

Consolidation des données statistiques par expertise et similarité pour la prévision des ventes

Andy MALO et Olivier MARTINEZ

Cubtek Conseils

Technopole Helioparc

2 Avenue du Président Pierre Angot

65053 Pau CEDEX – France

andy.malo@cubtek.fr, olivier.martinez@cubtek.fr

Eric VILLENEUVE et Laurent GENESTE

Laboratoire Génie de Production

Université de Toulouse – Ecole Nationale d'Ingénieurs de

Tarbes (ENIT)

47 Avenue d'Azereix – BP 1629

65016 Tarbes CEDEX – France

eric.villeneuve@enit.fr, laurent.geneste@enit.fr

Résumé – Cette communication a pour objectif de présenter des travaux sur des méthodes de prévisions de vente utilisant des données statistiques consolidées par des avis d'experts et des statistiques issues de cas similaires afin de piloter l'activité commerciale d'une entreprise disposant de plusieurs points de vente. Nous proposons un modèle de prévision intégrant les connaissances statistiques et expertes ainsi que les mesures de similarité pour affiner la prévision. Nous proposons également une méthode d'élicitation des avis d'experts sous forme de fonctions de croyance.

Mots-clés— Prévision des ventes, Aide à la décision, Élicitation d'avis d'experts, Similarité, Fonctions de croyance.

I. INTRODUCTION

La société CubTek Conseils développe actuellement une plateforme logicielle de pilotage de performance. Cette solution est aujourd'hui destinée aux sociétés de télécommunications mais l'objectif final est de proposer un outil utilisable pour différents secteurs.

Actuellement, les ventes de lignes mobiles et fixes et de matériels (accessoires, téléphones,...) sont en plein essor. Pour piloter l'activité commerciale de ce domaine, la prévision des ventes est un atout non négligeable. Les méthodes de prévision classiques [1] produisent des résultats convenables mais il est difficile d'améliorer la précision de ces prévisions. En effet, pour atteindre ce but, il est nécessaire de prendre en compte de nombreux paramètres tels que : la gestion de la concurrence, la gestion des offres commerciales, la typologie des points de vente et la gestion des aléas. De plus, lorsque l'on souhaite piloter l'activité plus finement (prévision des ventes d'un produit particulier par un vendeur ou un point de vente défini) la quantité de données disponibles est souvent insuffisante pour obtenir des statistiques fiables.

Cette communication a pour objectif de présenter des travaux sur des méthodes de prévisions de vente utilisant des données statistiques consolidées par des avis d'experts et des statistiques issues de cas similaires afin de piloter l'activité commerciale d'une entreprise disposant de plusieurs points de vente. Les connaissances permettant d'alimenter l'outil d'aide à la décision développé par la société Cubtek Conseils pour cette activité de prévision sont donc principalement issues de deux

sources : d'une part, les connaissances extraites de bases de données multidimensionnelles et, d'autre part, les connaissances des experts du domaine. Nous proposons ici d'élaborer un mécanisme d'hybridation permettant de coupler ces deux types de connaissances afin de les exploiter dans un même outil, de manière cohérente et complémentaire.

A. Projet

La société Cubtek Conseils est spécialisée dans le conseil l'étude et le développement d'applications dans le secteur de l'informatique décisionnelle. Dans ce cadre, la société développe son produit « Swifteo » qui est une plateforme décisionnelle collaborative, commercialisée en mode SaaS (Software as a Service – Logiciel en tant que service, c'est-à-dire que l'entreprise vend un abonnement périodique au logiciel associé à des services de maintenance et de support), et destinée au pilotage de la performance des points de ventes de télécommunications.

Le projet est découpé en trois phases : en premier lieu, nous réalisons la prévision de l'évolution des ventes en intégrant explicitement l'incertitude ; ensuite la prévision de l'impact d'une action choisie par l'utilisateur du logiciel sur l'évolution du point de vente ; et enfin, une aide à la décision de l'utilisateur dans le choix des actions à réaliser pour l'aider à mettre en place une véritable stratégie commerciale.

Dans ce contexte, nous nous intéressons spécifiquement, à la mise en place d'une base de données statistiques et à la capitalisation des connaissances émises par les experts pour déterminer les similarités entre les points de ventes et à leurs retours d'expériences pour agréger une base de données des événements perturbateurs. L'objectif final est de préparer le développement d'un module de prévisions pour un progiciel interactif d'aide à la décision dans le cadre d'une activité commerciale répartie sur un ensemble de points de vente.

Nous avons identifié trois tâches pour atteindre cet objectif :

- La première tâche consiste à choisir un formalisme de représentation de la connaissance. Actuellement, il existe plusieurs formalismes de représentation qui peuvent être envisagés en vue de cette hybridation. Ceux-ci doivent permettre de représenter l'incertitude de type aléatoire (inhérente aux données) et

l'incertitude de type épistémique (inhérente au savoir incomplet des experts). Nous avons décidé d'utiliser les fonctions de croyance car ce formalisme permet de prendre en compte ces deux types d'incertitude et de fusionner les informations en provenance de sources multiples. Il permet également de mettre en œuvre des mécanismes de similarité qui nous seront utiles pour retrouver les cas similaires et ainsi consolider les connaissances disponibles.

- La seconde tâche consiste à définir les mécanismes d'inférence du système de prévisions. Il s'agit ici, sur la base du choix d'un formalisme de représentation (première tâche), de définir les traitements nécessaires pour permettre l'aide à la décision. Il est donc nécessaire de construire un algorithme de prévision des ventes basé sur la représentation hybride proposée, c'est-à-dire intégrant les connaissances issues des données et les connaissances expertes. Un point essentiel concernant ces mécanismes d'exploitation sera leur capacité à préserver, autant que possible, de manière explicite l'information sur l'incertitude associée aux données et aux connaissances manipulées. Cette information sur l'incertitude sera restituée à l'utilisateur lui permettant ainsi de prendre une décision en étant pleinement informé.
- La dernière tâche consiste à utiliser les données de plusieurs points de vente. Une des limitations potentielles de l'approche envisagée est le manque éventuel de données concernant les ventes, ce qui induirait une incertitude relativement élevée et donc des résultats difficiles à exploiter en aide à la décision. C'est pourquoi nous pensons nous intéresser à la possibilité d'exploiter, pour un point de vente donné, à l'exploitation des données d'autres points de vente. Toutefois, afin que cette utilisation de données « externes » soit pertinente, il convient de vérifier que les points de vente concernés sont similaires.

Nous proposons donc de construire une mesure de similarité entre les points de ventes. Cette mesure reposera à la fois sur le profil du point de vente (dont les descripteurs sont à déterminer avec les experts) et sur le profil sociologique de la population associée au point de vente (issue des bases de données de l'INSEE). Cette mesure de similarité, comprise entre 0 et 1, permettra de qualifier les données issues d'autres points de vente. Si la similarité est maximale (valeur 1) alors les données pourront être exploitées comme si elles étaient issues du point de vente concerné. Si la similarité est faible alors les données ne devront pas être prises en compte.

B. Problématique

Afin d'atteindre les objectifs fixés, nous allons exploiter une base de données statistique issue des données du Système d'Informations de l'entreprise.

L'entreprise peut posséder plusieurs points de ventes. Les données que nous allons extraire seront réparties en fonction du type de produit, du vendeur associé, de la date de vente, du

code de l'agence, de la quantité de type de produits vendues, de la marge associée au type de produit vendu.

Généralement les méthodes de prévisions s'appuient sur une base de données statistique. Cependant, lorsqu'on souhaite réaliser des prévisions concernant un produit particulier pour une petite unité, les données sont rarement suffisantes pour fournir des statistiques significatives (par exemple les ventes de forfait Internet d'un point de vente de télécommunication d'une petite ville).

Ces imperfections concernant les données peuvent avoir deux causes principales. Tout d'abord, en ce qui concerne la période d'étude, celle-ci doit comporter des extractions de données sur une période minimum de quatre ans mais les données sont archivées sur des serveurs cryptés au bout de trois ans. D'autre part, en ce qui concerne la taille de l'échantillon périodique (par exemple le nombre de ventes par mois), l'idéal serait de disposer d'échantillons contenant au minimum une quarantaine d'observations par période pour obtenir des statistiques fiables, ce qui n'est pas le cas lorsque notre intérêt est un « petit » point de vente pour un produit particulier.

C'est pourquoi pour consolider les données statistiques nécessaires à la mise en œuvre d'une démarche de prévision, il faut intégrer les données expertes dans notre module de prévision. Ces données expertes sont issues des différents acteurs de l'entreprise répartie sur plusieurs points de vente.

Cependant l'hybridation dans une même base de connaissances d'éléments issus d'une base de données multidimensionnelle et d'éléments issus de l'expertise constitue une piste d'investigation peu abordée. Nous avons démontré récemment [2] l'intérêt pratique de telles approches. Le compromis à établir se situe entre la capacité descriptive du modèle (quelle représentation de l'incertitude inhérente aux données et aux experts) et la complexité, au sens algorithmique du terme, des mécanismes d'inférence à mettre en œuvre. Il s'agit donc, de préciser à quel point d'équilibre il est pertinent de se situer en vue d'une application logicielle interactive, c'est-à-dire procurant des temps de réponse suffisamment courts. De plus, la fusion d'informations issues de bases de données différentes et qualifiées selon une similarité entre les environnements de ces bases de données est, à notre connaissance, un axe scientifique original. Différentes techniques de fusions des données ont été proposées dans la littérature et dépendent naturellement du modèle d'incertitude choisi. De manière analogue, de nombreuses mesures de similarité ont été étudiées, en particulier dans le domaine du Raisonnement à Partir de Cas (ou Case Based Reasoning, CBR). Il s'agit ici d'utiliser ce corpus scientifique, de l'adapter à la problématique particulière et, surtout, d'intégrer de manière cohérente ces calculs de similarité avec la fusion des informations. Là encore, une étude de la complexité des traitements associés devra être réalisée afin de sélectionner l'approche la plus pertinente.

Les deux dernières tâches de notre objectif, qui concerne la mise en place d'une aide à la décision interactive, nécessitent la représentation des événements (événements externes subis ou décisions) pour permettre l'identification de leur impact sur les mesures de l'activité. Il s'agit donc de mémoriser, en lien avec

une base de données multidimensionnelle classique, les informations concernant les événements et les décisions de manière en en évaluer l'impact. Là encore, l'utilisation des données historisées et des connaissances expertes permet d'anticiper l'impact d'une décision potentielle et ainsi de proposer une aide de plus en plus sophistiquée au décideur.

II. DEMARCHE PROPOSEE

Pour répondre à la problématique énoncée, nous proposons une démarche destinée à effectuer une prévision la plus précise possible en tenant compte des contraintes identifiées précédemment. Cette démarche est décrite dans la figure 1.

La démarche s'appuie dans un premier temps sur l'exploitation d'une base de données statistique pour extraire les données correspondant au problème étudié (type de produit, entité concernée, acteur ciblé, ...). Une fois les statistiques extraites, il est nécessaire de faire un choix. Si les données sont en quantité suffisante, une première prévision peut être réalisée. Dans le cas contraire, il est nécessaire de compléter ces données en utilisant l'expertise et les mécanismes de similarité pour retrouver dans la base de données, les entités, acteurs ou produits similaires. Les données similaires sont alors extraites puis utilisées pour obtenir une prévision suffisamment fiable du point de vue des données utilisées.

Lorsque cette première prévision a été réalisée, il est nécessaire de vérifier la qualité des prévisions obtenues. Lorsque certaines valeurs prévues sont identifiées comme étant hors tolérance, il est nécessaire de déterminer la cause des variations constatées. Grâce à l'expertise, les événements perturbateurs sont identifiés (fermeture d'une agence pour travaux, sortie d'un nouveau produit, ...) et les données initiales sont corrigées pour tenir compte de l'impact de ces phénomènes. Pour finir, une nouvelle prévision est réalisée.

III. HYBRIDATION DES DONNEES

Comme nous venons de le voir, le cœur du mécanisme que nous proposons s'appuie sur une information statistique consolidée par la connaissance experte. Il est donc nécessaire de déployer des moyens permettant d'hybrider ces deux types de connaissances pour obtenir des connaissances plus nombreuses et/ou plus pertinentes. Nous allons donc présenter dans cette partie les méthodes permettant de formaliser les connaissances statistiques pour effectuer une prévision et les mécanismes permettant de formaliser la connaissance experte et de l'hybrider aux connaissances statistiques, nous devons consolider les statistiques par des connaissances expertes.

A. Formalisation des connaissances statistiques

Selon la démarche proposée, nous allons en premier lieu utiliser les méthodes de prévisions classiques alimentées par une base de données statistique. Nous décrivons dans cette partie différentes étapes permettant de construire la base de données en collectant les données pertinentes. Par la suite nous présentons le modèle de prévision choisi et son implémentation.

1) Collecte de données et création de la base de données statistique

Comme évoqué précédemment, le module de prévision ciblé doit s'intégrer dans le logiciel « Swifteo ». Ce logiciel possède une base de données qui lui permet de traiter les données envoyées par l'entreprise. Pour le module de prévision, nous utilisons une base de données paramétrée selon l'entreprise. Un processus de transformation permet de conserver uniquement les données utiles : la date de vente, le code vendeur, le code agence, le type de produit vendu, la quantité et la marge du type de produit vendu.

Cette démarche est décrite dans la figure 2.

Une fois les données extraites dans notre base de données, nous pouvons commencer à utiliser les méthodes de prévisions classiques.

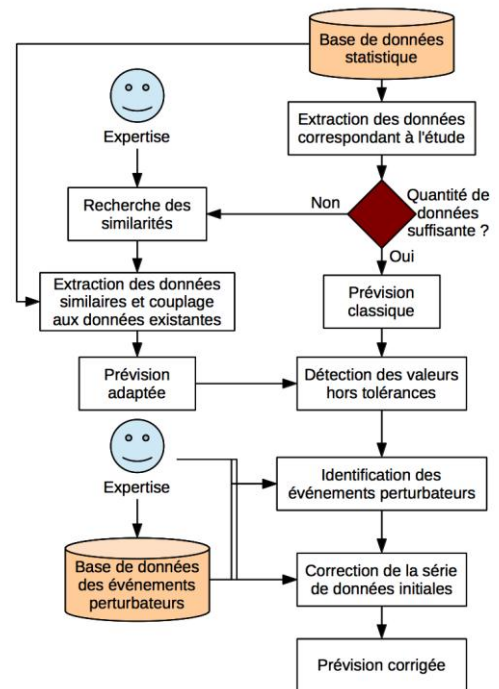


Fig. 1. : Démarche de prévision proposée

2) Méthode extrapolative sur une série chronologique

Cette méthode utilise les valeurs passées de la variable elle-même. Seul le passé de la variable est utilisé afin de prévoir son évolution future sans apport d'information extérieure. Ces méthodes sont dites endogènes car la résolution ne fait appel à aucune autre source de données. La prévision classique est souvent faite sous l'hypothèse de stabilité du système en vue de dégager un scénario tendanciel.

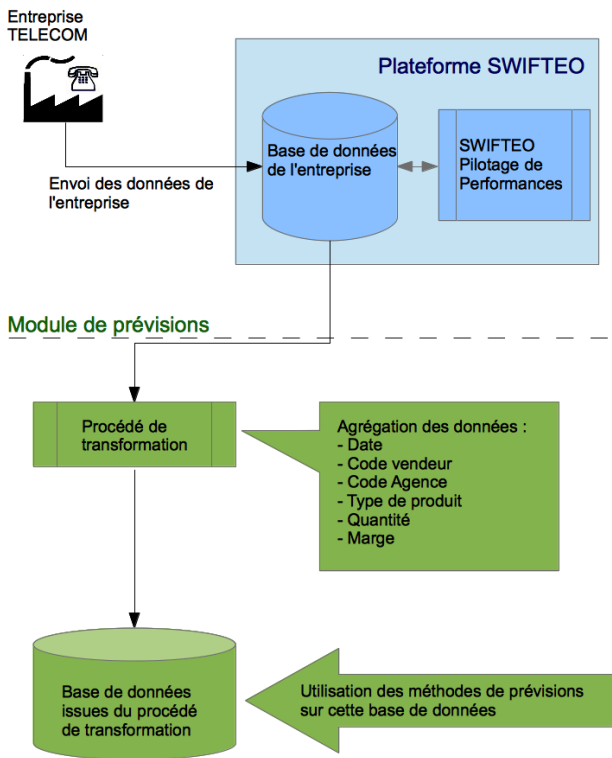


Fig. 2. : Création de la base de données statistique

a) Méthodologie de la prévision sur une série chronologique

Les différentes composantes d'une série chronologique sont [1]:

- La tendance T_t qui correspond à une variation observée sur une longue période et décrit les évolutions « lourdes » du phénomène.
- La cyclicité C_t qui représente des variations liées à la conjoncture, notamment aux périodes des phases du cycle économique (prospérité, crise, dépressions, reprise). Cette composante est assimilée à la tendance car elle concerne généralement des variations sur le long terme.
- La saisonnalité S_t qui représente le caractère répétitif à intervalles réguliers des variations des données autour de la tendance. Elle se caractérise par une périodicité p qui varie selon le domaine d'activité. Les causes de ces variations sont souvent multiples (factures climatiques, modes de vie, coutumes, ...) et se produisent généralement à la même date

Les fluctuations aléatoires R_t qui sont des variations généralement dues à un grand nombre de petites causes ou à des événements à caractère exceptionnel ou accidentel. Elles représentent, dans l'évolution de la série, la part qui ne peut être pris en compte par les autres composantes. C'est sur ce type de phénomènes, généralement considérés comme imprévisibles, que nous souhaitons agir en intégrant l'expertise.

Pour réaliser des prévisions plus précises, nous proposons une démarche destinée à effectuer les calculs nécessaires en

prenant en compte toutes les composantes d'une série chronologique et les modèles de prévisions comme illustrée sur la figure 3.

Dans un premier temps, nous désaisonnalisons la série chronologique afin d'obtenir une série désaisonnalisée. Cette dernière nous permettra de calculer les composantes de la série chronologique en fonction des deux modèles de prévisions classiques. De ces tendances, nous réalisons des projections auxquelles sont appliqués les coefficients saisonniers.

Il est possible de distinguer deux modèles de prévision permettant de prendre en compte les caractéristiques spécifiques à chaque série de données [1] :

- Le modèle additif est le plus simple. Il consiste à additionner les trois composantes T_t , S_t , R_t . Pour obtenir une série désaisonnalisée, on calcule pour chaque période l'écart e_t entre une moyenne glissante centrée et la série brute. On calcule ensuite la moyenne des écarts et enfin il faut normer les coefficients saisonniers. Il faut ensuite déterminer la série désaisonnalisée en retirant les coefficients saisonniers à la série brute. A partir de cette série désaisonnalisée, nous pouvons ensuite calculer la tendance et projeter cette tendance vers le futur. Enfin il ne nous reste plus que pour calculer une prévision en tenant compte de la saisonnalité à ajouter ce coefficient à notre équation de tendance.
- Le modèle multiplicatif est utilisé lorsque la saisonnalité est proportionnelle à la tendance. Il consiste à multiplier la saisonnalité avec la tendance et à y ajouter les fluctuations aléatoires pour obtenir la prévision. La méthode est la même que pour le modèle additif mais au lieu de déterminer des écarts entre la série brute et la moyenne glissante centrée, il convient de calculer les rapports entre ces deux valeurs. Ensuite il est nécessaire de faire la moyenne entre les rapports et de normer les coefficients saisonniers. Pour finir, la tendance est calculée puis projeté et on la multiplie aux coefficients saisonniers pour obtenir la prévision.

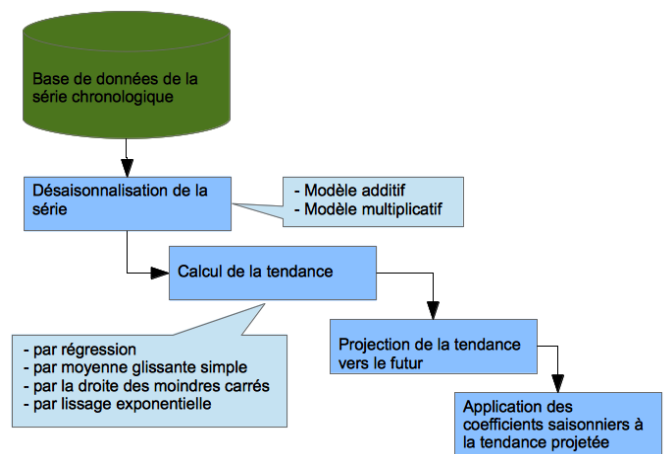


Fig. 3. : Workflow du module de prévisions

b) *Mise en prototype d'un module de prévisions de ventes à partir de la base de données statistiques*

Comme énoncé plus tôt, nous avons développé un module de prévision se basée uniquement sur la base de données de la série chronologique. Cette base de données (figure 3) correspond à la base de données issue du procédé de transformation (figure 2).

Nous avons dans un premier temps développé un démonstrateur ne tenant pas compte des données expertes. Ce démonstrateur nous servira de base pour valider les différents modèles de prévision. Notre étude consiste à intégrer les données expertes au démonstrateur pour consolider les données statistiques ; de ce fait, nous devons en priorité disposer d'un modèle statistique fiable pour ensuite prendre en compte la quantité de données et ajouter l'expertise dans les cas critiques.

B. *Elicitation des avis d'experts*

Comme nous l'avons évoqué précédemment, cette partie présente les moyens basés sur les fonctions de croyances mis en œuvre pour éliciter les avis d'experts et les intégrer au modèle de prévision.

1) *Modèle des Croyances Transférables*

Le MCT interprète une fonction de croyance comme la représentation de l'opinion d'un agent rationnel en se basant uniquement sur la croyance ou l'état de connaissance de l'individu, même si celle-ci est imprécise ou incomplète [3].

a) *Niveau crédal et niveau pignistique*

Le modèle des croyances transférables repose sur l'hypothèse que raisonnement dans l'incertain (niveau crédal) et prise de décision (niveau pignistique) sont deux tâches cognitives de nature différentes :

- Le niveau crédal correspond à la représentation et à la manipulation des états de croyances sans simplification (sans recours par exemple à l'utilisation de distribution équiprobable).
- Le niveau pignistique correspond à la prise de décision. Il s'agit alors de transformer les mesures subjectives de croyances non probabilistes en une mesure de probabilité dite pignistique (*BetP*). La distribution obtenue est équivalente à une distribution de probabilités.

La transformation pignistique est irréversible mais n'intervient que lors de la prise de décision et ne modifie pas le niveau crédal.

2) *Distribution de masse de croyance élémentaire (bba)*

La modélisation par les fonctions de croyance repose sur l'affectation d'une mesure subjective non-additive des chances de réalisation d'un événement. Contrairement à la théorie de Dempster-Shafer (DST) dont le MCT est originaire, cette représentation existe sans référence à une probabilité ou une famille de probabilités sous-jacentes.

La distribution de masse de croyance élémentaire m traduit un état de connaissance partiel sur la valeur d'une variable à valeurs dans Ω , le référentiel de définition nommé cadre de discernement. Soient w_i les éléments atomiques (singletons) et A_i des sous-ensembles de Ω , nous avons :

$$\Omega = \{w_1, w_2, w_3, \dots, w_n\} \text{ et } \forall i, A_i \subseteq \Omega$$

La distribution de masse de croyance est une distribution d'une masse unitaire sur les parties de Ω (power set en anglais), c'est à dire sur les sous-ensembles du cadre de discernement (pour un ensemble de taille n , cela peut représenter 2^n éléments. La notation est $P(\Omega)$ ou 2^Ω . La possibilité d'affecter des masses aux hypothèses composées, et donc de travailler sur 2^Ω plutôt que sur Ω , constitue un des avantages de cette théorie. C'est ce qui permet de représenter des connaissances imparfaites.

$m^\Omega(A)$ correspond à la « part » de croyance allouée par une source S à l'hypothèse $w \in A$ et à aucune hypothèse plus restrictive. m^Ω est alors donnée par la formule suivante :

$$m^\Omega = \left\{ \begin{array}{l} 2^\Omega \rightarrow [0,1] \\ \sum_{A \in 2^\Omega} m^\Omega(A) = 1 \end{array} \right\}$$

$$2^\Omega = \{A \subseteq \Omega\} = \{\emptyset, \{w_1\}, \dots, \{w_n\}, \{w_1 \cup w_2\}, \dots, \{\Omega\}\}$$

a) *Mesure de Crédibilité et de Plausibilité*

A partir d'une distribution de masse qui représente une croyance stricte que l'on ne peut préciser, il est possible d'obtenir de manière biunivoque des représentations équivalentes mais différentes sémantiquement. Les plus importantes sont les fonctions duales de crédibilité (notée Cr) et de plausibilité (notée Pl).

La crédibilité Cr traduit la croyance minimum en une information (vraisemblance minimale). Elle est donnée par la formule suivante :

$$\forall A \subseteq \Omega, Cr(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B)$$

La plausibilité Pl traduit la croyance maximum en une information (vraisemblance maximale). Il s'agit de toute la croyance qui ne contredit pas A . Elle est donnée par la formule suivante :

$$\forall A \subseteq \Omega, Pl(A) = \sum_{(B \cap A) \neq \emptyset} m(B)$$

3) *Méthode par élimination des hypothèses*

Nous proposons une méthode qui vise à utiliser un langage plus naturel pour exprimer les avis d'experts. Il existe plusieurs travaux étudiant ce problème. Wong et Linglas [4] ont proposé deux opérateurs binaires pour représenter les relations de préférence et d'indifférence. Yaghlane [5] a développé cette méthode en ajoutant une notion de contrainte permettant de modéliser la fonction de croyance la moins informative respectant ainsi le Principe du Minimum d'Information [6]. Bryson [7] a proposé une autre approche appelée « processus de discrimination qualitative ». Ce processus est composé de plusieurs étapes permettant de trier de manière itérative les propositions selon trois degrés de certitude puis de leur assigner une masse de croyance en comparant les hypothèses. L'étape finale de ce processus consiste à rendre cohérente la distribution de masses de croyance résultante. Toutes ces

méthodes utilisent un principe de comparaison des hypothèses deux à deux pour éliciter la connaissance experte. Nous souhaitons pour notre part développer une démarche permettant des comparaisons plus générales.

La méthode que nous proposons est proche de celle développée dans [7] mais elle est volontairement simplifiée pour faciliter le processus d'élicitation. Elle se présente sous la forme d'un processus itératif d'élimination des hypothèses du cadre de discernement associé à une échelle de regret pour quantifier la part de masse affectée à chaque hypothèse. L'objectif est de construire directement une distribution de masses de croyance consonante, c'est-à-dire que les éléments focaux de la distribution de masses sont des ensembles imbriqués. L'expert doit répondre à deux questions successives qui sont répétées à chaque itération.

La première étape consiste à poser à l'expert la question « Parmi toutes ces hypothèses, laquelle vous semble la moins vraisemblable? » en commençant par proposer le cadre de discernement complet. L'hypothèse choisie est alors éliminée des choix possibles pour la seconde itération. La deuxième étape consiste à demander à l'expert d'exprimer la difficulté qu'il a eu à faire son choix. Par exemple, il est possible d'imaginer une question de type « Sur une échelle de 0 à 9, évaluez la difficulté que vous avez eu à choisir l'hypothèse la moins vraisemblable (0 étant le niveau très facile, 9 le niveau très difficile) ». Cette deuxième question peut aussi être vue comme une évaluation du regret qu'il a à éliminer une hypothèse de la liste des hypothèses possibles. Lors de la seconde itération, l'expert doit répondre aux mêmes questions mais l'hypothèse choisie lors de la précédente itération ne fait plus partie de la liste des hypothèses disponibles. Ces deux opérations sont répétées jusqu'à ce qu'il ne reste plus qu'une hypothèse disponible ou jusqu'à ce que l'expert ne puisse plus départager les hypothèses.

La première question permet donc de choisir le chemin dans l'arbre de décision (figure 4) et la deuxième question permet d'affecter une masse aux hypothèses concernées en fonction du regret ressenti par l'expert. La masse est ensuite répartie, proportionnellement au regret exprimé par l'expert, entre l'ensemble incluant l'hypothèse choisie et l'ensemble sans cette hypothèse. Lorsqu'on passe à l'itération suivante, seule la masse du sous-ensemble concerné est remise en jeu en sachant que la masse initialement affectée à l'ensemble représentant l'ignorance totale est égale à un. L'exemple suivant illustre le principe de fonctionnement de cette méthode.

Exemple :

L'expert doit s'exprimer sur le cadre de discernement $\Omega = \{c_1, c_2, c_3\}$. La figure 4 Fig. 4. présente l'arbre des décisions possibles. Le tableau 1 présente le fonctionnement de la méthode du point de vue de l'expert et du point de vue du système.

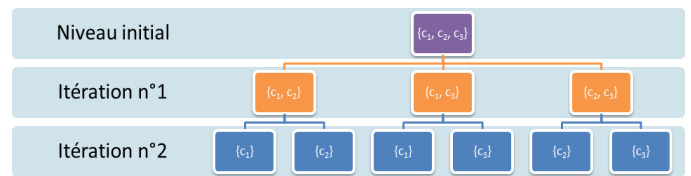


Fig. 4. : Arbre de décision

Cette élicitation a pour principal avantage d'utiliser des questions simples auxquelles l'expert peut répondre rapidement. Elle est également facile à informatiser. Cette méthode peut être améliorée en intégrant des notions complémentaires telles que l'hésitation (par exemple pour la première itération de l'exemple du tableau 1, l'expert n'arrive pas à déterminer s'il doit éliminer c_2 ou c_3). Cette notion d'hésitation pourrait être prise en compte en partageant la masse à affecter au sous-ensemble de cardinalité plus faible en deux et à suivre le processus avec deux chemins de décisions possibles. Ce type de mécanisme complémentaire n'est pas encore utilisé mais est une piste intéressante en perspective de notre étude.

C. Intégration de l'expertise dans le système de prévision

Maintenant que nous disposons d'une connaissance experte formalisée, il est nécessaire d'intégrer cette connaissance dans le système de prévision. Comme nous l'avons vu précédemment, le modèle de prévision que nous proposons présente deux points d'entrée pour l'introduction de l'expertise dans le système de prévision. Le premier point d'entrée se situe dans l'étape de recherche de similarité pour pallier le manque de données. Nous traiterons cet aspect dans la partie suivante.

Le second point d'entrée pour la connaissance experte dans le système de prévision est situé au niveau de la détection des événements et la correction de la série de données pour intégrer l'impact de ces événements. Deux cas de figure peuvent se présenter. L'expert peut constater par lui-même la présence d'un de ces phénomènes. Par exemple, des travaux d'aménagement de la boutique ont eu lieu pendant une semaine entraînant la fermeture de certains rayons et donc une baisse des ventes. Il lui est alors possible de signaler cet événement et d'en évaluer l'impact supposé sur les ventes. Cette évaluation, formalisée à l'aide des fonctions de croyance, est alors utilisée pour corriger la série de données utilisée pour la prévision en compensant artificiellement la baisse (ou la hausse) des ventes pour que l'impact des travaux ne perturbent pas les prévisions futures.

Le deuxième cas de figure se présente suite à la détection par le système d'une différence importante entre la prévision et la réalité. Un processus visant à déterminer les causes de cette erreur est alors lancé pour déterminer si un événement (ou une combinaison d'événements) a pu causer cet écart. Par exemple, le lancement d'un produit qui a eu un succès imprévu lié à un phénomène de « buzz » sur Internet. Dans ce cas, l'expert doit déterminer si l'événement est ponctuel ou s'il risque d'impacter les prévisions suivantes et adapter le modèle ou la série de données en conséquence.

Point de vue « Expert » (interface)	Point de vue « Système » (algorithme)																		
<p>Itération n°1 :</p> <ul style="list-style-type: none"> - Parmi les hypothèses suivantes, laquelle vous semble la moins vraisemblable ? c1, c2, c3 - L'expert choisit c3. - Sur une échelle de 0 à 9, évaluez la difficulté que vous avez eu à définir l'hypothèse la moins vraisemblable (0 étant le niveau très facile, 9 le niveau très difficile) : - L'expert choisit le niveau 1 sur 9. 	<p>Itération n°1 :</p> <ul style="list-style-type: none"> - Masse disponible : 1 - Élimination de c3. Il faut donc répartir la masse disponible entre $m(\{c1, c2\})$ et $m(\{c1, c2, c3\})$. - On affecte 1/10ème (niveau 1 sur dix niveaux possibles) de la masse disponible à l'ensemble avec l'hypothèse choisie et le reste à celui sans cette hypothèse. $m(\{c1, c2\}) = 0,9 \text{ et } m(\{\Omega\}) = 0,1$																		
<p>Itération n°2 :</p> <ul style="list-style-type: none"> - Parmi les hypothèses suivantes, laquelle vous semble la moins vraisemblable ? c1, c2 - L'expert choisit c2. - Sur une échelle de 0 à 9, évaluez la difficulté que vous avez eu à définir l'hypothèse la moins vraisemblable (0 étant le niveau très facile, 9 le niveau très difficile) : - L'expert choisit le niveau 6. 	<p>Itération n°2 :</p> <ul style="list-style-type: none"> - Masse disponible : 0,9 - Élimination de c2. Il faut donc répartir la masse disponible entre $m(\{c1\})$ et $m(\{c1, c2\})$. $m(\{c1\}) = 0,9 \times 0,4 = 0,36$ $m(\{c1, c2\}) = 0,9 \times 0,6 = 0,54$ <ul style="list-style-type: none"> - On affecte 6/10ème de la masse disponible à l'ensemble avec l'hypothèse choisie et le reste à celui sans cette hypothèse. 																		
<p>Fin : C'est terminé car il ne reste plus que c1.</p>	<p>La distribution de masse résultante est donc :</p> <table border="1"> <thead> <tr> <th></th> <th>{∅}</th> <th>{c1}</th> <th>{c2}</th> <th>{c1, c2}</th> <th>{c3}</th> <th>{c1, c3}</th> <th>{c2, c3}</th> <th>{Ω}</th> </tr> </thead> <tbody> <tr> <th>m</th> <td>0</td> <td>0,36</td> <td>0</td> <td>0,54</td> <td>0</td> <td>0</td> <td>0</td> <td>0,1</td> </tr> </tbody> </table>		{∅}	{c1}	{c2}	{c1, c2}	{c3}	{c1, c3}	{c2, c3}	{Ω}	m	0	0,36	0	0,54	0	0	0	0,1
	{∅}	{c1}	{c2}	{c1, c2}	{c3}	{c1, c3}	{c2, c3}	{Ω}											
m	0	0,36	0	0,54	0	0	0	0,1											

TABLEAU I. ILLUSTRATION DE LA METHODE D'ELICITATION DES AVIS D'EXPERT

Il est important de noter que tous les événements perturbateurs et leurs impacts sont capitalisés pour pouvoir détecter plus facilement les prochains phénomènes similaires et pour aider à en évaluer l'impact.

IV. MESURE DE SIMILARITE

Parmi les premiers objectifs opérationnels que nous avons identifiés comme étant essentiels à l'application des mécanismes développés dans ce projet de recherche, nous nous sommes intéressés à la comparaison des points de vente. En effet, il nous semble important d'intégrer les informations issues d'un point de vente dans la connaissance utilisée pour aider la décision d'un autre point de vente dans l'objectif de consolider la base de connaissances et ainsi de prendre une décision plus pertinente. Cependant, cette intégration nécessite de mesurer les différences et les ressemblances entre les points de vente pour pondérer la connaissance que l'on souhaite intégrer en fonction de la confiance du système dans ces nouvelles informations. Il convient donc de mettre en place une mesure de similarité permettant de comparer de manière chiffrée les points de vente.

L'objectif d'une mesure de similarité est de représenter numériquement les informations de ressemblances entre des entités (dans notre cas, des points de vente) [8]. Elle consiste à mesurer la proximité entre les différents éléments du domaine

généralement, mais pas nécessairement, au moyen d'une distance (au sens mathématique) [9].

La vision généralement admise considère la similarité et la distance comme étant deux fonctions duales permettant d'exprimer la proximité entre des éléments d'information [10]. Des entités ayant une ressemblance forte seront caractérisées par une mesure de similarité proche de 1 et une distance proche de 0. Inversement, des entités dissemblables auront une mesure de similarité proche de 0 et une distance proche de 1 (dans un cadre « normalisé » où toutes les mesures sont comprises entre 0 et 1).

Il est possible de distinguer plusieurs propriétés qui permettent de caractériser une mesure de similarité (et donc de distance) [11] telles que la réflexivité (une mesure de similarité est dite réflexive dans le cas où lorsqu'on compare deux entités identiques, la mesure vaut 1), la symétrie (une mesure de similarité est dite symétrique dans le cas où lorsque l'on compare l'entité β à l'entité γ , le résultat est le même si l'on compare γ à β) ou encore l'inégalité triangulaire et la monotonie.

Notre type de mesure de similarité se base sur les représentations incertaines. Celles-ci sont principalement basées sur la théorie des ensembles et ses extensions. Par conséquent, les mesures de similarités basées sur l'approche ensembliste sont transposables à ces différents cadres de représentation. Ainsi il existe des mesures de similarité

applicables à la théorie des possibilités [8] et la théorie des fonctions de croyance [12]. Cependant ces mesures permettent de déterminer une similarité locale (c'est-à-dire d'attribut à attribut), alors que notre objectif est de déterminer une similarité globale (c'est-à-dire de comparer des ensembles d'attributs)[10]. La mesure de similarité proposée par Beler [11] nous apparaît comme la plus pertinente car elle permet de faire une mesure globale dans un cadre de représentation tenant compte des incertitudes.

A l'heure actuelle, la définition des critères de comparaison entre les points de ventes est notre principal sujet d'étude. En effet, il existe de nombreux paramètres permettant de caractériser et donc de comparer des points de vente. Nous avons identifié trois catégories de critères permettant de construire une typologie :

- les critères concernant le point de vente d'un point de vue physique, c'est-à-dire la localisation (centre-ville, galerie commerciale, ...), la superficie, l'agencement, la fréquentation, la concurrence à proximité, ...
- les critères concernant la relation avec la clientèle du point de vente, c'est-à-dire la publicité, le type de clientèle (catégorie socio-professionnelle, âge, ...), l'accueil des clients, ...
- les critères concernant l'offre de produits proposée par le point de vente (liés à sa spécialité).

La principale difficulté réside dans le fait que certains paramètres ne peuvent être évalués que de manière subjective. En effet, s'il est facile d'estimer et de comparer la superficie d'un point de vente de manière objective, les critères cherchant à caractériser les relations avec la clientèle sont souvent plus subjectif et donc présente plus de risque d'erreur. Les fonctions de croyance sont donc un outil pertinent car elles permettent d'introduire la notion d'incertitude épistémique, c'est-à-dire l'incertitude liée à l'incomplétude et à l'imprécision qui seront ensuite répercutée sur la mesure de similarité et donc sur la pondération des données ajoutées au modèle de prévision pour le consolider.

V. CONCLUSION

Cette communication nous a permis d'introduire un processus de prévision s'appuyant sur la connaissance experte et sur les mécanismes de similarité pour consolider les données statistiques généralement utilisées pour ce type d'application. Nous avons également introduit une méthode d'élicitation des avis d'experts permettant de formaliser les connaissances issues de l'expertise dans le cadre de formalisation des fonctions de croyance. Nos objectifs actuels sont liés au

développement des aspects correspondant à la similarité pour pouvoir intégrer de manière efficace les données issues de points de ventes similaires. Les travaux présentés ici sont le préambule à un programme de recherche qui s'étendra sur trois ans. Ce projet a pour objectif d'obtenir une plateforme logicielle d'aide à la décision et au pilotage des performances intégrant la connaissance experte et la gestion des incertitudes.

VI. REFERENCES BIBLIOGRAPHIQUES

- [1] Bourbonnais R. et Usunier J.-C.. *Pratique de la Prévision des Ventes : Conception de Systèmes*, Economica, 1997.
- [2] Villeneuve E.. *Hybridation des retours d'expérience statistique et cognitif pour l'évaluation des risques : Application à la déconstruction des aéronaves*. Thèse de doctorat, Institut National Polytechnique de Toulouse, 2012.
- [3] Ramasso E.. *Reconnaissance de séquences d'états par le Modèle des Croyances Transférables : Application à l'analyse de vidéos d'athlétisme*. Thèse de doctorat, Université Joseph Fourier, Grenoble, 2007.
- [4] Wong S. K. M. et Lingras. P.. *Representation of qualitative user preference by quantitative belief functions*. *IEEE Transactions on Knowledge and Data Engineering*, volume 6: pages 72–78, 1994.
- [5] Yaghlane A., Denoeux T. et Mellouli K.. *Elicitation of expert opinions for constructing belief functions*. In *Uncertainty and intelligent information systems*, pages 75–88. World Scientific, 2008.
- [6] Hsia, Y. T.. *Characterizing belief with minimum commitment*. *International Joint Conference on Artificial Intelligence (IJCAI'91)*, volume 2: pages 1184–1189, 1991.
- [7] Bryson, N. K.-M. et Mobolurin A.. *A process for generating quantitative belief functions*. *European Journal of Operational Research*, volume 115: pages 624–633, 1999.
- [8] Ruet M.. *Capitalisation et réutilisation d'expériences dans un contexte multi acteur*. Thèse de doctorat, Institut National Polytechnique de Toulouse, 2002.
- [9] Bisson G.. *Why and how to define a similarity measure for object based representation systems*. *2nd International Conference on Building and Sharing Very Large-Scale Knowledge Bases*, pages 236–246, 1995.
- [10] Bergmann R.. *Experience Management, Lecture Notes in Computer Science*. Springer, 2002.
- [11] Béler C.. *Modélisation générique d'un retour d'expérience cognitif*. Thèse de doctorat, Institut National Polytechnique de Toulouse, 2008.
- [12] Jousselme A.-L., Grenier D., et Bossé E.. *A new distance between two bodies of evidence*. *Information Fusion*, pages 91–101, 2001.