Rivest, S., P. Gignac, J. Charron & Y. Bédard, 2004. Développement d'un système d'exploration spatio-temporelle interactive des données de la Banque d'information corporative du ministère des Transports du Québec. Colloque Géomatique 2004 - Un choix stratégique! Montréal, 27-28 octobre., Montréal, Canada

Développement d'un système d'exploration spatio-temporelle interactive des données de la Banque d'information corporative du ministère des Transports du Québec

Sonia Rivest[1], Pierre Gignac[2], Jacques Charron[3] et Yvan Bédard[1]

[1] Centre de recherche en géomatique (CRG)
et Département des Sciences géomatiques
Université Laval
Sainte-Foy (Québec) Canada
Téléphone : 418-656-2131
Télécopieur : 418-656-3607
sonia.rivest@scg.ulaval.ca

[2] Ministère des Transports du Québec (MTQ)
700, René-Lévesque Est
Québec (Québec) Canada
Téléphone : 418-643-3664
Télécopieur : 418-646-8366
pgignac@mtq.gouv.qc.ca

[3] KHEOPS Technologies
300, Saint-Sacrement, suite 114
Montréal (Québec) Canada
Téléphone : 514-285-1211
Télécopieur : 514-285-1177
jcharron@kheops-tech.com

Sonia Rivest possède un baccalauréat et une maîtrise en Sciences géomatiques de l'Université Laval. Elle possède une expérience de travail diversifiée (secteurs public et privé). Elle travaille présentement en tant que professionnelle de recherche au Centre de recherche en géomatique de l'Université Laval dans l'équipe des bases de données spatiales du Dr Yvan Bédard. Ses intérêts professionnels incluent les bases de données multidimensionnelles, les entrepôts de données spatiales ainsi que les outils SOLAP. Elle a participé à de nombreux projets d'application des concepts SOLAP dans différents domaines et différentes organisations.

Pierre Gignac possède une maîtrise en sciences politiques avec majeure en administration publique. Il cumule plus de trente ans d'expérience professionnelle, dont une vingtaine d'années dans des postes d'encadrement différents. Il œuvre comme chef du service de la gestion de l'information et des statistiques, à la Direction de la planification et des stratégies de l'information du ministère des Transports du Québec depuis novembre 2001. Parmi ses tâches actuelles, il élabore et voit à la mise en œuvre des orientations, politiques et normes ministérielles en matière de gestion de l'information. De plus, il procède, en collaboration avec les intervenants ministériels, à l'identification des besoins en données corporatives, à la réalisation et à l'exploitation d'un entrepôt de données (Banque d'information corporative ou BIC), de comptoirs de données et des outils d'exploitation corporatifs.

Jacques Charron est détenteur d'un baccalauréat et d'une maîtrise en Sciences géomatiques de l'Université Laval. Il est spécialisé dans les systèmes d'information à référence spatiale (SIRS) et oeuvre dans le domaine de la géomatique depuis plus de quatorze ans. Au cours des ans, Monsieur Charron a acquis une solide expérience en géomatisation des organisations et en développement de logiciels à titre d'analyste-concepteur et de chargé de projets. Monsieur Charron a fondé Les conseillers KHEOPS technologies en juin 1995 où il agit à titre de Président

et directeur général. Enfin, il a été l'instigateur du développement de la technologie JMap et il a grandement contribué à la conception de ce produit.

Dr Bédard est professeur en SIG et bases de données géospatiales depuis 1986. Il est membre du Centre de Recherche en Géomatique (CRG) où il effectue des travaux en modélisation de bases de données et en systèmes analytiques géodécisionnels. Il a été directeur du CRG pendant 7 années. Dr Bédard a un important palmarès en recherche fondamentale et appliquée, dont plusieurs projets impliquant des organisations gouvernementales et industrielles. Il a formé au-delà de 1000 professionels sur 3 continents et a publié comme auteur principal ou coauteur plus de 300 articles et conférences nationales et internationales. Il est aussi très actif dans le Réseau canadien de centres d'excellence en géomatique GEOIDE. Finalement, il est impliqué dans différentes associations et journaux scientifiques ainsi que dans des comités provinciaux et fédéraux, privés, publics et professionnels. Il est titulaire de la nouvelle chaire industrielle CRSNG du Canada en bases de données géospatiales décisionnelles.

Résumé

Cet article présente les résultats d'un projet de développement d'un logiciel d'exploration interactive des données et son application aux sujets décisionnels du ministère des Transports du Québec (MTQ). D'une part, le MTQ désirait faciliter les analyses de tendances ou de séries de données, supporter l'élaboration de projections et de stratégies, effectuer des suivis et évaluer les impacts de changements survenus ou à venir à partir des données de sa nouvelle Banque d'information corporative, laquelle vise à préserver et exploiter le patrimoine informationnel de l'organisation. D'autre part, le Centre de recherche en géomatique (CRG) réalisait depuis plusieurs années des recherches et des prototypes sur un outil d'accès et d'analyse de données spatio-temporelles appelé SOLAP (Spatial On-Line Analytical Processing, résultant du mariage entre « intelligence d'affaires » et « géomatique ») et pour lequel une entente de développement commercial existait avec KHEOPS Technologies. Le projet ici décrit présente les résultats des efforts de ces trois partenaires pour arriver à la mise en place du système.

Le projet a permis de développer un outil SOLAP générique et de l'implanter en tenant compte des fonctionnalités requises par le MTQ. Le SOLAP est défini comme : « un logiciel de navigation facile et rapide dans les bases de données spatiales qui offre plusieurs niveaux de granularité d'information, plusieurs thèmes, plusieurs époques et plusieurs modes de visualisation synchronisés ou non : cartes, tableaux et diagrammes » (Bédard, 2004). L'outil est basé sur les technologies de cartographie en ligne et d'informatique décisionnelle (entrepôts de données et OLAP). Le SOLAP présenté comprend deux sections : (1) la section administration permettant de configurer, à partir de l'entrepôt de données, les données multidimensionnelles et les données spatiales pour l'accès, ainsi que (2) la section visualisation qui permet l'affichage des données et la navigation d'une façon interactive, c'est-à-dire permettant l'utilisation d'opérateurs tels que le forage, le remontage et le forage latéral. La section visualisation offre une série de fonctionnalités novatrices telles que la synchronisation de la sémiologie graphique des affichages, l'utilisation de légendes interactives, la visualisation du cheminement d'exploration et l'accès aux métadonnées reliées aux données consultées.

Abstract

This paper presents the results of an interactive data exploration software development project and its application to the decisional subjects of the ministère des Transports du Québec (MTQ). On one hand, the MTQ wanted to facilitate trends or data series analyses, to support the elaboration of projections and strategies, to track and to evaluate the impacts of changes, past and future, from the data stored in its new Corporate Information Database. This database aims at preserving and exploiting the information assets of the organization. On the other hand, the

Center for Research in Geomatics (CRG) had been studying and prototyping, for many years, a new spatio-temporal data access and analysis tool called SOLAP (Spatial On-Line Analytical Processing, the result of a coupling between “business intelligence” and “geomatics”) and for which a commercial development agreement existed with KHEOPS Technologies. The current project presents the results of the efforts of the three partners to implement the system.

The project allowed to develop a generic SOLAP tool and to implement it taking into account the functionalities required by the MTQ. SOLAP is defined as: “a type of software that allows rapid and easy navigation within spatial databases and that offers many levels of information granularity, many themes, many epochs and many display modes synchronized or not: maps, tables and diagrams” (Bédard, 2004). The tool is based on on-line cartography and decisional technologies (Data Warehouses and OLAP). The presented SOLAP tool comprises two sections: (1) the administration section that allows to configure the multidimensional data and the spatial data for access; and (2) the visualization section that allows the interactive display and exploration of data, supporting the use of operators like drill-down, roll-up and drill-across. The visualization section offers a series of innovative functionalities like the synchronization of the graphical semiology from one display to another, the use of interactive legends, the visualization of the exploration paths and metadata access.

Introduction

Les organisations amassent d’énormes volumes de données à partir de leurs opérations courantes. Ces données sont habituellement collectées par des systèmes dits transactionnels, lesquels sont optimisés pour assurer la consistance des données et faciliter leur mise à jour. Ces optimisations rendent les données difficiles à exploiter par les gestionnaires et analystes qui nécessitent de l’information agrégée, des comparaisons rapides dans l’espace et le temps, l’extraction de tendances, et d’autres opérations complexes pour supporter leurs processus de décision. Parce que les systèmes transactionnels ne sont pas optimisés pour supporter les processus décisionnels, de nouveaux types de systèmes sont mis en place spécifiquement pour remplir les besoins décisionnels. Ces systèmes, au sein desquels l’entrepôt de données joue un rôle central, sont optimisés pour faciliter l’analyse et pour accroître la performance des requêtes complexes.

Le ministère des Transports du Québec (MTQ) suit aussi cette tendance. Depuis 2000, l’équipe du Service de la gestion de l’information et des statistiques, sous l’autorité de la Direction générale des services à la gestion et en collaboration avec d’autres services du ministère, travaille à la mise en place de la Banque d’information corporative (BIC). La BIC extrait régulièrement les données pertinentes des systèmes transactionnels ministériels et conserve à long terme l’information extraite et intégrée dans un dépôt central. Le contenu est adapté aux besoins des utilisateurs et conforme aux définitions standardisées. L’information doit être accessible facilement et rapidement, à l’aide d’interfaces conviviales et flexibles qui permettent aux utilisateurs d’exploiter le contenu selon les besoins du moment. Les interfaces de type OLAP (On-Line Analytical Processing) sont très bien adaptées à ce type de demande. Cependant, comme une grande partie de l’information contenue dans la BIC comprend une composante spatiale, une interface de type OLAP spatial (SOLAP) est plus appropriée pour permettre d’exploiter cette composante à son plein potentiel et ainsi combler les besoins d’analyse spatio-temporelle des utilisateurs. Quelquefois qualifiés de « Keyboardless-GIS », mais représentant beaucoup plus en terme de navigation dans les jeux de données multidimensionnels, les outils SOLAP possèdent une interface à l’usager intuitive qui permet aux utilisateurs non techniques de facilement accéder, visualiser et analyser leurs données.

Les outils SOLAP forment l’un des sujets principaux d’étude de l’équipe de recherche en bases de données spatiales du Centre de recherche en géomatique, dirigée par le Dr Yvan Bédard. Cette équipe a été mandatée par le MTQ pour développer un outil d’analyse spatio-temporelle permettant d’utiliser le plein potentiel de son nouvel entrepôt de données. Dans un esprit de

partenariat conduisant à une application fonctionnelle et à un nouveau produit, l'Université et son partenaire commercial KHEOPS Technologies ont également contribué en nature à la réalisation du projet.

Le présent article comprend un exposé des besoins du MTQ relatifs à l'exploration spatio-temporelle des données informationnelles. Ensuite, une révision des concepts OLAP et SOLAP est effectuée, puis les caractéristiques du logiciel SOLAP développé sont exposées. Finalement, un exemple d'analyse spatio-temporelle est présenté.

Le projet de Banque d'information corporative au ministère des Transports du Québec

Les directions et services du ministère des Transports du Québec, afin d'assumer adéquatement le rôle qui leur est confié, exploitent des données provenant de plusieurs systèmes et travaillent parfois avec plusieurs années d'historique. Les informations obtenues sont utilisées à de multiples fins comme pour :

- présenter une synthèse des activités du ministère (ex. rapport annuel);
- faire des projections, élaborer des stratégies ou des plans d'action pour les années à venir;
- effectuer des suivis des opérations, de données budgétaires ou d'autres aspects des activités du ministère (ex. suivi budgétaire, indicateurs de gestion);
- évaluer les impacts de changements survenus ou à venir (SGIS, 2002).

Des données sur la presque totalité des sujets d'affaires du MTQ sont utilisées en cours d'année pour répondre à des besoins de type informationnel. Afin de combler plus efficacement ces besoins informationnels, le Service de gestion de l'information et des statistiques (SGIS) a été créé en 1998 avec pour mandat de concevoir, de développer et de mettre en œuvre la fonction de gestion de l'information corporative. Au cours de l'année 2000, une conception administrative reliée à la mise en place d'une banque d'information corporative a été réalisée. Une architecture détaillée a été effectuée en 2001-2002. Au cours de la même période, le ministère a procédé au développement et à l'implantation d'une application de gestion des métadonnées, laquelle sera plus tard intégrée à la BIC. Depuis 2003, l'implantation de la BIC est en cours. Un premier comptoir de données est maintenant fonctionnel. Il s'agit du comptoir « Sécurité routière ». D'ici 2008, le plan actuel prévoit l'addition de quatre comptoirs, dont les thèmes sont les suivants : gestion des ressources, planification, modes de transport et utilisateurs externes.

L'architecture de la BIC suit une approche fédérée (Weldon, 1997), c'est-à-dire qu'un entrepôt central est alimenté à partir des différents systèmes transactionnels du ministère. Les données de cet entrepôt sont ensuite réparties dans un certain nombre de marchés de données (nommés « comptoirs de données » au MTQ). Les données de l'entrepôt et des différents comptoirs sont rendues disponibles aux différents types de logiciels clients (outils de rapports, OLAP, SOLAP) qui sont disponibles via l'Intranet corporatif de l'organisation. Environ 400 employés ont été ciblés comme éventuels utilisateurs de la BIC.

Les données d'une trentaine de systèmes transactionnels ministériels, tels que GSQ-Gestion des structures du Québec, GCH – Gestion des chaussées, PPR-Planification, programmation et suivi et BGR-Base géographique routière, sont intégrées à la BIC. Plusieurs des systèmes intégrés gèrent des objets possédant une composante spatiale géométrique. Pour permettre d'exploiter cette composante, plusieurs utilisateurs de la BIC ont fait ressortir le besoin de visualiser les résultats des requêtes ou les rapports qu'ils produisent d'une manière cartographique. D'autres utilisateurs ont fait ressortir le besoin de consulter la BIC directement à partir de cartes et de regrouper des résultats selon des découpages territoriaux.

Les analyses typiques réalisées à l'aide des données de la BIC comprennent plusieurs paramètres. Par exemple, dans un contexte de sécurité routière, les paramètres suivants peuvent être utilisés :

- la période des accidents à analyser;
- les types d'accidents;
- la méthode de calcul (ex. le taux d'accident ou le nombre d'accidents au km);
- le territoire à analyser;
- la longueur des sections sur lesquelles les calculs sont effectués (ex. 500 m, 1000 m);
- le débit des routes concernées, etc.

Ce type d'analyse, utilisant plusieurs paramètres, se prête parfaitement à l'analyse multidimensionnelle spatiale (SOLAP), où chacun des paramètres devient une dimension ou une mesure et où certains éléments des dimensions peuvent servir à une représentation cartographique interactive. Ces nouveaux concepts seront présentés à la section suivante.

Revue des concepts OLAP et SOLAP

Le terme OLAP a été défini au début des années 1990 comme « (…) the name given to the dynamic enterprise analysis required to create, manipulate, animate and synthesize information from exegetical, contemplative and formulaic data analysis models. This includes the ability to discern new or unanticipated relationships between variables, the ability to identify the parameters necessary to handle large amounts of data, to create an unlimited number of dimensions, and to specify cross-dimensional conditions and expressions » (Codd et al., 1993). D'autres définitions ont été proposées plus récemment, incluant : « Une catégorie de logiciels axés sur l'exploration et l'analyse rapide des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation » (Caron, 1998).

OLAP est basé sur l'approche des bases de données multidimensionnelles, qui introduit des concepts qui diffèrent des concepts reliés aux bases de données communément utilisées et qui sont qualifiées de transactionnelles. Ces concepts multidimensionnels incluent : les dimensions, les membres, les mesures, les faits et les cubes de données. Les dimensions représentent les thèmes ou les axes d'analyse. Une dimension contient des membres qui sont organisés hiérarchiquement en niveaux de granularité ou niveaux de détail. Les membres d'un niveau particulier peuvent être agrégés pour former les membres du niveau immédiatement supérieur dans la hiérarchie. Les dimensions peuvent être de différents types : temporel, spatial ou descriptif (ou thématique). Un exemple de dimension spatiale serait la dimension du découpage administratif québécois comprenant trois niveaux : les municipalités, les municipalités régionales de comté (MRC) et les régions administratives. « Montréal » serait alors un exemple de membre du niveau des municipalités. Les mesures sont des attributs, habituellement numériques, qui sont analysés en fonction des différentes dimensions. La valeur d'une mesure dépend des membres choisis dans les différentes dimensions utilisées. Par exemple, la mesure « nombre d'habitants » pourrait dépendre des membres choisis dans les dimensions « Découpage territorial », « Temps » et « Âge » dans un contexte d'étude démographique. Les différentes combinaisons de membres de dimensions et de mesures représentent les faits. Un cube de données, aussi appelé hypercube lorsque plus de trois dimensions sont utilisées, représente un ensemble de mesures agrégées selon un ensemble de dimensions (Thomsen et al., 1999). À l'intérieur d'un cube, les agrégations de mesures possibles sur toutes les combinaisons de dimensions possibles, peuvent être précalculées pour améliorer la performance des requêtes.

Une architecture OLAP générale comprend trois composantes : la base de données structurée

de façon multidimensionnelle, le serveur OLAP et le client OLAP. Le client permet à l'utilisateur de visualiser ses données en utilisant différents types de diagrammes et de tableaux. Il permet aussi à l'utilisateur d'explorer ses données à l'aide de différents opérateurs tels que le forage (visualiser un niveau plus détaillé à l'intérieur d'une dimension), le remontage (visualiser un niveau plus général à l'intérieur d'une dimension), le forage latéral (visualiser d'autres éléments au même niveau de détail) et le pivot (interchanger deux dimensions pour modifier le contenu des axes des diagrammes ou tableaux visualisés).

Une structure multidimensionnelle peut être implantée directement à l'intérieur d'un système de gestion de base de données multidimensionnel ou être simulée à l'intérieur d'un système de gestion de base de données relationnel. Lorsque le premier est utilisé, l'architecture OLAP est nommée « MOLAP » pour OLAP multidimensionnel. Lorsque le second est utilisé, l'architecture est nommée « ROLAP » pour OLAP relationnel. Une architecture mixte peut aussi être utilisée, elle est nommée « HOLAP » pour OLAP hybride. Lorsqu'une base de données relationnelle est utilisée, la structure multidimensionnelle est implantée en utilisant des schémas particuliers : schéma en étoile, en flocon, mixte ou en constellation. Ces schémas permettent de définir les tables de dimensions (stockant l'information des membres et leurs relations) et les tables de faits (stockant les faits multidimensionnels).

Il a été estimé que jusqu'à 80% des données d'un entrepôt possèdent une composante spatiale. (Franklin, 1992). Afin d'exploiter pleinement cette composante, les outils clients utilisés dans une architecture d'entrepôt de données ne sont pas adaptés et de nouveaux outils sont nécessaires. Il a été démontré que les outils OLAP possèdent un certain potentiel pour supporter l'analyse spatio-temporelle (Caron, 1998). Cependant, sans une interface pour visualiser et manipuler la composante géométrique des données spatiales, l'analyse peut demeurer incomplète. Les Systèmes d'information géographique (SIG) sont des candidats potentiels, mais malgré de puissantes capacités d'analyse spatio-temporelle, il est reconnu que les SIG actuels utilisés seuls ne sont pas conçus pour supporter de façon optimale des applications décisionnelles et que des outils alternatifs doivent être utilisés (Bédard et al., 2001). Parmi ces derniers, le couplage des SIG et des outils OLAP en une nouvelle famille d'outils, les outils SOLAP (OLAP spatiaux), offre une solution intéressante.

Les outils SOLAP peuvent être définis comme des « logiciels de navigation rapide et facile dans les bases de données spatiales qui offrent plusieurs niveaux de granularité d'information, plusieurs thèmes, plusieurs époques et plusieurs modes de visualisation synchronisés ou non : cartes, tableaux et graphiques statistiques » (Bédard, 2004). Ces outils :

- visent à supporter, de manière transparente, la façon dont les humains pensent et analysent;
- permettent aux utilisateurs de mettre l'accent sur les résultats de la navigation plutôt que sur le processus utilisé pour l'analyse;
- sont utilisés sans avoir à maîtriser un langage d'interrogation;
- procurent des temps de réponse pratiquement instantanés.

Les outils SOLAP supportent trois types de dimensions spatiales (Han et al., 1998) : les dimensions spatiales non géométriques, les dimensions spatiales géométriques et les dimensions spatiales mixtes. Les dimensions non géométriques utilisent une référence spatiale qui est nominale seulement. Les dimensions géométriques et mixtes associent une géométrie aux membres de tous les niveaux, ou aux membres de certains niveaux définis, respectivement. Cette association permet aux membres des dimensions géométriques et mixtes d'être visualisés sur des cartes, en utilisant une sémiologie graphique reliée aux différentes valeurs de mesures du cube. Les outils SOLAP supportent aussi les mesures spatiales (Han et al., 1998).

Caractéristiques de l’outil SOLAP développé

Notre équipe de recherche, en collaboration avec KHEOPS Technologies, a développé un outil générique SOLAP et appliqué celui-ci pour répondre aux besoins du ministère des Transport du Québec (MTQ). Cet outil SOLAP agira comme outil d’exploitation spatio-temporelle facile et rapide de la BIC. Il s'agit d'un projet typique de transfert technologique impliquant une équipe universitaire, un fournisseur de solution et un client innovateur.

L’outil SOLAP est développé comme une extension au logiciel de cartographie sur le Web JMap et en ce sens, peut être utilisé seul ou installé comme module d'exploitation pour tout jeu de données produit avec les principaux SIG et serveurs universels du marché. L’outil comprend deux volets : un module d’administration et un module de visualisation. Le module d’administration permet la définition, à l’aide d’assistants, des connexions aux bases de données, des composantes multidimensionnelles (dimensions, mesures, cubes), des éléments de représentation et des fonctionnalités cartographiques. Le module de visualisation, le client SOLAP, permet de visualiser les données multidimensionnelles sous forme de différents types de diagrammes (diagrammes à barres verticales et horizontales, camemberts, diagrammes de points, à lignes brisées et de surfaces), de tableaux et surtout de cartes dynamiques simples ou complexes, uni ou multi-échelles, uni ou multi-époques et directement compatibles avec l'environnement de production SIG (ex. ArcGIS) et serveurs universels (ex. Oracle Spatial).

L’extension est développée en Java et exploite aussi la technologie JSP (JavaServer Pages) pour les assistants de son module d’administration. Le SOLAP développé exploite une architecture ROLAP, sans serveur OLAP. Cette architecture permet l’exploitation des structures multidimensionnelles stockées dans des bases de données relationnelles, les tables de faits utilisées contenant les agrégats propres au multidimensionnel. L’utilisation d’une architecture MOLAP (avec serveur) sera supportée ultérieurement.

Les fonctionnalités de l’outil développé sont regroupées en fonctions générales, fonctions cartographiques, fonctions SOLAP, fonctions d’administration et fonctions de métadonnées. Les fonctions générales comprennent la gestion des profils d’usagers, les fonctions d’impression et d’exportation des affichages ainsi que les fonctions d’affichage du cheminement d’analyse et de retour en arrière. Les profils d’usagers comprennent entre autres l’environnement de travail, la définition des sémiologies graphiques personnalisées et la définition des mesures personnalisées. Les fonctions cartographiques reprennent les fonctions du logiciel JMap, avec ajout de fonctionnalités requises par le MTQ, par exemple, la mesure le long d’éléments linéaires. Les fonctions SOLAP permettent l’affichage (cartes, tableaux et diagrammes), la navigation dans ces affichages et la synchronisation de ces affichages (voir figure 4). Ces fonctions incluent aussi : la définition de seuils et de mesures personnalisées, la navigation dans la légende interactive (voir figure 5), la navigation temporelle à l’aide d’une ligne du temps interactive et le groupement de membres. Les fonctions d’administration, tel que mentionné, permettent la définition des éléments de connexion, des éléments multidimensionnels et des fonctionnalités cartographiques, en plus de permettre le filtrage des combinaisons valides (c’est-à-dire non vides) des membres de dimensions afin de faciliter la navigation. Les fonctions de métadonnées permettent l’accès aux différents types de métadonnées : métadonnées de la BIC (incluant les métadonnées spatiales) et métadonnées multidimensionnelles.

Dans l’interface de l’outil SOLAP, les opérateurs de forage, de remontage, de forage latéral et de pivot sont définis afin de tirer avantage de la structure multidimensionnelle et des différents niveaux de détail des données. Ces opérateurs sont fonctionnels dans les différents types d’affichage et il est à noter que, lorsque la synchronisation des affichages est activée, les données descriptives ainsi que les tableaux et graphiques statistiques s'ajustent automatiquement à l'élément visualisé (voir figure 4). Les figures 1 à 3 illustrent le fonctionnement des opérateurs de forage, remontage et forage latéral lorsqu’utilisés sur une carte dans le contexte du découpage administratif du MTQ.

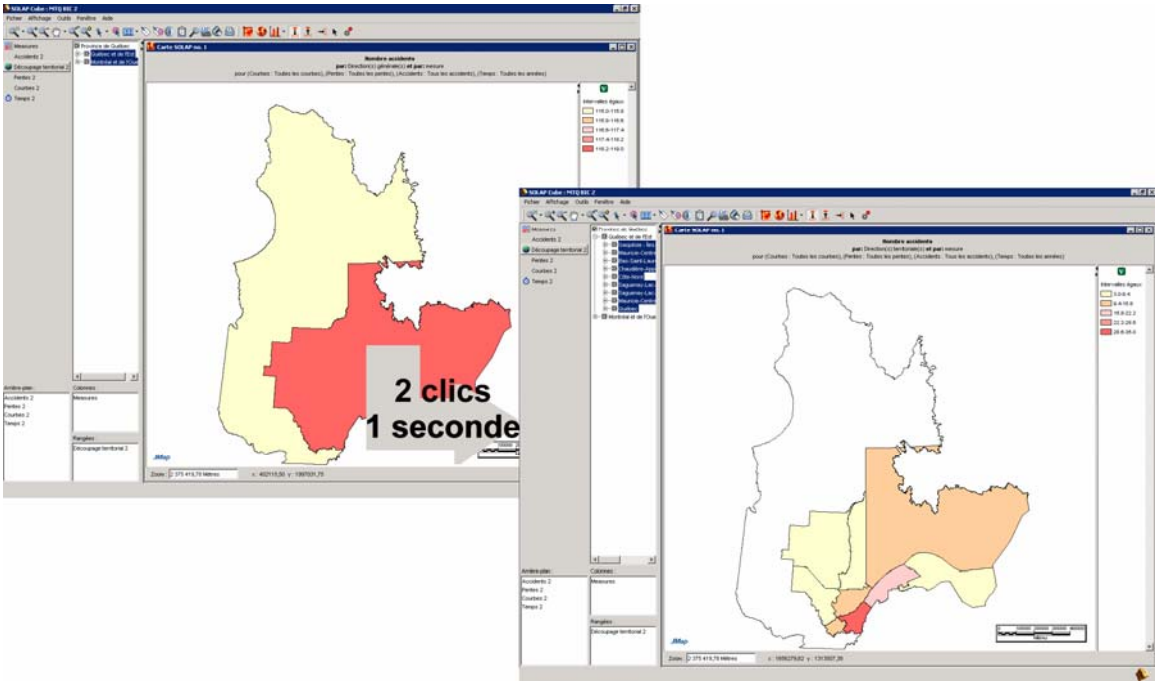

Figure 1. Un exemple d'opération de forage sur la Direction générale de Québec et de l'Est, montrant ses « enfants », les Directions territoriales.

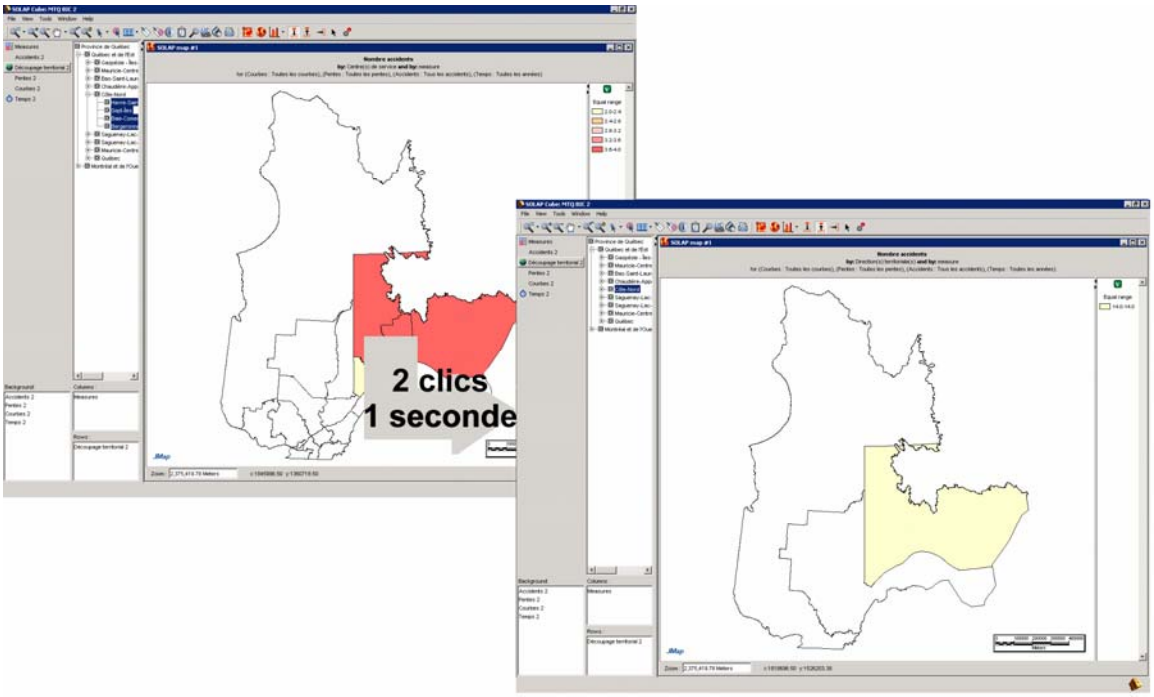

Figure 2. Un exemple d'opération de remontage sur un Centre de service, montrant son « parent », une Direction territoriale.

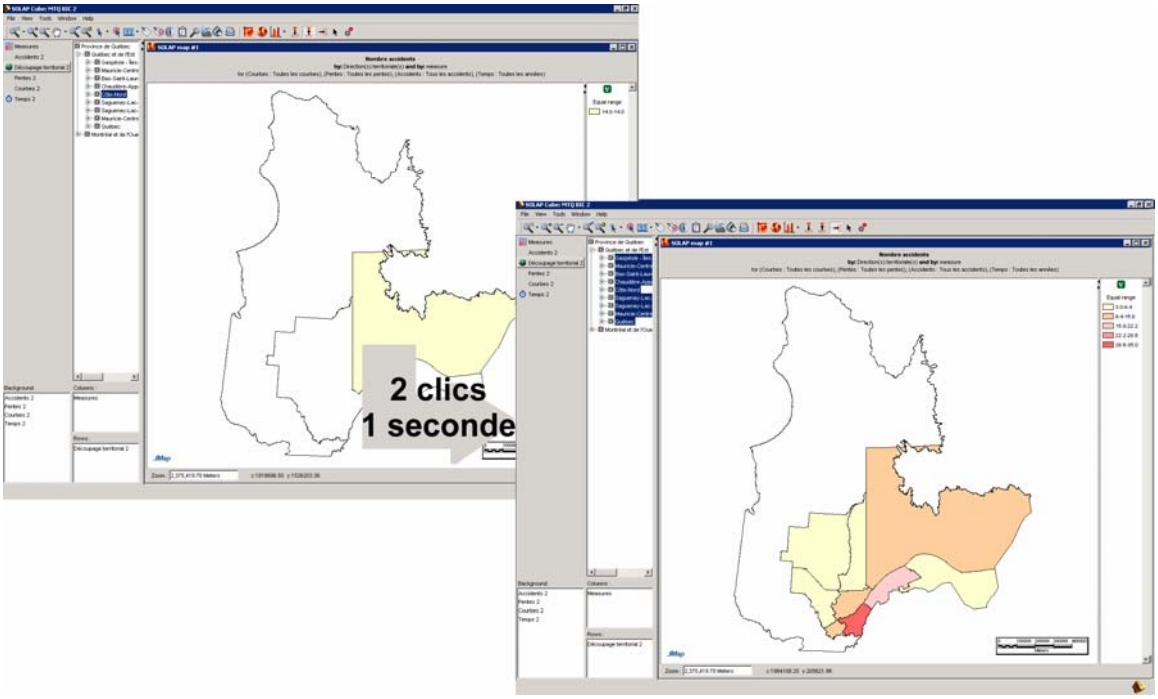

Figure 3. Un exemple d'une opération de forage latéral sur une Direction territoriale, montrant les éléments appartenant au même « parent », la même Direction générale.

L'outil SOLAP développé permet deux types de synchronisation : la synchronisation des opérations de forage d'un affichage à l'autre et la synchronisation de la sémiologie graphique d'un affichage à l'autre. Le premier type permet aux utilisateurs de visualiser, simultanément ou non (c'est-à-dire dans une ou plusieurs fenêtres) la même information, mais selon une perspective différente. Une opération exécutée dans un affichage peut être répercutée dans les autres affichages si désiré. La synchronisation de la sémiologie graphique d'un affichage à l'autre permet de conserver les mêmes couleurs, trames, etc. entre les cartes et avec les tableaux et diagrammes, facilitant ainsi l'identification et l'interprétation des données. La figure 4 présente un exemple de synchronisation des affichages et de la sémiologie graphique.

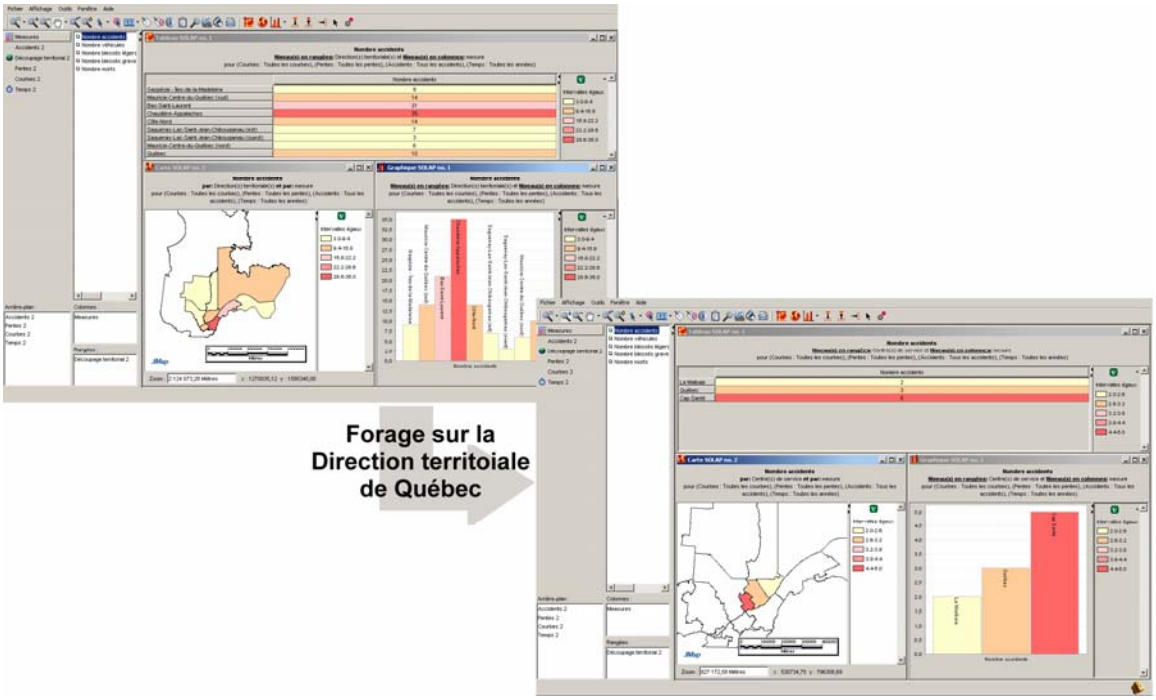

Figure 4. Une interface SOLAP présentant l'utilisation de la synchronisation des affichages et de la sémiologie graphique.

L'interface SOLAP présente des légendes interactives permettant l'utilisation d'opérations de forage sur la classification utilisée pour représenter les données. La figure 5 illustre ce type d'interactivité.

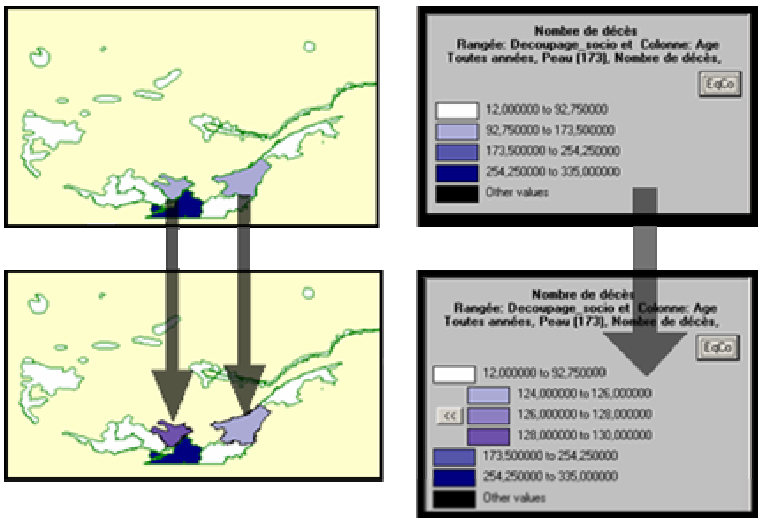

Figure 5. Une légende SOLAP présentant l'utilisation du forage sur la classification utilisée pour représenter les données. D'après (Pastor, 2004).

Exemple d’analyse

Cette section présente un exemple simple d’analyse afin de démontrer la facilité d’utilisation de l’outil SOLAP (données simulées). Un utilisateur désire visualiser le nombre d’accidents en fonction du genre d’accident et de la gravité d’accident pour les routes nationales de la Direction générale de Québec et de l’Est et de ses directions territoriales et en vérifiant comment varie ce nombre entre les années 1995 et 1999.

Supposons qu’un cube spécifique ait été défini pour réaliser cette analyse. La première étape, pour l’utilisateur, consiste donc à sélectionner, à l'aide de la souris, les différents membres de dimensions correspondant aux différents paramètres de départ :

- la mesure « Nombre d’accidents »
- le membre « Direction générale de Québec et de l’Est » de la dimension du découpage territorial
- le membre « 1995 » de la dimension du temps
- le membre « Nationales » de la dimension de classification fonctionnelle

Un clic de souris permet d’afficher rapidement une carte présentant le nombre d’accidents pour les routes nationales de la direction générale et pour l’année choisie. Par la suite, si l'utilisateur désire naviguer dans sa base de données pour avoir une vue plus générale ou comme ci-après plus de détails, un simple clic sur le polygone de la Direction générale permet d’effectuer une opération de forage et ainsi afficher les directions territoriales de l’Est. Cette nouvelle carte permet d’apprécier la répartition des nombres d’accidents sur les routes nationales pour les territoires de directions territoriales (voir figure 6). La sélection des membres « 1996 », « 1997 », « 1998 » et « 1999 » de la dimension du temps (une multisélection) permet d’afficher les cartes correspondant à ces années (une carte par année) pour une comparaison rapide des accidents de chacune de ces années. L’utilisateur peut ensuite afficher un tableau contenant cette même information (voir figure 7). Un forage dans la dimension des genres d’accidents permet de visualiser le nombre d’accidents pour chacun des différents genres d’accidents. Un forage dans la dimension des gravités d’accidents permet de visualiser le nombre d’accidents pour chacune des différentes gravités d’accident. Cet exemple sera illustré en détails par une démonstration de l’outil en kiosque lors de la conférence.

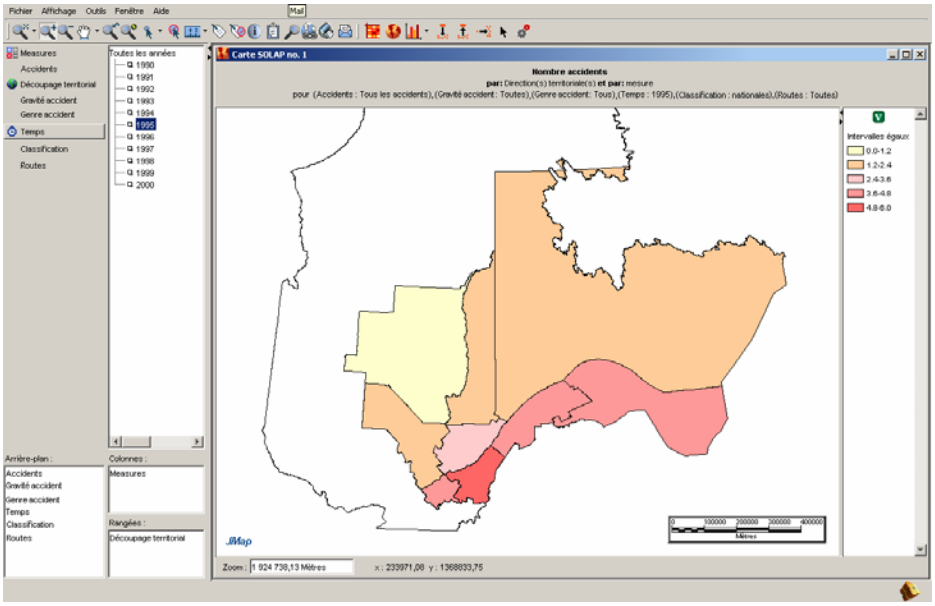

Figure 6. Carte de la répartition des nombres d'accidents sur les routes nationales pour les territoires de direction générales de l'Est, en 1995.

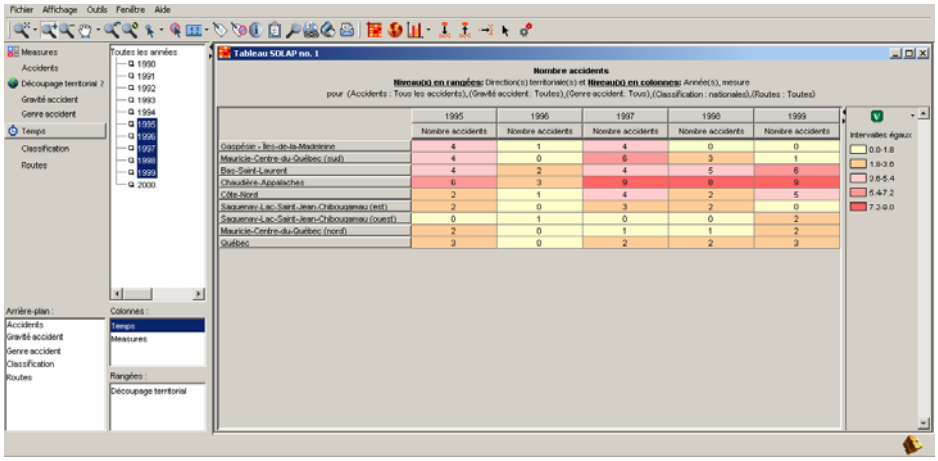

Figure 7. Tableau de la répartition des nombres d'accidents sur les routes nationales pour les territoires de directions territoriales de l'Est, de 1995 à 1999.

Cette analyse nécessite quelques clics de souris et quelques secondes seulement, voire aucune utilisation du clavier sauf pour les sélections multiples où la touche « shift » est nécessaire. Sans un tel module SOLAP, aucun SIG ne peut offrir une telle facilité et rapidité. L'alternative d'adjoindre un OLAP en tandem à un SIG représente ce que nous devions faire par le passé et une telle solution nécessitait beaucoup de programmation maison sans nécessairement atteindre la même flexibilité. Le développement d'un logiciel SOLAP générique adaptable à différentes applications permet donc d'améliorer la qualité du résultat tout en réduisant de façon majeure les efforts de développement et en éliminant presque les risque de malfonctionnement.

Conclusion

Cet article a présenté les résultats d'un projet de développement d'un outil SOLAP générique, c'est-à-dire applicable à divers domaines. Le SOLAP est défini comme : « un logiciel de navigation facile et rapide dans les bases de données spatiales qui offre plusieurs niveaux de granularité d'information, plusieurs thèmes, plusieurs époques et plusieurs modes de visualisation synchronisés ou non : cartes, tableaux et diagrammes ». Ce projet a débuté en septembre 2003 et se termine à l'automne 2004. Le résultat est l'implantation du logiciel, prêt pour la commercialisation, au MTQ comme outil d'exploitation spatio-temporelle des données de l'entrepôt de l'organisation.

Cette première version du logiciel (extension SOLAP de JMap) pourra ultérieurement profiter des résultats des recherches effectuées dans le cadre de la Chaire industrielle en bases de données décisionnelles géospatiales (octroyée par le Conseil de recherches en sciences naturelles et en génie du Canada et neuf partenaires) obtenue par le Dr Yvan Bédard et dont l'un des axes s'intéresse aux outils d'exploitation des bases de données décisionnelles.

Remerciements

L'équipe en bases de données spatiales du Centre de recherche en géomatique de l'Université Laval et KHEOPS Technologies tiennent à remercier le ministère des Transports du Québec

(MTQ) pour son soutien financier lors de ce projet. Nous tenons également à remercier le CRSNG et le Réseau de centres d'excellence en géomatique GEOIDE pour avoir supporté une part importante de la recherche en SOLAP.

Références

Bédard Y. 2004. Amélioration des capacités décisionnelles des SIG par l'ajout d'un module SOLAP (Spatial On-Line Analytical Processing). Université Aix-Marseille, École Polytechnique Universitaire de Marseille, Filière Génie Industriel et Informatique, 8 avril (Invited Professor).

Bédard, Y., Merrett, T., & Han, J. 2001. Fundamentals of spatial data warehousing for geographic knowledge discovery. In H. Miller & J. Han (Eds.), Geographic Data Mining and Knowledge Discovery. London: Taylor & Francis.

Caron, P.-Y. 1998. Étude du potentiel de OLAP pour supporter l'analyse spatio-temporelle. Mémoire de M.Sc., Département des sciences géomatiques, Faculté de foresterie et géomatique, Université Laval, 132 p.

Codd, E.F., Codd, S.B. & Salley, C.T. 1993. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. Hyperion white papers, 20 p.

Franklin, C. 1992. An Introduction to Geographic Information Systems: Linking Maps to Databases. Database, April, pp. 13-21.

Han, J., Stefanovic, N. & Koperski, K. 1998. Selective materialization: An Efficient Method for Spatial Data Cube Construction. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'98), Melbourne, Australia.

Pastor, J. 2004. Conception d’une légende interactive et forable pour le SOLAP. Mémoire de M.Sc., Département des sciences géomatiques, Faculté de foresterie et géomatique, Université Laval, 142 p.

SGIS. 2002. BIC-6000 – Besoins informationnels, 10 p.

Thomsen, E., Spofford, G. & Chase, D. 1999. Microsoft OLAP Solutions, John Wiley & Sons, 495 p.

Weldon, J.L. 1997. State of the Art: Warehouse Cornerstones, Byte, Janvier, p. 85.