

Réseau bayésien pour la gestion de l'obsolescence dans une base d'informations en vue de l'évaluation du risque de chute des personnes âgées

Salma Chaieb^{1,3,*}, Véronique Delcroix ^{2,**}
Ali Ben Mrad^{1,3,***}, Emmanuelle Grislin-Le Strugeon^{2,**}

1 : FSM, Université de Monastir, Skanes 5000 Monastir, Tunisie

2 : LAMIH-UMR CNRS 8201, Univ. de Valenciennes, 59304 Valenciennes, France

3 CES Lab, ENIS, Université de Sfax, 3038, Sfax, Tunisie

* salma.chaieb2@yahoo.com, *** benmradali2@gmail.com

** {veronique.delcroix,emmanuelle.grislin}@univ-valenciennes.fr

Résumé. L'évaluation périodique du risque de chute des personnes âgées requiert des informations fiables et nombreuses. Comme il n'est pas possible de recueillir régulièrement toutes ces informations, les observations sont faites au fil du temps et conservées, ce qui entraîne une problématique liée au vieillissement des informations. Cet article traite de la détection des informations obsolètes dans une base d'informations sur une personne âgée. Nous proposons une solution comportant un modèle de connaissances sur les personnes âgées sous forme d'un réseau bayésien et un module de raisonnement chargé de la détection et de la gestion des contradictions et des doutes sur les informations.

1 Introduction

La chute est la première cause de décès accidentel chez les personnes âgées de plus de 65 ans. L'évaluation régulière du risque de chute requiert des informations fiables et nombreuses sur la personne âgée et la collecte de ces informations est coûteuse en temps. Pour pallier ce problème, nous proposons un système chargé de collecter les informations au fil du temps, de les stocker et les gérer, en vue de fournir ces informations à la demande. Il s'agit donc d'un cas de la vie réelle où on ne dispose que d'informations incomplètes, incertaines ou incohérentes dans des situations de prise de décision. Ces informations sont très souvent variées et évolutives, risquent de devenir obsolètes et de contredire d'autres informations. Elles doivent ainsi être vérifiées continuellement afin de refléter fidèlement l'état et le comportement des personnes. L'objectif de ce travail est ainsi de maintenir un ensemble d'informations le plus fiable et complet possible, de façon à fournir à la demande des informations avec un degré de confiance élevé.

Pour cela, nous proposons un mécanisme de détection et de gestion des contradictions dans la base d'informations. La détection des observations obsolètes repose sur leur contradiction potentielle avec d'autres observations plus récentes et/ou plus

jugées plus fiables portant sur des variables différentes. Détecter ce type de contradiction nécessite de combiner des connaissances sur les dépendances entre les variables et du raisonnement dans l'incertain du fait des observations incomplètes d'une part et de l'incertitude stochastique sur les connaissances. Notre proposition est basée sur l'utilisation d'un réseau bayésien (RB) (Jensen, 1996; Naïm et al., 2011; Pearl, 1988) à même de combiner modèle de connaissance et de raisonnement dans l'incertain. Un RB est un graphe dirigé sans circuit G défini sur un ensemble de nœuds $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$, associé avec une distribution de probabilités P sur \mathbf{X} telle que $P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i))$ où $Pa(X)$ est l'ensemble des parents du nœud X dans G . Ce modèle représente un ensemble d'informations stables et générales : des *connaissances*. Il permet de calculer $P(V|Obs)$ où V est une variable non observée, et Obs désigne une *observation* c'est à dire l'instanciation d'une ou plusieurs variables. La distribution de probabilités *a posteriori* $P(V|Obs)$ représente l'état cognitif, ou la *croissance*, d'un observateur à propos de la variable V , dans une situation bien déterminée. Dans cet article, les observations ne sont pas synchrones, puisque nous considérons des informations recueillies au fil du temps. En conservant les observations du passé, nous ajoutons un degré d'incertitude supplémentaire puisque nous ne pouvons plus affirmer avec certitude que l'observation est encore vraie. La prise en compte du vieillissement des observations ne peut être gérée directement dans le RB. Il ne s'agit pas non plus d'observations incertaines (Pearl, 1988; Valtorta et al., 2002; Mrad et al., 2015). Pour gérer l'incertitude liée au vieillissement des observations, nous proposons des fonctions de *péremption*, qui permettent de calculer pour chaque observation un *degré de confiance*.

Cet article commence par un rapide état de l'art sur la problématique de l'obsolescence des informations. Il présente ensuite l'architecture du système et le modèle de raisonnement permettant la gestion de l'obsolescence. La dernière partie présente la réalisation de notre application.

2 Etat de l'art sur l'obsolescence des informations

Des solutions à la problématique de l'intégration de nouvelles données à une base de connaissance existante sont proposées par les algorithmes de bandit (Louède et al., 2015) dans le cadre des systèmes de recommandation en particulier. Il s'agit de mettre à jour l'ensemble des informations en supprimant celles les plus anciennes (dont l'âge dépasse un certain seuil). Des systèmes de reconnaissance automatique de cibles radar (Saidi et al., 2009) ou encore d'information en milieu hospitalier (Todoran, 2014) se basent sur l'étude de mesures de qualité de l'information en fonction de son évolution dans le temps (la qualité locale/en entrée et la qualité globale/en sortie). Dans notre proposition, comme dans les cas mentionnés, les informations traitées sont accompagnées de méta-informations permettant d'évaluer la qualité de ces dernières. Cependant, dans notre cas, il ne s'agit pas de déterminer la pertinence des observations "au moment où elles arrivent" comme cela est décrit dans ces cas, ni de prendre en compte uniquement les informations les plus récentes. Par ailleurs, le vieillissement des informations est un problème connu des bases de données dans lesquelles sont regroupés des enregistrements concernant un grand nombre d'entités, dont chacune est décrite par

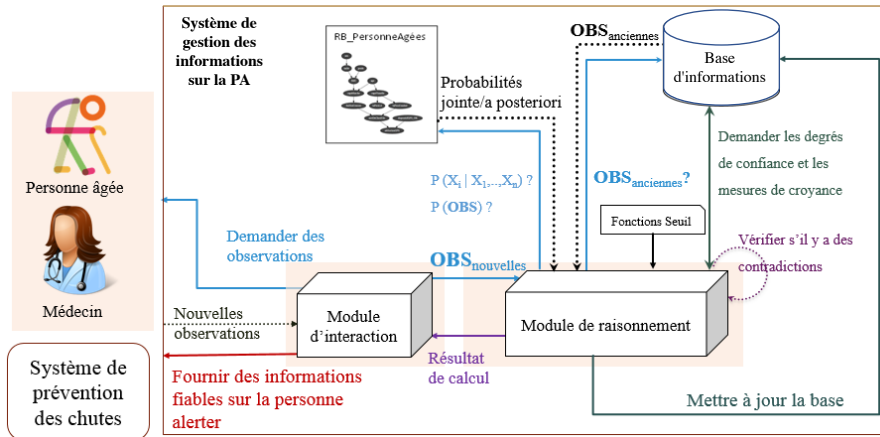


FIG. 1 – Schéma général du système chargé de la gestion des informations d'une PA

un nombre limité d'informations. A l'inverse, la base d'informations que nous considérons concerne une seule personne, mais regroupe de nombreuses informations qui sont en partie dépendantes les unes des autres (dépendance causale et/ou statistique). La problématique de cet article vise à prendre en compte ces dépendances pour y détecter d'éventuelles contradictions avec une approche probabiliste.

3 Gestion de l'obsolescence des informations

Le rôle du système proposé est de maintenir à jour une base d'informations sur une PA, sur la base d'informations acquises par des dispositifs externes et/ou par interaction avec la personne. Le système peut décider d'interagir avec la PA pour confirmer les observations. Les opérations suivantes sont effectuées de façon cyclique :

1. recueil de données sur la personne, constituées d'observations et de leurs dates ;
2. détection d'éventuelles contradictions entre ces nouvelles observations et les observations plus anciennes, et plus généralement avec l'ensemble des croyances sur l'état des variables pour lesquelles on n'a pas ou plus d'observations ;
3. mise à jour des informations sur la personne à partir des observations et du RB ;
4. interaction avec la personne en vue de réduire les contradictions.

L'architecture comporte quatre composantes (figure 1) : un modèle de connaissances, une base d'informations, un module de raisonnement et un module d'interaction.

Modèle de connaissances : RB qui encapsule les connaissances sur une population de personnes âgées. Dans ce travail, nous proposons un petit RB à treize nœuds (voir figure 2) qui vise à implémenter un démonstrateur.

Base d'informations : regroupe l'ensemble des variables du RB, les observations associées, les dates de saisie, les fonctions de préemptions et les degrés de confiance

(exemple en tableau 1). Seule la dernière valeur observée est conservée. Pour les variables qui n'ont pas ou plus de valeur observée, la base contient la valeur la plus probable sachant les observations récentes et la probabilité *a posteriori* associée. Ces mesures reflètent les croyances sur l'état de la PA.

Variable	Valeur observée	Date	Confiance	Le plus prob. obs	Prob. obs
laPersConduit	auMoins1ParSem	t_1	$Pm_1(t_1)=+$	-	-
faitSesCourses	oui	t_1	$Pm_2(t_1)=+$	-	-
dispositifGPS	ok	t_1	$Pm_3(t_1)=+$	-	-
capaVisuelle	-	-	-	correcte	0.98
sortDeChezElle	-	-	-	auMoins1ParSem	0.9
nbSortiesGPS	-	-	-	deuxOuPlus	0.73
capaMarche	-	-	-	normale	0.8
laPersLit	-	-	-	régulièrement	0.63
age	-	-	-	[60-63]	0.51
sexe	-	-	-	M	0.51
IMC	-	-	-	normale	0.38
taille	-	-	-	[160-170]	0.3
poids	-	-	-	[60-75]	0.19

TAB. 1 – Base d'informations : les observations (avec date et niveau de confiance) et les croyances sur les variables non observées (distributions de probabilités *a posteriori*).

Module de raisonnement : gère la cohérence de la base d'information. La présence d'une observation obsolète dans la base d'informations peut se manifester de deux façons. D'une part, quand une observation obsolète entre en contradiction avec des observations sur d'autres variables et qu'en conséquence l'observation simultanée de cet ensemble de valeurs est extrêmement improbable. Nous dirons que la base d'informations n'est pas dans un *état possible*. D'autre part, une observation obsolète peut générer un *doute* sur une variable non observée. Les activités d'inférence, de détection de contradictions, remise en cause des observations passées, suppression des observations obsolètes et vérification des doutes sont répétées jusqu'à ce que la base d'informations soit dans un *état stable*. Le module de raisonnement s'appuie sur les trois concepts suivants :

Une base d'informations est dans un *état stable* si et seulement si elle est dans un *état possible* et on n'a aucun *doute* sur ses variables n'ayant pas ou plus de valeur observée.

Une base d'informations est dans un *état possible* lorsque l'ensemble des observations $\mathbf{OBS} = \mathbf{obs}$ est tel que $P(\mathbf{OBS} = \mathbf{obs}) > \text{Seuil}(\mathbf{OBS})$, où $\text{Seuil}(\mathbf{OBS})$ est une fonction qui associe à un ensemble de variables une probabilité en dessous de laquelle un ensemble d'observations n'est "pas possible". Cette définition relaxe le sens strict du mot *possible* qui correspond généralement à une probabilité non nulle. La valeur du seuil dépend logiquement de la taille du domaine des observations et des conventions adoptées dans la définition du RB pour caractériser les situations locales impossibles dans les TPC (Tables de Probabilités Conditionnelles). Nous avons adopté une définition très simplifiée de la fonction seuil.

On dit qu'on a un *doute* sur une variable n'ayant pas ou plus de valeur observée si la croyance sur cette variable est fortement remise en cause par les nouvelles observations.

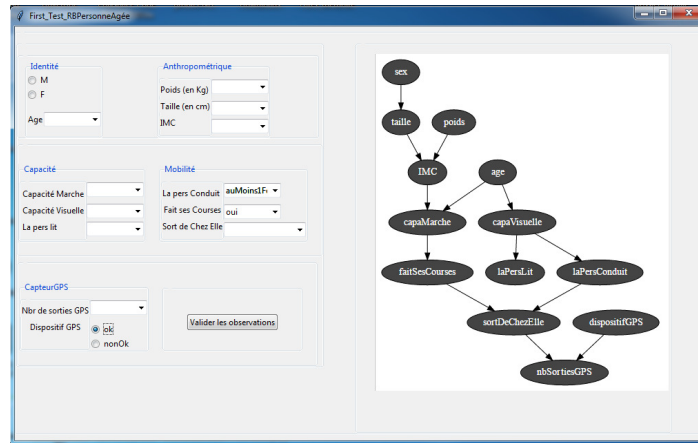


FIG. 2 – Interfaces graphiques pour la gestion des interactions.

Module d'interaction : gère en permanence l'arrivée des nouvelles informations et fournit des informations fiables sur des variables cibles à la demande.

4 Réalisation et résultats de l'application

Notre application a été implantée sous l'environnement Windows, avec le logiciel pyAgrum (Gonzales et al., 2014). Nous avons testé les méthodes proposées sur quelques scénarios à l'aide d'un premier démonstrateur (voir figure 2). Nous avons réalisé un premier test comparatif avec un raisonnement consistant à déclarer obsolète toute information dont le degré de confiance est sous un seuil fixé. Ceci peut conduire à deux écueils : supprimer une information qui n'est pas devenue obsolète; ne pas supprimer des informations obsolètes. Comparativement, notre proposition évite systématiquement le premier écueil et partiellement le second puisque c'est la contradiction avec une information plus récente qui permet de détecter l'obsolescence. Les résultats préliminaires trouvés sont encourageants mais doivent être vérifiés et améliorés afin d'améliorer la performance du système.

5 Conclusion et perspectives

Nous avons proposé un système chargé de la gestion de l'obsolescence des informations dans le cadre de l'évitement de la chute chez les personnes âgées. Le système détecte d'éventuelles contradictions entre de nouvelles observations et les informations précédemment recueillies sur la personne, et met à jour de la base d'informations à l'aide d'un modèle probabiliste de connaissances générales sur les PA sous forme d'un réseau bayésien. Nos perspectives à court terme concernent une meilleure définition formelle du problème de la gestion d'une base d'informations non synchrones et non

pérennes Par la suite, nous visons une amélioration des algorithmes proposés et des tests sur un RB plus complet sur la PA, défini par des experts du domaine.

Remerciements Ce travail est soutenu par le projet ELSAT2020 (Eco-mobilité Logistique Sécurité et Adaptabilité dans les Transports à l'Horizon 2020) et cofinancé par l'Union Européenne avec le Fond européen de développement régional, l'État et la Région Hauts de France. Les auteurs remercient le support de ces institutions.

Références

- Gonzales, C., L. Torti, et P.-H. Willemin (2014). Librairie agrum : a graphical universal model. *Revue d'intelligence artificielle-no 2(3)*, 1–10.
- Jensen, F. V. (1996). *An introduction to Bayesian networks*, Volume 210. UCL press London.
- Louède, J., M. Chevalier, A. Garivier, et J. Mothe (2015). Algorithmes de bandit pour les systèmes de recommandation : le cas de multiples recommandations simultanées. In *CORIA*, pp. 73–88.
- Mrad, A. B., V. Delcroix, S. Piechowiak, P. Leicester, et M. Abid (2015). An explication of uncertain evidence in bayesian networks : likelihood evidence and probabilistic evidence. *Appl. Intell.* 43(4), 802–824.
- Naïm, P., P.-H. Willemin, P. Leray, O. Pourret, et A. Becker (2011). *Réseaux bayésiens*. Editions Eyrolles.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems : networks of plausible inference*. Morgan Kaufmann.
- Saidi, M. N., A. Toumi, B. Hoeltzener, A. Khenchaf, et D. Aboutajdine (2009). Système automatique de reconnaissance de cibles radar. *6e Atelier Fouille de Données Complexes dans un processus d'extraction de connaissances*, 15.
- Todoran, I. G. (2014). *Etude dynamique de la qualité de l'information et des données d'un système d'information complexe*. Ph. D. thesis, Télécom Bretagne, Univ. de Rennes 1.
- Valtorta, M., Y.-G. Kim, et J. Vomlel (2002). Soft evidential update for probabilistic multiagent systems. *International Journal of Approximate Reasoning* 29(1), 71–106.

Summary

The frequent evaluation of the risk of fall for elderly people require reliable and abundant information about the person. Since it is not possible to gather regularly all the interesting pieces of information, observations are made over time and stored, which lead to a problem of information aging. This article deals with outdated information in a base of information about an elderly person. We propose a solution that includes a bayesian network-based model of knowledge and a reasoning module in charge of the detection and the management of contradictions and doubts about information.