réalisme agrégé des activités reste limité.

7.3. Expérimentations microscopiques

Le réalisme local de nos simulations peut être évalué par la méthode qualitative et la méthode quantitative. Nous allons présenter ici une première expérimentation qualitative. Une validation quantitative complète est actuellement en cours.

L'objectif est ici de vérifier de manière qualitative, pour un ensemble de simulations sélectionnées aléatoirement, que les activités simulées sont crédibles, et que malgré le cadre comportemental leur ayant été imposé (i.e. leurs emplois du temps leur ayant été attribué de manière automatique), nos agents sont toujours capables de s'adapter correctement à des modifications de leur environnement.

Le scénario d'activité considéré est celui d'un agent de type i_2 : « femme active dans la trentaine », pour un jour de type j_2 : « jour de semaine ». Nous allons générer 3 emplois du temps correspondant à cette situation, et les simuler. Chaque emploi du temps correspond à un déroulement possible d'une journée de semaine pour une femme active. Ils auront donc des similarités (travail, sommeil), mais également des différences (ménage, sorties, etc.). Chaque emploi du temps sera ensuite simulé une deuxième fois, mais avec l'activation d'un événement « panne de réveil », qui a pour conséquence de laisser dormir l'agent une heure de plus que d'habitude.

Nous présentons ci-dessous 3 couples de simulations dans lesquelles un individu alterne une journée « habituelle » sans événement, et une journée « inhabituelle » avec événement.

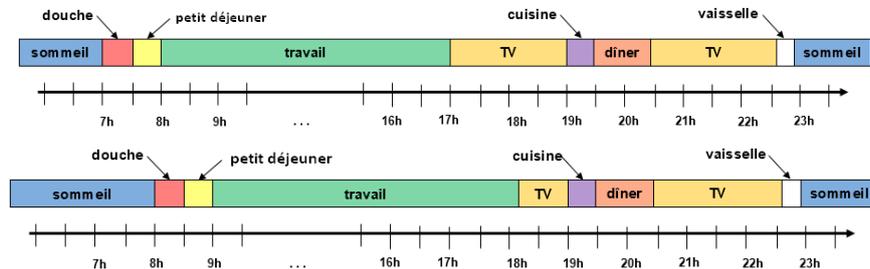


Figure 4a et b. Diagrammes d'activités (journées habituelle 1 et inhabituelle 1)

1)a) Journée habituelle (figure 4a) : l'individu se réveille à 7h, prend une douche, prend son petit déjeuner et se rend au travail à 8h. Il revient du travail à 17h et passe la soirée à la maison (télévision, cuisine, dîner, à nouveau télévision, puis vaisselle et sommeil à 23h).

b) Journée « inhabituelle » (figure 4b) : l'individu se réveille à 8h à cause de la panne de réveil. Il décale son emploi du temps du matin et part une heure en retard au travail. Il rentre une heure plus tard (18h 10). Une fois chez lui, il réduit le temps passé devant la télévision d'une heure environ (réduction du temps d'une activité).

2)a) Le deuxième individu a une journée « habituelle » similaire à la journée 1)a), si ce n'est qu'il se lève plus tôt, et qu'il occupe différemment sa soirée.

b) Sa journée « inhabituelle » se traduit par un déplacement de la douche dans la soirée (au lieu du matin), et une réduction du temps de travail.

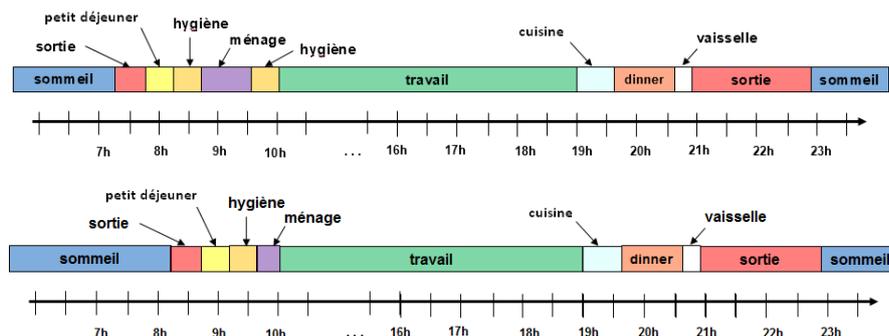


Figure 5. Diagrammes d'activités (journées habituelle 3 et inhabituelle 3)

3)a) Le troisième individu a une journée « habituelle » assez différente (voir figure 5).

b) Lors de la panne de réveil, il réduit le temps passé à faire du ménage et supprime un passage dans la salle de bain (voir figure 5).

Ces exemples tendent à montrer que nos agents sont capables de réagir à un événement et de s'adapter à un changement de situation en modulant leur emploi du temps. Ils peuvent décaler, supprimer, réduire, ou déplacer des activités. Aucune incohérence n'a été recensée dans les comportements des humains simulés. Cette expérimentation concernant l'adaptativité des agents n'est que préliminaire, mais elle donne un bon aperçu des capacités de nos agents.

8. Discussion

Avant de conclure, nous allons discuter plus en détails de la notion de réalisme local. Nous présenterons également quelques spécificités de notre travail liées à notre cadre applicatif : la simulation de consommation électrique.

8.1. Discussion sur le réalisme local

Différents problèmes apparaissent comme des freins à la validation du réalisme local. Tout d'abord, la coordination entre membres d'un même ménage peut difficilement être validée grâce aux enquêtes emploi du temps (la coordination entre individus n'est pas centrale dans l'enquête). Dans notre modèle, cette coordination est entièrement portée par le module de sélection de l'action (par exemple, activités

exclusives ou collectives, aide à la réalisation d'une action, etc.), et n'a été validée que via les simulations participatives.

Par ailleurs, Guibourdenche (2013) caractérise globalement l'activité humaine dans l'habitat comme un « contexte multi-préoccupationnel », c'est-à-dire un enchaînement d'actions, caractérisé par de nombreuses interruptions dans l'activité (voir figure 6).

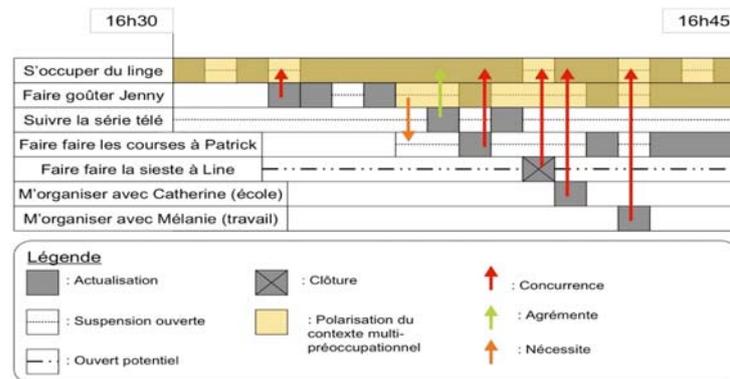


Figure 6. Exemple d'enchaînement d'actions

Ce constat soulève l'hypothèse que les humains planifient en partie ce qu'ils ont prévu de faire, mais qu'en pratique leur comportement effectif ne suit pas toujours cette planification en raison de nouvelles préoccupations non prévues initialement. Ce qui est intéressant dans cette vision c'est que notre système a un fonctionnement en partie comparable : nos agents ont un emploi du temps théorique (issus du générateur d'emploi du temps), mais leur comportement effectif ne le suit pas à la lettre (en raison de la présence du module de sélection de l'action). Le problème, c'est que nous construisons les emplois du temps théoriques de nos agents à partir du comportement effectif des humains (tel qu'il est présent dans l'enquête emploi du temps), et non à partir de leur planification initiale (qui est inconnue). Dans ce cadre, un comportement réaliste localement devrait-il correspondre ou ne pas correspondre aux comportements effectifs réels tels que décrits dans l'enquête emploi du temps ?

Comme nous visons une reproduction des activités et non du processus décisionnel, nous espérons obtenir des activités simulées similaires à celles enregistrées dans les enquêtes. Afin de vérifier cela, nous sommes actuellement en cours de réalisation d'une évaluation quantitative, basée sur de la classification (Darty *et al.*, 2014)). L'idée est de déterminer un certain nombre de paramètres, par exemple les durées des activités, rythmes, et périodes de réalisation, et de réaliser une classification des activités selon ces paramètres en mélangeant des activités réelles (provenant d'enquêtes emploi du temps), et simulées (provenant de notre plate-forme). Si les clusters obtenus sont homogènes, c'est-à-dire s'ils ne contiennent que des activités d'un seul type (réelles ou simulées), alors c'est qu'il

est possible de les différencier. Par contre, si les clusters obtenus sont hétérogènes, alors les activités réelles et simulées sont indiscernables les unes des autres par rapport aux paramètres considérés.

8.2. Simulation énergétique

Dans nos travaux, la consommation électrique des ménages est déduite de l'interaction entre les activités humaines et l'équipement électrique du ménage. Cependant, les enquêtes emploi du temps ne s'intéressent pas particulièrement aux activités consommatrices d'énergie. Par exemple, dans ces enquêtes il n'existe qu'une activité générique « hygiène personnelle », recouvrant de multiples comportements hétérogènes consommateurs d'énergie : « prendre une douche », « prendre un bain », « se laver les dents », etc. Nous avons été obligés de déterminer un pourcentage d'utilisation des différents appareils électriques pour chacune de ces activités « génériques », ce qui diminue la précision des consommations simulées.

Par ailleurs, les appareils électriques présents dans le ménage ont également un impact majeur sur les consommations (présence ou non d'un sèche-linge, d'un lave-vaisselle, type de chauffage, etc.), ainsi que la notion de confort qui est centrale dans le fonctionnement de SMACH (températures de consigne du chauffage en journée, pendant la nuit, en cas d'absence, etc.). Pour une grande partie de ces questions, nous avons pu nous appuyer sur l'expertise de notre partenaire industriel EDF, mais certains paramètres restent incertains (par exemple, les moyennes et écarts-type, dans la population de France métropolitaine, des capacités, puissances, efficacités, et températures de chauffage des chauffe-eau). Afin de vérifier la validité de l'ensemble des paramètres liés aux consommations électriques, nous avons réalisé en parallèle des travaux présentés dans cet article une calibration des consommations électriques. Cette calibration a eu lieu à la fois au niveau individuel lors des simulations participatives (Haradji et Sempé, 2015), et au niveau agrégé (par exemple par rapport aux courbes officielles ENEDIS²).

9. Conclusion

Dans cet article nous avons montré le manque, dans l'état de l'art, de méthodes permettant de simuler des activités réalistes à différents niveaux. Qui plus est, aucune méthode existante ne permet de réaliser des simulations de prospective de l'activité humaine sur de longues périodes temporelles. Nous avons proposé un modèle permettant de combler ce manque, c'est-à-dire une méthode qui permet de coupler réalisme individuel grâce à un SMA réactif, et réalisme agrégé grâce à des données statistiques sur l'activité humaine. Dans ce sens, notre approche peut être considérée comme une modélisation multi-agent hybride, couplant des activités prescrites de haut niveau (les emplois du temps), avec un mécanisme de sélection de l'action autonome définissant le comportement élémentaire des agents.

2. <http://www.enedis.fr/coefficients-des-profils>

À l'aide d'expérimentations qualitatives, nous avons montré que le SMA de base permet bien la simulation d'activités réalistes d'un point de vue individuel. Grâce à une expérimentation quantitative, nous avons montré que nous sommes également capables de simuler des activités réalistes au niveau agrégé. Nous sommes actuellement en cours de réalisation d'une étude quantitative visant à valider le réalisme local de nos simulations.

Le module de génération d'emplois du temps que nous avons présenté peut être réutilisé dans de nombreux autres domaines que celui de la simulation de consommation électrique. En effet, les activités recensées sont génériques, et peuvent s'adapter à bien d'autres applications. De plus, ces études ont été menées à une échelle mondiale, et respectent toutes un formalisme commun. Il est donc possible de les utiliser pour simuler des populations de nombreux pays du monde.

Au sein d'EDF, des experts de différents métiers (thermique du bâtiment, véhicule électrique, autoconsommation/autoproduction, etc.) s'appuient aujourd'hui sur les données de consommation produites par SMACH et réalisent des études pour des situations futures d'utilisation de l'électricité. En effet, relever des consommations électriques réelles est un processus coûteux (installation d'appareils de mesure), limité (on ne peut relever des consommations futures !), et pas forcément idéal (erreurs de mesure, pertes de données, etc.). Notre plate-forme permet quant à elle de générer des consommations électriques réalistes sur différentes échelles temporelles, pour n'importe quel type de population, tout en laissant la possibilité de réaliser des études de prospectives en générant des consommations électriques probables pour des scénarios futurs potentiels.

Bibliographie

- Amouroux E, Huraux T, Sempé F, Sabouret N, Haradji Y. (2014). SMACH: Agent-Based Simulation Investigation on Human Activities and Household Electrical Consumption. *Communications in Computer and Information Science* 449: 194–210.
- Amouroux E, Huraux T, Sempé F, Sabouret N. (2013). SMACH: Simuler l'activité Humaine Pour Limiter Les Pics de Consommation Électrique. *Journées Francophones Sur Les Systèmes Multi-Agents (JFSMA)*.
- Barot C, Lourdeaux D, and Lenne D. (2013). Using Planning to Predict and Influence Autonomous Agents Behaviour in a Virtual Environment for Training. *2013 IEEE 12th International Conference on Cognitive Informatics and Cognitive Computing*.
- Beckman R, Keith J, Baggerly A, and McKay M. (1996). Creating Synthetic Baseline Populations. *Transportation Research Part A: Policy and Practice* 30(6): 415–29.
- Bosse T, Hoogendoorn M, Klein M.C, Treur J, Van Der Wal C.N, Van Wissen A. (2013). Modelling Collective Decision Making in Groups and Crowds: Integrating Social Contagion and Interacting Emotions, Beliefs and Intentions. *Autonomous Agents and Multi-Agent Systems*, 1–33.
- Caillou P. (2010). Automated Multi-Agent Simulation Generation and Validation. *Principles and Practice of Multi-Agent Systems*, Springer-Verlag, 398–413.

- Chiou YS. (2009). Deriving US Household Energy Consumption Profiles from American Time Use Survey Data a Bootstrap Approach. *11th International Building Performance Simulation*.
- Darty K, Saunier J, and Sabouret N. (2014). Behavior Clustering and Explicitation for the Study of Agents' Credibility: Application to a Virtual Driver Simulation. *International Conference on Agents and Artificial Intelligence*: 82–99.
- Drogoul A, Amouroux E, Caillou P, Gaudou B, Grignard A, Marilleau N, Taillandier P, Vavasseur M, Vo D-A, Zucker J.D (2013). Gama: Multi-Level and Complex Environment for Agent-Based Models and Simulations. *2013 International Conference on Autonomous Agents and Multi-Agent Systems*, 1361–62.
- Drogoul A, Vanbergue D, and Meurisse T. (2003). Multi-Agent Based Simulation: Where Are the Agents? *Multi-Agent-Based Simulation II*, Springer Berlin, 1–15.
- Feldman MS, and Pentland BT. (2003). Reconceptualizing Organizational Routines as a Source of Flexibility and Change. *Administrative science quarterly* 48(1): 94–118.
- Gargiulo F, Ternes S, Huet S, and Deffuant G. (2010). An Iterative Approach for Generating Statistically Realistic Populations of Households. *PloS one* 5(1): e8828.
- Grosz BJ, and Kraus S. (1996). Collaborative Plans for Complex Group Action. *Artificial Intelligence* 86(2): 269–357.
- Guibourdenche J. (2013). *Préoccupations et Agencements Dans Les Contextes d'activité Domestique. Contribution à La Conception de Situations Informatiques Diffuses, Appropriables et Énergétiquement Efficaces*. Université Lumière-Lyon II.
- Haradji Y, Poizat G, and Sempé F. (2012a). Advances in applied human modeling and simulation *Human Activity and Social Simulation*. ed. V. G. Duffy. Boca Raton, FL : CRC Press.
- Haradji Y, Poizat G, and Sempé F. (2012b). L'activité Humaine et La Conception d'une Plateforme de Simulation Sociale. *47^e congrès de la Société d'Ergonomie de Langue Française (SELF)*.
- Haradji Y, and Poizat G. (2017). Simuler l'activité Humaine. In *Encyclopédie d'analyse Des Activités*, PUF, 255–81.
- Haradji Y, and Sempé F. (2015). *Valider La Simulation de Comportements Humains Dans SMACH : Principes Méthodologiques et Résultats*.
- Hubner JF, and Sichman JS. (2007). Developing Organised Multiagent Systems Using the MOISE+ Model: Programming Issues at the System and Agent Levels. *International Journal of Agent-Oriented Software Engineering* 1: 370–95.
- Lanquepin V, Carpentier K, and Lourdeaux D. (2013). HUMANS: A HUMAN Models Based Artificial ENVironments Software Platform. *Proceedings of the Virtual Reality International Conference: Laval Virtual*.
- Marsella S, and Gratch J. (2009). EMA: A Process Model of Appraisal Dynamics. *Cognitive Systems Research* 10(1): 70–90.
- Mohammadi B, and Saiac JH. (2003). *Pratique de La Simulation Numérique*. Dunod.
- Müller K, and Axhausen K. (2011). Population Synthesis for Microsimulation: State of the Art. *Transportation Research Board Annual Conference*.

- Navarro L, Flacher F, and Corruble V. (2011). Dynamic Level of Detail for Large Scale Agent-Based Urban Simulations. *The 10th International Conference on Autonomous Agents and Multiagent Systems*, 701–8.
- Paul R, Charles D, McNeill M, and McSherry D. (2010). MIST: An Interactive Storytelling System with Variable Character Behavior. *Joint International Conference on Interactive Digital Storytelling*, Springer, 4–15.
- Rao AS, and Georgeff MP. (1991). Modeling Rational Agents within a BDI-Architecture. *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, ed. James Allen and Richard Fikes and Erik Sandewall. Morgan Kaufmann publishers Inc., San Mateo, CA, USA, 473-484.
- Richardson I, Thomson M, Infield D, and Clifford C. (2010). Domestic Electricity Use: A High-Resolution Energy Demand Model. *Energy and Buildings* 42(10): 1878–1887.
- Sargent R G. (2013). Verification and Validation of Simulation Models. *Journal of Simulation* 7(1): 12–24.
- Sharma S, Otunba S. (2012). Collaborative Virtual Environment to Study Aircraft Evacuation for Training and Education. *Collaboration Technologies and Systems*: 569–74.
- Shendarkar A, Vasudevan K, Lee S, and Son YJ. (2008). Crowd Simulation for Emergency Response Using BDI Agents Based on Immersive Virtual Reality. *Simulation Modelling Practice and Theory* 16(9): 1415–29.
- Six L, Guessoum Z, Saunier J, and Ieng S. (2013). Towards a Truck-Driver Model Using a Hysteresis Based Analysis and Verification Approach. *The 2013 International Conference on Autonomous Agents and Multi-Agent Systems*, 1219–20.
- Stinson LL. (1999). Measuring How People Spend Their Time: A Time-Use Survey Design. *Monthly Lab. Rev.* 122: 12.
- Taillandier P, Vo D-A, Amouroux E and Drogoul A. (2010). GAMA: A Simulation Platform That Integrates Geographical Information Data, Agent-Based Modeling and Multi-Scale Control. *Principles and Practice of Multi-Agent Systems*, Springer, 242–58.
- Tambe M. (1997). Towards Flexible Teamwork. *Journal of artificial intelligence research* 7: 83–124.
- Tanimoto J, Hagishima A, and Sagara H. (2008). A Methodology for Peak Energy Requirement Considering Actual Variation of Occupants' Behavior Schedules. *Building and Environment* 43(4): 610–619.
- Thalmann D, and Musse SR. (2007). *Crowd Simulation*. London, UK, Springer.
- Traum D, Rickel J, Gratch J, and Marsella S. (2003). Negotiation over Tasks in Hybrid Human-Agent Teams for Simulation-Based Training. *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, 441–48.
- Ulicny B, Thalmann D. (2001). Crowd Simulation for Interactive Virtual Environments and VR Training Systems. *Computer Animation and Simulation*, Springer, 163–70.
- Widén J, Molin A, and Ellegård K. (2012). Models of Domestic Occupancy, Activities and Energy Use Based on Time-Use Data: Deterministic and Stochastic Approaches with Application to Various Building-Related. *Journal of Building Performance Simulation* 5(1): 27–44.
- Yamaguchi Y, and Shimoda Y. (2015). Evaluation of a Behavior Model of Occupants in Home Based on Japanese National Time Use Survey. In *Proceedings of BS2015*.

