

Détection précoce de tendances produits dans le cadre des activités commerciales de la grande distribution

Gaël Bardury*, Jean-Emile Symphor**

*Université des Antilles et de la Guyane
gbardury@gmail.com,
<http://www.univ-ag.fr>

**Université des Antilles et de la Guyane
jesympho@martinique.univ-ag.fr,
<http://www.univ-ag.fr>

Résumé. Dans ce papier, nous présentons une nouvelle approche qui permet la détection précoce de tendances "produits" dans le cadre des activités commerciales de la grande distribution. S'agissant d'un domaine où la concurrence est très vive entre les différentes enseignes avec des enjeux financiers colossaux, les stratégies commerciales ont pour principal objectif de fidéliser la clientèle pour limiter leur défection. C'est là qu'intervient la détection des changements de tendances produits, qui va permettre d'anticiper l'attrition de la clientèle. Déceler des tendances suffisamment tôt permettra aux décideurs de mettre en place des stratégies préventives efficaces à moindre coût. Notre objectif est donc d'analyser et de modéliser clairement les changements de tendances et leurs impacts potentiels globaux sur les achats des clients. Nous illustrerons notre approche sur des données réelles d'achats de clients d'une grande enseigne.

1 Introduction

Comprendre les changements de comportement des clients est un problème essentiel qu'affrontent les acteurs de la grande distribution alimentaire. Avec la concurrence féroce que se livre les acteurs de ce milieu, ne pas détecter assez tôt ces phénomènes, peut s'avérer risqué. La trace des articles achetés par un client est l'information la plus importante que possède les enseignes de la grande distribution pour détecter ces changements. En effet, les produits proposés constituent l'élément essentiel sur lequel repose la fréquentation ou non de la clientèle. L'étude de ces traces a été premièrement mise en lumière par Agrawal et Srikant (1994). Ils ont été les précurseurs de ce pan de recherche qu'est l'analyse du panier de la ménagère. Contrôler l'évolution de ce panier permettrait de déceler au plus tôt les événements qui font que les clients entament un processus d'attrition ou au contraire manifestent un intérêt grandissant pour les produits de l'enseigne. A titre d'exemple, la baisse de la qualité d'un ou plusieurs produits au cours du temps ou encore l'absence trop longue d'un produit recherché par la clientèle sont des facteurs qui peuvent engendrer une baisse de fréquentation voire une perte de clientèle. A l'inverse, une campagne promotionnelle sur de nouveaux produits "attractifs" peut élargir

la base d'achats des clients et du coup, améliorer la fidélisation vis à vis de l'enseigne. Par ailleurs, quand on sait que des associations fortes entre produits peuvent exister, par exemple en recherchant les règles d'association ou des liens de similarité, la baisse de qualité ou encore l'absence d'un produit peut provoquer un impact de baisse d'achat non seulement sur les produits directement concernés, mais aussi indirectement sur les produits associés. Cet incidence sur les produits associés est bien souvent peu ou mal voire pas du tout analysée. Song et al. (2001) évaluent les modèles issus des données au fil du temps. L'apparition ou la disparition d'un modèle, l'augmentation ou la diminution soudaine de leur support ou encore le changement d'une partie de ces modèles sont évalués au moyen de critères présentés dans leur article. Chen et al. (2005) ou encore Bottcher et al. (2009) ont présenté de même des études analogues. Sachant que les modèles qu'ils extraient tous, sont des règles d'associations, on peut se poser la question de savoir comment est-ce qu'ils les analysent. Il est connu que l'extraction de règles d'associations génèrent un problème au niveau du nombre de motifs extraits. Choi et al. (2005) et Chen (2007) utilisent diverses techniques afin de classer ce flot de règles extraites par ordre de pertinence. C'est ici que se situe le principal apport de notre article. Nous choisissons de représenter les motifs de façon globale et ainsi de constater plus aisément les changements de tendances.

Dans cet article, nous proposons une approche qui permet de détecter de façon précoce de nouvelles tendances produits afin, d'une part de limiter l'attrition des clients et d'autre part, d'améliorer leurs fidélisation. Dans la section 2, nous explicitons précisément notre méthodologie que nous illustrons en section 3 avec des cas réels extraits de la base de données achats et clients d'une enseigne de la grande distribution. Enfin, en section 4, nous présentons notre conclusion.

2 Notre approche

A notre sens, dans le paradigme de la grande distribution, quand on a l'objectif de contribuer à la mise en oeuvre de stratégies proactives pertinentes, il faut pouvoir apporter des réponses à trois phases. La première concerne la modélisation des liens entre articles fortement liés qui sera largement inspirée des réseaux sociaux. La deuxième phase concerne la recherche d'articles clés dans chaque cluster. Enfin, la dernière phase se focalise sur l'analyse de l'évolution des clusters, en prenant en compte particulièrement la question des cycles et de la saisonnalité, afin de détecter de véritables tendances fortes. En premier lieu, nous présenterons la façon dont nous construisons notre réseau d'articles. Puis, nous expliquerons comment déceler les articles pivots et enfin comment détecter les changements de tendances.

2.1 Construction du réseau

Les réseaux sociaux sont un ensemble d'entités sociales reliées entre elles par des liens pondérés ou non qui représentent des interactions. Raeder et Chawla (2009) proposent de modéliser le panier de la ménagère sous forme de réseau social qu'il appelle réseau d'articles. Ils construisent ce réseau de manière intuitive en considérant qu'un article (noeud) est relié à un autre article si ils ont été vendus simultanément. La transition aura un poids correspondant au nombre de ventes simultanées. Un algorithme de détection de communautés est proposé afin de détecter les meilleurs regroupements d'articles possibles au sein de ce réseau potentiellement

très dense. Nous proposons à l'inverse de créer ces clusters directement en calculant les indices de similarité de Jaccard entre les articles. Une transition existera entre deux articles si l'indice de similarité respecte le seuil fixé par l'utilisateur. En ne se fixant que le seuil de la similarité, on ne passe pas à côté des niches comportementales. En effet, le support, c'est-à-dire le nombre d'occurrences d'un article ou d'un ensemble d'articles, n'est pas un critère discriminant. De cette façon, si une minorité de clients commence à élaborer un schéma jamais vu, nous serons à même de le voir. Afin de garder, des informations sur une fenêtre variable, nous construisons notre réseau d'articles avec des flags attachés aux transitions. Les flags marqueront les périodes sur lesquelles la transition existe.

2.2 Diversité

Dans une enseigne, il est indubitable que des articles ont une plus grande importance que d'autres. Dans cette optique, (Liu et al., 2010) a introduit la notion de diversité. Cette mesure caractérise à quel point le voisinage d'un noeud est divers, même quand on ne dispose que de peu d'information sur les noeuds. Dans notre contexte, nous disposons de beaucoup d'informations sur les noeuds articles, comme leur nomenclature ou encore leur marque. Dans un magasin, les articles sont rangés selon leur rayon, leur famille puis sous famille d'appartenance. Si un article est souvent lié dans notre réseau à plusieurs articles qui n'appartiennent pas au même rayon, il se révèle donc d'une forte diversité. Prenons l'exemple du client qui achète un article à forte diversité. Il est fort probable qu'il cherche à acquérir des articles qui lui sont fortement liés. Il devra par conséquent, effectuer des déplacements dans le magasin. Ces déplacements sont la clé pour atteindre le potentiel de vente maximum du client. En effet, le client sera susceptible d'être tenté par des articles qui sont sur son parcours mais qu'il n'était pas venu acheter à l'origine. Il en découle donc un potentiel d'amélioration non négligeable de chiffres d'affaires. Déceler les variations de ventes inhabituelles de ces articles, entre autres, relève d'une importance capitale dans la fidélisation de la clientèle et de la pérennisation d'un magasin. Les articles à forte diversité, de par leur importance, seront les véritables garants de la stabilité des ventes de l'enseigne.

2.3 Saisonnalité et rythme d'achats

Pour mieux prendre en compte les aspects cycliques et saisonniers précisés en amont, nous nous fixons à la période mensuelle pour pouvoir analyser au mieux les tendances et les installations de celles-ci. Pour de multiples raisons, le mois est le compromis parfait :

- il pondère au mieux les différences entre les rythmes d'achat des clients en prenant simplement en compte le versement du salaire mensuel qui est le véritable garant des achats d'un client
- Pour deux articles fortement liés comme la confiture et le pain, clairement leur fréquence d'achat pour une même personne n'est la même. La cinétique de ces articles ne permet pas aux clients de les acheter tout le temps ensemble,
- ce choix nous évite les épiphénomènes, essentiellement, lors de la disparition ou l'apparition de connexions entre articles.

Le choix de regrouper les articles au fil des périodes grâce à la similarité en occultant les questions de supports pondèrent les variations saisonnières des ventes. En effet, les quantités vendues seules, pourraient amener des résultats biaisés et potentiellement erronés.

2.4 Détection de tendances

Pour un noeud, l'apparition ou la disparition, la diversité ou la banalité et l'évolution du nombre de voisins sont autant de paramètres qui indiqueront la stabilisation d'une tendance dans notre environnement. Nous détectons l'installation d'une tendance si on observe un changement du rythme d'apparitions de l'article sur un nombre de mois prédéfini. Exemple : Si un article est considéré comme un pivot et à un rythme d'apparition constant sur une période et qu'il cesse d'être pivot pendant un laps de temps défini. Nous déclenchons donc une alerte aux décideurs.

3 Application

Les exemples ont été réalisés à partir des données d'un hypermarché sur une période allant de janvier 2011 à septembre 2012. Par soucis de confidentialité, les articles seront uniquement identifiés par leur code. Dans les réseaux présentés ci-après, à l'intérieur du noeud est indiqué le code de l'article et sur les transitions, le nombre d'achats simultanés et l'indice de similarité des articles connectés. Nous fixons le seuil de détection à deux mois consécutifs de changement. Commençons par un exemple simple, sur l'apparition de plusieurs tendances. La figure 1 nous présente le réseau le plus dense que l'on peut assimiler au panier moyen sur le mois d'Avril 2012. La figure suivante 2 présente elle, le panier moyen après l'apparition de plusieurs tendances. On passe de 14 noeuds dont 6 pivots et 19 transitions à 30 noeuds dont 16 pivots et 122 transitions. Il est à savoir qu'il y a eu une campagne publicitaire dans ce laps de temps qui a été suivie et pérennisée. Il apparaît évident quand bien même les fluctuations saisonnières de chiffres d'affaires que l'élargissement du panel fréquent des clients est palpable et ne peut être que bénéfique pour l'enseigne.



FIG. 1 – Réseau d'articles le plus dense du mois d'avril 2012

Un second cas intéressant est celui de la disparition d'un article 009119 et de la compensation de celui-ci par un autre de même type 149290. La figure 3 nous présente le nombre de clients ayant acheté le produit 009119 et le produit de compensation 149190. En avril (point d'abscisse 15), nous constatons selon nos paramètres de départ pour notre algorithme, une tendance à la disparition du 009119. D'après la figure 3, on suppose que cela correspond à une rupture de stock. De plus, à la même période, nous constatons l'apparition du 149290. La figure 4 nous montre que l'article 09119 est un article pivot (en rouge). En regardant, l'article sur la figure 1, on constate le poids rédhibitoire de l'article 149290. Au total, l'article prendra 3 mois pour acquérir un statut d'article pivot au même titre que son homologue.

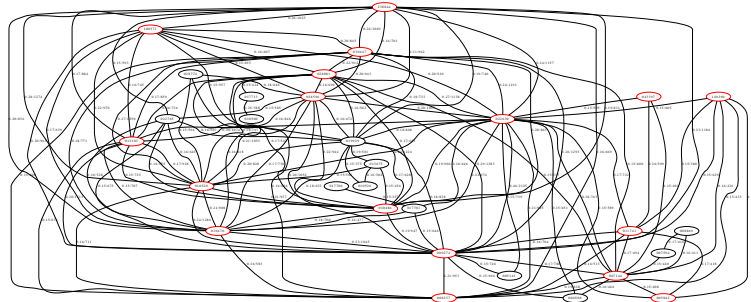


FIG. 2 – Réseau d’articles le plus dense du mois de juin 2012

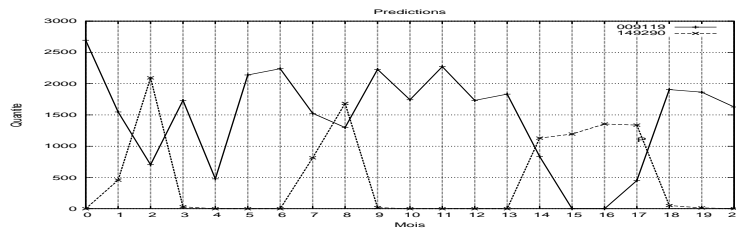


FIG. 3 – Évolution des ventes des articles 009119 et 149290

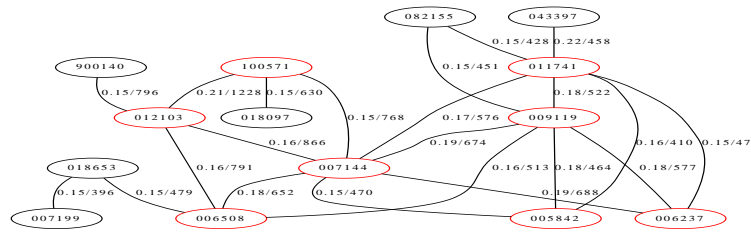


FIG. 4 – Réseau de Février 2012 comprenant le 009119

4 Conclusion

Dans ce papier, nous présentons une façon originale de détecter les tendances produits en analysant les connexions entre ceux-ci. En modélisant les relations entre articles sous forme de réseau et en marquant l’importance des articles, nous obtenons une vision globale des achats des clients à un moment donné. L’évolution de ces réseaux peut être interprétée de différentes façons et le but avant tout est de prendre des décisions assez tôt pour gérer au mieux les événements. Les décideurs commerciaux pour conforter la tendance devront gérer correctement le stock de l’article et de ces satellites dans le but de répondre à la demande. De plus, dans le cas d’une rupture inévitable, communiquer correctement sur un autre article de substitution afin

d'éviter le départ de la clientèle ou, dans le meilleur des cas, une simple non vente des produits liés à l'article en rupture.

Références

- Agrawal, R. et R. Srikant (1994). Fast algorithms for mining association rules. *Proceedings of the 20th International Conference on Very Large Databases 1*, 487–499.
- Bottcher, M., M. Spott, D. Nauck, et R. Kruse (2009). Mining changing customer segments in dynamic markets. *Expert Systems with Applications 36*, 155–163.
- Chen, M.-C. (2007). Ranking discovered rules from data mining with multiple criteria by data envelopment analysis. *Expert Systems with Applications 33*, 1110–1116.
- Chen, M.-C., A.-L. Chiu, et H.-H. Chang (2005). Mining changes in customer behavior in retail marketing. *Expert Systems with Applications 28*, 773–781.
- Choi, D. H., B. S. Ahn, et S. H. Kim (2005). Prioritization of association rules in data mining: Multiple criteria decision approach. *Expert Systems with Applications 29*, 867–878.
- Liu, L., F. Zhu, M. Jiang, J. Han, L. Sun, et S. Yang (2010). Mining diversity on social media networks. *Springer Science*.
- Raeder, T. et N. Chawla (2009). Modeling a store's product space as social network. *Proceedings of the 2009 international conference on advances in social network analysis and mining 1*, 164–169.
- Song, H. S., J. k. Kim, et S. H. Kim (2001). Mining the change of customer behavior in an internet shopping mall. *Expert Systems with Applications 21*, 157–168.

Summary

Throughout this paper, we will present a new approach that allows the precocious products trends detection within the framework of commercial activities of the mass-market retailing. As regards a field in which the competition is very vivacious between the different shop signs and which involves colossal financial wagers, commercial strategies have as main purpose to secure the loyalty of the customers in order to limitate their defection. The detection of the changes of the products trends occurs at that specific moment and will permit to look ahead the attrition of the clientele. Tracking down the trends early enough will allow to the decision makers to set up efficient preventive strategies at a lower cost. Therefore our goal is to analyze and to clearly model the variations of tendencies and their potential global impacts concerning the costumers' purchases. We will base our approach on actual data of clients purchases of a big sign/brand.