

Panorama de travaux autour de l'intégration de données spatio-temporelles dans les hypercubes

Anne Tchounikine, Maryvonne Miquel, Robert Laurini,
Taher Ahmed, Sandro Bimonte, Virginie Baillet

LIRIS – UMR CNRS 5205

Laboratoire d'InfoRmatique en Images et Systèmes d'information
Bâtiment Blaise Pascal, INSA, 7 avenue Capelle, 69621 Villeurbanne Cedex, France
prenom.nom@insa-lyon.fr

Résumé : Cet article présente un panorama des différents travaux qui sont effectués dans notre équipe autour de l'intégration des données spatio-temporelles dans les entrepôts de données et les hypercubes. Nous nous sommes plus particulièrement intéressés à la prise en compte des évolutions dans les dimensions spatiales, à la modélisation multidimensionnelle de données continues et à la conception d'interface de navigation dans des données multidimensionnelles spatio-temporelles. Ces quelques résultats nous permettent d'avancer de nouvelles perspectives dans la modélisation et la visualisation de données géo-spatiales dans les hypercubes.

1. Introduction

La modélisation multidimensionnelle et les technologies OLAP (On-Line Analytical Processing) permettent de réaliser une analyse rapide, intuitive et facile dans de grands volume de données [Inmon, 1992]. Ces données sont modélisées sous la forme d'hypercubes [Agrawal *et al*, 997] [Cabibbo et Torlone, 1998] dans lesquels les dimensions constituent des axes d'analyse indépendants, et les sujets d'analyse, ou faits, sont caractérisés par des mesures qui sont pré-calculées à l'aide de fonctions d'agrégations selon les différentes granularités définies par le schéma hiérarchique de chaque dimension. La figure 1 représente le modèle multidimensionnel qui peut être utilisé dans le cas d'une analyse de ventes. Les opérateurs OLAP (*Roll-up*, *Drill-down*, *Slice*, *Rotate* etc.) permettent de visualiser les mesures pour des ensembles de membres et des niveaux de granularité sélectionnés par l'utilisateur. Une requête OLAP dans un exemple d'analyse des ventes peut être la suivante : « Quel est le volume et le montant total des produits de type « ordinateur » vendus en 1999 dans le magasin M1 ? ».

Panorama de travaux autour de l'intégration de données spatio-temporelles

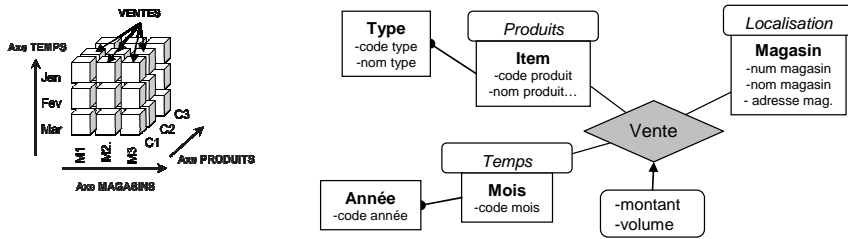




Figure 1 : Modèle multidimensionnel

Dans la majorité des cas d'applications, les modèles multidimensionnels intègrent une dimension temporelle, l'un des objectifs majeurs de l'analyse décisionnelle étant de mener des analyses d'impact et d'évolution sur de longues périodes. Un autre type d'information souvent utile à l'analyse décisionnelle est l'information spatiale. En effet, on estime qu'environ 80% des données transactionnelles présentent une composante spatiale [Franklin, 1992]. Très souvent encapsulée dans un champ de type « adresse » dans les bases OLTP (adresse d'un magasin, d'un client, d'un employé...), l'information doit alors être transformée, segmentée au cours du processus d'alimentation, de façon à faire apparaître une hiérarchie spatiale (par exemple : magasin < département < région). Les modèles multidimensionnels ainsi construits permettront d'analyser une mesure numérique en fonction de la localisation géographique des membres de dimensions [Bédard *et al*, 2001], [Marchand *et al*, 2003]. Ainsi la mesure « volume des ventes » pourra être analysée en fonction de la localisation du magasin dans notre exemple (figure 2). Dans cette figure, les pictogrammes  et  représentent la forme géométrique associée aux membres de la dimension spatiale. Les magasins sont représentés par un point, les départements et les régions par des polygones.

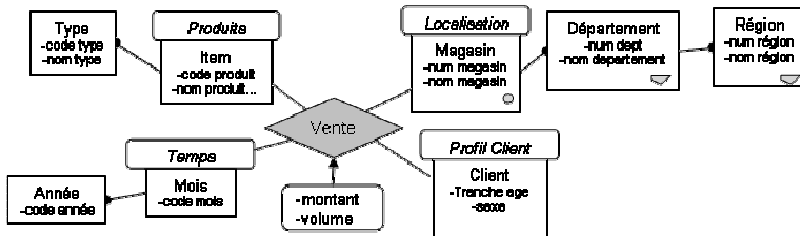


Figure 2 : Introduction d'une dimension spatiale

Une autre façon d'introduire la composante spatiale dans les entrepôts est de la transformer non pas en axe d'analyse, i.e. en dimension, mais en sujet d'analyse, i.e. en fait [Bédard *et al*, 2001], [Marchand *et al*, 2003], [Fidalgo *et al*, 2004] [Pedersen et Tryfona, 2001] [Stefanovic *et al*., 2000], [Malinowski et Zimanyi, 2004], [Jensen *et al*, 2004]. Dans ce cas, ce ne sont plus les ventes que l'on souhaitera analyser en fonction d'une région, mais les régions (ou les villes, ou les zones...) que l'on souhaite analyser en fonction des ventes

(figure 3). La mesure spatiale est alors le résultat d'une requête OLAP portant sur des dimensions spatiales ou non. Un exemple simple de requête est « Quelles sont les régions et leur superficie où se sont vendus des produits de type « informatique » en 1999 ? ».

Les travaux portant sur l'introduction de la composante spatiale dans les modèles multidimensionnels ont mené à la définition du concept de SOLAP : **Spatial On Line Analytical Processing** [Bédard, 1997]. Les domaines d'application du SOLAP sont nombreux : aménagement du territoire, gestion des risques, épidémiologie, météorologie, ...

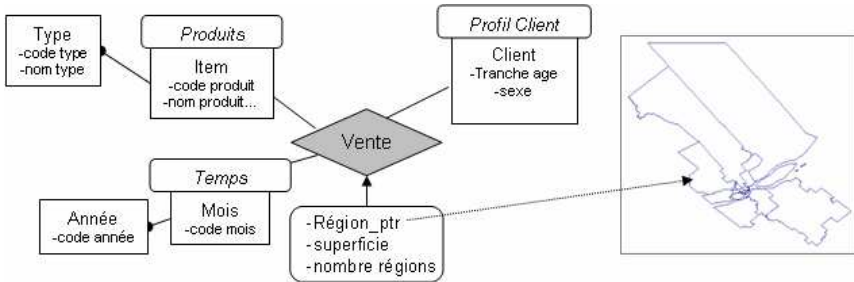


Figure 3 : Introduction d'une mesure spatiale

Notre équipe mène plusieurs travaux ayant trait à l'intégration des données spatio-temporelles dans les systèmes OLAP. Ces travaux portent sur les aspects modélisation et exploitation dans des systèmes SOLAP et nous en présentons une synthèse dans cet article. Dans la première partie, un ensemble de définitions du SOLAP et un panorama de la littérature sont donnés. Dans la deuxième section, nous présentons des travaux portant sur la conception de dimensions spatiales évolutives. La section trois présente un modèle multidimensionnel adapté à la gestion de données spatiales continues. Dans la section quatre, nous étudions la conception d'une interface SOLAP. La dernière section présente nos perspectives avec les travaux en cours.

2. Le SOLAP : définition et état de l'art

Le succès connu par les Systèmes d'Information Géographique (SIG) témoigne du besoin de gérer des représentations spatiales. Les SIG sont de puissants outils pour manipuler, interroger et visualiser les bases des données spatiales. Une donnée spatiale est généralement composée de trois parties : une partie géométrique (la forme par exemple ligne, point, polygone), une partie descriptive (par exemple le nom, la population d'une ville) et une partie métrique calculée (par exemple la superficie, le périmètre) [Rigaux *et al*, 2002]. La partie géométrique fournit des informations sur la position et la forme des objets étudiés et est représentée sous forme vectorielle ou de raster. Les données descriptives fournissent des informations qualitatives ou quantitatives sur les caractéristiques des objets. Les données métriques résultent d'un calcul effectué sur la partie géométrique.

Panorama de travaux autour de l'intégration de données spatio-temporelles

Dans une carte, chaque thème d'information est représenté par une couche, ensemble d'objets élémentaires de même nature. Une couche associe la représentation géométrique d'objets spatiaux et la table d'informations statistiques qui leur est associée. L'information y est stockée et représentée au niveau du pixel (représentation raster) ou au niveau des objets spatiaux élémentaires (représentation vectorielle). Les SIG fournissent en outre différentes fonctions d'analyse des données spatiales par exemple des fonctions de recherche, de proximité etc... [Cowen, 1988].

Jusqu'à présent, les propositions de solutions SOLAP reposent sur le couplage d'un serveur OLAP et d'un SIG, ce dernier étant uniquement utilisé pour le stockage et la visualisation des données géographiques [Kouba *et al*, 2000]. [Bédard, 1997] définit le SOLAP comme une plateforme visuelle supportant l'analyse et l'exploration rapides et faciles des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation via un affichage cartographique, tabulaire ou en diagramme statistique. [Stefanovic *et al*, 2000] définissent l'entrepôt de données spatiales comme une collection de données spatiales et non-spatiales intégrées, orientées sujet, non volatiles, historisées, résumées et disponibles pour l'interrogation et l'analyse. L'intégration de la donnée spatiale peut s'effectuer au niveau des dimensions, les problématiques posées sont alors souvent centrées autour de la définition des hiérarchies spatiales. L'intégration de la donnée spatiale peut aussi se faire au niveau des mesures, les problèmes posés concernent alors principalement le traitement de l'agrégation, dans ses aspects sémantiques et programmatiques. Dans tous les cas, une problématique commune aux deux types de modèles est la définition d'interface de navigation et d'exploitation cartographique adaptée et l'extension des algèbres OLAP à l'aide d'opérateurs spatiaux.

La prise en compte de données spatio-temporelles dans des applications décisionnelles nécessite l'adaptation de la technologie OLAP. [Rivest, 2003] présentent les concepts, les caractéristiques d'un SOLAP et les domaines d'application possibles. Ils distinguent comme [Stefanovic *et al*, 2000] et [Bédard, 1997] trois types des dimensions spatiales : non géométriques (textuelle), géométriques et mixtes. Les deux dernières permettent la représentation, la visualisation et le requêtage cartographiques. Dans [Li, 2003] une dimension représentant les opérateurs topologiques est adjointe au modèle multidimensionnel, pour exploiter les relations spatiales entre les objets spatiaux. [Papadias *et al*, 2002] montrent l'intérêt de pouvoir introduire des dimensions dont les hiérarchies ne sont pas définies lors de la modélisation. Les auteurs proposent plusieurs solutions d'indexation et développent des index multi-arbres qui permettent de définir des dimensions spatiales dynamiques en intégrant des pré-agrégations. Cette approche définit un nouveau modèle de données pour les entrepôts de données spatio-temporels.

L'ensemble de ces travaux confirment l'importance et l'aspect innovateur du SOLAP. La définition des concepts de base du SOLAP est un thème de recherche ouvert : définir un framework conceptuel, logique et physique pour le SOLAP, tracer le périmètre des analyses multidimensionnelles et/ou spatiales qui peuvent être menées grâce au SOLAP sur un entrepôt des données spatiales, restent des sujets majeurs.

3. Dimensions spatiales évolutives

1.1 Problématiques

Un exemple des difficultés de conception d'un modèle multidimensionnel intégrant des données spatio-temporelles concerne la prise en compte de l'évolution de la structure et/ou des membres de ces dimensions spatiales. Nous illustrons ici ce problème sur un cas d'application pris dans le domaine de la foresterie [Miquel *et al.*, 2002].

Un inventaire forestier consiste à partitionner la surface de la forêt étudiée en zones appelées peuplements, qui présentent des caractéristiques forestières homogènes (essence d'arbres, âge, densité, hauteur, etc.). Le résultat d'un inventaire est donc un ensemble de peuplements auxquels sont associés une géométrie et les attributs qui le caractérisent. Le modèle du cube est donné en figure 4. Il fait apparaître une table de faits contenant les mesures *surface* et *volume* ainsi que 5 dimensions : *Essence*, *Age*, *Densité*, *Temps* et *Découpage*, cette dernière étant la dimension spatiale qui décrit un peuplement.

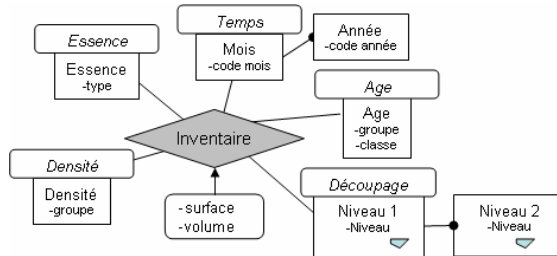


Figure 4 : Modèle multidimensionnel appliqué à l'analyse de données forestières

La modélisation multidimensionnelle doit supporter la richesse informationnelle des objets géospatiaux notamment en prenant en compte l'hétérogénéité liée à la nature des objets spatiaux et à l'évolution sémantique des données descriptives. Les objets spatiaux pris en compte évoluant au cours du temps, ils ne peuvent être utilisés comme objets de référence. Il en découle une hétérogénéité des données géométriques lorsque le découpage de l'espace évolue au cours du temps. Les données descriptives évoluent en fonction de la législation, du mode d'acquisition, de la variation des domaines de valeurs des attributs et de leur codage.

L'exemple précédent montre que l'hétérogénéité des données géospatiales doit être traitée au niveau géométrique et descriptif. Pour le premier type de données, cela consiste à trouver une organisation de la surface du territoire en entités spatiales invariantes dans le temps. Ces entités formeront les membres du niveau le plus fin de la dimension spatiale. Le deuxième type de données pose un problème plus classique d'intégration de données hétérogènes avec la contrainte de faire cohabiter dans le même modèle des données détaillées temporellement non comparables et des données agrégées temporellement comparables.

1.2 Un modèle multidimensionnel évolutif

Nous avons étendu le modèle M3 que nous avons proposé dans le cadre d'une structure multidimensionnelle multiversion [Body *et al*, 2003], pour la prise en compte des évolutions de la dimension spatiale et des dimensions sémantiques dans les applications SOLAP [Miquel et Tchounikine, 2003]. Le modèle M3 permet de définir des dimensions dont les liens et les membres sont associés à des intervalles de temps de validité. Cette méthode de construction de dimensions par des graphes temporels peut être appliquée aux dimensions descriptives lorsque leur sémantique est évolutive dans le temps. Les relations de mapping sont définies entre membres d'une dimension descriptive et permettent de passer d'une version de structure à une autre. Ces relations de mapping sont également utilisées pour effectuer la conversion entre les découpages correspondant aux différents périodes d'acquisition des données. Pour pouvoir extraire les relations de mapping entre les découpages des différents inventaires, nous avons choisi de passer par l'intermédiaire d'une référence spatiale invariante dans le temps. Cette méthode repose sur une structuration en mosaïque du territoire selon un mode matriciel. La surface du territoire est alors représentée à l'aide de cellules régulières. On obtient ainsi un découpage de référence fixe, un zone de l'espace devenant un ensemble de cellules. Ce maillage de la surface permet de déduire la relation de mapping exact reliant chaque zone dans sa version i à sa version $i+1$.

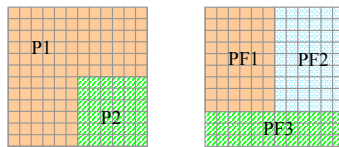


Figure 5 - Exemple de l'utilisation du maillage de la surface

Ainsi, la figure 5 représente une forêt initialement composée de 2 peuplements P1 et P2. Lors de l'inventaire suivant, trois peuplements sont détectés PF1, PF2 et PF3. Une relation de mapping est définie par un tuple $\langle Id_from, Id_to, F, F^{-1} \rangle$ où :

- Id_from est l'identifiant de la version de membre avant évolution.
- Id_to est l'identifiant de la version de membre après évolution.
- F est un ensemble de couples $\langle fm_k, cf_k \rangle$ où fm_k est une fonction de mapping précisant comment la mesure m_k doit être mappée. cf_k est l'indice de confiance associée à cette fonction fm_k . Les indices de confiance (em : *Exact Mesure*, am : *Approximate Mesure*) sont des mesures ajoutées à la table de faits qui permettent de déterminer la qualité du mapping.
- F^{-1} est l'ensemble de couples $\langle fm'_k, cf'_k \rangle$ définissant le passage inverse de la version Id_to vers la version Id_from .

A partir du maillage représenté, on en déduit les relations de mapping suivantes pour une mesure x :

$$\langle P1, PF1, \{(x \mapsto \frac{1}{2}x, em)\}, \{(x \mapsto x, em)\} \rangle ;$$

$$\langle P1, PF2, \{(x \mapsto \frac{1}{3}x, em)\} \{(x \mapsto \frac{2}{3}x, em)\} \rangle ;$$

$$\langle P1, PF3, \{(x \mapsto \frac{1}{6} x, em)\} \{(x \mapsto \frac{1}{2} x, em)\} \rangle;$$

$$\langle P2, PF2, \{(x \mapsto \frac{1}{2} x, em)\} \{(x \mapsto \frac{1}{3} x, em)\} \rangle;$$

$$\langle P2, PF3, \{(x \mapsto \frac{1}{2} x, em)\} \{(x \mapsto \frac{1}{2} x, em)\} \rangle$$

Ces relations de mapping traduisent donc que l'on peut déduire les mesures associées à PF1, PF2 et PF3 à partir des mesures de P1 et P2 par les relations suivantes :

$$PF1 = \frac{1}{2} P1 ; PF2 = \frac{1}{3} P1 + \frac{1}{2} P2 ; PF3 = \frac{1}{6} P1 + \frac{1}{2} P2$$

De même, les mesures associées à PF1, PF2 et PF3 seront représentées selon P1 et P2 en utilisant les fonctions de mapping inverses :

$$P1 = PF1 + \frac{2}{3} PF2 + \frac{1}{2} PF3 ; P2 = \frac{1}{3} PF2 + \frac{1}{2} PF3$$

On notera qu'il n'y a pas de relation de mapping entre P2 et PF1, ce qui traduit qu'il n'y a aucun lien entre ces deux versions de membres comme le montre la figure 5. Ainsi, une mesure peut être représentée sur différents découpages territoriaux.

La table de faits multiversion regroupe les différentes mesures selon les dimensions spatiale et descriptives précédemment définies et selon les différents modes temporels de présentation. Ces derniers interviennent au niveau de la table de faits multiversion comme une dimension supplémentaire. Notre architecture se compose donc de trois parties : un entrepôt temporel qui contient les données temporellement consistantes et les métadonnées (dont les relations de mapping), l'entrepôt multiversion qui introduit la dimension *Modes Temporels de Présentation* et implémente la table de faits multiversion et enfin le cube créé par le serveur OLAP. La prise en compte des évolutions temporelles accroît les capacités de navigation de l'utilisateur en lui donnant la possibilité de choisir le type de présentation qu'il désire, par exemple mapper les données les plus récentes dans un découpage géographique plus ancien afin de mieux évaluer les évolutions. Toutes les informations recueillies sont accessibles quelle que soit la représentation choisie, les données étant alors mappées ou d'origine.

4. Modélisation d'hypercubes continus

1.3 Problématique

Les hypercubes de données utilisés pour explorer et analyser des données historisées comportent par définition des données discrétisées correspondant aux mesures, et ce quelque soit le niveau de granularité des dimensions. Or, les dimensions spatiales et temporelles peuvent représenter des axes d'analyse continus, notamment pour des applications portant sur des phénomènes naturels (météorologie, études maritimes, pollution ...). Nous proposons d'introduire au sein des structures multidimensionnelles des mécanismes basés sur des

Panorama de travaux autour de l'intégration de données spatio-temporelles

interpolations permettant de restituer des données continues afin de donner à l'utilisateur la vision d'un hypercube continu.

Deux niveaux de représentations sont définis :

- le premier niveau physique concerne l'implémentation pour laquelle la représentation des données reste discrète,
- le deuxième niveau est l'univers de l'utilisateur pour lequel les données sont continues. Des fonctions d'interpolation assurent le passage d'un niveau à l'autre (figure 6).

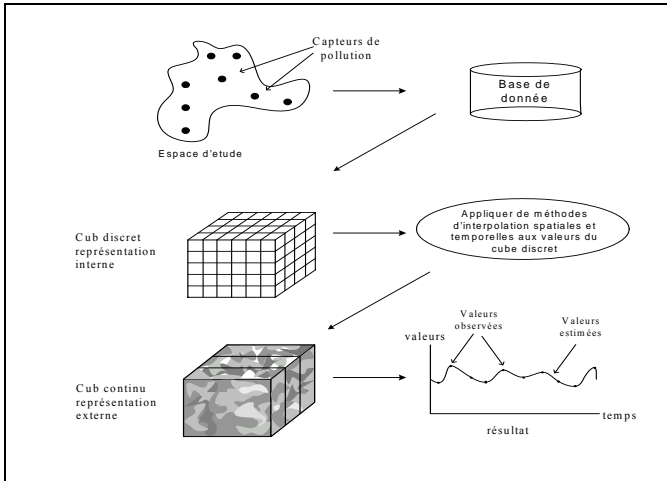


Figure 6 : Etapes de construction d'un hypercube continu

Les données représentant le phénomène continu (la pollution dans cet exemple) sont mesurées par plusieurs capteurs distribués régulièrement ou non sur l'espace d'étude. La continuité des membres de chaque niveau hiérarchique des dimensions est représentée dans le modèle du cube par un pictogramme (figure 7).

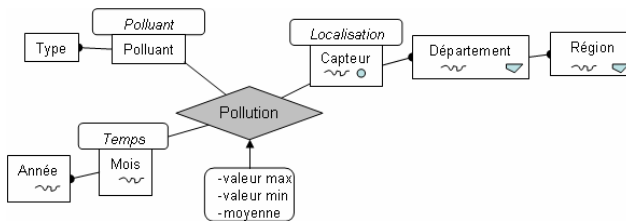


Figure 7 : Hypercube pour l'analyse de la pollution

1.4 Modélisation multidimensionnelle continue

A partir de cette base de données, un hypercube est construit pour une analyse multidimensionnelle. Les axes d'analyse continus sont intégrés au cube par une représentation discrète. Des méthodes d'interpolation spatiale et temporelle sont introduites pour supporter la vision continue. Nous avons expérimenté plusieurs méthodes d'interpolation, notamment pour la continuité spatiale telles que la méthode IDW, le krigeage ou les splines [Arnaud, 2000], [Wackernagel, 2003]. Il est évident que la pertinence des méthodes d'interpolation conditionne la qualité des données continues produites et pour l'instant, ce problème reste ouvert notamment pour des phénomènes naturels complexes tels que la diffusion de la pollution. Les résultats d'une requête reprennent à la fois des données connues et des données estimées. Un modèle formel et des opérateurs ont été définis [Ahmed, 2005]. Dans ce modèle, la notion de cube de base discret *discCb* reprend la définition des cubes de base introduite par [Vassiliadis, 1998]. Cette définition est étendue aux cubes de base continus *contCb* par l'intégration de fonctions d'interpolation et finalement un cube *C* est défini comme 4-tuple $\langle D, L, contCb, R \rangle$ où *D* est une liste de dimensions comprenant *M*, *L* est le niveau de dimension, *R* représente les données de cellules et *contCb* est le cube de base continu à partir duquel *C* est construit. Nous distinguons deux types de dimensions : les dimensions discrètes composées de niveaux hiérarchiques regroupant un nombre fini de membres et les dimensions continues qui représentent les axes d'analyse présentant un aspect continu [Ahmed et al., 04] [Ahmed, 05].

Les opérateurs classiques OLAP conviennent pour naviguer dans les dimensions discrètes d'hypercube continu, toutefois dans les dimensions continues, ces opérateurs doivent être étendus. Les fonctions d'agrégation auront une signification différente selon que la requête porte sur une dimension discrète ou continue. Ainsi, par exemple, la somme de mesures pour une région spécifique ou un intervalle de temps spécifique est représentée par le calcul de l'intégrale selon l'espace ou le temps. Nous avons étendu les opérateurs discrets aux données continues à partir de la notion de champs continus [Gordillo, 2001].

5. Une interface SOLAP

Nous avons voulu reprendre les concepts du SOLAP tels qu'ils étaient définis dans la littérature et implémentés dans les applications de l'équipe d'Yvan Bédard afin de disposer d'un atelier permettant d'intégrer de nouvelles fonctionnalités telles que les hypercubes continus.

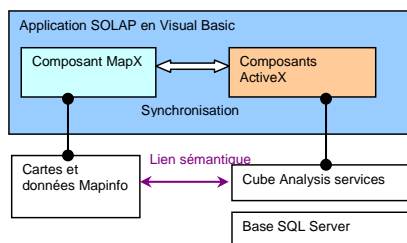


Figure 8 : Architecture logicielle de GèOlap

Panorama de travaux autour de l'intégration de données spatio-temporelles

Le prototype de système SOLAP développé est appelé GéOlap [Baillot, 2004]. GéOlap utilise Analysis Services comme serveur OLAP, Mapinfo comme serveur de carte et des composants ActiveX et MapX pour le développement de l'interface (figure 8).

GéOlap est une interface dynamique et intégrée qui permet de se connecter, puis de naviguer dans un cube maintenu par Analysis services et comportant une ou plusieurs dimensions spatiales. Chaque niveau hiérarchique de la dimension spatiale est représenté par un layer cartographique dans le SIG. La réponse à une requête OLAP est un nouveau layer calculé dynamiquement et représentant les membres de la dimension spatiale au grain souhaité et les mesures sous forme graphique (valeur, densité de points, dégradé de couleur, histogramme...). L'interface permet de naviguer indifféremment et de façon synchronisée dans le cube à travers l'arborescence des dimensions et des faits, des tableaux multidimensionnels, des représentations graphiques et/ou une carte. Chacune de ses fenêtres d'interrogation/visualisation offre une barre d'outils de navigation OLAP classique (*roll-up*, *drill-down*, *pivot*, ...). Dans la fenêtre de visualisation cartographique, les outils de forage sont appliqués sur les objets géométriques sélectionnés graphiquement. Cette fenêtre offre de plus une fonctionnalité de recentrage automatique sur la zone d'intérêt. La figure 9 représente un exemple de navigation dans un cube de données de pollution. Ces données reprennent l'exemple de la pollution (figure 7) avec une dimension spatiale *Localisation* hiérarchisée en niveaux *Ville*, *Département* et *Région*.

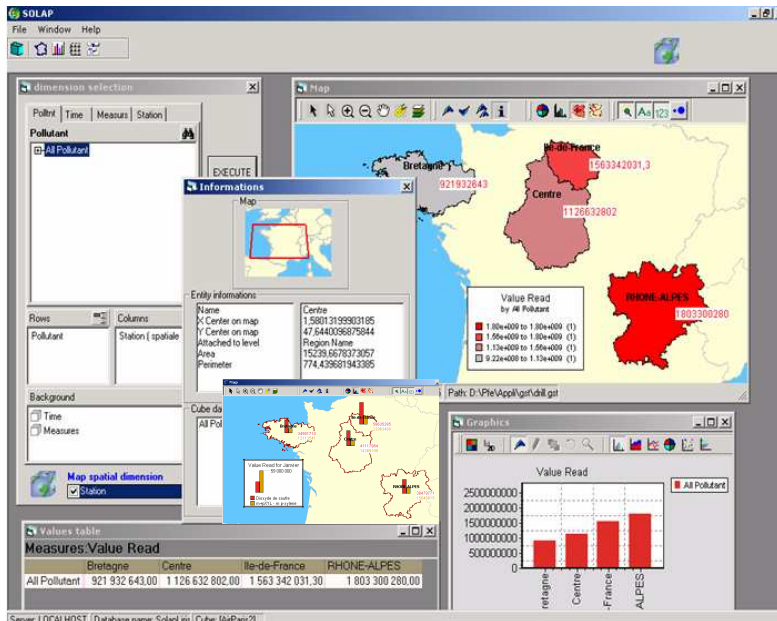


Figure 9 : Interface de navigation de GéOlap

Les autres dimensions sont le *Temps*, et le *Type de Polluants*. Les mesures de concentration de différents polluants dans les régions de France sont représentées dans la carte sous forme d'histogrammes et de couleurs (valeur moyenne, valeur minimum, maximum).

6. Conclusion et perspectives

Dans les travaux de notre équipe exposés dans cet article, nous avons repris les concepts du SOLAP en les adaptant pour prendre en compte les évolutions de la dimension spatiale et introduire la notion de dimension continue, dans le temps et/ou dans l'espace. Actuellement, nos réflexions s'orientent dans deux directions. La première concerne l'étude de modèles formels permettant de modéliser les mesures spatiales. Comme nous l'avons évoqué dans la partie 1, la définition de mesures spatiales pose de nombreux problèmes. En effet, nous observons que la plupart des modèles proposés définissent la mesure spatiale comme un pointeur vers un objet géométrique, associé à des mesures numériques tels que la superficie ou le nombre de zones qui deviennent alors, implicitement, des mesures dérivées. De ce fait, l'indépendance des différentes mesures est remise en question, et une réflexion et une formalisation des fonctions d'agrégation associées à ces mesures dérivées deviennent indispensables. Un modèle multidimensionnel permettant de définir non pas une mesure spatiale avec un ensemble de mesures dérivées disjointes mais un véritable *fait spatial*, avec ses attributs descriptifs, géométriques et métriques reste à construire.

Notre deuxième axe de recherche concerne l'exploitation et la visualisation des données spatio-temporelles dans un hypercube. Nous étudions notamment l'utilisation de langages sémiologiques, principalement les chorèmes, pour l'interrogation et la visualisation cartographique dynamique des résultats de la navigation. Cette nouvelle forme d'interaction utilisateur-machine va être intégrée dans notre prototype GéOlap et sera utilisée dans le cadre du projet international Corila [Corila, 2005] qui vise à étudier les évolutions de paramètres environnementaux mesurés dans la lagune de Venise (température, biodiversité, pollution, courants...). Cette application a pour vocation d'être utilisée par un ensemble de chercheurs de différentes communautés, dans différents instituts. C'est pourquoi GéOlap est actuellement en cours d'évolution pour offrir ses fonctionnalités sous la forme de service web.

Références

- [Agrawal et al., 1997] R. Agrawal, A. Gupta, S. Sarawagi. Modeling multidimensional databases. *Proceeding of the 13th International Conference on Data Engineering, (ICDE'97)*, Birmingham, UK, April 7-11, pp. 232-243. 7-11 1997.
- [Ahmed, 2005] T. Ahmed. Multidimensional Structures Dedicated to Continuous Spatiotemporal Phenomena. *Proceeding of the 22nd British National Conference on Databases (BNCOD)*, 5-7 juillet 2005, Sunderland, England, UK

- [Ahmed et al., 2004] T. Ahmed, M. Miquel, R. Laurini. Continuous Multidimensional Structures for Continuous Phenomena Analysis, *Actes du congrès SETIT, Sciences Electroniques, Technologies de l'Information et des Télécommunications, IEEE*, Edité par MS Bouhelel et B. Solaiman, 2004
- [Arnaud et Emery, 2000] M. Arnaud et X. Emery. Estimation et interpolation spatiale. Méthodes déterministes et méthodes géostatistiques. Editions Hermès Sciences Publications, 2000
- [Baillot, 2004] V. Baillot. *Spécifications et développement de GéOlap*. Rapport technique, Juin 04
- [Bédard, 1997] Y. Bédard. Spatial OLAP. 2^{ème} Forum annuel sur la R-D, Géomatique VI : Un monde accessible, 1997, Montréal.
- [Bédard et al., 2001] Y. Bédard, T. Merrett, J. Han. Fundamentals of Spatial Data Warehousing for Geographic Knowledge Discovery. In *Geographic Data Mining and Knowledge Discovery* edited by H. Miller and J. Han, Research Monographs in GIS series edited by Peter Fisher and Jonathan Raper, Taylor & Francis, 2001
- [Body et al., 2003] M. Body, M. Miquel, Y. Bédard, A. Tchounikine. Handling Evolutions in Multidimensional Structure. *Proceeding of the IEEE International Conference on Data Engineering, ICDE*, March 5-8 2003, Bangalore, Inde
- [Cabibbo et Torlone 1998] Cabibbo et R. Torlone. A Logical Approach to Multidimensional Databases. *Proceedings of the 6th International Conference on Extending Database Technology (EDBT'98)*, Valencia, Spain, 1998.
- [Corila, 2005] Rapport d'avancement du projet Corila, 2005
- [Cowen, 1988] D.J. Cowen. GIS versus CAD versus DBMS: What are the differences ? *Fotogrammetric Engineering and Remote Sensing*, 1988, vol. 54, p. 1551-1555.
- [Fidalgo et al., 2004] R.N. Fidalgo, V. Times, J. Silva, F.F. Souza. GeoDWFrame: A Framework for Guiding the Design of Geographical Dimensional Schemas. *Proceeding of the International Conference on Data Warehousing and Knowledge Discovery*, Zaragoza Espagne, Septembre 2004, p. 26-37.
- [Franklin, 1992] C. Franklin. An Introduction to Geographic Information Systems: Linking Maps to databases. Database. pp. 13-21. 1992.
- [Gordillo, 2001] S. Gordillo. *Modeling and manipulation of spatiotemporal phenomena*. PhD. Thesis, Claude Bernard University– Lyon I, France, 12 October 2001.
- [Inmon, 1992] W. H. Inmon. Building the data warehouse. Edition John Wiley and sons. 1992.
- [Jensen et al., 2004] C. S. Jensen, A. Kligys, T.B. Pedersen, I. Timko. Multidimensional data modeling for location-based services. *The VLDB Journal*, vol. 13 n° 1, 2004, p. 1-21.
- [Kouba et al., 2000] Z. Kouba, K. Matoušek, P. Mikšovský. On Data Warehouse and GIS integration. *Proceeding of the 11th International Conference and Workshop on Database and Expert Systems Applications*, Greenwich, Anghelter, Septembre 2000, Springer-Verlag, p. 604-613.
- [Li et al., 2003] : Y. Li, Y. Chen, F.Y. Rao. The Approach for Data Warehouse to Answering Spatial OLAP Queries. *Lecture Notes in Computer Science*, vol. 2690, March 2003, Springer-Verlag Berlin Heidelberg, p. 270-278.
- [Malinowski et Zimányi, 2004] E. Malinowski, E. Zimányi. Representing spatiality in a conceptual multidimensional model. *Proceeding of the 12th annual ACM International workshop on Geographic information systems*, Washington DC New York USA, 2004, ACM Press , p. 12-22.

- [Marchand et al., 2003] P. Marchand, A. Brisebois, Y. Bédard, G. Edwards. Implementation and evaluation of a hypercube-based method for spatio-temporal exploration and analysis. *Journal of the International Society of Photogrammetry and Remote Sensing*, Vol. 59, 2003, p. 6-20.
- [Miquel et al, 2002] M. Miquel, Y. Bédard, A. Brisebois. Conception d'entrepôts de données géospatiales à partir de sources hétérogènes, exemple d'application en foresterie. *Revue ISI-NIS, Special Issue Data warehousing*, Edition Hermès Science, Paris, Volume 7-n°3/2002
- [Miquel et Tchounikine 2003] M. Miquel et A. Tchounikine. Extension du modèle M3 aux évolutions temporelles dans les applications SOLAP. *Revue des Nouvelles Technologies de l'Information (RNTI)*, Editions Cepaduès, 2003
- [Papadias et al, 2002] D. Papadias, Y. Tao, J. Zhang, N. Mamoulis, Q. Shen, J. Sun. Indexing and retrieval of historical aggregate information about moving objects. *IEEE Data Engineering Bulletin*, 25(2): 10-17, June, 2002.
- [Pedersen et Tryfona, 2001] T. B. Pedersen, N. Tryfona, Pre-Aggregation In Spatial DataWarehouses. *Proceeding of the 7th International Symposium on Spatial and Temporal Databases*, California USA, Juillet 2001, p. 460-480.
- [Rigaux et al, 2002] : P. Rigaux, M. Scholl, A. Voisard. Spatial databases with application to GIS. San Francisco, CA, Morgan Kaufmann Publishers Inc., 2002.
- [Rivest et al, 2003] : S. Rivest, Y. Bédard, M.J. Proulx, M. Nadeau. SOLAP: A new type of user interface to support spatiotemporal multidimensional data exploration and analysis. *Proceeding of the of ISPRS workshop on Spatial, Temporal and Multi-Dimensional Data Modeling and Analysis*, Ville de Québec, Québec, Canada. 2-3 October, 2003.
- [Stefanovic et al, 2000] : N. Stefanovic, J. Han, K. Koperski. Object-based selective materialization for efficient implementation of spatial data cubes. *IEEE Transaction on Knowledge and Data Engineering*, Vol. 12, No. 6 pp. 938 - 957 November - December 2000.
- [Vassiliadis, 1998] : P. Vassiliadis. Modelling Multidimensional Databases, Cubes and Cube Operations. *Proceeding of the 10th Int. Conf. on Scientific and Statistical Database Management (SSDBM)*, Capri, Italy, 1998.
- [Wackernagel, 2003] H. Wackernagel. Multivariate Geostatistics. An Introduction with Applications. *Third Edition, Eds. Springer-Verlag, 2003*

Summary

This article presents a survey of various work that are carried out in our team and deal with the integration of the spatio-temporal data in data warehouses and hypercubes. We focus on the taking in account of evolutions in spatial dimensions, on the multidimensional modelling of continuous data and on the design of navigation interface for spatio-temporal data. These results enable us to advance new prospects for multidimensional models and visualization requirements for spatio-temporal data.