
De la Conception d'un Entrepôt de Données Spatiales à un Outil Géo-Décisionnel pour une Meilleure Analyse du Risque Routier

Khalissa Derbal Amieur* — **Ibtissem Frihi**** — **Kamel Boukhalfa***
— **Zaia Alimazighi***

(*) Université des Sciences et de la Technologie Houari Boumediene (USTHB)

Email : {kderbal,kboukhalfa,zalimazighi}@usthb.dz

(**) Université M'hamed Bougara, Boumerdes (UMBB)

Email : i.frihi@gmail.com

RÉSUMÉ. Fléau des temps modernes, le phénomène des accidents de la circulation obère lourdement les ressources de la société. La route fauche tous les ans des milliers de vies et handicape des dizaines de milliers de blessés. Ce nombre important de victimes nécessite l'étude de ce phénomène pour en extraire tous les facteurs d'influence et permettre ainsi de proposer des mesures préventives. La plupart des études actuelles ignorent l'intégration de l'information géographique et notamment l'objet spatial Route dans l'analyse de ce phénomène. Nous présentons dans ce travail, une solution géo-décisionnelle basée sur un entrepôt de données spatiales et un outil géo-spatial intégrant l'objet spatial route comme un axe d'analyse du risque routier. Nous présentons dans la partie expérimentation les principales fonctionnalités de notre outil.

ABSTRACT. Plague of modern times, the phenomenon of the traffic accidents hampers heavily resources of the society. The road mows every year thousands of deaths and tens of thousands wounded. This frightening number of victims requires a study of this phenomenon in order to extract all indicators affecting this problem and allow proposing preventive measures. Most current studies ignore the integration of geographic information including the geographic object "Road" in the analysis of this phenomenon. We present in this work, a geospatial decision making solution based on a spatial data warehouse and a geospatial tool integrating the geographic object "road" as an analysis dimension. Some functionalities of our developed geodecisional tool are described in experiment section.

MOTS-CLÉS : Risque Routier, Entrepôt de Données Spatiales, Outil Géospatial

KEYWORDS: Road Risk, Spatial Data Warehouse, Geospatial Tool

1. Introduction

Les entrepôts de données sont généralement destinés aux décideurs d'entreprises souvent non spécialistes de l'informatique (Royer *et al.* 2012). Ils doivent donc être munis d'interfaces graphiques adaptées à ce type d'utilisateurs. L'objectif de ces entrepôts est de servir de support pour la prise de décision en permettant une meilleure exploitation des informations contenues dans les systèmes opérationnels des entreprises. Cependant, ces données ont parfois, *un aspect géographique* et servent souvent à représenter et à analyser certains phénomènes sur une carte. Par conséquent, l'intégration de l'information géographique dans un entrepôt de données devient de plus en plus indispensable au processus de prise de décision (Bimonte 2007).

L'analyse d'un phénomène nécessite l'exploration des différents facteurs qui interviennent dans sa description. Ces phénomènes sont généralement caractérisés par une complexité du point de vue quantité d'informations manipulées et instabilité dans le temps. Les utilisateurs d'une telle information notamment les décideurs ont besoin donc d'une synthèse sous forme d'indicateurs perceptibles et facilement analysables en vue d'une prise de décision rapide et efficace. Ceci est rendu possible grâce aux différents types de diagrammes auxquels s'ajoute l'information cartographique. Parmi les phénomènes étudiés : *gestion du trafic routiers* (Bauzer-Medeiros *et al.* 2006), *feux de forêts* (Miquel *et al.* 2002a), *protection de l'environnement* (Bimonte 2007), *études pour l'agriculture* (Nilakanta *et al.* 2008, Bimonte *et al.* 2012b), les accidents de la route, etc. Ce dernier phénomène constitue un enjeu humain considérable pour la société. Chaque année, en Algérie et partout dans le monde on estime à des milliers de personnes tuées et d'autre blessées dans des accidents de la route. En effet, en Algérie, le Centre National de Prévention et de Sécurité Routière (CNPSR¹), chargé de l'étude et de l'analyse de ce phénomène, publie périodiquement des rapports et des statistiques sur ce phénomène. Le tableau 1, représente une étude comparative des statistiques du nombre de décès durant les huit premiers mois de 2011 et 2012². Ces statistiques indiquent clairement l'état alarmant de ce phénomène.

	Zone urbaine	Zone rurale	Totale
Janvier- Août 2011	509	2516	3025
Janvier- Août 2012	477	2578	3055
Différence	-32	62	30
%	-6,29	2,46	0,99
Moyenne quotidienne (2011)	02	10	12
Moyenne quotidienne(2012)	02	11	13

Tableau 1. *Le nombre de morts en Algérie (Les Huits premiers mois de 2011-2012)*

Le risque routier est devenu de plus en plus un problème très inquiétant enregistrant une nette évolution en Algérie et à travers le monde. Ce dernier, ne cesse de susciter l'intérêt de l'ensemble des décideurs de différents organismes publics et

1. <http://www.cnpsr.org.dz/>

2. <http://www.cnpsr.org.dz/dwld/stat%208%20mois%202012.pdf>

socio-économiques car le taux d'accidents est préoccupant et dépasse les efforts fournis dans tous les pays du monde (Bauzer-Medeiros *et al.* 2006, Bertini *et al.* 2005). Toutes ces considérations, nous ont motivé à étudier ce phénomène et à proposer des solutions décisionnelles contribuant à son analyse. Le risque routier, résulte d'un ensemble de facteurs énumérables connus par tous comme *le conducteur, le véhicule, la route et les conditions de circulation* (Houria 2010, Bauzer-Medeiros *et al.* 2006). Étant donnée que la route est identifiée en tant qu'objet géographique, le *facteur route* est donc porteur de cette information en plus de la zone pouvant être de différentes nature (rurale, urbaine, etc.) et constitue généralement le fond de la carte. Une route est constituée d'un ensemble de tronçons caractérisés par leurs localisations (coordonnées spatiales) et un ensemble de données descriptives. Toutefois l'analyse de ce phénomène est complexe et nécessite une étude étalée dans le temps en mettant en évidence la route comme un objet géographique visualisable sur la carte. Cela nécessite d'intégrer cette composante géographique dans l'entrepôt de données afin de produire des analyses avec des résultats visualisables sur la même carte (Bimonte *et al.* 2010, Bimonte *et al.* 2012a). Plusieurs travaux ont été consacrés à l'étude du problème du risque routier dans le contexte des systèmes d'information géographique (Bauzer-Medeiros *et al.* 2006, Houria 2010, Pirotti *et al.* 2011). Cependant, peu de travaux ont traité ce problème dans le contexte des entrepôts de données spatiales.

Les rapports publiés par le CNPSR sont d'une importance extrême pour les décideurs en risque routier, néanmoins ils présentent quelques insuffisances. Les statistiques publiées sont élaborées par rapport à des critères présentant une grande corrélation comme la nature de la zone (rurale ou urbaine), le parc de véhicules, le nombre de décédés et de blessés par tranches d'âges, par sexe, etc. Ces statistiques sont reportées sur des tableaux ou sur des histogrammes indépendants, ce qui rend l'analyse du phénomène par prise en compte simultanée de ces différents facteurs, ce qui est souvent indispensable, complexe voire impossible. De plus, ces rapports négligent certains facteurs prépondérants du phénomène de risque routier (acteurs du phénomène) comme les causes d'accidents. En effet, les informations concernant le conducteur et l'état de la route par exemple, reportés généralement par les autorités (gendarmerie, protection civile etc.), ne sont pas impliquées directement dans l'étude. Par ailleurs, la représentation symbolique de l'objet route sur la carte est d'un impact réel, car elle permet de prendre en charge plusieurs critères de visualisation implicites et explicites à savoir : la nature de la route (double sens, autoroute, à double circulation, etc), la zone concernée (nord, sud, rurale, urbaine, montagnaise, etc). Cependant, l'information géographique est généralement sous-employée dans le processus décisionnel notamment dans les entrepôts de données, d'où la nécessité de couplage entre les systèmes OLAP classiques et les Systèmes d'Informations Géographiques (SIG) afin d'aboutir à un système OLAP Spatial (SOLAP) au sein d'un entrepôt de données spatiales (Bimonte 2007, Bimonte *et al.* 2012a).

Pour remédier aux insuffisances citées ci-dessus et contribuer à l'analyse de ce phénomène mondial, nous proposons la conception d'un Entrepôt de Données Spatiales (EDS) intégrant l'objet spatial route et supportant un outil géo-décisionnel que

nous avons mis en œuvre. Ce dernier permettra de consolider l'analyse du phénomène par la carte interactive et faciliter ainsi la prise de décision.

Le présent papier est organisé comme suit. Dans la section 2, nous passons en revue quelques concepts de base liés aux entrepôts de données spatiales. La section 3 sera consacrée à l'exploration de quelques travaux de recherche liée à la problématique abordée. La section 4 propose notre démarche de conception de l'EDS. La section 5 présente les différents types de requêtes permettant de répondre aux besoins des décideurs. La section 6 présente l'outil décisionnel que nous avons mis en œuvre. Nous terminons par une conclusion et des perspectives.

2. Concepts de Base

2.1. Entrepôt de données spatiales

L'Entrepôt de Données (ED) a été formalisé en 1990 par Bill Inmon. Il s'agit d'une collection de données orientées sujet, intégrées, et contenant des informations historiées, non volatiles et exclusivement destinées aux processus d'aide à la décision (Inmon 1992). En effet, dans un ED, il est nécessaire de conserver l'historique de la donnée. Ainsi, une même requête effectuée à deux mois d'intervalle en spécifiant la date de référence de la donnée, donnera le même résultat.

Un Entrepôt de Données Spatiales (EDS) est une collection de données spatiales et non spatiales : orientées-sujet, intégrées, variantes dans le temps, et non volatiles dédiées à la prise de décision spatiale (Stefanovic *et al.* 2000). Un EDS est une reformulation d'un ED conventionnel à la seule différence d'intégrer des données de nature spatiale. Ces dernières imposent de nouveaux opérateurs de navigation et une reformulation des concepts classiques de dimensions et de mesures pour prendre en compte la composante spatiale (Ruiz *et al.* 2009).

Une dimension correspond à l'axe de l'analyse, elle est décrite par une ou plusieurs hiérarchies. Un fait modélise le sujet de l'analyse, il est formé de mesures correspondantes aux informations de l'activité analysée. Par exemple, la gestion des commandes peut être analysée selon les dimensions : Client, Magasin, et Temps. Les dimensions se présentent sous forme d'une liste d'éléments organisés de façon hiérarchique. Dans les entrepôts de données spatiales, les dimensions et les mesures peuvent contenir des composantes spatiales. Il existe trois types de dimensions spatiales : (1) les dimensions géométriques, (2) les dimensions non géométriques et (3) les dimensions mixtes (voir figure 1).

– *Géométriques* : chaque élément de la dimension est associé à un objet géométrique.

– *Non géométriques* : dans ce cas, la dimension contient uniquement des données non spatiales. Les données n'ont que la valeur nominale de la référence spatiale (nom du pays, nom de l'état, nom de la ville).

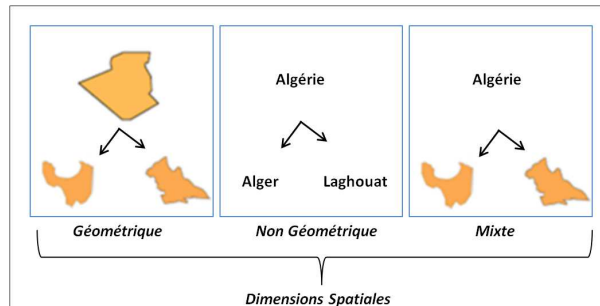


Figure 1. Les trois types de dimensions spatiales.

– *Mixtes* : qui contiennent en même temps des données géométriques et non géométriques.

Les mesures d'un fait sont généralement numériques et fournissent une description quantitative du fait. Par exemple, pour la gestion des commandes, les mesures peuvent être *la quantité du produit commandé* et *le montant de la commande*. Un fait est associé à une ou plusieurs combinaisons de membres des dimensions. Certaines mesures peuvent être calculées à partir d'autres mesures ou propriétés de membres. Elles sont appelées *mesures dérivées* (Blaschkaand *et al.* 1998). Les mesures spatiales sont le résultat d'un calcul obtenu par l'application d'un opérateur spatial, elles peuvent être sous les trois formes suivantes : *objets géométriques* (polygones, cercles, points etc.), *valeurs métriques* (surface, distance, longueur) ou encore de *pointeurs spatiaux*. Dans ce dernier cas, les formes géométriques sont stockées dans un fichier à part et ceci lorsque la technologie utilisée ne permet pas de les stocker dans la structure de l'entrepôt données.

2.2. Dimension spatiale et hiérarchie

Un EDS relationnel hérite de plusieurs composantes des entrepôts de données classiques, tels que les tables de faits, de dimensions, de mesures numériques et les hiérarchies qui agrègent ces mesures selon plusieurs niveaux de granularité. Les dimensions spatiales doivent être organisées selon plusieurs niveaux de granularité ou selon différents niveaux de détails pour supporter les opérations SOLAP comme roll-up et drill-down. SOLAP définit des hiérarchies de dimensions spatiales par le biais d'inclusions ou de relations topologiques d'intersection (Chen *et al.* 2011, Praing *et al.* 2009) et ne refèrent pas la sémantique des relations hiérarchiques. Cependant, les objets géographiques peuvent être liés par des relations thématiques, spatiales et de généralisation cartographique (Weibel *et al.* 1999, Bernier *et al.* 2007, McMaster *et al.* 1992, Sabine *et al.* 1995, Sabine 1998, Bimonte 2007, Bimonte *et al.* 2012a, Bedard 2012).

La généralisation est la sélection et la représentation simplifiée de détail adapté à une échelle et/ou à un objectif de la carte (Derbal *et al.* 2012). La généralisation est utile pour différentes tâches, comme par exemple pour réduire la quantité inutile d'informations, elle assure les fonctionnalités du zoom intelligent, etc. Une hiérarchie est spatiale s'il y a au moins un niveau qui contient la composante spatiale. Celui-ci est dit niveau spatial. Une hiérarchie spatiale peut être totalement spatiale si tous les niveaux sont spatiaux, ou partiellement spatiale s'il y a au moins un niveau non spatial.

2.3. OLAP Spatial

Les systèmes SOLAP sont des technologies utilisées dans le domaine de l'informatique décisionnelle ou Business Intelligence (BI) (Bimonte *et al.* 2010, Bimonte *et al.* 2012a). Ils permettent aux décideurs d'explorer et d'analyser une grande quantité de données géo-référencées, généralement par le biais d'interfaces utilisateur interactives de navigation. Ils permettent d'organiser l'information selon le modèle spatio-multidimensionnel. SOLAP permet l'analyse des données numériques et spatiales selon plusieurs dimensions, qui sont organisées en hiérarchies thématiques et spatiales. Cette technologie est appliquée dans plusieurs domaines d'application (par exemple, des risques environnementaux, santé, etc). L'un des premiers chercheurs qui a introduit ce concept est *Yvan Bédard* et son équipe. Ils ont défini OLAP spatial (SOLAP) comme *une plate-forme visuelle construite pour supporter l'analyse spatiotemporelle rapide et facile et l'exploration de données dans une approche multidimensionnelle à plusieurs niveaux d'agrégation disponible sur les cartes, les tableaux et les diagrammes* (Bedard *et al.* 2005, Bedard *et al.* 2007). Plusieurs travaux de recherche ont été consacrés à l'étude des outils SOLAP pour la combinaison des différentes technologies OLAP et SIG, modélisation de cubes de données spatiales, indexation spatiale, agrégation spatiale, la matérialisation des cubes de données géo-spatiales (Han *et al.* 1998, Stefanovic *et al.* 2000, Lu *et al.* 2000).

3. Travaux Connexes

Les systèmes décisionnels ont été développés pour répondre aux préoccupations des entreprises telles que l'exploitation efficace de l'important volume d'informations, provenant soit de leurs systèmes opérationnels, soit de leur environnement extérieur. La plupart de ces systèmes reposent sur un espace de stockage centralisé, appelée entrepôt de données ; son rôle est d'intégrer et de stocker l'information utile aux prises de décisions et de conserver l'historique des données pour supporter les analyses effectuées. L'évolution de la cartographie numérique a fait que cette information soit disponible, et son usage s'est répandu dans tous les domaines. La carte est devenue désormais, un moyen incontournable dans l'interprétation de ces phénomènes, elle facilite la perception de leurs propriétés spatiales et aide à leur compréhension. (Han *et al.* 1998, Stefanovic *et al.* 2000) ont été les premiers à proposer un modèle multidimensionnel en étoile pour les entrepôts de données spatiales. Les dimensions dans

ce cas, peuvent être spatiales, non spatiales ou bien des données numériques. Les auteurs dans (Miquel *et al.* 2002b), traitent le problème d'intégration des données hétérogènes (temporelle ou spatiale) dans un entrepôt de données afin d'en extraire des connaissances géographiques. Ceci peut être effectué par l'exploration des données détaillées associées à une période et des études temporelles sur des données intégrées et comparatives. Deux approches ont été proposées pour modéliser les données spatio-temporelles dans des structures multidimensionnelles. La première consiste à utiliser un cube unique pour intégrer toutes les données, et la deuxième consiste à créer un cube spécifique pour chaque période à analyser. Dans (Malinowski *et al.* 2004), les auteurs visent à améliorer l'analyse et la conception de l'EDS et OLAP (SOLAP) en intégrant les systèmes de composantes spatiales dans un modèle multidimensionnel à savoir les dimensions spatiales et les mesures spatiales. Plusieurs scénarios ont été présentés à savoir les dimensions spatiales avec une hiérarchie spatiale. En outre, plusieurs dimensions peuvent partager une partie de la hiérarchie spatiale, lorsque plus d'une dimension spatiale est représentée dans un modèle multidimensionnel, une relation topologique entre ces deux est nécessaire. Cette relation est fondée sur un prédicat spatial. Dans ce papier, une mesure peut être représentée par une géométrie, ou bien comme le résultat d'opérateurs spatiaux. Dans ce type d'application spatio-multidimensionnelle, la table de faits définit une jointure spatiale entre les dimensions spatiales. Cette table de faits permet d'avoir une vision multidimensionnelle des relations spatiales. Elle autorise le calcul des mesures spatiales pour les membres de tous les niveaux des hiérarchies des dimensions, comme par exemple les parties des autoroutes qui passent par des villes ou les parties des autoroutes qui passent par des pays.

(Bertini *et al.* 2005) a établi un entrepôt à des fins de stockage et de visualisation. La visualisation se base sur l'utilisation de couleurs permettant de différencier la congestion du trafic de sa fluidité. Elle relève de la fonction de suivi du trafic routier et ne concerne pas les fonctionnalités de décision et de contrôle. (Bauzer-Medeiros *et al.* 2006) a proposé une démarche multidisciplinaire pour le traitement de masses de données spatio-temporelles dans le domaine du trafic routier. Cette démarche intègre un entrepôt de données à des fonctionnalités d'analyse et de représentation multi-échelle numérique et symbolique. La plupart des systèmes d'information pour le trafic routier utilise des bases de données classiques pour la gestion des données, ce qui limite les fonctionnalités offertes aux utilisateurs. Le travail que nous présentons dans ce papier, vient compléter ces fonctionnalités classiquement connues dans les systèmes décisionnels par une visualisation cartographique afin d'aider à mieux interpréter, analyser et par conséquent à prendre de bonnes décisions.

Sandro Bemonte (Bimonte 2007) a proposé un modèle formel *GeoCube* et une algèbre associée, afin d'introduire les aspects sémantiques et spatiaux de l'information géographique dans l'analyse multidimensionnelle. L'algèbre proposée permet de fournir les opérateurs de forage et de coupe, et des opérateurs de navigation au sein de la hiérarchie de mesures. Cette algèbre permet, grâce aux opérateurs qui modifient dynamiquement la structure de l'hypercube, de concilier analyse OLAP et analyse spatiale.

Ainsi un prototype web *GeWolap* conforme au modèle proposé est réalisé. Les données proviennent de l'étude environnementale effectuée dans *la lagune de Venise*.

4. Le Modèle Conceptuel de l'EDS Proposé

Afin de fournir aux décideurs un moyen d'analyse efficace, dans leur processus d'analyse d'un phénomène aussi important que le risque routier, nous avons intégré la donnée géographique dans un entrepôt de données supportant le système décisionnel développé. Nous proposons dans cette section le modèle conceptuel de l'EDS mis en œuvre. Comme nous l'avons déjà évoqué, dans cette étude nous nous sommes basés sur des rapports d'analyse publiés actuellement par le CNPSR et qui présentaient des insuffisances marquantes (voir section 2). De ce fait, notre modèle proposé prend en charge ces insuffisances particulièrement par l'intégration de données spatiales à travers la dimension *route* qui constitue un facteur prépondérant dans l'analyse. Les décideurs peuvent ainsi tirer profit des avantages de la carte et peuvent par conséquent exploiter la carte de façon interactive. Des exemples sont présentés dans la section consacrée à la mise en œuvre de notre système (voir section 6). En effet, les SIG actuels se concentrent sur l'aspect topographique, en revanche notre système ajoute à cet aspect des mesures extraites de l'entrepôt et qui seront affichées sur la carte. Nous menons une conception orientée objet basée sur le langage UML.

Les risques routiers sont apparus d'un fait dommageable pour un usager de la route, il peut s'agir de risque d'accident, de blessure ou de mort. Un accident de la route est un choc qui a lieu sur le réseau routier entre un engin roulant (automobile, moto, vélo, etc.) et toute autre chose ou personne et qui engendre des blessures humaines et/ou des dégâts matériels. Ces dégâts sont soit des dégâts occasionnés aux véhicules, à un élément de la route (chaussée, panneaux, barrières de protection, etc.) ou à un élément extérieur à celle-ci (bâtiment, mobilier urbain, cabine de téléphone). De nombreux facteurs contribuent aux risques de collision tel que le type de véhicule, la vitesse de circulation, l'environnement routier, l'expérience du conducteur, le sexe mais aussi son comportement et son état physique. Deux méthodes ont été utilisées afin de collecter les besoins des utilisateurs de notre entrepôt de données : (1) une collecte orientée source, qui se limite à identifier les besoins de l'ED à partir de l'ensemble des données disponibles au niveau des sources et (2) une collecte orientée besoins qui identifie les besoins des utilisateurs et des décideurs. L'analyse des besoins effectuée a été en majeure partie basée sur des documents fournis par le CNPSR. Nous avons donc identifié les acteurs et les objets qui interviennent dans un accident.

Notre entrepôt est modélisé par un schéma en fonction de neige (voir figure 2), où les dimensions sont organisées en hiérarchie comme la dimension *Temps*, *Route* et *Conducteur*. Chacun des membres appartient à un niveau hiérarchique (ou niveau de granularité) particulier. Par exemple, la hiérarchie de la dimension *Temps* commence par le moment de l'accident représenté par l'attribut *ID_Temps*, ensuite l'heure, le jour, le mois et année. Notre modèle est composé d'une grande table de faits et un ensemble de tables de dimensions reliées par des clés étrangères. La table de faits dans

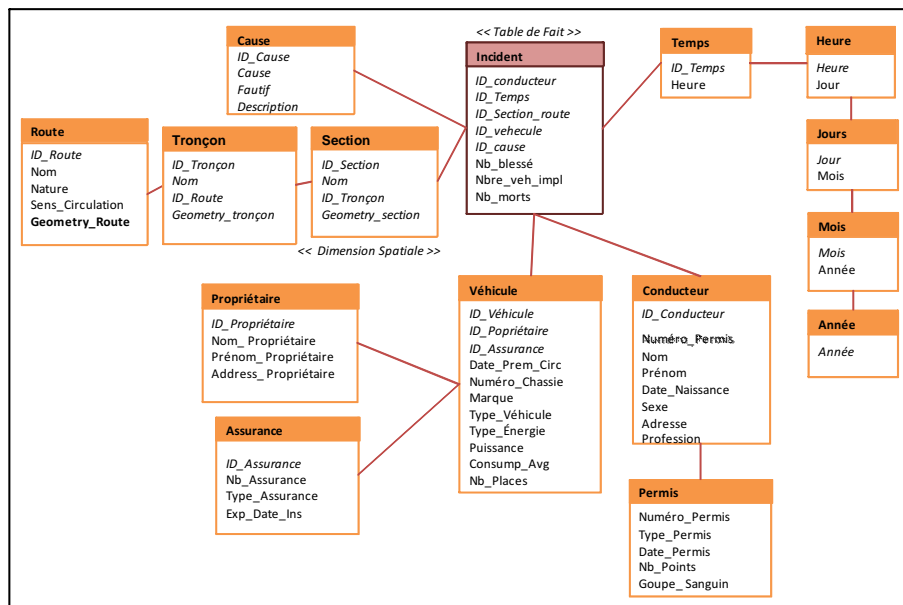


Figure 2. Le modèle conceptuel de l'EDS développé

notre cas, est représentée par la table *Incident*. Elle contient les informations omniprésentes (mesures) dans la description d'un accident tel que *le nombre de victimes, le nombre de blessés, le nombre de véhicules impliqués dans l'accident*. L'information géographique est interprétée par la dimension *Route*. Cette dernière est décrite par des informations descriptives (*nom, nature, sens de circulation*) et par l'information géométrique qui est la géométrie de l'objet (il s'agit dans ces cas d'une poly-ligne représentée en mode vecteur par un ensemble de points $\{(x_i, y_j)\}$). La dimension *Route* est décrite via une hiérarchie géographique, où le niveau le plus bas est représenté par une *Section*. Un ensemble de sections forment un tronçon et un ensemble de tronçons forment une *Route*. La table de dimension *Cause* contient des informations descriptives *Cause, Description* et *Fautif* afin de distinguer entre les conducteurs ayant provoqué l'accident et les victimes.

Les informations relatives aux conducteurs peuvent être divisées en deux catégories : (1) *des informations standards* à savoir nom, prénom, adresse, numéro de permis et (2) *des informations utiles à la prise de décision* à savoir la profession, date d'obtention du permis et le genre. Les informations concernant les véhicules sont également classées en deux catégories : (1) *des informations de description générale* telles que : immatriculation, nom, prénom et adresse du propriétaire du véhicule et numéro de châssis du véhicule et (2) *et d'autres utiles à la prise de décisions* telles que : le type, date de sortie de première circulation, marque, énergie et puissance, etc.

5. Traitement de Requêtes SOLAP

Le défi majeur auquel est confronté le traitement des requêtes SOLAP est de réduire le temps de réponse aux requêtes ad hoc de type *windows query*, que l'utilisateur génère lors de la manipulation de la carte, tout en minimisant le nombre de jointures entre la table de faits et les tables de dimension. Nous proposons dans cet article, deux types de requêtes (1) *un ensemble de requêtes prédéfinies* et *des requêtes de type windows-query*. L'ensemble des requêtes prédéfinies permet de répondre aux besoins liés à l'usage de l'information géographique. En effet, à l'issue de notre étude conceptuelle et en se basant sur les rapports mis en ligne par CNPSR, nous avons recensé un répertoire de questions auxquelles peut répondre notre entrepôt de données. Ces requêtes sont utilisées afin de faciliter le processus de prise de décision. Parmi ces requêtes nous citons : *les routes les plus dangereuses marquées par un grand nombre de victimes d'accidents* (plus de détails sont donnés dans la section 6). Répondre efficacement à ces requêtes nécessite l'appel à l'information cartographique via une carte affichée à l'écran permettant de visualiser les routes ou les tronçons de routes impliqués dans l'incident. Cela est réalisé à travers *des requêtes de type windows-query* ou l'utilisateur peut interroger directement la carte : *nombre de blessés d'une ville sélectionnée sur la carte, nombre de morts sur une route particulière, un tronçon particulier*, etc.

6. Mise en Œuvre de l'Outil Géo-Décisionnel

Le développement de notre prototype géo-décisionnel, dédié au risque routier a été réalisé dans l'environnement *Oracle 11g* qui intègre *Oracle Spatial* via ses différents composants : *WebLogic Server* et *Jdeveloper* (voir figure 3) : (i) L'outil de visualisation *Oracle MapViewer* permet d'afficher la carte géographique et les résultats des différentes requêtes spatiales exécutées sur la carte (les exemples sont donnés dans les paragraphes suivants), (ii) L'outil *Oracle MapBuilder* a été utilisé pour charger les données géographiques dans le SGBD et la construction de la carte géographique, et cela par la création la carte de base, des styles et des thèmes (iii) *Jdeveloper* a été utilisé pour la réalisation de l'interface client et l'implémentation des différentes fonctionnalités du système.

Le jeu de données spatiales utilisé est en mode vecteur, il représente une région géographique située au nord d'Algérie caractérisée de zone urbaine dense, il s'agit de la commune de *Dar-El-Beida* dans la capitale *Alger*. Deux thèmes sont représentés dans notre entrepôt de données, *le thème routier* défini à un haut niveau de détail et *le thème bâti* défini un niveau moins détaillé car il sert de point de référence sur la carte. La route est représentée par des tronçons représentant un objet géographique stocké avec sa description sémantique et ses coordonnées.

Ce prototype décisionnel s'articule autour de deux acteurs principaux : *l'administrateur* et *le décideur*. Dans la figure 3, nous présentons l'architecture globale de l'application développée basée sur un EDS. Ce dernier et les méta-données de la carte

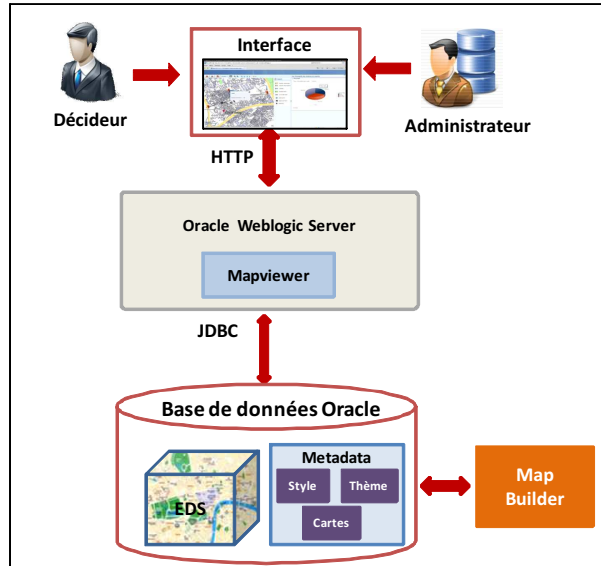


Figure 3. L'architecture de l'outil Géo décisionnel

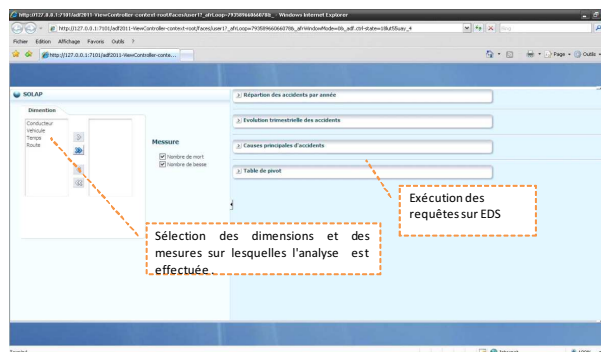


Figure 4. Présentation de l'interface principale de l'outil Géo-decisional proposé

sont stockés dans la *SGBD oracle spatiale*, le composant *Mapviewer* qui est déployé dans *Weblogic Server* se charge de l'affichage de la carte au client.

Les fonctionnalités de notre outil sont présentées à travers un ensemble d'interfaces. Deux catégories d'utilisateurs sont considérées, l'administrateur et le décideur. L'accès au système est assuré par un système d'authentification classique avec nom

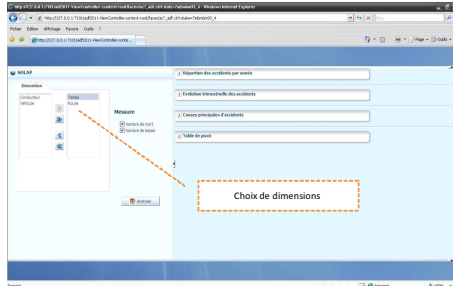


Figure 5. Sélection des Dimensions

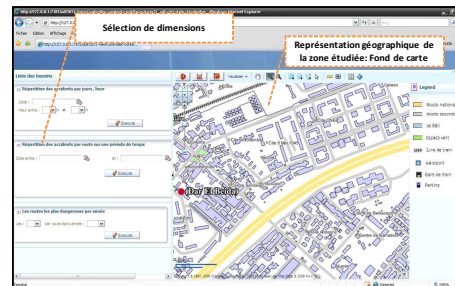


Figure 6. Exemple de requêtes générées

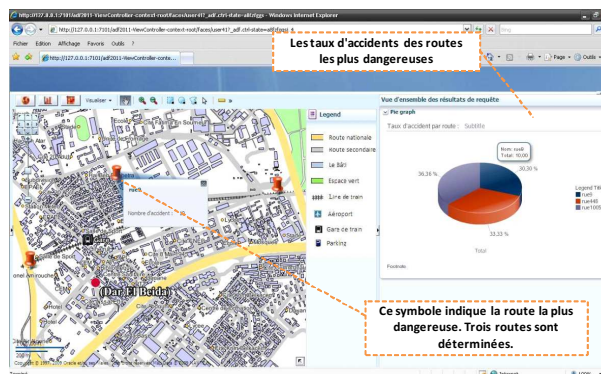


Figure 7. Affichage du résultat de la requête sur la carte ou en camembert

d'utilisateur et mot de passe. L'interface principale de notre application-décideur est présentée par la figure 4.

Fonctionnalités Destinées au Décideur :

Comme l'illustre la figure 4, le décideur peut exécuter des requêtes (présentées dans la partie de droite de l'interface) ou bien choisir une analyse plus détaillée, en choisissant les dimensions et les mesures sur lesquelles il veut effectuer l'analyse (sur le côté gauche de l'interface).

L'outil permet également au décideur de choisir les dimensions sur lesquelles il veut exécuter ces requêtes. La figure 5, représente les requêtes proposées selon deux dimensions : *temps* et *route* choisis par le décideur.

Par exemple, si l'utilisateur sélectionne deux dimensions : *Route* et *Temps*, l'outil permet de générer la requête appropriée. La figure 6, présente trois requêtes générées Q_1 , Q_2 et Q_3 comme suit :

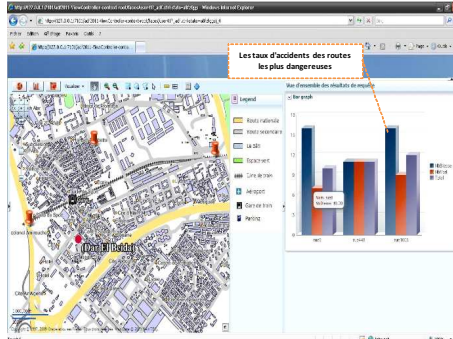


Figure 8. Affichage dans un Bar graphe les routes les plus dangereuses

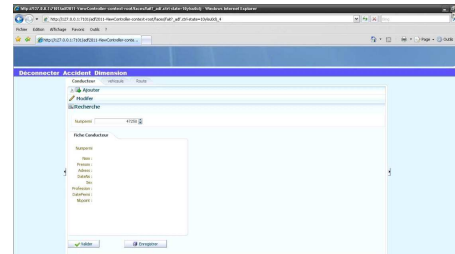


Figure 9. Recherche spécifique sur les tables de dimensions

- Q_1 : La répartition des accidents par année.
- Q_2 : La distribution des accidents par route sur une période de temps.
- Q_3 : Les routes les plus dangereuses par année.

Comme une étude de cas, nous avons considéré la requête Q_3 : *Les routes les plus dangereuses par année*. Notre outil géo-décisionnel permet au décideur de visualiser le résultat (les routes ciblées) de la requête sous différents types affichage :

- Affichage sur la carte avec une présentation sur un Diagramme en camembert comme l'illustre la figure 7.
- Affichage dans une table
- Affichage sous forme de bar graphe (voir figure 8)

Fonctionnalités Destinées à l'administrateur :

L'administrateur effectue des tâches liées à l'administration et la personnalisation de EDS sur lequel est basé notre outil décisionnel. Il peut ajouter des requêtes prédéfinies pour être exploitées par les utilisateurs. Il peut aussi modifier ou supprimer des requêtes existantes. L'administrateur peut mettre à jour les informations stockées dans l'EDS à savoir : l'ajout de données, la suppression de données ou la recherche d'une information particulière. A travers l'interface présentée par la figure 9, l'administrateur peut effectuer la recherche d'une donnée particulière dans l'entrepôt de données et ainsi vérifier sa validité. Pour cela certaines informations doivent être introduites par l'administrateur. Par exemple, effectuer une recherche par le numéro de permis pour afficher les informations du conducteur. L'administrateur peut aussi intervenir au niveau de la table des faits afin de compléter des informations concernant les accidents comme la section ou le tronçon de route où est survenu l'accident. En effet, lors de l'alimentation via l'ETL, seule la route peut être mentionnée. Nous donnons la possibilité à l'administrateur de compléter cette information en précisant la section de route où l'accident est survenu.

7. Conclusion

Nous avons présenté dans ce papier, la conception d'un entrepôt de données spatiales supportant l'outil Géo-décisionnel mis en œuvre, dédié au risque routier. La complexité de l'étude et l'analyse de ce phénomène est particulièrement due au nombre de facteurs intervenants ainsi qu'à leurs interactions. Bien que ces facteurs soient connus et énumérables, leur prise en compte simultanée lors de l'analyse influe considérablement sur le résultat présenté. En effet, les rapports publiés par CNPSR, contiennent des informations marquées par une richesse sur le plan statistique mais manquent d'orientation d'analyse où différents tableaux présentent des statistiques par willaya, ville et commune selon un facteur ou deux. De plus, nous avons constaté l'absence du facteur route, cette information géographique qui par sa caractéristique cartographique visualisable aidera le décideur dans son processus d'analyse. L'entrepôt de données spatiales que nous avons proposé intègre l'objet géographique route comme dimension spatiale indispensable pour une meilleure analyse du risque routier.

Notre outil permet de répondre à certaines requêtes décisionnelles complexes comme : les routes marquées par le plus d'accidents, les routes les plus meurtrières (celles qui marquent le nombre maximum sur un bilan périodique), routes faisant le plus de victimes etc. Il permet aussi la présentation des résultats sur une carte pour une meilleure analyse et exploitation. Ainsi il fournit une interface interactive permettant la formulation des requêtes sur la carte.

La mise en œuvre de cet outil a requis l'utilisation de beaucoup de technologies et d'outils intégrés dans la chaîne d'Oracle. Cependant, quelques pistes techniques et de recherche restent à explorer. Nous pouvons citer : (1) l'enrichissement de l'outil par l'implémentation des opérations d'agrégations sur les mesures, par exemple, suite à un zoom arrière, (2) le calcul automatique de mesures pour un ensemble de tronçons de route constituant la route entière et (3) l'adaptation des approches de généralisation définies dans le cadre des SIG aux les entrepôts de données spatiales (Bimonte *et al.* 2012a).

8. Bibliographie

- Bauzer-Medeiros C., Carles O., Devuyst F., Huguency B., Joliveau M., Jomier G., Manouvrier M., Naija Y., Scemama G., Stean L., « Vers un entrepôt de données pour le trafic routier », *EDA*, 2006.
- Bedard Y., « Origines, évolution et défis de la technologie SOLAP », *Conférencier invité*, 2012.
- Bedard Y., Rivest S., Proulx M.-J., « Spatial On-Line Analytical Processing (SOLAP) : Concepts, Architectures and Solutions from a Geomatics Engineering Perspective, Dans : Robert Wrembel & Christian Koncilia (ed(s)), *Data Warehouses and OLAP : Concepts, Architectures and Solutions*, Chap. 13 », *IBM Press-Idea Group*, 2007.
- Bedard Y., Rivest S., Proulx M.-J., Badard T., « Merging hypermedia GIS with spatial on-line analytical processing : Towards hypermedia solap », *Geographic Hypermedia : Concepts and Systems*, 2005.

- Bernier E., Bédard Y., « A Data Warehouse Strategy for on-Demand Multiscale Mapping », *Generalisation of Geographic Information*, 2007.
- Bertini R., Matthews S., Hansen S., Delcambre A., Rodriguez A., « ITS archived data user service in Portland, Oregon : now and into the future », *Intelligent Transportation Systems, 2005. Proceedings. IEEE*, p. 568 - 573, 2005.
- Bimonte S., Intégration de l'information géographique dans les entrepôts de données et l'analyse en ligne : de la modélisation à la visualisation., PhD thesis, Institut National des Sciences Appliquées de Lyon, 2007.
- Bimonte S., Bertolotto M., Gensel J., Boussaid O., « Spatial OLAP and Map Generalization : Model and Algebra. », *IJDWM*, 2012a.
- Bimonte S., Bouilil K., Chanet J.-P., Pradel M., « Definition and Analysis of New Agricultural Farm Energetic Indicators Using Spatial OLAP », *ICCSA (2)*, p. 373-385, 2012b.
- Bimonte S., Fatto V. D., Paolino L., Sebillo M., Vitiello G., « A Visual Query Language for Spatial Data Warehouses », *Geoinformation and Cartography*, 2010.
- Blaschkaand M., Sapia C., Hofing G., Dinter B., « Finding your way through multidimensional data models », *Proceedings of the 9th International Workshop on Database and Expert Systems Applications*, p. 198 -203, 1998.
- Chen T., Schneider M., « The Neighborhood Configuration Model : A Framework to Distinguish Topological Relationships between Complex Volumes », *ER Workshops*, p. 251-260, 2011.
- Derbal K., Boukhalfa K., Alimazighi Z., « A Muti-Representation and Generalisation Based Webmapping Approach Using Multi-Agent System Predicates », *ICWIT*, p. 83-92, 2012.
- Han J., Koperski K., Stefanovic. N., *An efficient two-step methodforclassification of spatial data*, Proc.InternationalSymposiumon Spatial Data Handling, 1998.
- Houria B., « Maîtrise du Risque Routier en Milieu Professionnel Cas d'une Entreprise du Transport Routier en Algérie », *17e Congrès de Maitrise des Risques et de Sûreté de Fonctionnement 5-7 octobre 2010 La Rochelle*, 2010.
- Immon W. H., *Building the data warehouse*, WILLEY, 1992.
- Lu C.-T., Tan X., Chawla S., Shekhar S., Vatsavai R. R., « MapCub :Avisualisation tool for spatial data warehouses », *Geographic Data Mining and Knowledge Discovery,London : Taylor and Francis*, 2000.
- Malinowski E., Zimanyi E., « Representing Spatiality in a Conceptual Multidimensional Model », *ACM GISp*. 12-21, 2004.
- McMaster R. B., K S. S., « Generalization in Digital Cartography », *Washington : Association of American Geographers*, 1992.
- Miquel M., Bédard Y., Brisebois. A., « Conception d'entrepôt de données géospatiales à partir de sources hétérogènes. Exemple d'application en foresterie. », *Ingénierie des Systèmes d'information*, 2002a.
- Miquel M., Bedard Y., Brisebois A., Pouliot J., Marchand P., Brodeur J., « Multidimensional spatiotemporal data warehouse in acontext of evolving specifications », *Symposium on Geospatial Theory, Processing and Applications, Ottawa*, 2002b.
- Nilakanta S., Scheibe K., Rai A., « Dimensional issues in agricultural data warehouse designs », *Computers and Electronics in Agriculture*, vol. 60, n° 2, p. 263-278, 2008.

- Pirotti F., Guarnieri A., Vettore A., « Collaborative Web-GIS Design : A Case Study for Road Risk Analysis and Monitoring », *T. GIS*, vol. 15, n° 2, p. 213-226, 2011.
- Praing R., Schneider M., « Topological feature vectors for exploring topological relationships », *Int. J. Geogr. Inf. Sci.*, vol. 23, n° 3, p. 319-353, 2009.
- Royer K., Bellatreche L., Le-Mouel A., Schmitt G., « Un Entrepôt de Données pour la Gestion des Véhicules Electriques : Retour d'Expérience », *EDA*, p. 118-127, 2012.
- Ruiz C. V., Times V. C., « A Taxonomy of SOLAP Operators », *SBBD*, p. 151-165, 2009.
- Sabine T., Hierarchical structures in map series, PhD thesis, Technical University Vienna, 1998.
- Sabine T., U F. A., « A Multi-Scale DAG for Cartographic Objects », *ACM/ASPRS*, 1995.
- Stefanovic N., Han J., Koperski K., « Object-based selective materialization for efficient implementation of spatial data cubes. », *IEEE Transactions on Knowledge and Data Engineering*, 2000.
- Weibel R., Dutton G., « Generalising spatial data and dealing with multiple representations », *Geographical Information Systems : Principles and Technical Issues*, 1999.