

Détermination du niveau de consommation des abonnés en téléphonie mobile par la théorie des ensembles flous

Rachid EL MEZIANE (*), Ilham BERRADA (*), Ismail KASSOU (*), Karim BAINA (*)

Laboratoire Al Khawarizmi - ENSIAS - BP 713 - Agdal - Rabat - Maroc
(*){meziane, iberrada, kassou, baina@ensias.ma}

Résumé. La détermination du niveau de consommation chez les clients est essentielle pour tout objectif de segmentation stratégique et de churn. Nous présentons sur un cas réel l'utilisation de la théorie des ensembles flous pour la définition d'une fonction d'appartenance permettant d'évaluer, de manière précise, le niveau de consommation, des abonnés en téléphonie mobile.

1 Contexte

Notre travail s'inscrit dans le contexte d'un projet de fouille de données mis en oeuvre à Maroc Telecom et visant à mieux connaître la clientèle de la téléphonie mobile. Le niveau de consommation d'un abonné est souvent calculé à partir de la durée facturée qui s'avère insuffisante pour la plupart des cas. En effet, deux abonnés peuvent avoir la même durée d'appel pour des services différents mais sans avoir le même degré de consommation. D'où la nécessité d'introduire d'autres critères dans la détermination du niveau de consommation.

2 Problématique et approche de résolution préconisée

La problématique à laquelle on s'intéresse consiste à établir une échelle de mesure permettant de quantifier les niveaux de consommation afin discriminer entre les abonnés (Viertl, R. (2005)). L'approche de résolution proposée comporte trois étapes principales. Son originalité réside dans l'utilisation de la théorie des ensembles flous à travers la définition expérimentale d'une fonction d'appartenance (Mitaim, S. et B, Kosko. (2001)).

Dans une première étape, on attribue un score aux abonnés par rapport aux critères de type catégoriels (trafic, produits, services, plage horaire) caractérisant le niveau de consommation. La binarisation de chaque modalité de ces critères induit la création de plus de 60 variables indicatrices dans notre exemple qui traite 2 millions d'enregistrements. Afin de réduire la taille de ces indicatrices, l'Analyse des Correspondances Multiples (ACM) a été utilisée fournissant ainsi 10 facteurs expliquant 80,62% d'inertie totale.

L'objectif de l'étape 2 est la segmentation des abonnés par produits et services afin de discriminer entre les abonnés en se basant sur le comportement d'utilisation des produits et services. Les facteurs obtenus par l'ACM ont été utilisés comme variables d'entrée des différents algorithmes non supervisés (K-means, Two Step, Réseau de Kohonen) qui ont été comparés. Le réseau de Kohonen a été plus concluant en terme d'homogénéité entre les classes. Les inerties intra classes de chaque facteur ont ensuite été utilisées dans l'étape 3 comme indicateur de la variation du niveau de consommation au sein de chaque classe. Un tel indicateur a permis d'établir une mesure du niveau de consommation tenant compte de la durée facturée en appliquant la théorie des ensembles flous (Masson, M. H. (2003)).

Détermination du niveau de consommation des abonnés en téléphonie mobile

La fonction d'appartenance que nous proposons, dans l'étape 3, pour quantifier l'ensemble flou $A = \ll \text{niveau de consommation en téléphonie mobile} \gg$, représente une mise à l'échelle exponentielle de la durée facturée. Cette fonction d'appartenance donnée par l'équation (1), présente non seulement l'avantage de ne pas masquer les optima locaux de la distribution de la durée facturée mais aussi l'avantage d'être paramétrée par des critères liés aux produits et services.

$$f_A(d) = \sum_{j \in K} I_j(d) \sum_{i \in F_j} (d)^{p x_{ij}} \quad (1)$$

où d : la durée facturée, $K = \{C_1, C_2, C_3, C_4\}$: l'ensemble des classes issues du réseau de Kohonen, F_j : l'ensemble des facteurs contribuant à la construction de la classe j , $j \in K$, x_{ij} : l'inertie du facteur i appartenant à F_j , I_j : la fonction indicatrice de la classe j , p : une constante à ajuster (dans notre cas $p : 0.01$ suite aux différents tests numériques réalisés).

3 Résultat et discussion

La fonction d'appartenance définie par l'équation (1) a permis de distinguer entre les 4 classes des produits et services et fournit une précision au niveau de l'interprétation du niveau de consommation. En effet, la **Classe C₂ de niveau de consommation moyennement élevé**, comporte des abonnés ayant utilisé la majorité des produits et services mais **en faible durée facturée**. La **Classe C₃ de niveau de consommation élevé** regroupe des abonnés ayant utilisé la majorité des produits et services mais avec **des durées facturées plus importantes**. La **Classe C₄ de niveau de consommation faible** se caractérise par l'utilisation des produit et services par un trafic national, en faible durée facturée. La **Classe C₁** se caractérise par un **niveau de consommation moyennement faible** et regroupe des abonnés utilisant des produits et services par un trafic international mais **en faible durée facturée**. Il en résulte que le niveau de consommation ne devrait pas être basé uniquement sur la durée facturée mais doit tenir compte aussi des produits et services.

Références

- Masson, M. H. (2003). Apports de la théorie des possibilités et des fonctions de croyance à l'analyse de données imprécises. *Thèse de l'Université de Technologie de Compiègne*.
- Mitaim, S. et B, Kosko. (2001). The Shape of Fuzzy Sets in Adaptive Function Approximation. *IEEE transactions on fuzzy systems*, vol. 9, no. 4.
- Viertl, R. (2005). Fuzzy Models for precision measurements. Department of statistics and probability theory, *Vienna University of technology*, Austria.

Summary

The aim of this paper is to provide a tool allowing the decision maker of a wireless operator to determine the level of consumption in a less subjective way in order to find the knowledge hidden in their database. Our approach uses the fuzzy sets theory with an experimental membership function.