

**République Algérienne Démocratique et Populaire**  
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**  
**Université Amar Telidji Laghouat**  
**Faculté des sciences et sciences de l'ingénieur**  
**Département d'informatique**



**MEMOIRE DE FIN D'ETUDES**  
**Pour l'obtention du diplôme d'Ingénieur d'Etat en Informatique**  
**OPTION : Systèmes d'information avancées**

Thème

**Un Système d'information d'aide à la décision**  
**Pour la régulation du trafic routier**  
**De la ville de Laghouat**

**Présenté par :**

- **Maïcha Mohamed el Habib**
- **Hammouti Kamel**

**Encadré par :**

- **Nardjes Hamini**

**Promotion : 2010/2011**

**N° d'ordre : .....**

# **Remerciements**

*Nous remercions Dieu tout Puissant de nous avoir permis de mener à terme ce projet.*

*Nos remerciements vont tout spécialement à nos familles, qui ont sus nous supporter et encourager tout au long de notre vie, ainsi que pour leur aide inestimable, leur patience et leur soutien indéfectible.*

*Nous remercions Mlle.Hamini Nardjes pour son encadrement, son encouragement, sa disponibilité, ses conseils qui ont été très bénéfiques.*

*Nous tenons aussi, à remercier tout les enseignants qui ont contribué de près ou de loin à notre formation.*

*Nous remercions Mr.YAGOUBI, pour l'honneur qu'il nous a fait en acceptant la présidence de ce jury. Qu'il trouve donc ici l'assurance de notre profonde gratitude.*

*On remercie vivement l'ensemble des membres du Jury qui nous fait grand Honneur d'avoir accepter d'évaluer ce travail.*

*Pour finir, et afin de n'oublier personne (amis, membre de la famille et tous ceux qui nous sont chers) nous utiliserons la formule : « Merci à... ».*

## ***Dédicaces***

### **A ma très chère mère ZOHRA,**

*Affable, honorable, aimable : Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et de prier pour moi.*

*Tu as fait plus qu'une mère puisse faire pour que ses enfants suivent le bon chemin dans leur vie et leurs études.*

*Je te dédie ce travail en témoignage de mon profond amour. Puisse Dieu, le tout puissant, te préserver et t'accorder santé, longue vie et bonheur.*

### **A la mémoire de mon Père KOUIDER**

*Aucune dédicace ne saurait exprimer l'amour, l'estime, le dévouement et le respect que j'ai toujours eu pour vous.*

*Je vous dédie aujourd'hui ma réussite. Que Dieu, le miséricordieux, vous accueille dans son éternel paradis.*

### **Je dédie ce travail aussi a**

*Ma sœur Siham, et mes frères Hamid et Youcef, pour leurs encouragements et leur amour,*

*Mes chers amis, pour tous ce qu'on a partagé ensemble,*

*Toutes les personnes proches que je n'ai pas citées*

*Mohamed el Habib*

## **Dédicaces**

*Je dédie ce modeste travail à :*

*Mes parents MOHAMED & NACIRA, qui n'ont jamais cessé de m'encourager et me soutenir,*

*Mon frère : Mohammed, et mes sœurs :Wafa et Sara,*

*Mes grand-mères & grand-père Puisse Dieu, le tout puissant les préserver et les accorder santé, longue vie et bonheur.*

*Mes tantes et oncles paternelle, et spécialement a Mostafa, Houari, Farida et Amar avec Mohamed amine*

*Mes tantes et oncles maternelle, spécialement Mohamed, Boubakar, Lahcen et Zakari*

*Mes amis : Mostafa, Yacine, Omar, Redha, Oussama, Youcef, Khalil, Walid et Habib*

*Tous les membres de ma famille,  
Ceux qui me sont chers,*

*Kamel*

## ***Résumé***

L'informatique décisionnelle apporte des solutions nouvelles pour la modélisation, l'interrogation et la visualisation de données dans un objectif d'aide à la décision. Les entrepôts de données spatiales et leurs outils d'exploitation les SOLAP (Spatial On-Line Analytical Processing) sont les nouvelles technologies sur la piste du géo-décisionnel, les outils issus de ces technologies doivent offrir la rapidité nécessaire pour des utilisateurs qui exigent des réponses en temps réel.

Dans ce contexte, ce mémoire présente la démarche multidisciplinaire que nous avons adoptée pour construire un système d'information d'aide à la décision pour la régulation du trafic routier. L'architecture du système, le schéma de l'entrepôt de données ainsi que les différentes représentations des séquences spatio-temporelles, stockées dans l'entrepôt, y sont détaillés.

***Mots-clés: Géodécisionnel, Entrepôt de données spatiale, SOLAP.***

## ***Abstract***

Business intelligence provides new solutions for modeling, querying and visualization of data in an objective decision support. The spatial data warehouses and their operating tools the SOLAP (Spatial on- Line Analytical Processing) are the new technologies on the trail of geo decision-making, tools from these technologies must provide the necessary speed for users who need answers in the second.

In this context, this thesis presents the multidisciplinary approach which we have adopted to construct an information system for decision support for road traffic regulation. The system architecture, the schema of data warehouse and the different representations of spatio-temporal sequences stored in the warehouse, therein detailed.

***Keywords: Geodecision, Spatial Data Warehouse, SOLAP.***

---

---

# Table des matières

## **CHAP I : Introduction Générale**

1.1. Contexte	02
1.2. Problématique	03
1.3. Objectif	03
1.4. Plan de travail	04
1.5. Contenu des chapitres	06

## **CHAP II : Système d'Information d'Aide à la Décision**

2.1. Introduction	08
2.2. Aide a la décision ?	09
2.3. Système d'Information d'Aide à la Décision (SIAD)	10
2.4. Des systèmes d'aide à la décision aux entrepôts de données	12
2.5. Entrepôt et magasins de données	13
2.6. Conclusion	15

## **CHAP III : Les Entrepôt de données**

3.1. Concepts principaux des entrepôts de données	17
3.1.1. Entrepôts et bases de données	18
3.1.2. Dimensions et hiérarchies	19
3.1.3. Faits et mesures	20
3.1.4. Hypercube	22
3.1.5. Requêtes multidimensionnelles	22
3.2. Architecture des systèmes à entrepôts de données	23
3.3. Entrepôt de données Spatiale	23
3.3.1. Serveur SOLAP	23
3.3.1.1. Fonctionnalités d'un serveur SOLAP	24
3.3.1.2. Opérateurs SOLAP	24
3.3.1.3. Implémentation physique d'un serveur SOLAP	25
3.3.2. Exemples d'applications SOLAP	27
3.4. Intégration des données spatiales dans un entrepôt	28
3.5. Conclusion	28

## **CHAP IV: Conception**

4.1. Introduction	30
4.2. Présentation du Projet	30
4.3. La Wilaya de Laghouat, Géographiquement	31
4.4. La DTP	31
4.5. Étude de l'existant	32
4.6. Buts du SIAD routier	32
4.7. Conception du SIAD routier	33

---

---

4.7.1. Réalisation de la Carte routière	33
4.7.2. Logicielle utilisée	34
4.7.3. Conception de l'entrepôt de données spatiale	35
4.7.4. Structure du schéma	37
4.7.5. Les hiérarchies	38
4.7.6. Détails du fait	38
4.7.7. Détails des dimensions	39
4.7.8. Schéma relationnel détaillé	40
4.8. Conclusion	42
<b>CHAP V: Réalisation</b>	
5.1. Introduction	44
5.2. Paramètres machines	44
5.3. Architecture technique de la solution	44
5.4. Logiciels utilisés	45
5.5. La Carte	45
5.6. L'entrepôt de données spatiale	46
5.7. Le système obtenue	47
5.8. Testes Et exemples	52
5.9. Conclusion	53
<b>CHAP VI: Conclusion générale</b>	
6.1. Conclusion générale	55
6.2. Perspectives	56
<b>Références</b>	58
<b>Annexe 1 : Code Source de l'application</b>	61
<b>Annexe 2 : Le Projet CADDY</b>	65

---

---

## Liste des Figures

<b>Figure 1.1</b> : Plan de travail	04
<b>Figure 2.1</b> : Représentation systémique d'une organisation	09
<b>Figure 2.2</b> : Le SIAD dans le SI	10
<b>Figure 2.3</b> : Le système d'Aide à la Décision	11
<b>Figure 2.4</b> : Entrepôt et magasins de donnée	14
<b>Figure 3.1</b> : Hiérarchie des produits a) Schéma b) Instance	19
<b>Figure 3.2</b> : Application multidimensionnelle a) Schéma b) Hypercube	20
<b>Figure 3.3</b> : Architecture à trois niveaux d'un système d'entrepôt de données.	23
<b>Figure 3.4</b> : a) Schéma en étoile b) Schéma en flocon c) Schéma en constellation	26
<b>Figure 3.5</b> : Une application sur les accidents sur le réseau routier	27
<b>Figure 3.6</b> : Une application en santé environnementale	27
<b>Figure 4.1</b> : Localisation de la wilaya de Laghouat	31
<b>Figure 4.2</b> : Etapes de la conception de l'EDS routier	33
<b>Figure 4.3</b> : Structure du schéma de l'entrepôt	37
<b>Figure 4.4</b> : Représentation des hiérarchies	38
<b>Figure 4.5</b> : Schéma relationnel de l'entrepôt	41
<b>Figure 5.1</b> : Architecture technique de la solution	44
<b>Figure 5.2</b> : Capture de la carte routière de la commune de Laghouat	45
<b>Figure 5.3</b> : La carte routière de la commune de Laghouat avec MapInfo 9.0	46
<b>Figure 5.4</b> : Les tables de l'EDS avec Paradox 7.0	47
<b>Figure 5.5</b> : La page d'accueil de l'application	48
<b>Figure 5.6</b> : La page principale de l'application	48
<b>Figure 5.7</b> : La page de l'analyse	49
<b>Figure 5.8</b> : La page de résultats de l'analyse	50
<b>Figure 5.9</b> : La page d'informations sur les routes	51

---

---

# Liste des Tableaux

<b>Tableau 3.1:</b> <i>Différences entre SGBD et entrepôts de données</i>	18
<b>Tableau 4.1 :</b> <i>Les paramètres de la projection</i>	34
<b>Tableau 4.2 :</b> <i>Préparation pour la Conception de l'entrepôt</i>	36
<b>Tableau 4.3:</b> <i>Tableau représentatif des détails du fait</i>	38
<b>Tableau 4.4 :</b> <i>Tableau représentatif des détails des dimensions</i>	40

---

# *CHAPITRE I* **Introduction Générale**

- 1.1 Contexte
- 1.2 Problématique
- 1.3 Objectif
- 1.4 Plan de travail
- 1.5 Contenu des chapitres

### 1.1 Contexte

Le trafic routier est un phénomène complexe d'une part en raison du nombre élevé d'acteurs qui y participent, d'autre part à cause du caractère très maillé du réseau sur lequel il se déroule. Le phénomène le plus marquant dans le trafic routier est la congestion qui alimente de nombreuses discussions d'usagers. Depuis une cinquantaine d'années, des théoriciens du trafic cherchent à comprendre et quantifier les mécanismes mis en œuvre.

Tout cela est dû à l'accroissement des accidents de la circulation, le nombre élevé de décès en plus de l'augmentation du trafic routier ces dernières années et les embouteillages, surtout pendant les heures de pointe, Et l'incapacité de ces routes pour accueillir les voitures.

Les systèmes d'informations décisionnels sont nés d'un besoin des entreprises à fournir aux décideurs des moyens d'accéder aux données de leurs propres systèmes dans le but de piloter leurs activités. Dès lors qu'il s'agit de faire du reporting ou de l'analyse de données pour arriver à fournir des tableaux de synthèse il fallait mettre en place des requêtes complexes, coûteuses en temps de réponse et en ressource informatique. Les structures des entrepôts de données, des bases de données multidimensionnelles et les outils d'exploration, par leur nature, sont construites pour supporter des analyses complexes et une découverte des connaissances. La fonctionnalité des OLAP (*On-line Analytical Processing*) est caractérisée par l'analyse multidimensionnelle et dynamique de données consolidées qui supportent les activités analytique et navigationnelle d'un utilisateur final. Les technologies de la Business Intelligence, comme les tableaux de bord, l'OLAP, le forage de données, sont disponibles commercialement depuis déjà une décennie, et l'open source s'attaque depuis quelques années à la gamme d'outils d'aide à la décision avec des solutions aujourd'hui très matures et en perpétuel progrès.

Le réseau routier de la wilaya de Laghouat possède un volume significatif de données complexes et a besoin de systèmes efficaces pour consulter et analyser les tendances pour la gestion durable des routes. De plus ils doivent permettre d'accéder aux informations appropriées plus rapidement. C'est tout naturellement, que la DTP (Direction des Travaux Public) doit s'intéresser à cette technologie.

Notre travail consiste à développer un système d'information d'aide à la décision pour la régulation du trafic routier de la commune de Laghouat, Ce système devrait permettre à la

DTP (Direction des Travaux Public) de faire un constat du taux d'amélioration engendré par les nouvelles installations, des projections sur l'avenir et de ce fait aider à mieux planifier ou à améliorer l'installation routière.

### 1.2 Problématique

Afin d'éviter les erreurs des années passées concernant la mise en place des feux de signalisation et les désagréments causés par la mauvaise gestion de nos routes (voir 2<sup>ième</sup> paragraphe du contexte), nous devons contribuer par ce modeste travail concernant la création d'un système d'information d'aide à la décision pour la régulation du trafic routier de la commune de Laghouat.

Le plan de gestion actuellement utilisé par les secteurs concernés ne permet pas de suivre le rythme de développement témoigné par ce secteur, d'autant que l'urbanisation conduit à la pompe d'une énorme quantité de données qui ne peuvent absorber les systèmes d'information actuellement utilisés par ces secteurs, et face à ça ces systèmes sont incapables d'analyser les données et prendre une durée infinie pour le faire.

Pour une analyse facile et aussi rapide que possible les données doivent être résumées et regrouper dans un seul système, la meilleure approche pour arriver à cette solution est un EDS (Entrepôt de données Spatiales) et SOLAP (*Spatiale On-line Analytical Processing*) parce que l'EDS permet d'organiser et d'historier les données pour la prise de décisions, d'autre part solap offrent des affichages sous forme de cartes, tableaux et graphes statistiques en tous genres ce qui facilite l'analyse.

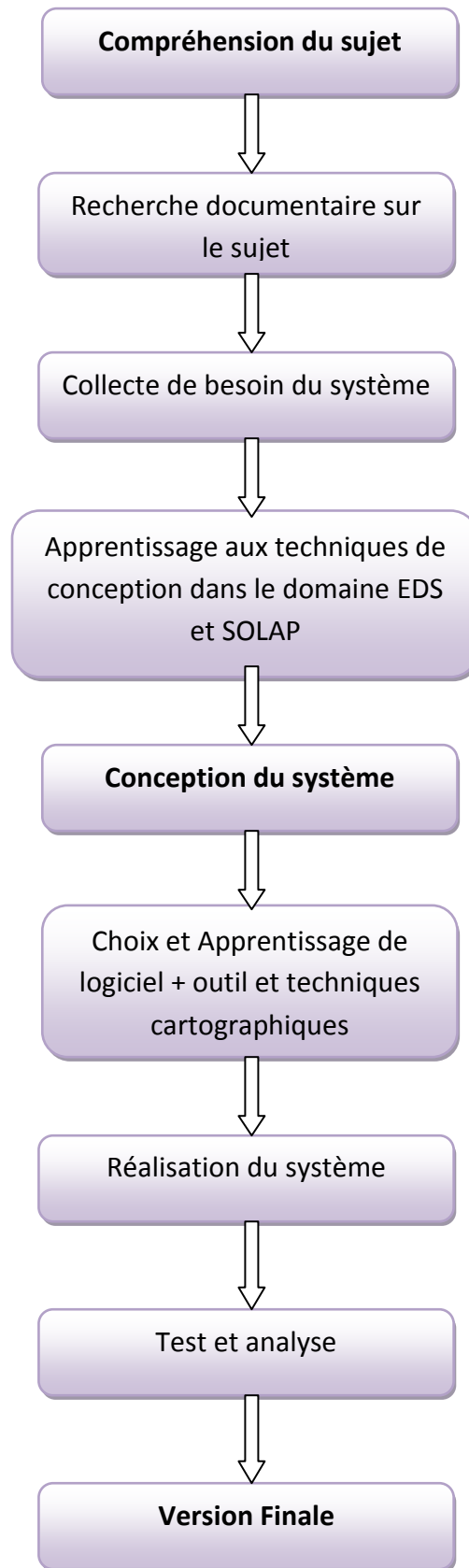
### 1.3 Objectifs

Notre objectif principal est de concevoir un système d'information d'aide à la décision pour résoudre les problèmes liés aux trafics routiers, en proposant une application dotée d'une interface interactive, intuitive avec des fonctionnalités d'analyse selon plusieurs critères (zone géographique, temps...).

Nos objectifs spécifiques sont donc:

- Réalisation de la carte routière de la commune
- la conception d'une interface SOLAP

## 1.4 Plan du Travail



*Figure 1.1 : Plan du travail*

## Introduction Générale

---

La première étape Compréhension du sujet consiste en premier lieu à rassembler les connaissances sur le data warehousing et de comprendre le trafic routier.

La deuxième étape sert à faire une recherche approfondie sur les documents et les publications existantes afin de rassembler le plus d'idées possibles sur le géodécisionnel pour faciliter la suite du travail.

La troisième étape Collecte des besoins du système, cette étape est la plus difficile car elle comprend le contact avec les personnes (personnel de la direction concernée).

Nous avons constaté qu'il y'avait beaucoup de travail à faire pour établir le dialogue entre l'informaticien et le personnel des entreprises, parmi les techniques que nous avons poursuivi été une présentation de notre projet sous forme de conférences, dans lequel nous avons défini les objectifs de notre système et ce qui va apporter à la direction pour clarifier un peu l'utilité de se SAD, mais on a remarqué qu'il y a un manque de confiance aux outils informatique dans l'administration.

La Conception du système, consiste comme son nom l'indique à mettre en place les idées et les bases sur lesquelles va reposer le système, c'est aussi l'étape la plus délicate du travail car d'elle dépend le bon déroulement du projet et surtout le temps nécessaire pour le mener à bout car si elle est soigneusement effectuée la réalisation du projet sera plus efficace et les résultats de meilleure qualité.

L'étape de la réalisation du système est une implémentation des concepts mis au point qui va permettre de faire une première série de tests afin de mieux les adapter ou de les changer.

Lorsque les principes sont bien établis et validés, la version finale du logiciel peut alors être mise à la disposition des décideurs des services concernés .

### 1.5 Contenu des chapitres

Après cette introduction, nous structurons en cinq chapitres ce mémoire, en présentant respectivement une généralité sur le décisionnel, une vue globale sur les entrepôts de données, notre travail en deux chapitres, et une conclusion.

Le Chapitre 2 vise à présenter et à définir les différents concepts des technologies décisionnelles servant de support à nos travaux. Nous étudions les composants d'un Système d'Information d'Aide à la Décision (SIAD), leurs architectures ainsi que Les principes de la modélisation dimensionnelle seront décrits car ils ont été utilisés dans la mise en place de l'entrepôt de données.

Le chapitre 3 introduit les principaux concepts des entrepôts de données spatiales et de l'analyse en ligne. Nous présentons les principes de la modélisation multidimensionnelle car nous nous appuyerons sur ces concepts pour le développement de notre SIAD. Le chapitre 3 présente également les concepts principaux de l'OLAP Spatial.

Le quatrième chapitre décrit nos travaux relatifs à la conception des entrepôts. Cet espace de stockage de données sources nécessaires aux processus d'aide à la décision doit permettre de stocker les données évoluant dans le temps de manière détaillée ou archivée.

La réalisation du SIAD est présentée dans le cinquième chapitre qui sert aussi à finaliser la présentation de nos travaux. Notamment, il permet d'explicitier les principales fonctionnalités des outils que nous avons développés.

Enfin dans le Chapitre 6 une conclusion générale clôture ce travail, présente un bilan de notre travail et les perspectives de recherche que nous envisageons pour compléter ce Système.

# *CHAPITRE II*

## **Système d'Information d'Aide à la Décision**

- 1.1 Introduction
- 1.2 Aide a la décision ?
- 1.3 Système d'Information d'Aide à la Décision (SIAD)
- 1.4 Des systèmes d'aide à la décision aux entrepôts de données
- 1.5 Entrepôt et magasins de données
- 1.6 Conclusion

## 2.1. Introduction

La mondialisation et la concurrence qu'elle engendre rendent le pilotage d'une organisation de plus en plus complexe. Cette complexité est liée non seulement à l'augmentation du nombre de paramètres à prendre en compte mais également à la nécessité de prises de décisions rapides afin d'être réactifs à l'évolution de la concurrence et de la demande des clients. L'efficacité de ces prises de décisions repose sur la mise à disposition d'informations fiables, pertinentes et d'outils facilitant cette tâche. Les systèmes traditionnels, dédiés à la gestion quotidienne d'une organisation, s'avèrent inadaptés à une telle activité [Codd et al., 1993 ; Inmon, 1996 ; K.imball & Ross, 2002]. Face à ce besoin est né le secteur de l'informatique décisionnelle.

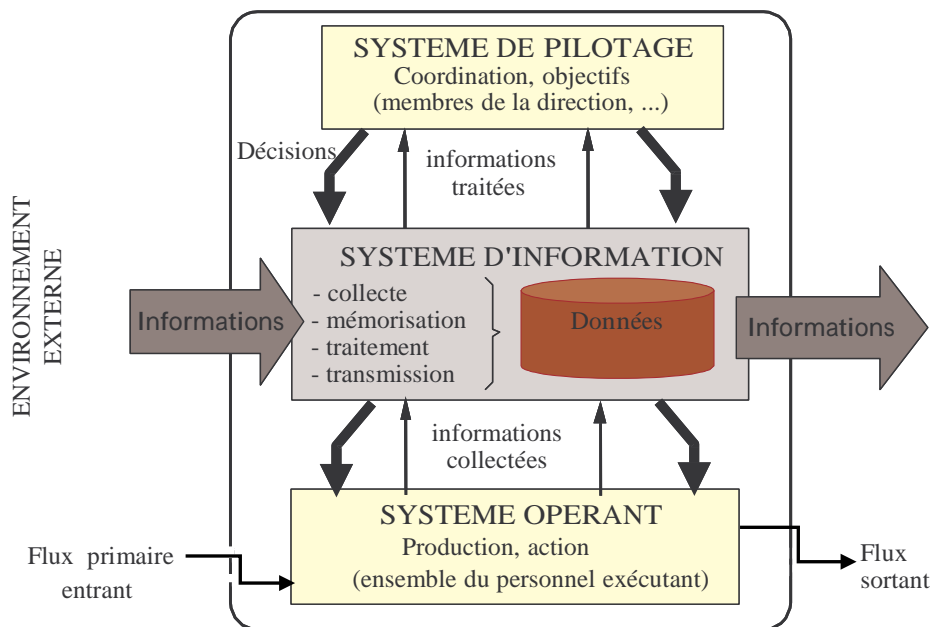
Les systèmes d'aide à la décision (**Decision Support System**) sont destinés à faciliter les prises de décision au niveau de l'entreprise. Un système d'aide à la décision bien conçu est un logiciel interactif qui aide les décideurs à dégager des informations utiles à partir de données brutes, de documents, de connaissances personnelles et de modèles métier afin d'identifier et résoudre des incidents et prendre des décisions.

Généralement, une application d'aide à la décision regroupe les informations suivantes:

- Ensemble des informations en cours (y compris celles provenant de sources de données propriétaires ou relationnelles, de cubes, entrepôts de données et data marts)
- Comparaison des valeurs du Taux d'une période à l'autre
- Conséquences des différentes décisions possibles, compte tenu de l'expérience passée.

## 2.2. Aide a la décision ?

La modélisation systémique de toute organisation se décompose en trois sous-systèmes : Système Opérant (SO), Système d'Information (SI) et Système de Pilotage (SP). Le SO représente l'activité productrice de l'organisation étudiée. Cette activité consiste à transformer les flux primaires (matières, finance, personnel..) pour répondre aux besoins des clients. Le SP regroupe l'ensemble du personnel d'encadrement qui effectue les tâches de régulation, de pilotage et d'adaptation de l'organisation à son environnement [Mélèse, 1972]. Le SI permet de collecter, mémoriser, traiter et restituer les différentes données de l'organisation afin de permettre au SP d'effectuer ses fonctions tout en assurant son couplage avec le SO [Nanci & Espinasse, 2001]. L'activité du SO produit des informations stockées dans le SI ; après traitement, la transmission de ces informations vers le SP permet à ce dernier de connaître l'activité du SO (flèches "informations" dans la Figure 2.1). Les décisions du SP seront répercutées vers le SI puis vers le SO pour permettre au SP d'en maîtriser le fonctionnement (flèches "décisions" dans la Figure 2.1).



**Figure 2.1:** Représentation systémique d'une organisation [Mélèse, 1972]

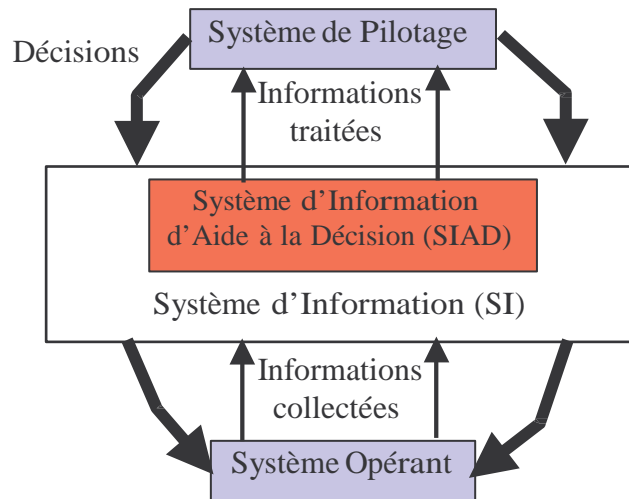
Pour répondre aux besoins des décideurs, il est nécessaire de synthétiser, réorganiser et historiser les données de production du SI afin d'en déterminer une sous-partie relative à l'aide à la décision. La suite de ce mémoire se centre sur cet aspect. Notamment, dans les sections suivantes, nous définissons les concepts de système d'information d'aide à la décision, de système d'aide à la décision, d'entrepôts et de magasins de données.

### 2.3. SYSTEME D'INFORMATION D'AIDE A LA DECISION (SIAD)

Par analogie à la définition précédente d'un SI, nous proposons la définition du Système d'Information d'Aide à la Décision (SIAD) suivante :

Définition : Un SIAD est la partie d'un système d'information permettant d'accompagner les décideurs dans le processus de prise de décision. Les fonctions d'un SIAD permettent de

- collecter, intégrer, synthétiser et transformer les données opérationnelles d'un SI,
- mémoriser de manière adaptée les données décisionnelles,
- traiter ces données (alimentation, rafraîchissement, pré-calculs.),
- restituer de manière appropriée ces données afin de faciliter la prise de décision.



**Figure 2.2 : Le SIAD dans le SI**

De nos jours, l'ensemble des outils informatiques permettant de supporter un SIAD est qualifié de Business Intelligence (BI) ou de Système d'Aide à la Décision (SAD). Un SAD vise à exploiter les données opérationnelles d'une organisation afin de faciliter la prise de décision pour un pilotage éclairé. Afin d'être plus explicite, nous proposons la définition suivante :

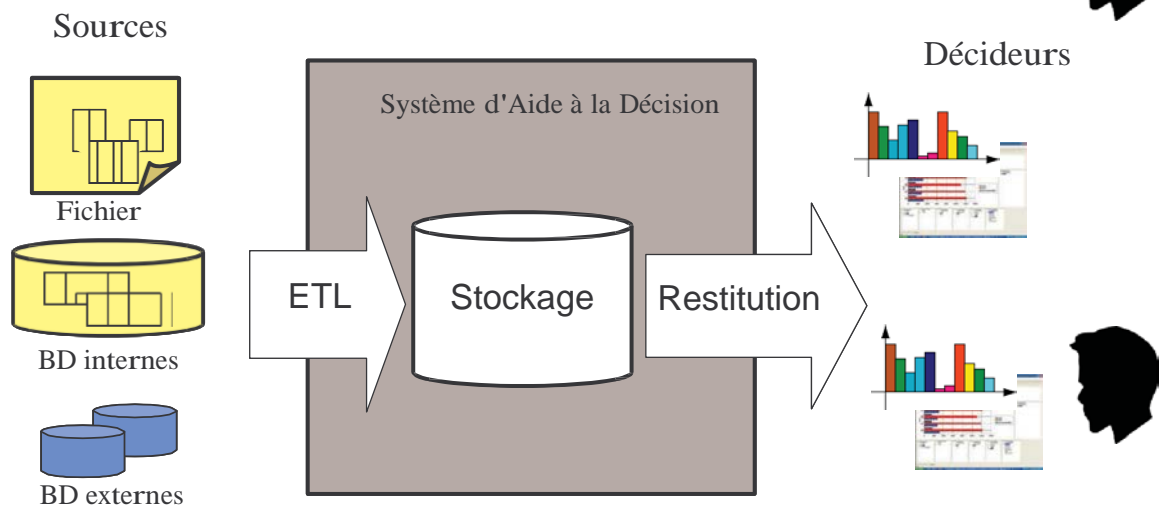
Définition : Un Système d'Aide à la Décision (SAD) regroupe l'ensemble des outils informatiques (matériels et logiciels) permettant :

- D'extraire, de transformer et de charger les données opérationnelles,
- De constituer un ou des espaces de stockage de données décisionnelles,
- De manipuler ces données au travers d'outils d'analyse ou d'interrogation destinés au pilotage des organisations.

La plupart des travaux déclinent ces applications informatiques en trois catégories :

- Extraction, transformation et chargement (ou ETL acronyme de "Extraction Transformation Loading") des données opérationnelles (hétérogènes et disparates) pour alimenter et rafraîchir le système d'aide à la décision,
- Stockage et traitement des données décisionnelles,
- Restitution des données sous une forme adaptée aux utilisateurs (interrogations ou analyses décisionnelles).

Nous pouvons schématiser ces différents outils dans la figure suivante :



**Figure 2.3** : Le système d'Aide à la Décision

## 2.4 DES SYSTEMES D'AIDE A LA DECISION AUX ENTREPOTS DE DONNEES

De nos jours, les entrepôts de données constituent une solution adéquate pour construire un système d'aide à décision [Widom, 1995 ; Inmon, 1996]. Un entrepôt de données (ED) est défini comme étant "une collection de données intégrées, orientées sujet, non volatiles, historisées, résumées et disponibles pour l'interrogation et l'analyse" [Inmon, 1996]. Cette définition met l'accent sur les caractéristiques suivantes :

- **Intégrées** : les données alimentant l'entrepôt proviennent de sources multiples et hétérogènes. Les données des systèmes de production doivent être converties, reformatées et nettoyées de façon à avoir une vision globale dans l'entrepôt.
- **Orientées sujet** : contrairement aux systèmes de production structurant les données par processus fonctionnel, les données d'un ED s'organisent par thèmes d'analyse. L'intérêt de cette organisation est de disposer de l'ensemble des informations utiles sur un thème, le plus souvent transversales aux structures fonctionnelles et organisationnelles d'une entreprise. Cette orientation sujet permet de mettre en avant les indicateurs de performance pour chaque thème d'analyse.
- **Non volatiles** : après intégration, transformation et synthèse des données opérationnelles dans un ED, les seules actions que peuvent effectuer des décideurs sont des interrogations et des analyses décisionnelles (pas de mise à jour).
- **Historisées** : l'alimentation et le rafraîchissement d'un ED consiste en l'intégration des données opérationnelles à différents points d'extraction. Cette intégration de données à des dates différentes permet de conserver "l'historisation" des données qui est vitale pour toute prise de décision.
- **Résumées** : les informations issues des sources doivent être transformées mais surtout agrégées pour faciliter le processus de prises de décision.
- **Disponible pour l'interrogation et l'analyse** : afin d'améliorer les performances d'une organisation, les décideurs doivent pouvoir consulter et analyser les données contenues dans un ED au travers d'outils interactifs

## 2.5 ENTREPOT ET MAGASINS DE DONNEES

D'après la définition de [Inmon, 1996], l'ED doit permettre d'extraire, de transformer et de stocker un grand volume de données opérationnelles et, en même temps, de répondre à des requêtes utilisateurs concernant un thème d'analyse spécifique. En fait, cette définition regroupe deux problématiques que nous avons identifiées comme suit dès le début de nos travaux :

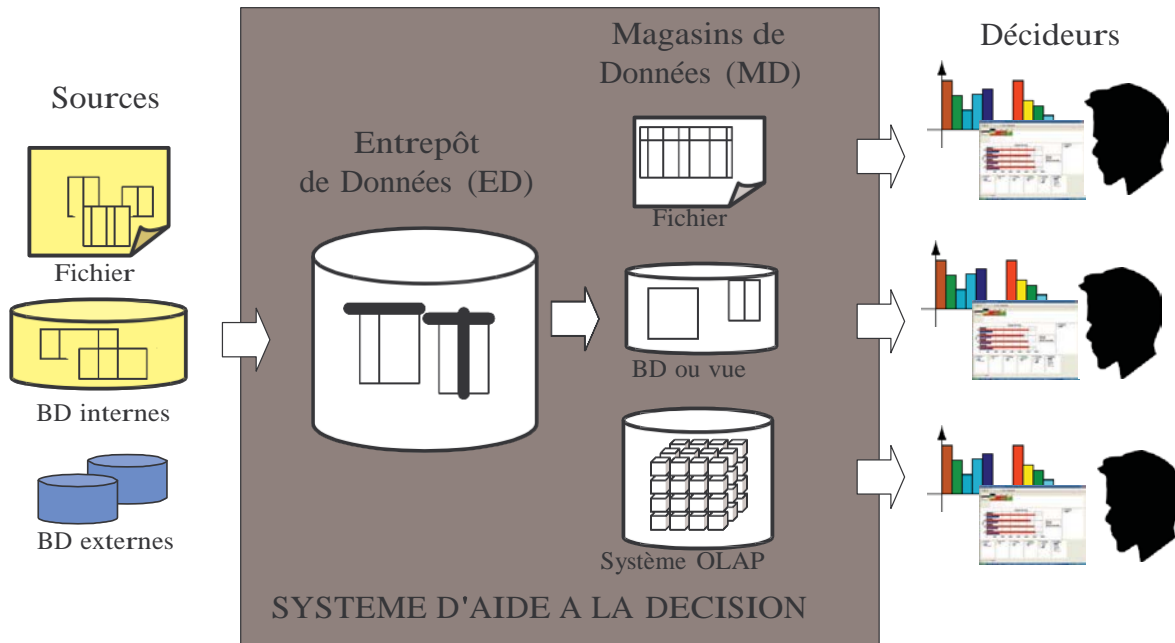
- la gestion efficace des données "historisées", "centralisées" (intégration des sources),
- la définition d'un sous-ensemble de données autour d'un thème particulier afin de répondre aux besoins spécifiques de décideurs.

Aussi, l'architecture des systèmes d'aide à la décision que nous proposons est basée sur une dichotomie d'espaces de stockage : l'entrepôt et les magasins de données [Ravat et al. 1999] [Ravat & Teste, 2000c ; Teste, 2000].

Définition : Un Entrepôt de Données (ED) est l'espace de stockage centralisé d'un extrait des sources pertinent pour les décideurs. Son organisation doit faciliter la gestion des données et la conservation des évolutions nécessaires pour les prises de décision.

Définition : Un Magasin de Données (MD) est un extrait de l'ED adapté à un thème d'analyse particulier et organisé selon un modèle adapté aux outils d'analyse et d'interrogation décisionnelle.

Dans la figure suivante, nous schématisons l'architecture des SAD telle que nous l'avons définie précédemment [Ravat & Teste, 2000a ; Ravat & Teste, 2000b].



**Figure 2.4 :** Entrepôt et magasins de donnée

Cette dichotomie des espaces de stockage a servi de guide pour nos travaux de développement.

### **2.6 Conclusion**

Dans cette première section, nous avons défini les différents concepts servant de support à nos travaux sur l'aide à la décision. A partir de la représentation systémique d'une organisation, nous avons identifié le concept de SIAD qui est la partie d'un SI permettant d'accompagner un ou plusieurs décideurs dans le processus de prise de décision. Un Système d'Aide à la Décision (SAD), partie informatisée d'un SIAD, regroupe l'ensemble des outils informatiques capables d'extraire les données opérationnelles afin de les transformer en informations pertinentes pour les décideurs. D'un point de vue architectural, nous avons identifié deux espaces de stockage des données dans un SAD : l'entrepôt (espace de stockage centralisé) et les magasins de données (espace de stockage extrait d'un ED et centré sur un thème d'analyse particulier).

# *CHAPITRE III*

## **Entrepôts de données & entrepôts de données Spatiales**

- 3.1 Concepts principaux des entrepôts de données
  - 3.1.1 Entrepôts et bases de données
  - 3.1.2 Dimensions et hiérarchies
  - 3.1.3 Faits et mesures
  - 3.1.4 Hypercube
  - 3.1.5 Requêtes multidimensionnelles
- 3.2 Architecture des systèmes à entrepôts de données
- 3.3. Entrepôt de données Spatiale
  - 3.3.1 Serveur SOLAP
    - 3.3.1.1 Fonctionnalités d'un serveur SOLAP
    - 3.3.1.2 Opérateurs SOLAP
    - 3.3.1.3 Implémentation physique d'un serveur SOLAP
  - 3.3.2 Exemples d'applications SOLAP
- 3.4 Intégration des données spatiales dans un entrepôt
- 3.5 Conclusions

Les Systèmes d'Aide à la Décision (SAD) sont des systèmes d'information flexibles et interactifs qui aident les décideurs dans l'extraction d'informations utiles pour identifier et résoudre des problèmes et pour prendre des décisions [Alter, 1980]. Cette connaissance est obtenue à partir de données brutes, de connaissances personnelles et de modèles analytiques. Les SAD présentent ces informations, provenant de différentes sources, dans un environnement unique, uniforme et familier à l'utilisateur. Ils combinent, uniformisent et synchronisent les bases de données, les modèles d'analyse et les techniques de visualisation, en permettant de comparer différents résultats et de concevoir et valider des hypothèses. Parmi les systèmes d'aide à la décision, les systèmes d'entrepôts de données sont probablement les plus utilisés dans le monde académique et industriel.

Dans ce chapitre, nous décrivons les concepts principaux des entrepôts de données et de l'analyse multidimensionnelle et présentons les caractéristiques principales de leur modélisation formelle. Ensuite, nous décrivons les architectures des systèmes d'entrepôts de données.

### 3.1 Concepts principaux des entrepôts de données

Bill Inmon définit le Data Warehouse, dans son livre considéré comme étant la référence dans le domaine "Building the Data Warehouse" [Inmon, 2002] comme suit:

***« Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles et évolutives dans le temps, organisées pour le support d'un processus d'aide à la décision. »***

Les paragraphes suivants illustrent les caractéristiques citées dans la définition d'Inmon.

**Orienté sujet :** le Data Warehouse est organisé autour des sujets majeurs de l'entreprise, contrairement à l'approche transactionnelle utilisée dans les systèmes opérationnels, qui sont conçus autour d'applications et de fonctions telles que : cartes bancaires, solvabilité client..., les Data Warehouse sont organisés autour de sujets majeurs de l'entreprise tels que : clientèle, ventes, produits.... Cette organisation affecte forcément la conception et l'implémentation des données contenues dans le Data Warehouse. Le contenu en données et en relations entre elles diffère aussi. Dans un système opérationnel, les données sont

essentiellement destinées à satisfaire un processus fonctionnel et obéit à des règles de gestion, alors que celles d'un Data Warehouse sont destinées à un processus analytique.

**Intégrée** : le Data Warehouse va intégrer des données en provenance de différentes sources. Cela nécessite la gestion de toute incohérence.

**Evolutives dans le temps** : Dans un système décisionnel il est important de conserver les différentes valeurs d'une donnée, cela permet les comparaisons et le suivi de l'évolution des valeurs dans le temps, alors que dans un système opérationnel la valeur d'une donnée est simplement mise à jour. Dans un Data Warehouse chaque valeur est associée à un moment « Every key structure in the data warehouse contains - implicitly or explicitly -an element of time » [Inmon, 2000].

**Non volatiles** : c'est ce qui est, en quelque sorte la conséquence de l'historisation décrite précédemment. Une donnée dans un environnement opérationnel peut être mise à jour ou supprimée, de telles opérations n'existent pas dans un environnement Data Warehouse.

**Organisées pour le support d'un processus d'aide à la décision** : Les données du Data Warehouse sont organisées de manière à permettre l'exécution des processus d'aide à la décision (Reporting, Data Mining...).

### 3.1.1 Différences entre Entrepôts et bases de données

Le tableau 3.1 résume ces différences entre les systèmes de gestion de bases de données et les entrepôts de données [DG01].

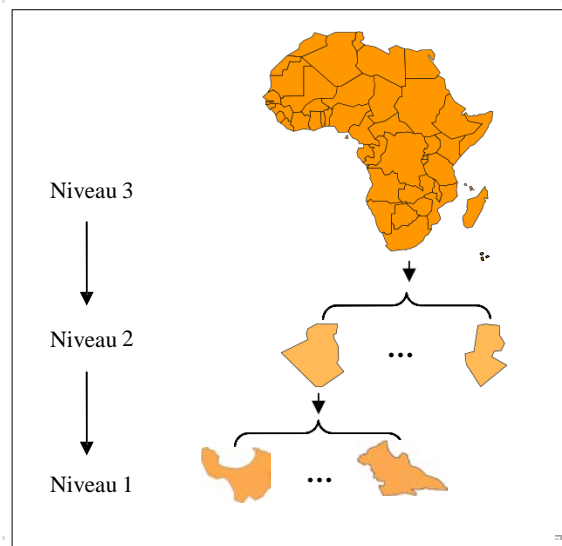
	SGBD	Entrepôts de données
<b>Objectifs</b>	<i>Gestion et production</i>	Consultation et analyse
<b>Utilisateurs</b>	<i>Gestionnaires de production</i>	Décideurs, analystes
<b>Taille de la base</b>	<i>Plusieurs gigaoctets</i>	Plusieurs teraoctets
<b>Organisation des données</b>	<i>Par traitement</i>	Par métier
<b>Type de données</b>	<i>Données de gestion (courantes)</i>	Données d'analyse (résumées, historisées)
<b>Requêtes</b>	<i>Simple, prédéterminées, données détaillées</i>	Complexes, spécifiques, agrégations et <i>group by</i>
<b>Transactions</b>	<i>Courtes et nombreuses, temps réel</i>	Longues, peu nombreuses

**Tableau 3.1:** Différences entre SGBD et entrepôts de données

### 3.1.2 Dimensions et hiérarchies

Les dimensions représentent les axes de l'analyse multidimensionnelle. Elles sont organisées en schémas hiérarchiques. Un schéma de hiérarchie, composé par plusieurs niveaux, représente différentes granularités ou degrés de précision de l'information. Un exemple de dimension peut être la dimension temporelle qui organise le temps grâce à une hiérarchie des niveaux jour, mois, années, ou encore une dimension représentant une classification de pays décrite par une hiérarchie avec les niveaux « pays » et « wilaya » (*Figure 3.1a*). L'instance d'une dimension est un ensemble de membres. Ces membres sont connectés par des liens hiérarchiques en accord avec le schéma hiérarchique. Nous supposons que chaque hiérarchie contient un niveau (All), avec un seul membre. Un exemple d'une instance de la hiérarchie des pays est montré en *Figure 3.1b*, où par exemple les wilayas « w1 » et « w2 », appartenant au «niveau1», sont liés au membre « Algérie » du «niveau2».

Chaque niveau de la dimension peut présenter des attributs [Hüsemann et al. 2000] qui ne sont pas utilisées pour la définition du schéma hiérarchique, par exemple un produit peut présenter un attribut représentant, le prix, la couleur, etc. Ces attributs peuvent être utilisés dans l'analyse multidimensionnelle.

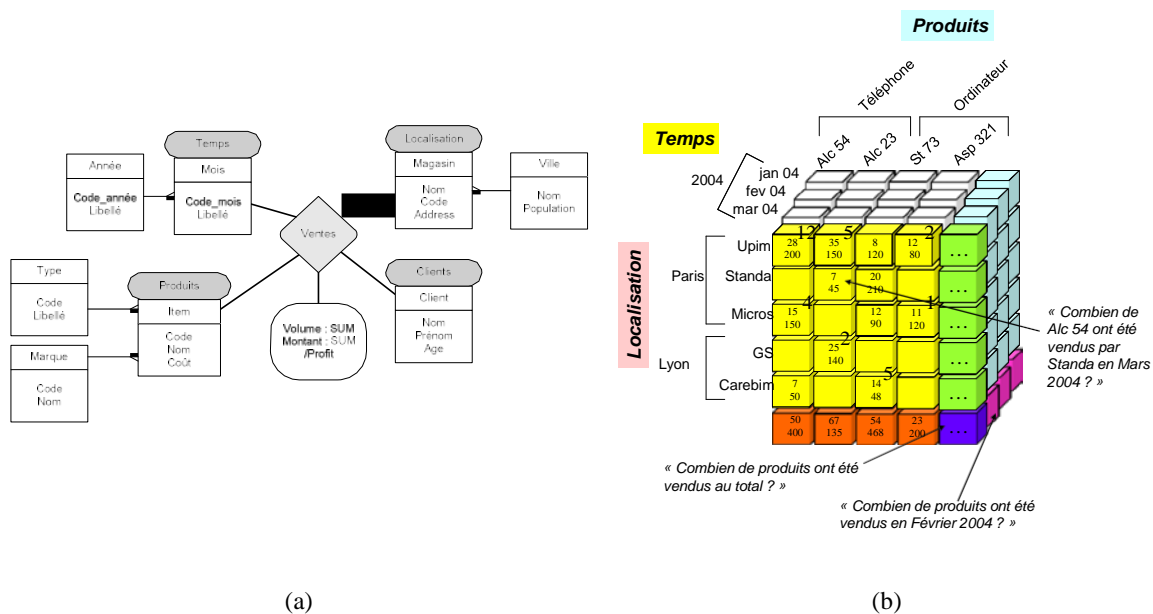


**Figure 3.1:** Hiérarchie des produits a) Schéma b) Instance.

### 3.1.3 Faits et mesures

Un fait est un concept relevant du processus décisionnel et, typiquement, modélise un ensemble d'événements d'une organisation. Un fait est décrit par plusieurs mesures. Les mesures représentent usuellement des valeurs numériques qui fournissent une description quantitative du fait. Un fait est associé à une ou plusieurs combinaisons de membres des dimensions. Enfin, certaines mesures peuvent être calculées à partir d'autres mesures ou propriétés de membres. Elles sont appelées mesures dérivées [Blaschka et al. 1998].

Ainsi, une analyse multidimensionnelle portant sur un fait « ventes » d'un ensemble de magasins pourra être réalisée en définissant comme mesures « le volume des produits vendus », « le montant de la vente », et la mesure dérivée « profit », et comme dimensions « le temps », « les magasins », et « les produits » vendus. La *Figure 3.2a* représente le schéma de cette application grâce au modèle conceptuel multidimensionnel présenté dans [Malinowski et Zimányi, 2004]. Ce modèle permet d'examiner le volume et le montant totaux des produits vendus pour chaque mois et chaque année dans chaque magasin et dans chaque ville.



**Figure 3.2 :** Application multidimensionnelle a) Schéma b) Hypercube. [Malinowski et Zimányi, 2004]

A chaque combinaison des niveaux des dimensions correspond un niveau différent de détail des mesures. Dans les niveaux moins détaillés des dimensions les mesures sont agrégées en utilisant les fonctions d'agrégations. Les fonctions classiques pour agréger les mesures sont les mêmes que les opérations SQL "COUNT", "SUM", "MIN", "MAX" et "AVG".

Il est fondamental dans un modèle multidimensionnel de contrôler que le processus d'agrégation soit correct, afin de garantir une analyse multidimensionnelle exacte. Dans les systèmes OLAP, l'opérateur d'agrégation le plus utilisé est la somme. La précision de l'agrégation dépend de la sémantique de la mesure et de la structure de l'application multidimensionnelle. Par exemple, sommer des concentrations de pollution n'a pas de sens, en revanche le maximum est significatif. Il est possible de sommer la population des villes d'un pays, mais sommer la population d'une ville dans le temps est erroné, car les mêmes habitants seront comptés plusieurs fois. Cette problématique est connue dans l'OLAP sous le nom de problème d' « additivité » [Kimball, 1996]. Une mesure est dite :

- Additive si l'on peut lui appliquer la somme sur toutes les dimensions.
- Semi-additive si la somme a du sens seulement sur certaines dimensions.
- Non-additive si elle n'est sommable sur aucune dimension.

Notons que la mesure peut être prise en compte plusieurs fois dans le processus d'agrégation, peut dépendre du type de hiérarchie, par exemple lorsque le modèle présente de hiérarchies non strictes ou multiples [Horner et al. 2004]. Dans les bases de données statistiques, cette problématique est appelée « summarizability » [Lenz et Shoshani, 1997]. La « summarizability » représente la situation dans laquelle le résultat d'une agrégation pouvait être calculé en utilisant les agrégations précédentes.

Une agrégation correcte implique :

- Eviter de prendre en compte plusieurs fois la même mesure
- Respecter la sémantique de l'agrégation : il s'agit de fournir un contrôle sur le type d'agrégation en prenant en compte la sémantique de la mesure.

### 3.1.4 Hypercube

L'instance d'un modèle conceptuel multidimensionnel est un hypercube. Un hypercube contient dans les cellules les valeurs des mesures et ses axes sont les membres des différentes dimensions. Ensuite, ce cube de base est rapetissé avec des cellules qui contiennent l'agrégation des valeurs de mesures pour chaque combinaison de membres des niveaux moins détaillés. Un exemple d'hypercube pour l'application de la *Figure 3.2a* est montré en *Figure 3.2b*. Sur les axes du cube de base on trouve les membres des niveaux des dimensions (temps, client, localisation et produit) et dans les cellules les valeurs des deux mesures (volume et montant).

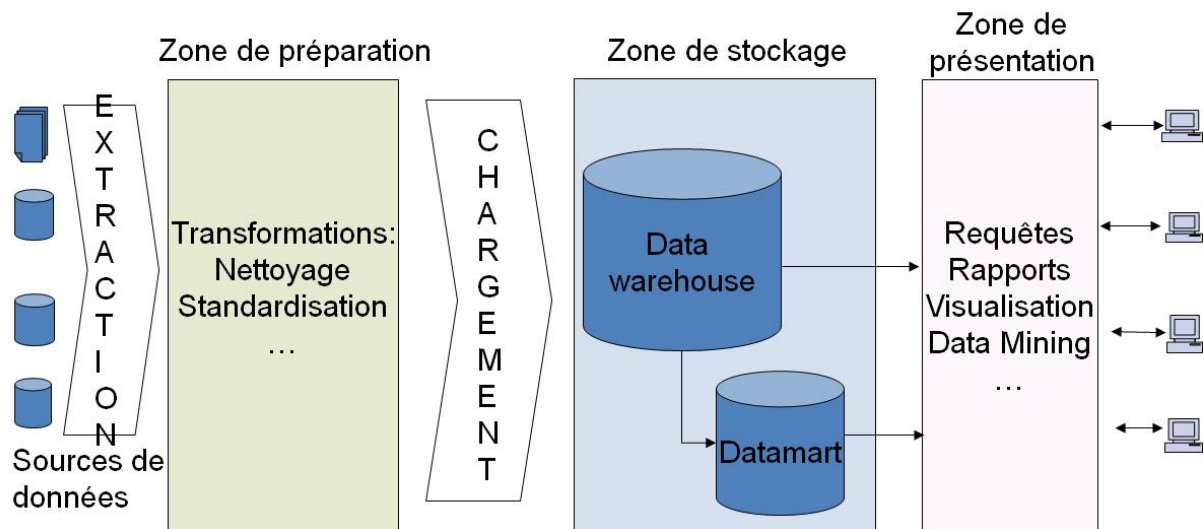
### 3.1.5 Requêtes multidimensionnelles

Le processus décisionnel multidimensionnel consiste en l'exploration de l'hypercube. L'utilisateur parcourt les données de l'hypercube selon les différents axes d'analyses à la recherche d'informations utiles, dans un processus fortement interactif, itératif et constructif, qui comprend des étapes de formulation des hypothèses, expérimentation et analyse [Tang et al. 2003]. Les utilisateurs interagissent itérativement avec le modèle multidimensionnel pour formuler, modifier et valider leurs hypothèses. Les chemins d'analyse sont imprédictibles, contrairement aux données qui sont définies lors de la conception de l'application. Chaque résultat d'analyse est la conséquence des résultats précédents. Chaque étape du processus d'analyse est représentée par une navigation dans l'hypercube, ou par une requête multidimensionnelle. Ces requêtes utilisent les opérateurs OLAP, ces derniers sont bien détaillées dans la section (3.3.1.2).

Un exemple de requête multidimensionnelle portant sur l'application de *Figure 3.2a* est : « *Quels sont le volume et le montant de chaque produit vendu par le magasin Carebim pour chaque année ?* ». Cette requête utilise à la fois l'opérateur de slice et celui de roll-up.

### 3.2 Architecture des systèmes d'entrepôts de données (ED Classique)

Les architectures des systèmes d'entrepôts de données spatiales sont classiquement des architectures à trois niveaux, comme montré en *Figure 3.3*, constituées par un entrepôt de données, un serveur SOLAP et un client SOLAP.



*Figure 3.3* : Architecture à trois niveaux d'un système d'entrepôt de données.

### 3.3 Entrepôt de données Spatiale

Le premier niveau est un SGBD. Les données sont extraites à partir des bases de données transactionnelles, nettoyées et transformées avec des outils ETL (Extract-Transform-Load ou en français extraction, transformation et alimentation), et intégrées dans l'entrepôt de données. Le SGBD contient aussi un ensemble de métadonnées concernant les sources de données, les mécanismes d'accès, les procédures de nettoyage et d'alimentation, les utilisateurs, etc.

#### Données Spatiales :

Aussi connu sous le nom de *données géospatiales* ou *d'information géographique*, elles représentent les données ou informations qui identifie l'emplacement géographique des caractéristiques et des limites sur la Terre, tels que les caractéristiques naturelles ou construites, les océans, et plus encore. Les données spatiales sont souvent stockées sous forme de coordonnées et de topologie, et sont des données qui peuvent être mappés. Les données spatiales sont souvent accessibles, manipuler ou analysées par les systèmes d'information géographique

### 3.3.1 Serveur SOLAP (Spatial On-Line Analytical Processing)

La technologie SOLAP peut être définie comme "un type de logiciel qui permet la navigation rapide et facile dans les bases de données spatiales et qui offre plusieurs niveaux de granularité d'information, plusieurs thèmes, plusieurs époques et plusieurs modes d'affichage synchronisés ou non : cartes, tableaux et diagrammes " [Bédard 2004].

#### 3.3.1.1 Fonctionnalités d'un serveur SOLAP

Un serveur SOLAP fournit aux utilisateurs une vue multidimensionnelle des données qui peuvent être analysées grâce à un ensemble d'opérateurs SOLAP (Roll-Up, Drill-Down, etc.). De plus, le serveur SOLAP permet de gérer de façon transparente les données spatiales et non spatiales, agrégées et variantes dans le temps destinées à l'analyse d'un sujet en particulier.

#### 3.3.1.2 Opérateurs SOLAP

Les opérateurs OLAP permettent d'explorer les données multidimensionnelles en utilisant les concepts de dimensions et hiérarchies.

Un panorama des opérateurs OLAP proposés dans la littérature est présenté par Rafanelli en [Rafanelli, 2003]. Les plus communs sont :

- Les opérateurs de forage
  - **Roll-up** permet de monter dans les hiérarchies des dimensions, et d'agréger les mesures.
  - **Drill-Down** est l'inverse du Roll-Up et permet de descendre dans une hiérarchie.
- Les opérateurs de coupe
  - **Slice** utilise un prédicat défini sur les membres des dimensions pour couper une partie de l'hypercube limitant le champ d'analyse et permettant à l'utilisateur de se concentrer sur des aspects particuliers du phénomène. En utilisant la terminologie de l'algèbre relationnelle, l'opération de slice est l'équivalent de la sélection.
  - **Dice** réduit la dimensionnalité de l'hypercube en éliminant une dimension. Cette opération est équivalente à la projection de l'algèbre relationnelle.

- **Drill-Accross** met en relation plusieurs hypercubes pour comparer leurs mesures. En effet dans un processus d'exploration et d'analyse, comparer plusieurs phénomènes est fondamental pour aboutir à une connaissance finale.

Corréler plusieurs hypercubes pour avoir une vision unique des différentes mesures est donc nécessaire dans le processus d'analyse multidimensionnel. L'opération de drill-accross fusionne plusieurs hypercubes en utilisant les axes d'analyse en commun.

### 3.3.1.3 Implémentation physique d'un serveur SOLAP

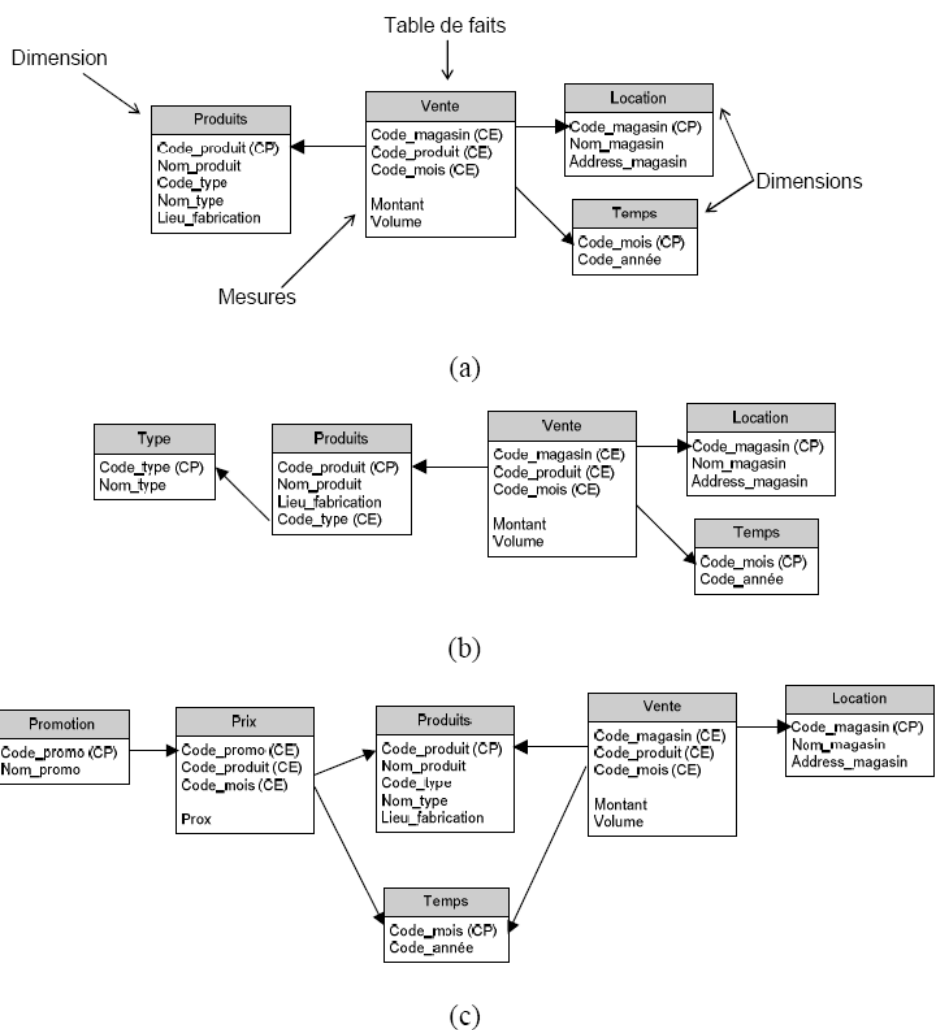
Les approches principales pour l'implémentation de serveurs SOLAP sont Multidimensional OLAP (MOLAP), Relational OLAP (ROLAP) et Hybrid OLAP (HOLAP).

Les serveurs **MOLAP** extraient les données de l'entrepôt de données Spatiale et les mémorisent en utilisant des structures de données particulières et ils appliquent des techniques d'indexation et de hachage pour localiser les données lors de l'exécution des requêtes multidimensionnelles.

Les serveurs **ROLAP** [Kimball, 1996] utilisent la technologie des bases de données relationnelles pour mémoriser les données. Pour obtenir de performances acceptables, ces systèmes utilisent des structures d'indexation particulières comme l'index bitmap, et les vues matérialisées [Winter, 1998].

La modélisation logique d'une base de données multidimensionnelle selon l'approche ROLAP ne systématise pas l'utilisation de la 3<sup>ème</sup> forme normale, contrairement aux systèmes OLTP. Le modèle logique le plus utilisé est le schéma en étoile [Kimball, 1996]. Le schéma en étoile (*Figure 3.4a*) se constitue d'une table de faits et de tables de dimensions. Chaque entrée dans la table de faits représente un fait ou une cellule du cuboïde de base, et elle est liée, à travers des clés étrangères, à des dimensions. Les tables de dimensions sont des tables dénormalisées contiennent des attributs descriptifs et des attributs qui représentent les hiérarchies des dimensions. Par exemple, un produit est caractérisé par un code, par un nom, qui représente un attribut descriptif, et un type. Le type est utilisé pour regrouper les produits en catégories différentes. La dénormalisation améliore les performances des requêtes car elle réduit le nombre de jointures. Une variante du schéma

en étoile est le schéma en flocon (*Figure 3.4b*). Il présente des dimensions partiellement ou totalement dénormalisées. La normalisation est utilisée pour réduire la redondance, surtout dans le cas où chaque niveau hiérarchique présente différents attributs et/ou quand la dimension est constituée de nombreux niveaux. Enfin, classiquement, un entrepôt de données est formé par différents hypercubes liés entre eux par des dimensions. La structure logique utilisée dans ce cas est le schéma en constellation. Il s'agit de plusieurs tables des faits qui partagent des tables des dimensions et qui peuvent être vu comme une collection d'étoiles (schéma en galaxie ou constellation de faits) (*Figure 3.4c*). Cette représentation logique permet d'utiliser l'opération de drill-across.

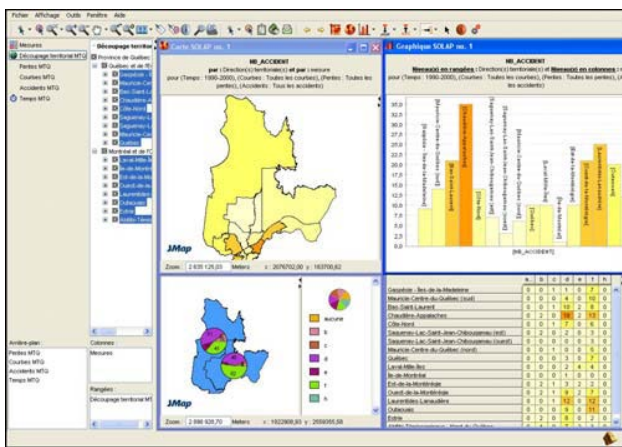


**Figure 3.4 :** a) Schéma en étoile b) Schéma en flocon c) Schéma en constellation.

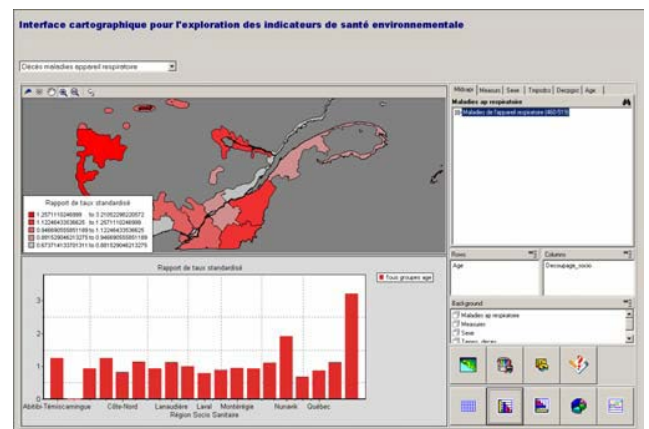
En général, les serveurs **MOLAP** [Thomsen, 1997] sont plus rapides et nécessitent moins d'espace de stockage, car les données sont stockées par le serveur SOLAP de façon à garantir ces aspects. Les serveurs ROLAP permettent le passage à l'échelle et sont plus adaptés pour des mises à jour très fréquentes.

Enfin, il existe une autre typologie qui combine les deux technologies le Hybrid OLAP. Selon cette approche, une partie des données sont stockées en relationnel et l'autre en utilisant des techniques particulières.

### 3.3.2 Exemples d'applications SOLAP



**Figure 3.5 :** Une application sur les accidents sur le réseau routier (JMap Spatial OLAP Extension et Oracle 10g) : Visualisation de la fréquence des accidents par découpage territorial (en haut) et selon les types d'accidents (en bas).



**Figure 3.6 :** Une application en santé environnementale (ProClarity et KMapx). Visualisation des cas de décès des maladies respiratoires.

La *figure 3.5* illustre une autre application en transport permet d'analyser les données relatives aux différents types d'accidents en fonction de leur position sur le réseau routier et des caractéristiques de celui-ci, le tout en fonction de différentes périodes [Rivest et al 2004].

Une application en santé environnementale permet d'explorer les relations entre les états de santé et les phénomènes environnementaux, comme l'incidence des maladies respiratoires en fonction de la qualité de l'air pour rapidement valider ou invalider une hypothèse. La *figure 3.6* présente cette application.

### 3.4 Intégration des données spatiales dans un entrepôt

Comme cité ci-dessus un entrepôt de données spatiales s'approvisionne en données à partir d'autres sources, ces dernières contiennent des données qui peuvent différer dans le format, dans le système de référence, dans l'unité de mesure etc. Pour pouvoir stocker ces données dans une même structure elles doivent avoir les mêmes propriétés, pour cela on procède à une intégration avec des outils spécialisés dans l'intégration de données spatiales.

### 3.5 Conclusions

Dans ce chapitre, nous avons traité le sujet des entrepôts de données qui étendent les concepts et la technologie traditionnelle des bases de données pour créer des systèmes d'aide à la décision, Les ED ont à leur tour étaient étendus pour donner naissance aux EDS dont nous avons vu les concept et le potentiel dans les exemples d'applications sur l'environnement, la santé et les accidents sur le réseau routier.

En utilisant l'architecture d'un entrepôt de données Spatiale, nous avons expliqué les différents composants qu'il intègre, comme les diverses sources, les types de données et les différents outils pour arriver à la visualisation de l'information. Nous avons décrit les différents modèles multidimensionnels pour la construction d'un entrepôt de données, ainsi que les différentes opérations pour la manipulation des données multidimensionnelles.

# CHAPITRE IV

## Conception

- 4.1 Introduction
- 4.2 Présentation du Projet
- 4.3 La Wilaya de Laghouat, Géographiquement
- 4.4 La DTP
- 4.5 Étude de l'existant
- 4.6 Buts du SIAD routier
- 4.7 Conception du SIAD routier
  - 4.7.1 Réalisation de la Carte routière
  - 4.7.2 Logicielle utilisée
  - 4.7.3 Conception de l'entrepôt de données spatiale
  - 4.7.4 Structure du schéma
  - 4.7.5 Les hiérarchies
  - 4.7.6 Détails du fait
  - 4.7.7 Détails des dimensions
  - 4.7.8 Schéma relationnel détaillé
- 4.8 Conclusion

### 4.1. Introduction

La conception d'un entrepôt de données est une tâche complexe et délicate. Elle se compose de plusieurs étapes. Dans un premier temps, nous devons analyser l'ensemble des sources de données internes et externes. Cette analyse sert aussi bien à la sélection de l'ensemble de données à stocker dans l'entrepôt, qu'à la sélection des outils requis pour l'extraction et la transformation de ces données avant leur stockage. La deuxième étape consiste à organiser ces données à l'intérieur de l'entrepôt. Pour ce faire, nous devons concevoir le modèle multidimensionnel à utiliser.

Ce chapitre présente dans une première partie, la description ou bien la réponse à la question « *A qui s'adresse ce système ?* » et les principaux Buts pour lesquels a été conçue. Nous définissons ainsi dans une deuxième partie, un modèle multidimensionnel qui est composé de trois classes : Cube, Dimension et Hiérarchie.

### 4.2. Présentation du projet

Ce projet consiste à réaliser un système d'analyse du taux de trafic (Véhicule\Heure), du taux d'accidents (Accident\Mois), du taux de violation du code de la route par composant puis secteur ainsi que du taux de croissance du nombre de véhicule sur la route sur une longue période. Ce système devrait permettre à la DTP (Direction des Travaux Public) de faire un constat du taux d'amélioration engendré par les nouvelles installations, et essentiellement ce système doit répondre aux questions citées dans le point **4.6**, et de ce fait aider à mieux planifier ou à améliorer l'installation routière.

### 4.3. La Wilaya de Laghouat, Géographiquement

La wilaya de Laghouat est une wilaya algérienne ayant pour chef-lieu la ville du même nom. Située au cœur du pays à 400 km au sud de la capitale Alger, la wilaya s'étend sur une superficie de 25 052 km<sup>2</sup> et compte 484 252 habitants. Région pastorale de l'Algérie, elle possède également le plus grand gisement de gaz naturel d'Afrique.

Le climat est continental aride avec des moyennes de -5 °C l'hiver et de plus de 40 °C l'été. La wilaya est elle-même divisée en dix daïras et 24 communes.



**Figure 4.1 :** Localisation de la wilaya de Laghouat

Avec tous ces chiffres nous pouvons constater que cette ville a un trafic routier remarquable, et dans ce stade les services concernés veillent.

### 4.4. La DTP

La Direction des travaux publics (DTP) de la wilaya de Laghouat a entrepris un vaste programme devant répondre essentiellement aux problèmes de congestion de la circulation.

Il a pour objectifs, la densification du réseau routier par de nouvelles réalisations ou l'augmentation des capacités existantes (élargissement, dédoublement, double voie, passerelle,...), le traitement des points de congestion, l'amélioration des zones carrefours, le maintien d'un niveau de service appréciable du réseau.

Durant ces deux dernières années, la DTP réceptionnera des nouveaux projets. Il s'agit des carrefours, de la deuxième section de la RN1, et de nombreux double voix a travers toute les routes principales de la wilaya.

L'idée de réaliser un système pour la régulation du trafic routier devrait permettre à la DTP de faire un constat du taux d'amélioration engendré par les nouvelles installations, des projections sur l'avenir et de ce fait aider à mieux planifier ou à améliorer l'installation routière.

### **4.5. Étude de l'existant**

La gestion des projets routiers au niveau de la DTP de Laghouat est documentaire, les projets antérieurs sont sous format papier (fichier.DOC) et les cartes sont sous forme de fichier Autocad ou papier (Format A1), chaque fois qu'un projet de construction ou d'entretien sur une route se lance, un rapport contenant toute les informations concernant ce projets est créer, en plus coté informatique, la DTP ne dispose pas d'un logiciel d'archivage des données routières de ce fait, ils ont besoin d'un système (Outil) dynamique qui synthétise ces données, et donne une vue générale sur les informations utiles.

### **4.6. Buts du SIAD routier**

Le système interactif d'aide à la décision pour la régulation du trafic routier, doit répondre aux questions suivantes :

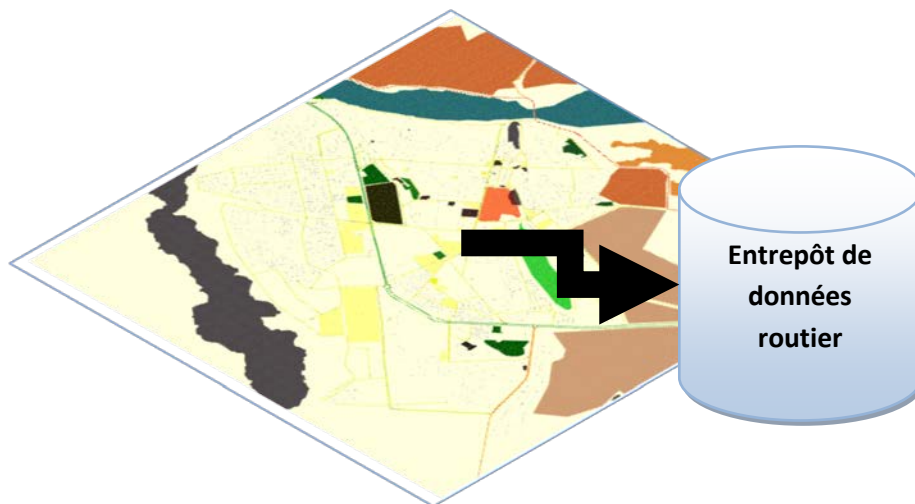
- Quelles sont les routes qui ont un taux de trafic élevé/bas par une unité de temps ou par zone/quartier ?
- Quelles sont les routes ou bien les tronçons routiers les plus/moins fréquentés ?
- Quelles sont les périodes (jour, heures, mois, saison) les plus embouteillées ?
- Quelle est le taux d'accident sur telle route dans telle période sur une condition météorologique quelconque (beau-temps/mauvais-temps) ?
- Quelle est la corrélation taux de trafic, temps de parcours ?
- Et bien plus d'autres questions qui peuvent être posé par les décideurs pour renforcer leurs décision.

### 4.7. Conception du SIAD routier

Pour réaliser une tel solution (*SAD pour la régulation du trafic routier*), la structure d'entrepôt de données spatiale est la plus adapté par ce qu'elle supporte l'analyse et elle offre un temps de réponse rapide par rapport aux structures classiques. Une carte détaillée des routes de la commune de Laghouat a été réalisée afin d'alimenter la grande partie de l'entrepôt, plus une interface SOLAP interactifs été crée pour afficher les résultat d'une analyse d'une façon simple et compréhensible.

#### 4.7.1. Réalisation de la Carte routière

La carte routière que nous avons réalisée joue le rôle d'une base de données cartographique, elle contient les données détaillées sur les routes de la ville de Laghouat, et comme signalé plus haut elle représente la base de données source qui alimentera la grande partie de l'entrepôt de données spatiales.



**Figure 4.2 :** Etapes de la conception de l'EDS routier

Sa réalisation est passé par plusieurs étapes, après des recherches intensives sans résultats apparent sur une carte routière, nous avons fini par l'idée de réaliser cette carte nous mêmes. Sur le site OpenStreetMap<sup>1</sup>, nous avons trouvé ce que nous cherchions, mais il était un défi pour nous de collecter des quarantaine de poses de format 1820x1712 chacune.

<sup>1</sup> <http://www.openstreetmap.fr>

Ensuite nous avons procédé au mosaïquage une opération cartographique qui consiste à assembler les images une à une pour former une seule carte de 14560x14000 tant que résolution d'image, le fichier obtenue est de taille de 236873 Ko.

Le travail où nous avons passé beaucoup de temps c'été la vectorisation manuelle de l'image de la carte, une autre opération cartographique qui s'agit de tracer les différentes routes avec la souris, pour obtenir un ensemble de coordonnées retrassant leur formes, ensuite nous l'avions géoréférencier en utilisant quelques informations cartographique nécessaire pour réaliser un telle travail, le tableau suivants illustre ces informations.

<b>La projection</b>	<b>Long\Lat WGS 84 (World Geodesic System 1984)</b>
<b>Left longitude</b>	<b>2.844858</b>
<b>Right longitude</b>	<b>2. 894554</b>
<b>Top latitude</b>	<b>33.824347</b>
<b>Bottom latitude</b>	<b>33.777059</b>

*Tableau 4.1 : Les paramètres de la projection*

### 4.7.2. Logicielle utilisée

La base de données cartographique est réalisée avec le leader des SIG, MapInfo professionnelle 9.0.

MapInfo<sup>2</sup> Professional est un Système d'Information Géographique à l'origine Bureautique créé dans les années 1980 aux État-Unis. C'est un logiciel qui permet de réaliser des cartes en format numérique. MapInfo est conçu autour d'un moteur d'édition de cartes qui permet la superposition de couches numériques. Il permet de représenter à l'aide d'un système de couches des informations géo-localisées : points, polygones, image raster ... Il incorpore un grand nombre de formats de données, de fonctions cartographiques et de gestion de données... Un système de requêtes cartographiques adapté permet la conception des cartes et bases de données cartographiques. MapInfo est ouvert vers le Web et les

---

<sup>2</sup> <http://fr.wikipedia.org/wiki/MapInfo>

globes virtuels ; il permet de publier sur le web des cartes réalisées sur un PC, de faire de la cartographie interactive, d'incorporer des informations des globes virtuels...

MapInfo Professional est un logiciel destiné aux chargés d'étude et d'aménagement territorial, aux chargés d'études d'implantation, de géomarketing, aux analystes des réseaux physiques et commerciaux.

Connexe et complémentaire au logiciel desktop MapInfo Professional, il existe une vaste gamme de modules permettant de faire des traitements de localisation par géocodage automatique d'adresses, de l'optimisation des déplacements routiers, des analyses de risques locaux (Crédit, Assurance,...), des analyses géomarketing et socio-démographique, de l'enrichissement de fichiers d'adresses à l'aide d'informations localisées, ainsi que de la diffusion de cartes et données sur support WEB.

### 4.7.3. Conception de l'entrepôt de données spatiale

Plusieurs chercheurs ou équipes de recherche ont essayé de proposer des démarches pour la réalisation d'un projet Data Warehouse, ces démarches se croisent essentiellement dans les étapes suivantes :

- Modélisation et conception du Data Warehouse,
- Alimentation du Data Warehouse,
- Mise en œuvre du Data Warehouse,
- Administration et maintenance du Data Warehouse,

Pour notre projet nous avons choisis pour l'analyse, la méthode de *Monsieur R. Laurin*<sup>3</sup> car elle est à la fois facile et très efficace.

Le but est de déceler les axes d'analyses (les dimensions) avec leurs attributs ainsi que les éléments à analyser (les faits). La meilleure façon de ce faire, est l'étude approfondie de ce qui se passe dans l'entreprise : documents échangés, rapports périodiques, interviews des personnes clés, étude des besoins. Il faut vraiment faire un travail d'acteur, et rentrer dans la peau de chaque utilisateur, savoir comment les analystes organisent leurs raisonnements, savoir ce que voient les décideurs avant de décider, connaître les indicateurs de bonne santé de l'entreprise et de la concurrence. Un vrai travail de fourmi.

---

<sup>3</sup> Méthode enseignée à l'université Sherbrooke par Monsieur R. Laurin

## Conception

Une manière très pratique de modéliser un cas en décisionnel se fait comme suit :

	Temps	Zone géographique	Météo	Annotation
	Heure	Zone	Meteo	Annotation
	Jour	Quartier		
	Mois	Route		
	Année			
<b>Analyse</b>				
Taux_trafic				
Taux_accident				
Temps_parcours				
Taux_embouteillage				

**Tableau 4.2** : Préparation pour la Conception de l'entrepôt

Explications : le tableau précédent a été rempli pendant la phase d'analyse, en posant des questions aux décideurs du type :

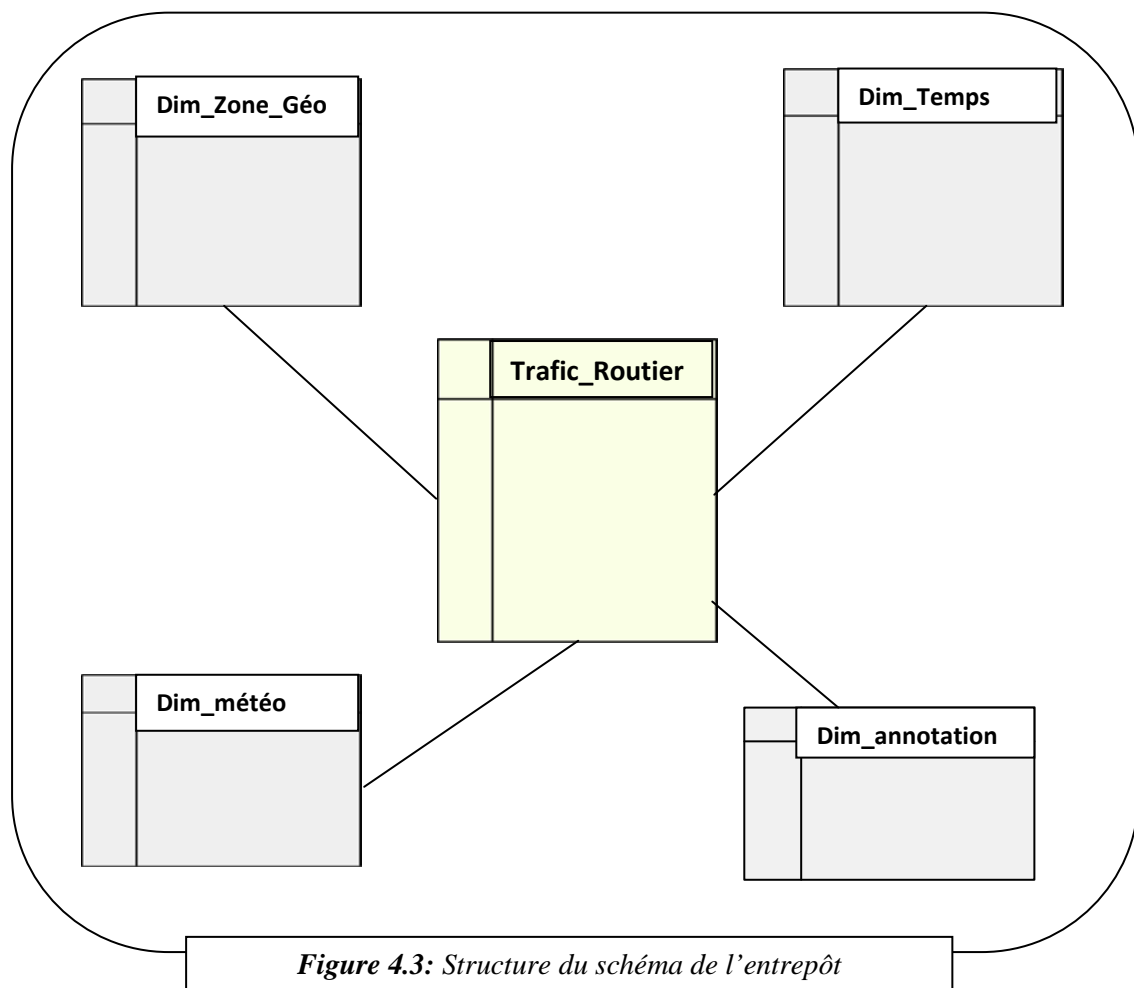
- Que voulez vous analyser (la dernière ligne du tableau) ?
- Quels sont vos critères d'analyse (la première ligne du tableau) ?
- Jusqu'à quel niveau de détail voulez vous aller (les cellules à l'intérieur) ?

La conception d'un ED Spatiale suit les mêmes étapes que celles d'un ED classique, à savoir Structure du schéma qui représente un premier croquis de la structure de l'entrepôt, l'Identification des hiérarchies qui se fait en étudiant les valeurs des différentes données de l'entrepôt, le Détail du fait et des dimensions sous la forme d'un tableau qui contient les différents champs, Schéma relationnel détaillé, le Schéma global et la création physique de l'entrepôt.

### 4.7.4. Structure du schéma

L'analyse se fait spécifiquement sur des valeurs des Taux au niveau des routes, quartiers et zones de la communes en utilisant les paramètres temporelle et météorologiques, comme cité auparavant c'est alors toutes les données que se soit cartographiques ou spatio-temporelle seront contenues dans l'entrepôt de données, le reste sera estimées. On a donc 4 dimensions de types différents, une dimension spatiale Dim\_Zone\_Géo, une dimension nominale Dim\_Météo, une dimension temporelle Dim\_Temps et une dimension Dim\_Annot qui représente tout les autre acteurs qui influent sur le trafic routier.

L'entrepôt de données spatiales est conçu avec des tables de type Paradoxe7.0



#### 4.7.5. Les hiérarchies

Les hiérarchies identifiées sont les suivantes :

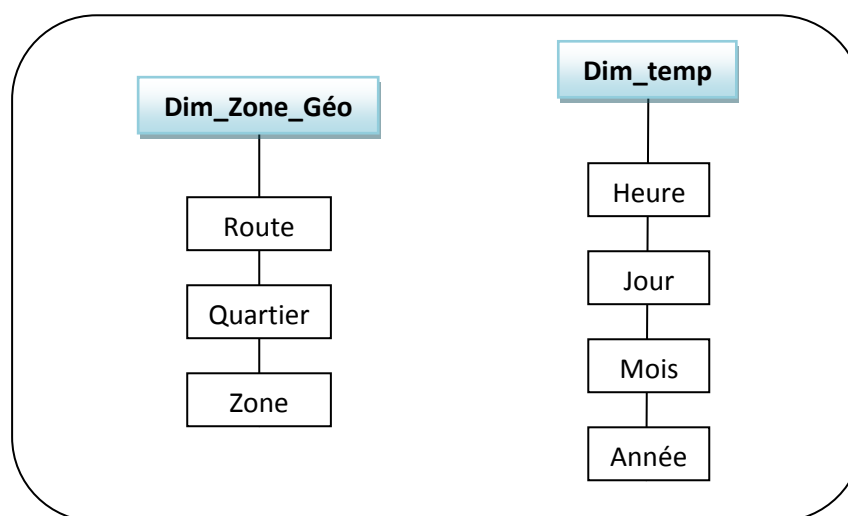


Figure 4.4: Représentation des hiérarchies

#### 4.7.6. Détails du fait

L'identifiant du fait est constitué de la concaténation des identifiants des trois dimensions qui l'entourent, trois champs de type clés étrangères vont donc être ajoutés à la table fait.

Field Name	Type	Désignation	Formule d'extraction
<b>Fait_Trafic</b>			
<b>Id_trafic</b>	Long integer	Identificateur Trafic (key)	Auto incrémenté
<b>Id_routier</b>	Long integer	Identificateur routier	Routier.id_routier
<b>Id_heure</b>	Long integer	Identificateur heure	Heure.id_heure
<b>Id_meteo</b>	Short integer	Identificateur météo	Meteo.id_meteo
<b>Taux_traffic</b>	Number	Le taux de trafic sur chq route	Estimation (Veh/Km)
<b>Taux_accident</b>	Number	Le taux d'accident	Estimation (Acc/mois)
<b>Taux_viol_cdr</b>	Number	-----	-----
<b>Taux_embouteillage</b>	Number	Le taux d'embouteillage	Estimation (%)
<b>Temps_de_parcours</b>	Number	Le temps d'attente moyen	Estimation (Sec/Km)

Tableau 4.3: Tableau représentatif des détails du fait

4.7.7. Détails des dimensions

Field Name	Type	Désignation	Formule d'extraction
<b>Dim_Routier</b>			
<b>Id_routier</b>	Long integer	Identificateur routier	Route.tab (MapInfo)
<b>Nom_routier</b>	Alphabet	Nom de la route	Route.tab (MapInfo)
<b>Longueur</b>	Number	La longueur de la route	ObjectLen(obj, "km")
<b>Largeur</b>	Short integer	La largeur de la route	Route.tab (MapInfo)
<b>Etat_routier</b>	Alphabet	L'état de la route	Etat actuel 'bonne'
<b>Date_constr</b>	Date	La date de construction rte	Estimation
<b>Date_suppr</b>	Date	-----	-----
<b>Date_dern_maint</b>	Date	La date de dernier maint	Estimation
<b>Type_routier</b>	Alphabet	Type de la route	3 type routière
<b>Geometry_X</b>	Alphabet	Coordonnée X de la route	CentroidX(obj)
<b>Geometry_Y</b>	Alphabet	Coordonnée Y de la route	CentroidY(obj)
<b>Id_quartier</b>	Short integer	Identificateur du Quartier	Quartier.id_quartier
<b>Dim_quartier</b>			
<b>Id_quartier</b>	Short integer	Identificateur Quartier	Quart.tab (MapInfo)
<b>Nom_quartier</b>	Alphabet	Nom du quartier	Quart.tab (MapInfo)
<b>Geometry_X</b>	Alphabet	Coordonnée X du Quartier	CentroidX(obj)
<b>Geometry_Y</b>	Alphabet	Coordonnée Y du Quartier	CentroidY(obj)
<b>Id_zone</b>	Short integer	Identificateur de la Zone	Zone.id_zone
<b>Dim_Zone</b>			
<b>Id_zone</b>	Short integer	Identificateur Zone	Zone.tabe (MapInfo)
<b>Nom_zone</b>	Alphabet	Nom de la zone	Zone.tabe (MapInfo)
<b>Geometry_X</b>	Alphabet	Coordonnée X de la zone	CentroidX(obj)
<b>Geometry_Y</b>	Alphabet	Coordonnée Y de la zone	CentroidY(obj)
<b>Dim_Heure</b>			
<b>Id_heure</b>	Long integer	Identificateur de l'heure	Auto incrémenté
<b>Heure</b>	Time	L'heure « HH :MM :SS »	Chargement direct
<b>Id_jour</b>	Short integer	Identificateur du jour	Jour.id_jour

## Conception

Dim_Jour			
<b>Id_jour</b>	Short integer	Identificateur Jour	Auto incrémenté
<b>Jour</b>	Short integer	Le numero du jour (calndar)	1 – 31
<b>Jour_semaine</b>	Alphabet	Le nom du jour	Jour dans semaine
<b>Id_mois</b>	Short integer	Identificateur du mois	Mois.id_mois
Dim_Mois			
<b>Id_mois</b>	Short integer	Identificateur Mois	Auto incrémenté
<b>Mois</b>	Alphabet	Nom du mois	Mois dans l'année
<b>Id_annee</b>	Short integer	Identificateur de l'année	Annee.id_annee
Dim_Année			
<b>Id_annee</b>	Short integer	Identificateur Annee	Auto incrémenté
<b>Annee</b>	integer	L'année	manuel
Dim_Météo			
<b>Id_meteo</b>	Short integer	Identificateur Meteo	Auto incrémenté
<b>Description</b>	Alphabet	Beau\Mauvais temps	Manuel

*Tableau 4.4 : Tableau représentatif des détails des dimensions*

### 4.7.8. Schéma relationnel détaillé

La modélisation en flocon existe pour des raisons de performances. En effet, des dimensions de plusieurs millions de lignes peuvent poser des problèmes de lenteur lors de l'exploitation des données.

Le principe de la modélisation en flocon est de créer des hiérarchies de dimensions, de telle manière à avoir moins de lignes par dimensions.

Le schéma d'une modélisation en flocon de notre entrepôt pourrait être comme suit :

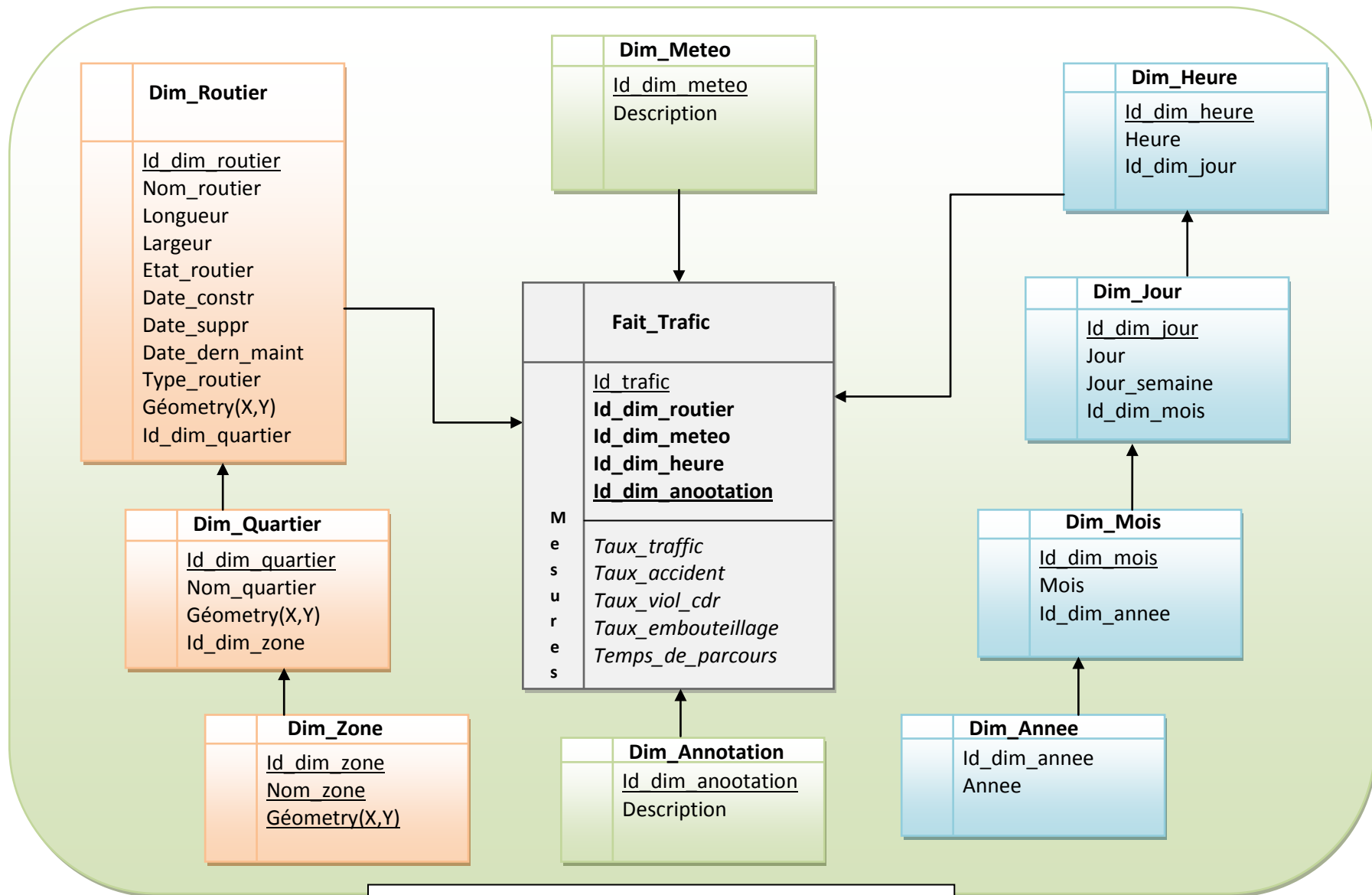


Figure 4.5: Schéma relationnel de l'entrepôt

### 4.8. Conclusion

Dans notre collecte de donnée nous avons utilisé toutes les astuces que nous avons appris dans le module de génie logiciel, en plus on a essayé des nouvelle technique pour profiter du help des habitants de la commune de Laghouat tel que le célèbre réseau sociale *Facebook*. Par ce qu'il été difficile d'obtenir des information auprès des entreprises travaillants dans ce secteur. Donc nous étions obligé d'utiliser des données routières estimées, puisqu'il faut faire un travail de sensibilisation pour avoir des bonnes résultat.

Néanmoins, la conception de ce système est achevée, et en comparaison avec des travaux antérieur comme Le projet CADDY [voir **Annexe2**], nous pouvons dire que notre conception est complète.

le Chapitre V qui vient juste après, présente l'étape de réalisation ainsi que les testes sur l'application finale.

# *CHAPITRE V*

## **Réalisation**

- 5.1 Introduction
- 5.2 Paramètres machines
- 5.3 Architecture technique de la solution
- 5.4 Logiciels utilisés
- 5.5 La Carte
- 5.6 L'entrepôt de données spatiale
- 5.7 Le système obtenue
- 5.8 Testes Et exemples
- 5.9 Conclusions

## 5.1. Introduction

Ce chapitre présente toutes les étapes que nous avons suivies pour réaliser le système d'aide à la décision pour la régulation du trafic routier de la commune de Laghouat, allant de la description des paramètres machines sur lesquels ce SIAD va fonctionner, passant par les techniques et logiciels utilisés, jusqu'aux tests et exemples sur l'application finale.

## 5.2. Paramètres machines

Cette partie décrit les infrastructures déjà en place, en effet c'est une étape à ne pas négliger, car la diversité des sources et leurs plateformes techniques pourront engendrer des problèmes de compatibilité.

Le SIAD sera installé sur des machines INTEL ayant de préférence les caractéristiques :

- Processeur : Core2Duo
- Mémoire RAM : 1Go +
- Espace Disque : + de 512Mo
- Système d'Exploitation : Windows Xp 32bits
- Résolution graphique : 1024x768 +

## 5.3. Architecture technique de la solution

La figure suivante illustre l'architecture technique de la solution proposée :

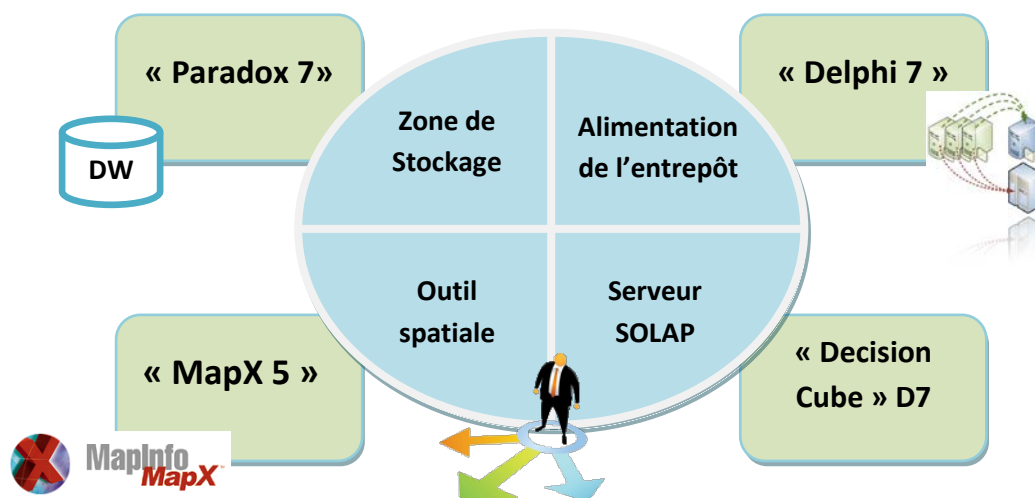


Figure 5.1 : Architecture technique de la solution

### 5.4. Logiciels utilisés

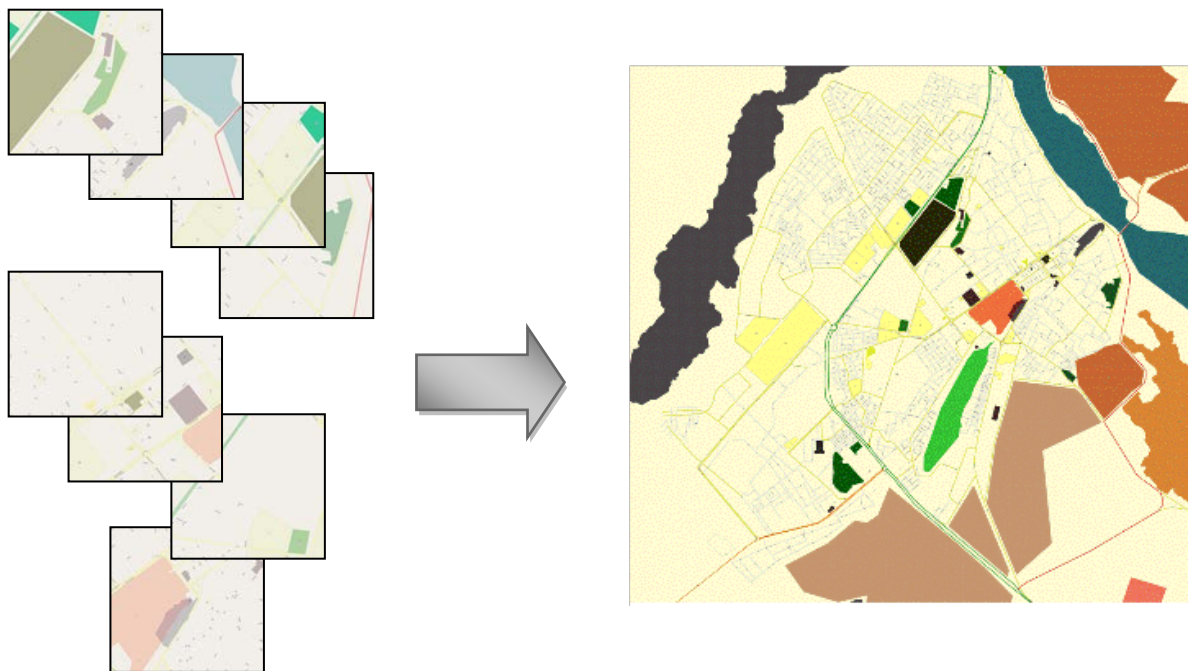
Nous avons choisi de travailler avec **Mapinfo** pour la réalisation de la carte, en effet il satisfait à tous nos besoins.

De plus pour la réalisation de l'interface utilisateur nous avons choisi l'environnement de développement **Delphi** car c'est l'un des outils qui supporte le composant cartographique **MapX 5**, il sera complété par **DécisionCube** pour la manipulation des résultats de l'analyse. Et pour le côté esthétique de l'interface nous avons utilisé le célèbre logiciel d'infographie **Adobe Photoshop CS2**.

Pour l'entrepôt de données nous avons décidé de travailler avec des tables de type **Paradox** puisqu'il est délivré avec Delphi, mais ce type de table a une certaine capacité de stockage (la taille de la table  $\approx$  80Mo).

### 5.5. La Carte

Tout d'abord nous avons capturé plusieurs poses à partir du site web, puis nous les avons assemblées comme montre la figure 5.2, ensuite nous avons géoréférencé la carte globale avec la même projection que celle du site internet et le résultat de cette technique est montré dans la figure 5.3.



**Figure 5.2 :** Capture de la carte routière de la commune de Laghouat

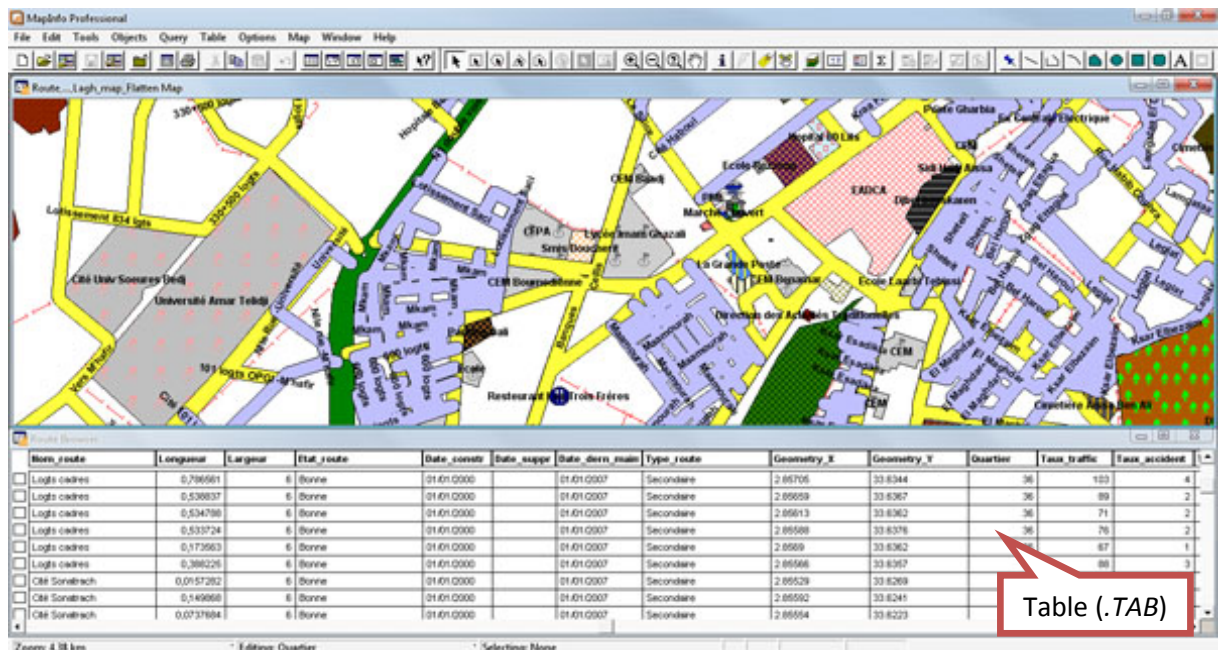


Figure 5.3 : La carte routière de la commune de Laghouat avec MapInfo 9.0

La Carte réalisée est constituée de 4 couches superposées, la première couche c'est la couche Route, la deuxième c'est la couche Quartier, la troisième c'est la couche Zone et la dernière couche pour les autres objets figurant sur la carte (faisant figure de repères).

Chacune de ces couches génère un fichier .TAB (Voir la Figure Précédente) qui sera exploité par le composant MapX pour visualiser la carte dans l'application finale.

## 5.6. L'entrepôt de données spatiale

Comme nous l'avons annoncé plus haut, nous avons implémenté l'entrepôt sur des tables de type *paradox* comme il est montré dans la figure 5.4.

L'alimentation de l'entrepôt se produit manuellement car un événement sur le réseau routier peut se produire à n'importe quel moment, mais elle peut aussi être déclenchée à tout moment ou selon une autre politique de mise à jour (chaque semaine, chaque mois etc.).

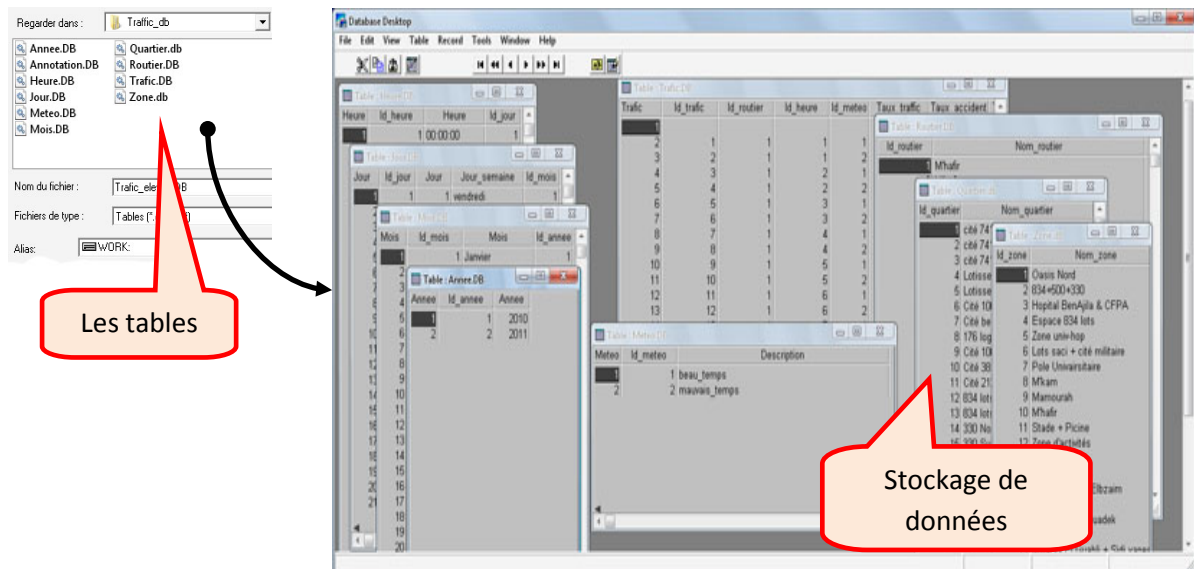


Figure 5.4 : Les tables de l'EDS avec Paradox 7.0

### 5.7. Le Système obtenue

Nous avons choisi de travailler avec le langage de programmation Delphi car il offre toutes les fonctionnalités dont nous avons besoin en plus d'être le langage qui supporte la bibliothèque spatiale *MapX*, cette dernière fait partie de la même gamme de produit que *MapInfo* le logiciel SIG que nous avons utilisé pour créer le carte.

L'interface est construite de manière à être intuitive et facile à utiliser suivant les règles régissant les interfaces SOLAP. Elle se compose de trois types d'affichage de données carte, diagramme et tableau croisé, elle comporte aussi des outils de navigation classiques et des outils d'analyse (navigation SOLAP), une fenêtre de navigation plus grande, une fenêtre d'affichage des données attributaires (Nom, Longueur etc.), la liste des options d'analyse disponibles (Taux de trafic, Nombre d'accidents par route quartier et zone, etc.) et l'ensemble des couches affichées avec possibilité d'inhiber ou d'afficher chaque couche.

La page d'accueil

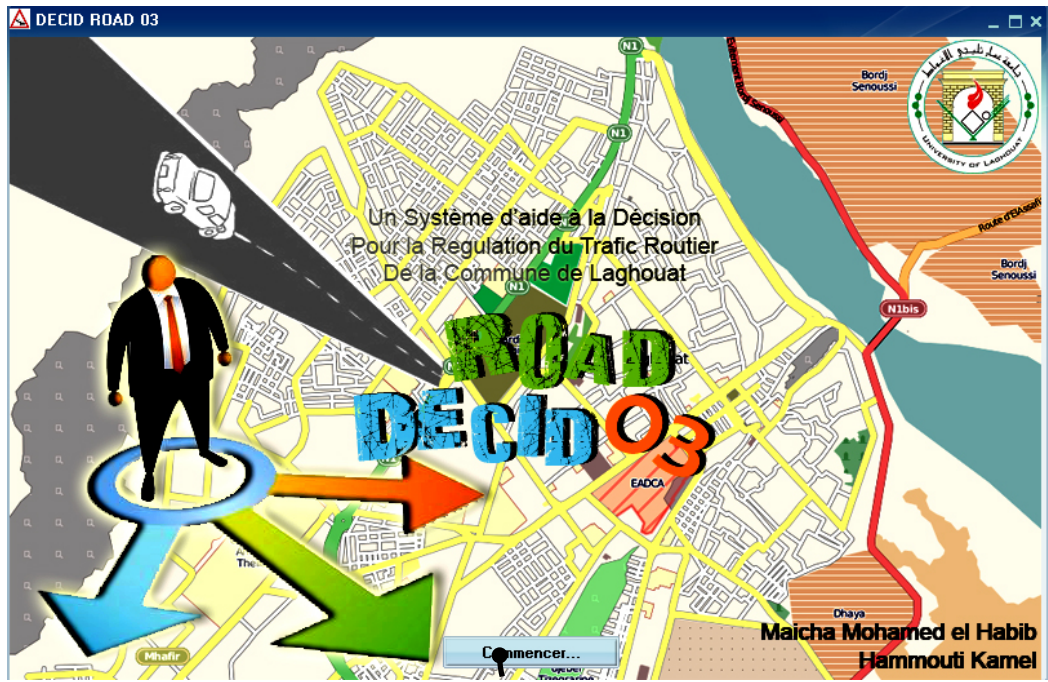


Figure 5.5 : La page d'accueil de l'application

La page principale

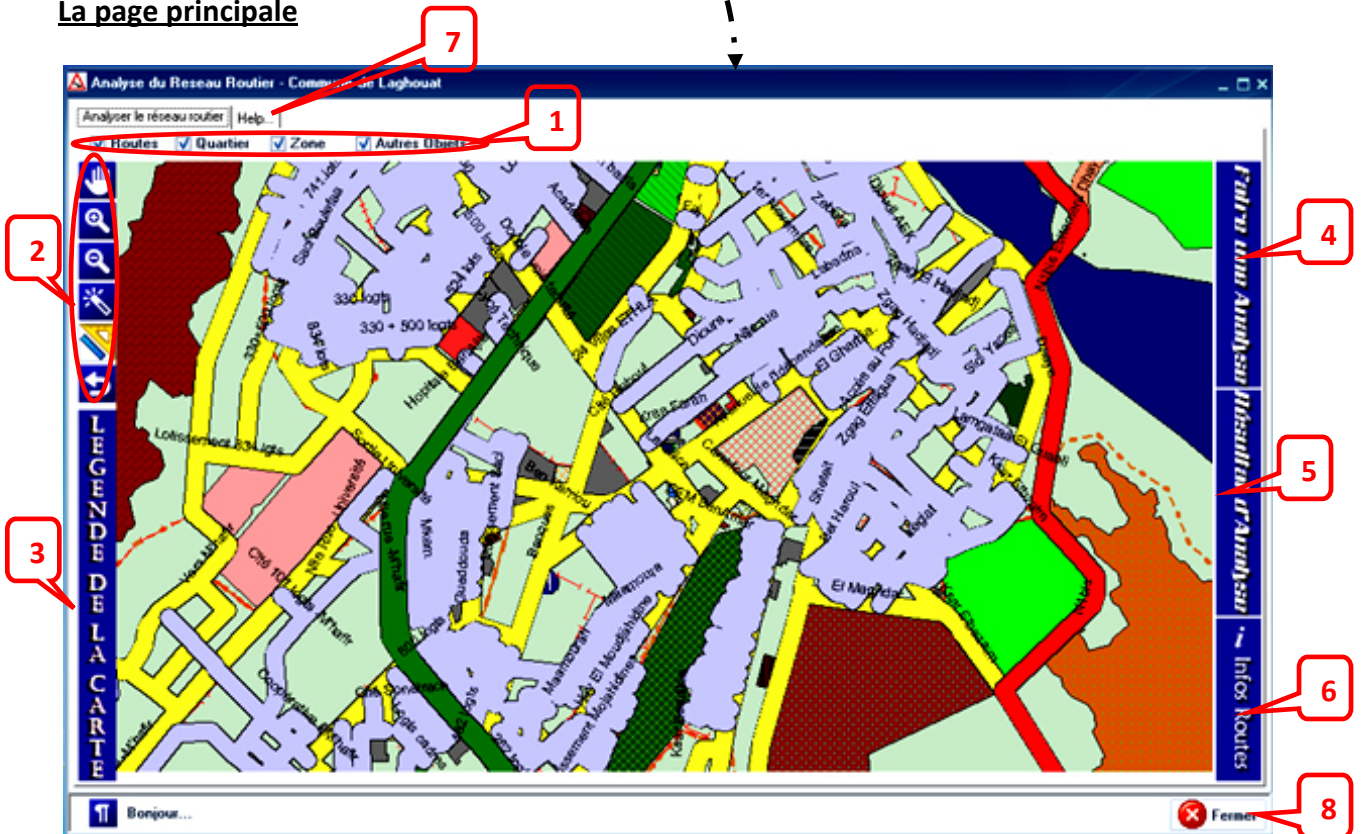


Figure 5.6 : La page principale de l'application

## Réalisation

- 1) L'ensemble des couches affichées avec possibilité d'afficher ou masquer chaque couche.
- 2) L'ensemble d'outils de manipulation de la carte.
- 3) La légende de la carte.
- 4) Afficher le volet de l'analyse.
- 5) Afficher le volet de résultats d'analyse.
- 6) Afficher le volet des informations sur les routes.
- 7) Le Help.
- 8) Quitter l'application.

### Le Volet d'analyse

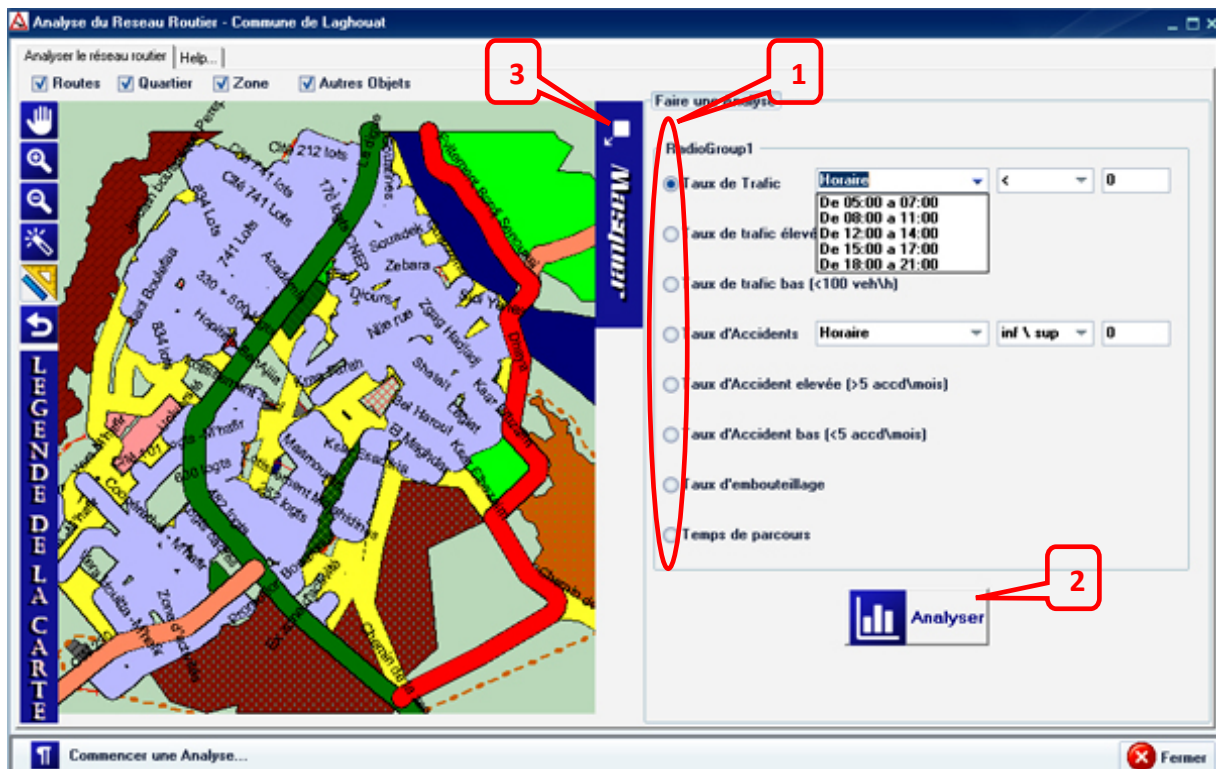


Figure 5.7: La page de l'analyse

- 1) L'ensemble des options d'analyse.
- 2) Démarrer l'analyse.
- 3) masquer le volet de l'analyse.

La Volet diagramme et tableau croisé

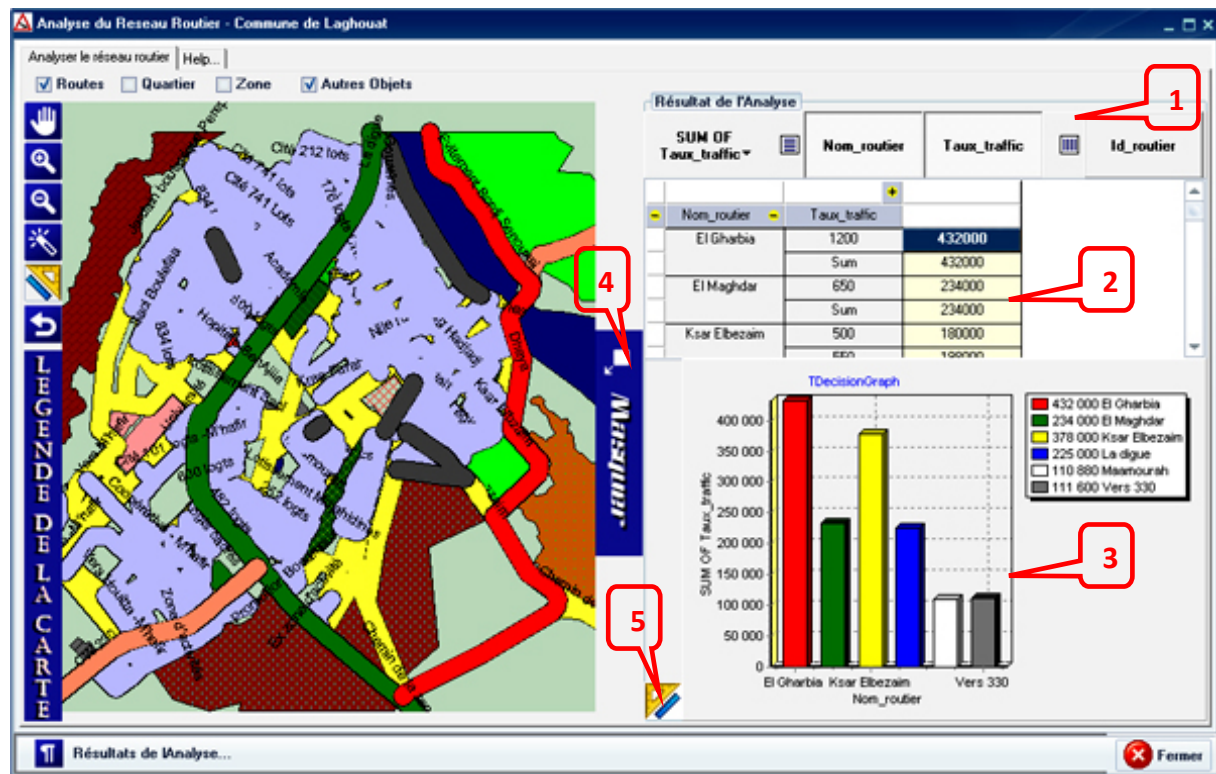


Figure 5.8: La page de résultats de l'analyse

- 1) Les dimensions ouvertes lors d'une analyse.
- 2) Le tableau croisé du résultats d'une analyse.
- 3) Les résultats sous forme graphique.
- 4) Masquer le volet résultats de l'analyse.
- 5) Agrandir \ minimiser le diagramme statistique.

La page d'informations sur les routes

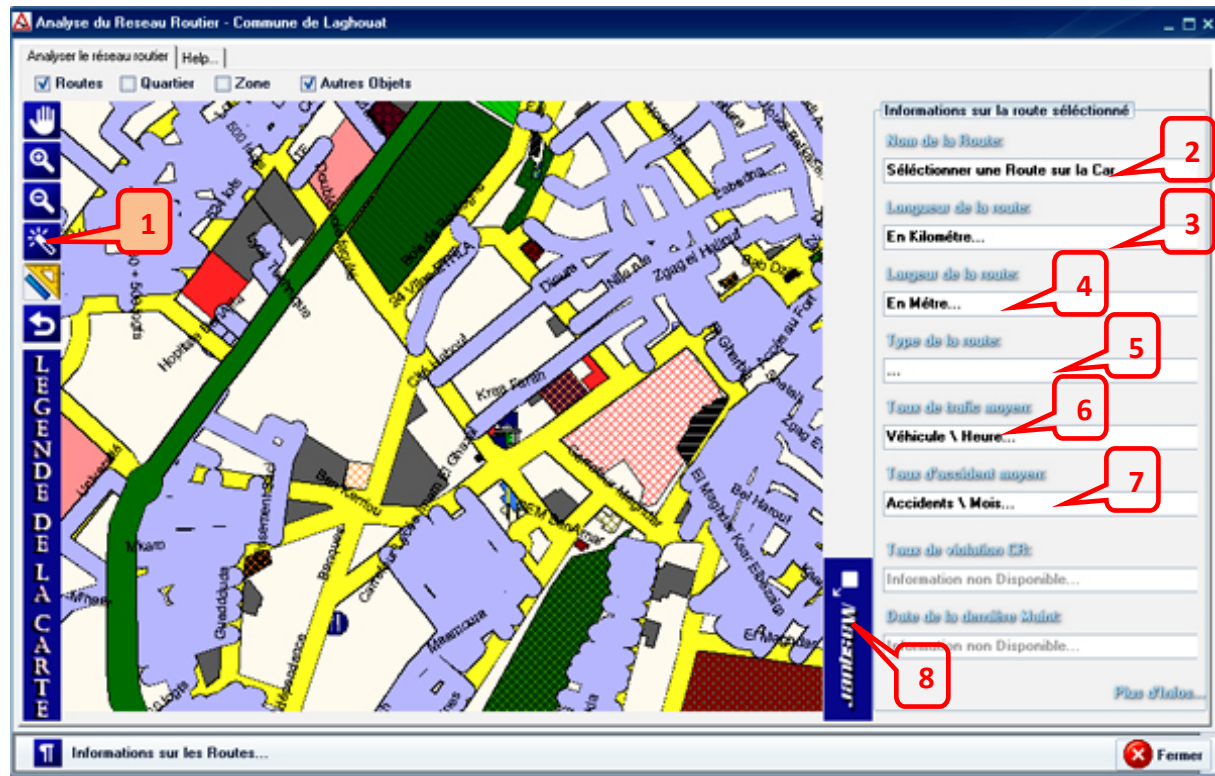


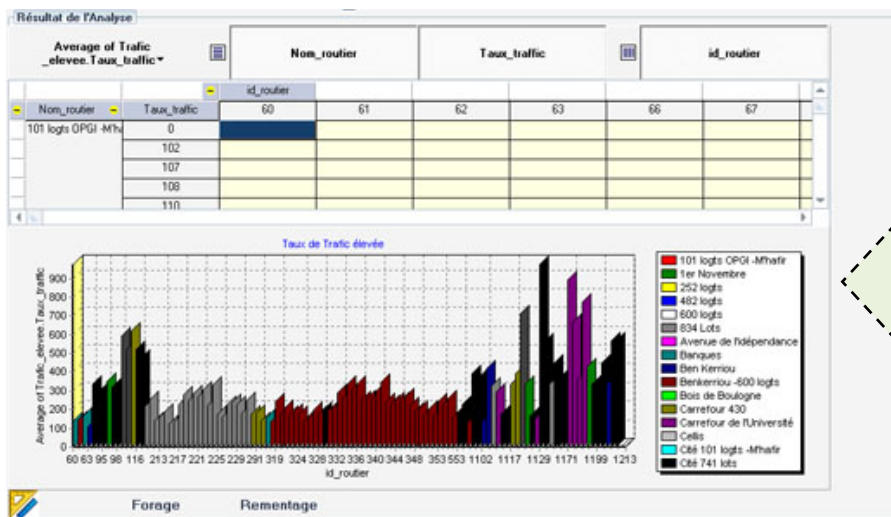
Figure 5.9 : La page d'informations sur les routes

- 1) L'outil de sélection d'une routes pour afficher ses informations.
- 2) Le nom de la route.
- 3) La longueur de la route.
- 4) La largeur de la route.
- 5) Le type de la route.
- 6) Le taux de trafic moyen en Véhicule \ Heure.
- 7) Le taux d'accident moyen en Accident \ Mois.
- 8) Masquer le volet des informations sur les routes.

### 5.8. Tests & exemples

Nous prenons comme exemple de teste sur l'application, l'analyse par taux de trafic supérieur a 100 veh\h. En aura les résultats sous forme du tableau croisé et diagramme a bars et d'un autre coté les routes qui ont un taux de trafic > 100 veh\h sont apparues sur la carte avec une couleur différente (Noirs).

On choisit la 2<sup>ème</sup> option dans la liste des options d'analyse.



En haut nous trouverons les dimensions ouvertes par la requête, au milieu Ilya le tableau croisé des résultats, et en bas se trouve les résultats sous forme graphiques (diagramme a bars).

Cette figure illustre Les résultats de cette exemple d'analyse sur la carte routière, où Les tançons avec une couleur **noire** représente les routes ayant un **trafic > 100 veh\h**.



### **5.9. Conclusion**

Ce chapitre montre l'importance du travail que nous avons fait, parce qu'il n'y avait aucune carte des routes dans la ville de Laghouat. nous avons décidé de concevoir nous-mêmes cette carte qui a pris plus de trois mois dans le temps de réalisation.

Ensuite, le chargement de l'entrepôt de données qui été un processus lourd vu le volume important des tables.

mais tout cela ne pouvait pas nous empêcher de terminer la réalisation de DECID-ROAD 03, les résultats des testes effectués sur l'application illustrés a la fin de ce chapitre, montre bien que DECID-ROAD 03 peut aider a mieux planifier la gestion routière de la ville de Laghouat.

# *CHAPITRE VI*

## **Conclusion Générale**

- 6.1 Conclusion
- 6.2 Perspectives

### 6.1. Conclusion générale

Tout au long de ce modeste travail nous avons pu nous introduire dans le domaine de l'informatique décisionnelle et aussi dans le domaine des applications cartographiques plus précisément dans celui des entrepôts des données spatiales. D'un autre coté nous avons appris un nouvel art , très tendance et d'actualité à savoir le Trafic routier dans son ensemble d'acteurs concernés qui à fait de lui un domaine d'étude complexe.

Autre acquis très précieux est celui de la manière de mener un projet, d'enquêter et de tout mettre en œuvre pour valider une hypothèse. Mais comme tout autre projet informatique nous avons rencontré de gros problèmes et de nombreux obstacles essentiellement dus au manque des sources d'informations et la majorité sont liés a la non sensibilisation des personnes à l'importance des Systèmes d'aide à la décision.

Pour finir cette formidable expérience, et avant de citer les perspectives du projet, nous pouvons dire que ce travail nous a permis d'acquérir une très bonne expérience professionnelle et d'évoluer dans des domaines intéressants qui nous était totalement inconnus. Sur un plan personnel nous aimerons dire que cette thèse nous a ouvert l'appétit sur la recherche et nous a appris la persévérance, ***“et que la découverte est une quête longue mais vitale même si les résultats ne sont qu'une pierre dans un édifice ”<sup>1</sup>***.

---

<sup>1</sup> **N.Hamini**, mémoire de magister informatique

### 6.1. Perspectives

Comme un projet *Data Warehouse* n'est jamais complètement terminé, nous pouvons citer les perspectives et les développements suivants :

- ✓ Changer le support de l'entrepôt vers un SGBD commerciale qui supporte les *Data Warehouse* Oracle par exemple avec son extension Oracle-Spatiale.
- ✓ Concevoir un module de chargement puissant et flexible avec par exemple des Capteurs, pour enrichir le *Data Warehouse* d'une manière régulière.
- ✓ Suivre le déploiement actuel et recueillir les correctifs et remarques des utilisateurs.
- ✓ Utilisation des méthodes et algorithmes de Data Mining pour une meilleure exploitation des données.
- ✓ Continuer le développement de ce système pour le rendre un système ***intelligent*** d'aide à la décision .
- ✓ D'autres échanges de données avec d'autres partenaires externes doivent être automatisés pour enrichir le *Data Warehouse*.



**Bibliographie**

## Références

---

**[Alter, 1980]** ALTER Steven. Decision Support System: Current Practices and Continuing Challenges. Massachusetts: Addison-Wesley Publishing Co., 1980, 316 p.

**[Bédard 2004]** Y. Bédard, 2004. Amélioration des capacités décisionnelles des SIG par l'ajout d'un module SOLAP. Université de Provence, Centre de Mathématiques et Informatique, LSIS, Marseille, 8 avril.

**[Blaschka et al. 1998]** BLASCHKA Markus, SAPIA Carsten, HÖFLING Gabriele et DINTER Barbara. Finding your way through multidimensional data models. In : WAGNER Roland. 9th International Workshop on Database and Expert Systems Applications, 24-28 Aout, 1998, Vienne, Autriche. Washington, DC, USA : IEEE Computer Society, 1998, 198-203 p.

**[Codd et al., 1993]** E. F. Codd, S. B. Codd, C.T. Salley, "*Providing OLAP to user analysts: an IT mandate*", Rapport technique, E.F. Codd and associates, 1993.

**[DG01]** A. Doucet and S. Gangarski. Entrepôts de données et Bases de Données Multidimensionnelles, Chapitre 12 du livre : Bases de Données et Internet, Modèles, langages et systèmes. Editions Hermès, 2001.

**[Horner et al. 2004]** HORNER John, SONG Il-Yeol et CHEN Peter P. An Analysis of Additivity in OLAP Systems. In : SONG Il-Yeol et DAVIS Karen C. 7th ACM international workshop on Data warehousing and OLAP, 12-13 Novembre, 2004, Washington, DC, USA. New York, NY, USA : ACM Press, 2004, 83-91 p.

**[Hüsemann et al. 2000]** HÜSEMANN Bodo, LECHTENBÖRGER Jens et VOSSSEN Gottfried. Conceptual data warehouse modeling. In : JEUSFELD Manfred A., SHU Hua, STAUDT Martin et VOSSSEN Gottfried. econd Intl. Workshop on Design and Management of Data Warehouses, June 5-6, 2000, Stockholm, Suède. CEUR-WS.org, 2000, Vol. 28, 6.

**[Inmon, 1996]** W. H. Inmon, "*Building the Data Warehouse*", John Wiley and Sons, New York, NY, deuxième édition, ISBN: 04771-14161-5, 1996.

**[Inmon, 2000]** B. Inmon; What is a Data Warehouse; Article; <http://www.billinmon.com>; 2000.

**[Inmon, 2002]** W. H. Inmon ; « Building the Data Warehouse Third Edition » ; Wiley Computer Publishing 2002.

**[Kimball, 1996]** KIMBALL Ralph. The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. New York : John Wiley & Sons, 1996, 374 p.

**[Kimball & Ross, 2002]** ball & Ros s, 2002] R. Kimball, M. Ross, "*The Data Warehouse Toolkit: the complete guide to dimensional modelling*", (deuxieme edition), Wiley, 2002.

**[Lenz et Shoshani, 1997]** LENZ Hans-Joachim et SHOSHANI Arie. Summarizability in OLAP and Statistical Databases. In: IOANNIDIS Yannis E., HANSEN David M.. 9th International Conference on Scientific and Statistical Databases Management, 11-13 Aout, 1997, Olympia, Washington, USA. Los Angeles, CA : IEEE Computer Society 1997, 132-143 p.

**[Malinowski et Zimányi, 2004]** MALINOWSKI Elzbieta et ZIMÁNYI Esteban. OLAP Hierarchies: A Conceptual Perspective. In : PERSSON Anne et STIRNA Janis. 16th International Conference Advanced Information Systems Engineering, 7-11 Juin, 2004, Riga, Latvia. Berlin Heidelberg : Springer, 2004, 477-491 p. (Lecture Notes in Computer Science 3084).

**[Mélèse, 1972]** Mélèse *"L'analyse modulaire des systèmes de gestion, AMS"*, Editions Hommes et Techniques, Paris, 1972.

**[Nanci & Espinasse, 2001]** D. Nanci, B. Espinasse *"Ingénierie des Systèmes d'information : Merise, Deuxième génération"*, quatrième édition, Vuibert, ISBN 2-7117-8674-9,2001.

**[Rafanelli, 2003]** RAFANELLI Maurizio. Operators for Multidimensional Aggregate Data. In : Multidimensional databases: problems and solutions. Hershey, PA, USA : IGI Publishing, 2003, 116-165 p.

**[Ravat et al. 1999]** F. Ravat, O. Teste, G. Zurfluh *"Towards Data Warehouse Design"*, 8th International Conference on Information and Knowledge Management - (CIKM'99), Kansas City (Missouri, USA), ACM Press - ISBN 1-58113-146-1, p. 359-366, November 2-6 1999.

**[Ravat & Teste, 2000a]** F. Ravat, O. Teste *"Object-Oriented Decision Support System"*, Contribution a l'ouvrage < Enterprise Information Systems II », Kluwer Academic Publishers - ISBN 0-7923-7177, sélection des meilleurs articles de la 2nd International Conference on Enterprise Information Systems (ICEIS'00), p. 42-48, July 2001.

**[Ravat & Teste, 2000b]** F. Ravat, O. Teste *"A Teinporal Object-Oriented Data Warehouse Model"* 11th International Conference on Database and Expert Systems - (DEXA'00), London (UK), Springer Verlag, Lecture Notes in Computer Science N°1873 p. 583-592, September 2000.

**[Ravat & Teste, 2000c]** F. Ravat, O. Teste *"An Object Data Warehousing Approach: a Web Site Repository"*, 4th East-European Conference on Advances in Databases and Information Systems- DASFAA-ADBIS'00, Prague (Czech Republic), ISBN 80-85863-56-1, p. 128- 137, September 5-8, 2000.

**[Rivest et al 2004]** S. Rivest, P. Gignac, J. Charron, Bédard, Y, 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, 27-28 octobre, Montréal, Canada.

## Références

---

**[Tang et al. 2003]** TANG Diane, STOLTE Chris et BOSCH Robert. Design Choices when Architecting Visualizations. In : 9th IEEE Symposium on Information Visualization, 20-21 Octobre 2003, Seattle, WA, USA. Piscataway, N.J : IEEE Computer Society, 2003.

**[Teste, 2000]** O. Teste " *Modelisation et manipulation d'entrepots de donnees historisees*", These de doctorat de l'Universite Paul Sabatier, Decembre 2000.

**[Thomsen, 1997]** THOMSEN Erik. OLAP Solutions: Building Multidimensional Information Systems. New York,USA : Wiley, 1997, 608 p.

**[Widom, 1995]** J Widom., "*Research problems in data warehousing*", 4th International Conference on Information and Knowledge Management (CIKM'95), Baltimore - USA, pp 25-30, November 29 -2 December 2, 1995.

**[Winter, 1998]** WINTER Richard. Databases: back in the OLAP game. Intelligent Enterprise Magazine, 1998, Vol. 1, 60–64 p.

# *ANNEXE 1*

**Chargement & interrogation  
de l'entrepôt**

## 1. Chargement de l'EDS

D'habitude pour la création des tables de l'entrepôt on utilise des requête SQL classiques, dans notre projet nous avons les créé d'une manière automatique avec l'outil visuel de Paradox (Database Desktop). Pour ce qui concerne le chargement de l'entrepôt nous avons programmé nous même notre programme de chargement sous Delphi, et voici quelque procédure de chargement :

```

1 procedure TForm1.Button4Click(Sender: TObject);
2 var i,j,k,n,id_route: Integer;
3     idh: string;
4
5 const
6   tabh2 : array[1..4] of string = ('05:00:00', '06:00:00', '07:00:00', '07:00:00') ;
7   tabh3 : array[1..4] of string = ('08:00:00', '09:00:00', '10:00:00', '10:00:00') ;
8   tabh4 : array[1..4] of string = ('11:00:00', '12:00:00', '13:00:00', '13:00:00') ;
9   tabh5 : array[1..4] of string = ('14:00:00', '15:00:00', '16:00:00', '17:00:00') ;
10  tabh6 : array[1..4] of string = ('18:00:00', '19:00:00', '20:00:00', '21:00:00') ;
11 begin
12   for i:=1 to 200 do
13   begin
14     id_route := strtoint(map1.DataSets[1].Value[1,13]);
15     for j:=1 to 24 do
16     begin
17       for k:=1 to 2 do
18       begin
19         idh := heure.FieldByName('Heure').Text;
20         traffic.Fields.DataSet.Insert;
21         traffic.FieldByName('id_traffic').value := traffic.RecordCount+1;
22         traffic.FieldByName('id_routier').value := 1;
23         traffic.FieldByName('id_heure').value := heure.FieldByName('id_heure').Text;
24         traffic.FieldByName('id_meteo').value := k;
25         for n:=1 to 4 do
26         begin
27           if idh = tabh2[n] then
28           begin
29             if traffic.FieldByName('id_meteo').value = 1 then
30               traffic.FieldByName('Taux_traffic').value := cell(id_route/3)
31             else traffic.FieldByName('Taux_traffic').value := cell(id_route/6);
32           end
33           else
34             if idh = tabh3[n] then
35             begin
36               if traffic.FieldByName('id_meteo').value = 1 then
37                 traffic.FieldByName('Taux_traffic').value := cell(id_route/2)
38               else traffic.FieldByName('Taux_traffic').value := cell(id_route/4);
39             end
40             else
41               if idh = tabh4[n] then
42               begin
43                 if traffic.FieldByName('id_meteo').value = 1 then
44                   traffic.FieldByName('Taux_traffic').value := cell((id_route*3)/2)
45                 else traffic.FieldByName('Taux_traffic').value := cell((id_route*3)/4);
46               end
47               else
48                 if idh = tabh5[n] then
49                 begin
50                   if traffic.FieldByName('id_meteo').value = 1 then
51                     traffic.FieldByName('Taux_traffic').value := cell(id_route/3)
52                   else traffic.FieldByName('Taux_traffic').value := cell(id_route/6);
53                 end
54                 else
55                   if idh = tabh6[n] then
56                   begin
57                     if traffic.FieldByName('id_meteo').value = 1 then
58                       traffic.FieldByName('Taux_traffic').value := id_route
59                     else traffic.FieldByName('Taux_traffic').value := cell(id_route/2);
60                   end;
61                 end;
62             end;
63             heure.Next;
64           end;
65           heure.First;
66         end;
67       traffic.ApplyUpdates;
68       showmessage('Enregistrement TRAFIC effectué');
69     end;

```

Figure 1. Procédure de chargement de Fait Traffic

```

1 procedure TForm1.charg_route(Sender: TObject);
2 var i : integer;
3 begin
4
5 Routier.Last;
6 Routier.Append;
7 for i:=1 to 1237 do
8 begin
9 //Routier.first;
10 Routier.Fields.DataSet.Insert;
11 Routier.FieldName('Id_routier').Value := Map1.DataSets[1].Value[i,1];
12 Routier.FieldName('Nom_routier').Value := Map1.DataSets[1].Value[i,2];
13 Routier.FieldName('Longueur').Value := Map1.DataSets[1].Value[i,3];
14 Routier.FieldName('Largeur').Value := Map1.DataSets[1].Value[i,4];
15 Routier.FieldName('Etat_routier').Value := Map1.DataSets[1].Value[i,5];
16 Routier.FieldName('Date_constr').Value := Map1.DataSets[1].Value[i,6];
17 Routier.FieldName('Date_suppr').Value := Map1.DataSets[1].Value[i,7];
18 Routier.FieldName('Date_dern_maint').Value := Map1.DataSets[1].Value[i,8];
19 Routier.FieldName('Type_routier').Value := Map1.DataSets[1].Value[i,9];
20 Routier.FieldName('Geometry_X').Value := Map1.DataSets[1].Value[i,10];
21 Routier.FieldName('Geometry_Y').Value := Map1.DataSets[1].Value[i,11];
22 Routier.FieldName('Id_quartier').Value := Map1.DataSets[1].Value[i,12];
23 end;
24 Routier.ApplyUpdates;
25 showmessage('Enregistrement effectué');|
26 end;

```

*Figure 2. Procédure de chargement de la dimension Routier*

## 2. Requêtes SQL

### 2.1. La création des vues

Dans un entrepôt de données il est fortement recommandé d'utiliser des vues sur les tables pour un accès rapide aux données et de ne pas laisser les utilisateurs attendre pendant un grand temps. Dans notre projet nous avons créé des vues matérialisées et voici un bout de code SQL illustrant une vue de la Mesure *Traffic* :

```

1 /*
2
3 Answer: :WORK:Trafic_18_21.db
4 AuxTables: True
5 RunMode: Default
6 Alias: WORK
7 LiveAnswer: FALSE
8
9 */
10 CREATE MATERIALIZED VIEW Trafic_18_21 AS
11 SELECT Routier.Id_routier, Routier.Nom_routier, Quartier.Nom_quartier, Zone.Nom_zone, Trafic.Taux_traffic
12 FROM Trafic
13 INNER JOIN Routier
14 ON (Trafic.id_routier = Routier.Id_routier)
15 INNER JOIN Quartier
16 ON (Routier.id_quartier = quartier.id_quartier)
17 INNER JOIN Zone
18 ON (Quartier.id_zone = Zone.id_zone)
19 INNER JOIN Heure
20 ON (Trafic.id_heure = Heure.id_heure)
21 WHERE ( Trafic.id_meteo = 1 ) AND (Heure.heure between '18:00:00' and '21:00:00')
22
23

```

*Figure 3. Vue matérialisé du Fait traffic*

## 2.2. Les requêtes de l'analyse

Le code suivant montre la fonctionnement de la première option de liste d'analyse :

```

...
case RadioGroup1.ItemIndex of
-1:
begin
ShowMessage('Vous n"avez rien coché, Veuillez sélectionner une options parmi la liste svp ! ');
end;
0:
begin
case ComboBox1.ItemIndex of
-1:
begin
showmessage('Vous n"avez rien choisi !');
end;
0:
begin
DecisionQuery1.SQL.Add('SELECT Id_routier, Nom_routier, Taux_traffic, SUM(
Taux_traffic ), COUNT(
Taux_traffic ) FROM Trafic_5_7 WHERE Taux_traffic'
+ComboBox2.Items.Strings[ComboBox2.ItemIndex]+edit1.text+' GROUP BY Id_routier,
Nom_routier, Taux_traffic' );
DecisionQuery2.SQL.Add('SELECT Nom_quartier, Taux_traffic, SUM( Taux_traffic ),
COUNT( Taux_traffic ) FROM Trafic_5_7 WHERE Taux_traffic'
+ComboBox2.Items.Strings[ComboBox2.ItemIndex]+edit1.text+' GROUP BY Nom_quartier,
Taux_traffic' );
DecisionQuery3.SQL.Add('SELECT Nom_zone, Taux_traffic, SUM( Taux_traffic ),
COUNT( Taux_traffic ) FROM Trafic_5_7 WHERE Taux_traffic'
+ComboBox2.Items.Strings[ComboBox2.ItemIndex]+edit1.text+' GROUP BY Nom_zone,
Taux_traffic' );
end;
1:
begin
.....
end;
2:
begin
.....
end;
3:
begin
.....
end;
end;
...

```

# ANNEXE 2

## Exploitation de données brutes de trafic routier urbain issues d'un réseau de capteurs géoréférencés

*Claudia Bauzer-Medeiros, Florian Devuyst, Marc Joliveau, Geneviève Jomier,*

**Institute of Computing (IC) - University of Campinas  
Caixa Postal 6176 13084-971 Campinas, SP - Brazil  
cmbm@ic.unicamp.br**

**Grande Voie des Vignes F-92 295 Châteney-Malabry Cedex  
{florian.de-vuyst, marc.joliveau}@ecp.fr**

**Université Paris Dauphine - Laboratoire LAMSADE  
Place du Maréchal de Lattre de Tassigny 75775 Paris Cedex 16  
genevieve.jomier@lamsade.dauphine.fr**

**Résumé.** Cet article Les données de trafic provenant de réseaux de capteurs ont été la source de nombreuses recherches reliées aux SIT (Systèmes d'Information en Transport). Ces données se représentent généralement par de grands ensembles de séries spatio-temporelles corrélées. Ce papier présente une nouvelle approche pour manipuler des données brutes issues de capteurs fixes géoréférencés. Notre travail est fondé sur la combinaison de méthodes analytiques pour préparer les données des capteurs et sur la proposition d'une architecture pour un système d'information dédié au trafic routier. Il a été conduit dans le cadre d'un projet utilisant des données réelles générées par 1000 capteurs pendant 3 ans dans une grande agglomération française.(ville et couronne).

## 1. Introduction

De par la place prédominante occupée par le transport dans notre environnement socio économique et les enjeux d'environnement et de santé publique, de nombreuses recherches actuelles visent à fournir des outils de précision d'aide à la gestion et d'aide à la décision.

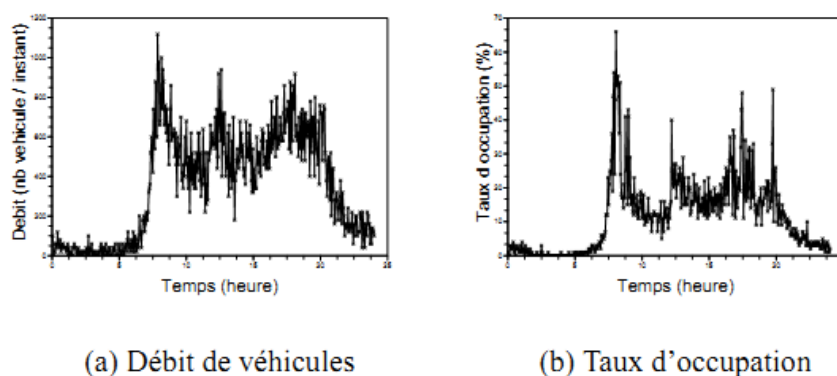
Les travaux peuvent par exemple concerner l'étude des comportements des usagers, la prévision et le traitement des congestions ou l'influence de la météo sur le trafic. Ce papier se rapporte au traitement, à l'extraction de connaissances et au stockage de données de trafic routier urbain issues d'un réseau de capteurs géoréférencés.

## 2. Contexte du problème

Nos recherches s'inscrivent dans le cadre du projet ACI Masse de données **CADDY** (Contrôle de l'Acquisition de Données, stockage et modèles DYnamiques). Ce projet réunit une équipe scientifique pluridisciplinaire de différents instituts. Le but de CADDY est de développer un outil d'aide à la décision pour la gestion du trafic routier. Les capteurs utilisés sont géoréférencés, fixés le long des voies de circulation. Ils collectent différentes informations sur le trafic durant la journée à fréquence régulière.

Deux variables principales sont mesurées, produisant deux séries spatio-temporelles distinctes (mais interdépendantes) :

- le débit de véhicules ( $q$ ), i.e., le nombre de véhicules passé devant le capteur durant une période donnée (3 minutes dans notre cas) ;
- le taux d'occupation ( $t$ ), i.e., l'espace moyen entre les véhicules durant un intervalle de temps donné. Un taux d'occupation de 100% signifie qu'il n'y a aucun espace entre les véhicules, alors que 0% indique qu'il n'y a pas de véhicules.



*Figure 1. Débit et taux d'occupation journalier pour un capteur donné.*

La figure 1 illustre les séries temporelles journalières de débit et de taux d'occupation mesurées à un capteur donné, un jour de semaine. Nos données sont collectées depuis 1000 capteurs sur 3 ans, les prises de mesure s'effectuant toutes les 3 minutes. Ceci représente un total de 420 106 valeurs. Nos données proviennent du système de supervision du trafic CLAIRE [SCE 04] développé par l'INRETS (Institut National de REcherche sur les Transports et leur Sécurité) au laboratoire GRETIA (G. Scemama). CLAIRE modélise le réseau routier urbain par l'intermédiaire d'un graphe orienté, où chaque arc correspond à une portion de route. Les localisations des capteurs mesurant le trafic sont associées à ces arcs.

### **3. Résumés spatio-temporels de données issues de capteurs**

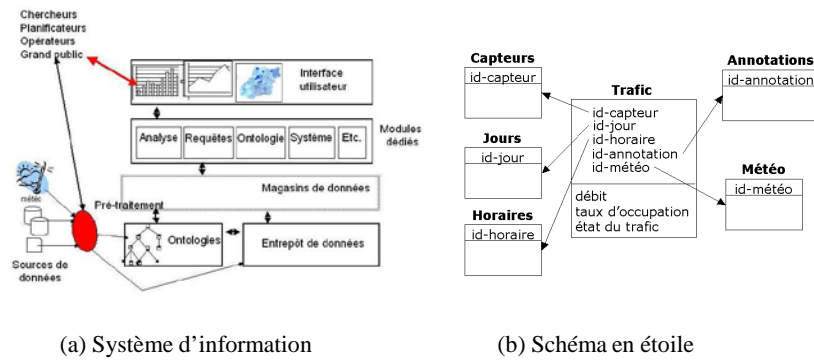
Cette section présente nos solutions pour traiter les données spatio-temporelles issues de capteurs, afin de les résumer. On commence en présentant une méthode qui diminue très fortement la dimension des séries temporelles (de débit et de taux d'occupation) pour l'ensemble des capteurs tout en conservant au mieux l'information au sens de l'énergie. Ensuite, une adaptation de cette méthode à un ensemble de données contenant des valeurs manquantes est proposée. Finalement, on combine sans perte les informations contenues dans les deux séries (débit et taux d'occupation) à l'aide d'une variable à valeurs continues dans l'intervalle  $[0; 1]$  pour décrire le comportement du trafic à un capteur. Cet intervalle peut être divisé en plusieurs classes représentant des états du trafic définis par des experts (par exemple "saturé" ou "calme").

### **4. Extraction de connaissances sur les données de trafic**

La méthode STPCA et l'introduction de la variable d'état de circulation  $E$  présentés dans la section précédente ont permis de préparer les données et notamment de passer de séries temporelles incomplètes à des résumés symboliques. Ces éléments combinés à d'autres outils permettent aussi de fouiller les données et d'en extraire des connaissances.

### **5. Un système d'information pour le trafic routier**

Nous présentons l'architecture du système d'information pour le trafic routier que nous proposons. Celle-ci est construite autour d'un entrepôt de données dont le schéma en étoile est illustré sur la figure 2(a). Ce modèle ajoute notamment au schéma standard des ontologies au niveau du stockage des données, comme on peut le voir dans la partie inférieure de la

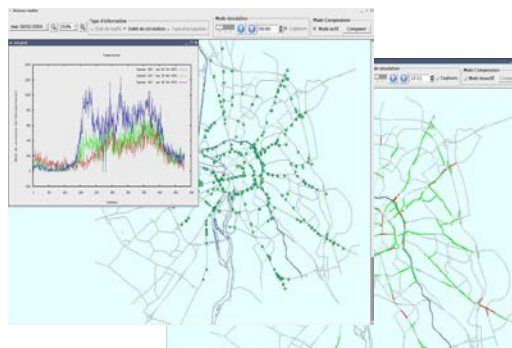


**Figure 2.** Architecture du système d'information pour le trafic routier et schéma en étoile de l'entrepôt de données proposés.

Figure 2(b). L'ensemble des ontologies permet d'organiser les définitions et la terminologie utilisées par le domaine de l'application, aide à la construction et à la maintenance de l'entrepôt. Il est également utilisé pour la création et la maintenance des magasins de données.

## 6. Conclusion

Ce papier présente les recherches effectuées dans un projet pluridisciplinaire relié au traitement de données de trafic routier urbain issues d'un réseau de capteurs. Il combine la description de méthodes analytiques pour réduire la dimension des données, traiter les valeurs manquantes, proposer des résumés symboliques des différentes variables mesurées et les recherches sur une structure complexe permettant de stocker ces données et de les manipuler de manière efficace. Une partie du travail élaboré au sein du projet CADDY a conduit à la construction d'un prototype logiciel, illustré sur la figure 3 permettant d'explorer visuellement les données des capteurs dans les deux dimensions spatiale et temporelle.



**Figure 3.** Copie d'écran du prototype.