

République Algérienne Démocratique et Populaire
**Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique**

Université Hassiba Benbouali de Chlef
Faculté des Sciences
Département d'Informatique



POLYCOPIE

Système d'information spatial

Pour :

Master : M2 Ingénierie des systèmes

D'informations Avancées

Par :

GHERBAOUI Radhwane

Dernière mise à jour janvier 2023

????

Préambule

Ce polycopié est le résultat des années d'enseignement du module « Systèmes d'information géographique » à l'université Hassiba Benbouli Chlef (UHBC). Ce module est dispensé pour les étudiants en master 2 en informatique dont l'option « ISIA ».

Dans les systèmes d'information géographique, plusieurs disciplines sont impliquées telles que l'intelligence artificielle, l'imagerie, la simulation. ;etc. De ce fait, nous avons insisté dans cette polycopie sur la structuration des données géographiques et certain traitement qui sont rarement impliqués dans les autres domaines.

Cette polycopie est principalement destinée aux étudiants en master 2 et aux ingénieurs de l'informatique afin d'être capable de réaliser et de développer des prototypes de SIG. L'objectif est de familiariser les futures spécialistes en informatique sur les concepts de base pour la modulation et l'implémentation des SIG.

Pour une amélioration continue de cette polycopie, je vous prie de me contacter pour toute suggestion, critique, remarque ou correction sur l'email suivant : r.gharboui@univ-chlef.dz

La dernière mise à jour de cette polycopie est faite en mars 2023.

Table des matières

Préambule	1
------------------	----------

Introduction générale	7
------------------------------	----------

Chapitre 1

Présentation des Systèmes d'Information Géographique

1.1 Introduction	10
1.2 Définitions	10
1.3 Multidisciplinarité des SIG	12
1.4 Objectifs généraux	13
1.5 Composantes d'un SIG	15
1.6 Acquisition des données Géographiques	16
1.7 Conclusion	19

Chapitre 2

Structures des données spatiales

2.1 Introduction	23
2.2 Base de données géographique	23
2.2.1 les avantages d'utilisation des bases de données	24
2.2.2 Les modèles de base de données	25
2.3 Les modes de présentation des données géographiques	28
2.4 Structure des données dans le mode vectoriel	30
2.4.1 Structures non topologiques	30
2.4.2 Structures topologiques	31
2.5 Structure spécifique à l'information altimétrique	37
2.5.1 Structure en réseaux triangulaires irréguliers : RTI (en anglais, Tri- angulated Irregular Network : TIN)	37

2.5.2	Structure topologique pour les courbes de niveau	39
2.6	Mode raser	43
2.7	Conclusion	48

Chapitre 3

Analyse des données géographiques
--

3.1	Introduction	58
3.2	Notion d'analyse des données spatiales	58
3.3	Types des données géographiques	59
3.4	Implémentation des SIG	63
3.5	Exemple d'application	66
3.5.1	Les requêtes attributaires	68
3.5.2	Les requêtes spatiales	68
3.6	Conclusion	69

Conclusion générale	72
----------------------------	-----------

Bibliographie	73
----------------------	-----------

Table des figures

1.1	Exemple des fonctionnalités métier.	11
1.2	Structure d'un SIG.	12
1.3	Multidisciplinarité des SIG.	13
1.4	Schéma représentatif d'un SIG.	16
1.5	Matériel du levé du terrain.	17
1.6	Matériel GPS.	17
1.7	Matériel Photogrammétrie digitale.	18
1.8	Matériel de télédétection.	18
2.1	Le modèle hiérarchique d'une paye.	26
2.2	Une carte M composée de deux régions.	26
2.3	Le modèle hierarchique de la carte M.	26
2.4	Le modèle relationnel de la carte M.	27
2.5	Les classes du modèle orienté objet de la carte M.	27
2.6	Exemple de formalisation d'une carte M_1 selon le modèle inspiré du domaine de la théorie de graphe.	29
2.7	Structure séquentielle de la carte M_1	31
2.8	Structure séquentielle avec dictionnaire de points de la carte M_1	31
2.9	Nœuds de la carte M_1 selon la structure DIME. (a) la carte M_1 ; (b) fichier des noeuds.	33
2.10	Arrêtes de la carte M_1 selon la structure DIME. (a) la carte M_1 ; (b) fichier des arrêtes.	34
2.11	Nœuds de la carte M_1 selon la strucutre POLYVRT. (a) La carte M_1 ; (b) Le fichier des nœuds.	35
2.12	Fichier des points intermédiaires et fichier des zones de la structure POLYVRT de la carte M_1 . (a) le fichier des points intermédiaires ; (b) le fichier des zones.	35
2.13	Fichier chaine de la structure POLYVRT de la carte M_1	36
2.14	Construction des triangles à partir des points.	37
2.15	Structure en réseaux triangulaires irréguliers.	38
2.16	Sauvegarde des données par triangles.	38
2.17	Sauvegarde des données par sommets.	38
2.18	Projection des points sur le plan (XY).	39
2.19	Courbe de niveau.	39
2.20	Le fichier « entité » de la carte précédente.	41

2.21	Le fichier des coordonnées.	42
2.22	Le fichier de voisinage inférieur.	43
2.23	Le fichier de voisinage supérieur.	43
2.24	Les trois types de maillage fréquemment utilisés. De gauche à droite : Maillage carré, Maillage triangulaire et Maillage hexagonal.	44
2.25	La méthode RLC (Run Length Coing). A gauche, une image en format BitMap. A droite, l'image en format RLC.	45
2.26	La méthode des blocs carrés appliquée sur la matrice de l'image précédente.	45
2.27	La méthode de Tétra-arbre. (a) La matrice des valeurs des bloc. (b) Codage des blocs. (c) représentation des blocs sous forme d'une arbre (la hiérarchie). (d) L'enregistrement de Tétra-arbre dans le mémoire.	47
2.28	Carte C	49
2.29	Le modèle hiérarchique de la carte C	50
2.30	Le modèle hiérarchique de la carte C	50
2.31	La structure séquentielle simple de la carte C	51
2.32	la carte C	52
2.33	la carte C-2	52
2.34	Fichier des noeuds de la carte C	53
2.35	Fichier des arrêtes de la carte C	53
2.36	Fichier des zones de la carte C.	53
2.37	Fichier des noeuds de la carte C	54
2.38	Fichier des points intermédiaires de la carte C	54
2.39	Fichier des chaînes de la carte C	54
2.40	Carte des courbes de niveaux C-3	55
2.41	Carte C-4 des courbes de niveaux.	55
2.42	Carte des courbes de niveaux C-2.	56
2.43	Les trois fichiers de la sstructure des courbes de niveaux. (a) Fichier des voisins inférieurs. (b) Fichier des voisins supérieurs. (c) Fichier des points. .	56
3.1	Exemple d'une base de données composée de deux tables : la table parcelle et la table propriétaire.	60
3.2	Les composantes de l'information géographique.	60
3.3	Les deux parties de l'information sémantique.	61
3.4	deux tables représentant l'information sémantique et l'information géomé- trique du même endroit.	62
3.5	le modèle entité-association correspond aux segments droits.	63
3.6	le modèle-entité association correspondant aux segments obliques.	63
3.7	Exemple de couplage entre les SGBD-R et SI spatial.	64
3.8	La table commune.	66
3.9	La table Route.	67
3.10	La table Bâtiment.	67
3.11	La table Wilaya.	68
3.12	Table Bâtiment	70
3.13	Table Route	70
3.14	Table Daira	70

3.15 Table Wilaya	70
-----------------------------	----

Introduction générale

Les systèmes d'information géographique (SIG) comme son nom l'indique sont des systèmes d'information dédiés pour l'information à référence géographique. L'évolution des systèmes d'information a permis l'évolution des SIG de telle sorte qu'ils sont considérés actuellement comme des systèmes évolués. De ce fait, les caractéristiques et les objectifs ne cessent pas d'être de plus en plus nombreux. Aux débuts des années soixante-dix, l'architecture de ce système est enfermée sur une unité de traitement, un disque de stockage, un écran d'affichage et un scanner d'acquisition. Plus récemment, cette architecture connaît grande évolution de telle sorte que les disques de stockage et de traitement sont remplacés par des serveurs entiers. Les moyens d'acquisition de l'information sont développés pour qu'ils soient des serveurs des données, des satellites, des radars, des avions..etc. Alors que les outils d'affichage dépassent largement un simple d'écran pour qu'ils soient dans les véhicules (GPS), les terminaux mobiles, les panneaux d'affichage des réseaux électriques...etc.

L'évolution des systèmes d'information géographique ne restreint pas seulement sur la partie matérielle, mais les outils logiciels sont aussi évolués. Des disciplines entières sont intervenues dans les SIG telles que l'imagerie, la géométrie, les systèmes d'aide à la décision, la simulationetc. Les SIG parrait dans plusieurs domaines incluant des établissements le cadastre, les communes, les centres de météoect, et aussi des entreprises telles que les sociétés d'entreprise, transport maritime et routier, les usines de véhiculesetc. Le besoin des établissements et des entreprises aux systèmes d'information géographique ne cesse pas d'agrandir, jour après jour, ces systèmes sont exploités dans de nouveaux domaines.

Les SIG sont avant tout des systèmes d'information. Donc, la réalisation de ce système profite des systèmes de gestion des bases de données (SGBD) standard et des modèles existants pour réduire les coûts de développement. Par conséquent, des extensions des SGBD ont été adoptées pour supporter les données spatiales. Pour faciliter la conception des SIG, une donnée géographique est subdivisée en deux parties, une partie attributaire sollicitée par les SGBD et une autre partie spatiale qui a une référence géographique. De ce fait, d'autres approches couplent les SGBD avec des systèmes de gestion spatiale pour réduire en plus le coût de développement. L'émergence des SIG de plus en plus dans plusieurs domaines et disciplines a conduit les chercheurs à concevoir des modèles qui sont complètement nouveaux pour faciliter la modélisation et le traitement des données de caractère géographique.

Cette manuelle est organisée de la façon suivante. Dans le chapitre 1, des définitions qui encadrent les systèmes d'information géographique sont présentées. Dans ce chapitre, l'architecture et les objectifs, pour lesquels les SIG sont conçues, seront illustrés. Les méthodes d'acquisition des géographiques seront détaillées. Dans le deuxième chapitre, les modes de présentations et de structuration de données géographique seront expliqués avec des exemples. Le troisième chapitre est consacré pour l'analyse de l'information géographique.

1

Présentation des Systèmes d'Information Géographique

Sommaire

1.1	Introduction	10
1.2	Définitions	10
1.3	Multidisciplinarité des SIG	12
1.4	Objectifs généraux	13
1.5	Composantes d'un SIG	15
1.6	Acquisition des données Géographiques	16
1.7	Conclusion	19

1.1 Introduction

Dans ces dernières décimées, les systèmes d'information géographique connaissent une évolution sans précédentes. Cette évolution est due essentiellement grâce au développement des capacités des traitements des calculateurs, les technologies de communication et d'échange d'information, et les moyennes d'acquisition de l'information spatiale. En effet, les besoins des entreprises et les sociétés aux outils de traitement spatiales restent le facteur le plus important qui poussent les concepteurs à réaliser tels systèmes. Dans ce chapitre, nous allons introduire les lecteurs dans les systèmes d'information géographique en exposant les définitions dans la littérature encadrant la notion de ces systèmes. Une description des composantes principales et les fonctions associées sont illustrées et aussi les objectifs pour lesquels ces systèmes sont conçus.

1.2 Définitions

Plusieurs définitions ont été proposées dans la littérature identifiant la notion d'un " système d'information géographique.

- Définition de Clark (1986) [7] : « *"Système de développement et d'élargissement du domaine d'utilisation des cartes"* ».
- Définition de Aronoff(1989) [1] : « *"Ensemble de procédures utilisées pour conserver et traiter de l'information à référence géographique"* ».
- Définition de Burrough (1998) [4] : « *"Ensemble d'outils puissants permettant l'acquisition, le stockage, l'édition aussi bien que la transformation et l'analyse des données spatiale relatives au monde réel".* »
- Définition de Dider [8] : « *"Ensemble de données repérées dans l'espace, structuré de façon à pouvoir en extraire commodément des synthèses utiles à la décision"* ».
- Définition de Comité Fédéral de Coordination Inter-agences pour la Cartographie Numérique aux États-Unis (1988) : « *Un SIG est un système informatique de matériels, de logiciels et de processus conçu pour permettre : la collection, la gestion, la manipulation, l'analyse, la modélisation, l'affichage des données à référence spatiale, afin de résoudre des problèmes complexes d'aménagement et de gestion* ».
- Définition de Konecny [11] : « *Un SIG, dans une définition étroite, est un système informatique pour la saisie, la manipulation, le stockage et la visualisation des données spatiales numériques. Dans une définition plus large, il est un système numérique d'acquisition, de gestion, d'analyse, de modélisation et de visualisation de données spatiales aux fins de la planification, l'administration et le contrôle de l'environnement naturel et aux applications socio- économiques* »

D'après ces définitions, on conclut qu'un système d'information géographique est un système d'information qui opère des traitements sur des données dans l'espace. De ce fait, les concepts de l'espace sont reformulés dans ces définitions par carte, géographie ou espaces. Autrement dit, les systèmes d'information géographique sont des systèmes d'information élargie pour le traitement des informations spatiales. Les définitions de Burrough et du comité fédéral américain incluent les composants fondamentaux qui doivent contenir une architecture d'un SIG qui sera bien détaillée dans les sections suivantes.

Un SIG peut aussi être défini par les réponses aux questions Où ?, Quoi ?, Comment ?, Quant ? et Si comme il est indiqué dans l'exemple de la figure 1.1.

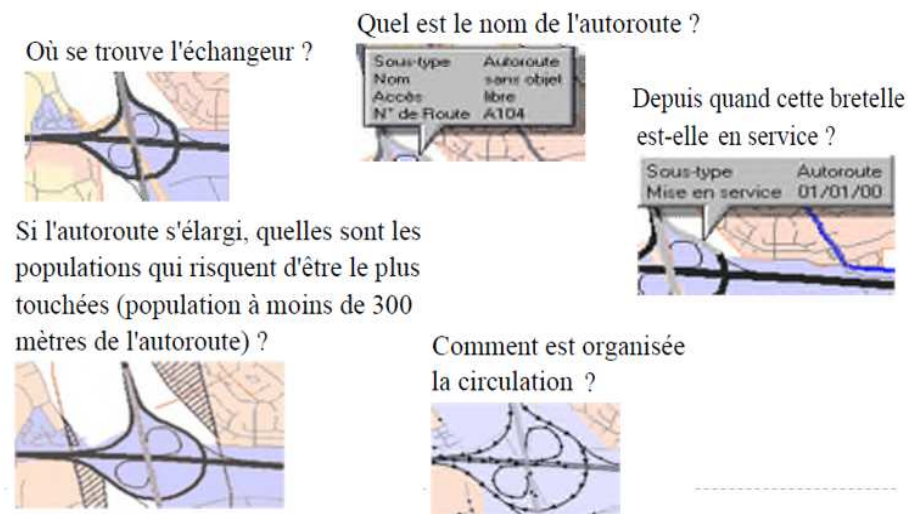


FIG. 1.1 – Exemple des fonctionnalités métier.

Les fonctions d'un SIG peuvent aussi être résumées par, ce qui est appelé, les cinq « A ».

1. Abstraction : Modélisation de l'information
2. Acquisition : récupération des informations existantes.
3. Archivage : stockage des données d'une façon à les retrouver et interroger facilement.
4. Analyse : réponse aux requêtes. l'analyse peut faire appel à des sous-système d'aide à la décision ou d'intelligence artificielle. Le sous-système d'analyse affecte la qualité de SIG ;
5. Affichage : restitution graphique, les cartes et les données attributaires sont affichées clairement de telle sorte que les utilisateurs potentiels peuvent tirer facilement des connaissances sur les zones ou les objets visés par la requête.

Ces cinq fonctions représentent effectivement les modules principales d'une structure d'un SIG (voir 1.2).

Pour mieux comprendre les SIG, la nature des données connue sous le nom de l'information géographique(IG) dont le traitement est subi par les SIG doit être précisée.

L'IG est une donnée référencée spatialement apportant les connaissances suivantes.

1. le nom de lieu,
2. les coordonnées géographiques,
3. l'adresse postale ou autre.

Dans le cas général, IG peut être partagée entre plusieurs clients et serveurs, dupliqué, enrichie et liées avec d'autres IG. L'information géographique subie toute les fonctionnalités de « cinq A » d'un système d'information géographique. Elle est présentée par les deux parties suivantes.

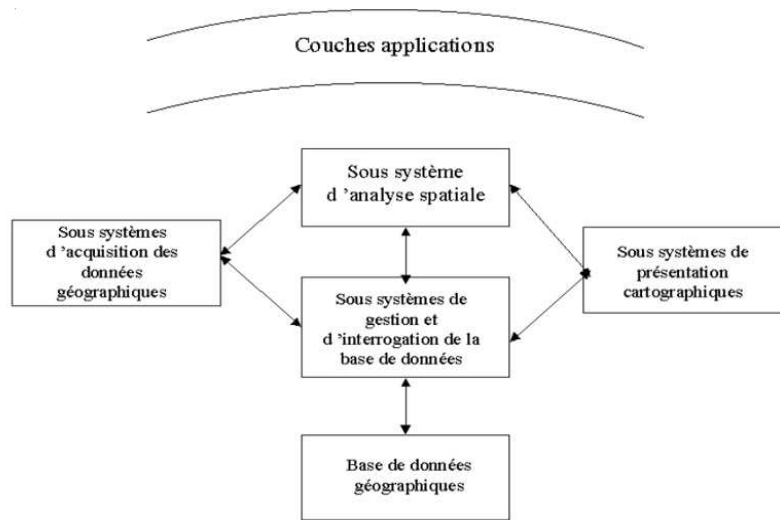


FIG. 1.2 – Structure d'un SIG.

1. Par sa nature et son aspect (les caractéristiques).
2. Par son emplacement sur la terre(ou sur un autre astéroïde!).

la première partie de l'IG est connue sous le nom "information attributive", cette partie est traitée fréquemment par les systèmes de gestion des bases de données. La deuxième partie représente la partie spatiale de l'IG.

1.3 Multidisciplinarité des SIG

Un SIG peut être vu comme étant un outil d'analyse de l'information dans le domaine spatial. Il permet l'analyse simultanée de plusieurs paramètres spatialisés sur des ensembles de données multi source et multimédia. Quand il existe un certain nombre de cartes mono disciplinaires sur une zone d'étude bien identifiée, une combinaison de l'information multidisciplinaire est indispensable pour une supervision totale(voir Fig.1.3). Cette opération peut être effectuée de deux manières : manuelle et assistée par l'ordinateur [3]

1. Méthode manuelle : c'est une procédure classique mise en place par Mc Harg (1969). Elle se fait simplement par superposition des copies transparentes des cartes sur une table lumineuse. Les inconvénients de cette méthode sont principalement :
 - Le nombre limité de cartes qui peuvent être superposées simultanément pour une analyse visuelle.
 - Les difficultés d'obtenir à une même échelle les différentes copies transparentes.
2. Méthode assistée par ordinateur : Les données sont alors sauvegardées sous forme numérique. Des programmes réalisés pour des opérations d'analyse, offrant la possibilité de concevoir ce que Mc Harg faisait avec la superposition des transparents, mais d'une manière plus rapide et reproductible.

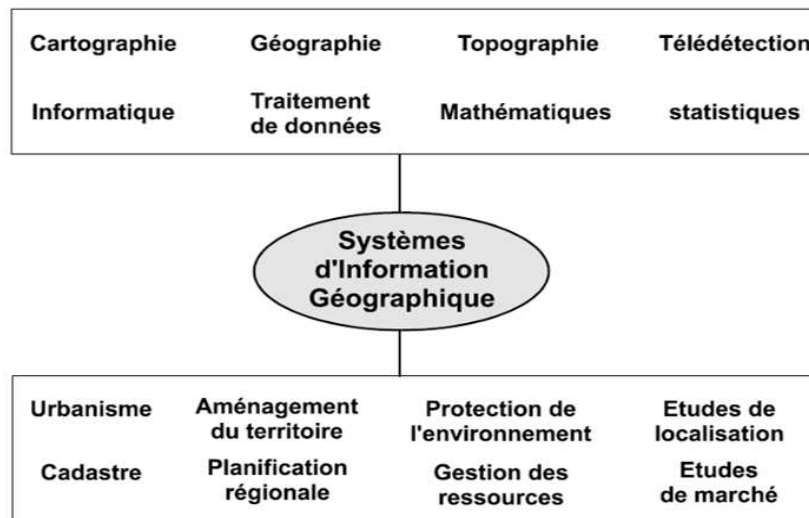


FIG. 1.3 – Multidisciplinarité des SIG.

1.4 Objectifs généraux

Les objectifs pour lesquels les SIG sont conçus sont multivarié faisant intervenir plusieurs axes de recherche.

Saisie et stockage numérique de plans et de cartes : Le premier et principal objectif des SIG est le stockage numérique de données géographiques, bi ou tridimensionnelles. Mais la différence entre un système qui va conserver des objets, avec une description aussi bien graphique que descriptive, et un système qui va seulement conserver un dessin sans contenu sémantique est marquante.

Structuration de l'information : Comme tout système de gestion de bases de données, un SIG qui gère une base de données, une modélisation du monde réel et une structuration de l'information est nécessaire. Cette structuration est souvent difficile à concevoir, car elle s'exerce sur des objets qui peuvent avoir des formes multiples, aussi bien graphiques que descriptives. Elle est fortement relative aux fonctions appliquées sur ces objets.

Calculs métriques, calculs techniques, positionnement et projections géographiques : les calculs métriques interviennent dans les calculs des mesures des distances et surfaces alors que les calculs techniques comprennent la visibilité, les volumes et la recherche opérationnelle. Les SIG permettent de calculer facilement les surfaces, distances et volumes à partir des données de localisation des objets. Les calculs et les changements de projections géographiques sont facilement accessibles. La recherche opérationnelle (essentiellement les calculs de chemins dans des graphes) trouve dans les SIG toutes les données dont elle a besoin.

Gestion et traitement des collections d'objets : C'est l'un des objectifs primordiaux des SIG. Après que l'information est structurée, elle doit être saisie et gérée par le système. Souvent, les SIG délèguent la gestion des données descriptive aux SGBD

relationnels classiques (comme ACCESS, ORACLE, SQL Server, DBase, etc.), et ne préoccupent eux-mêmes que par la localisation des objets et les liens entre la partie graphique et la partie descriptive des objets.

Ces tâches sont assurées par une bonne gestion des flux d'information, des mises à jour, des modifications et essentiellement pour la partie graphique des objets.

Gestion administrative et partage de données entre utilisateurs : Lorsque les données sont distribuées sur plusieurs utilisateurs, comme c'est souvent le cas pour les applications administratives de type cadastre. Les SIG sont conçus pour la gestion de ce partage et l'optimisation d'accès aux données entre utilisateurs.

Gestion et analyse spatiale : Les SIG doivent gérer tout type d'objet géographique, du point au pixel, en passant par les zones, les réseaux, etc. Le but est d'atteindre la construction d'une base de données géoréférencées, en reliant de différents objets de la base, quels que soient les types de ces objets.

Cette connectivité entre les objets doit assurer l'analyse spatiale, autrement dit, la prise en considération la localisation dans l'analyse des données. Un nombre important des processus faisant intervenir la localisation des objets sont implantés dans les SIG (sélections d'objets sur des critères de distances, recherche opérationnelle, agrégations spatiales et changements d'échelle, géo jointures, interpolations, vectorisations, classifications par proximité, etc.).

Statistique et géostatistique : La construction d'une base de données géographique a souvent pour but l'étude de tous les objets d'un territoire dans toutes ses composantes, et le SIG doit donc assurer l'accès facile au calcul statistique, qu'il soit exploratoire ou méthodologique. Certains SIG contiennent un module statistique. D'autres sont couplés via des interfaces avec un logiciel spécialisé. L'utilisation de méthodes de la géostatistique doit également être l'un des objectifs du SIG, puisqu'en gérant la localisation, il facilite considérablement l'utilisation de ces méthodes d'analyse ou d'interpolation spatiales.

Gestion spatio-temporelle : Introduire le temps dans les SIG assure la faisabilité des requêtes mêlant espace et temps, de façon à être capables de gérer à la fois l'historique d'un objet et l'état d'un ensemble d'objets à une date donnée. Les SIG doivent aussi gérer les évolutions des objets géographiques dans le temps. Mais les réalisations concrètes sont peu répandues, car la gestion de l'historique des évolutions ou changement de la localisation d'un grand nombre d'objets est complexe.

Simulation et modélisation : La mise en place d'un modèle pour la simulation d'un processus est actuellement considérée comme un objectif des SIGs. Une interface entre le programme de modélisation ou de simulation et la base de données géographiques doit exister, et doit se préoccuper de l'accès à l'information spatiale dont a besoin le programme d'application.

Téledétection, géoréférencement et traitement d'image : Les SIG ont vocation à gérer tout type d'objet géographique. La téledétection aérienne ou spatiale offre une source favorable de données géographiques. Les sources d'acquisition des données doivent être prises en charge. De ce fait, toutes les méthodes et techniques

intervenant dans la discipline de l'imagerie doivent être utilisées. Ces techniques assurent le bon géo-référencement et facilitent l'accès et le traitement des données qui sont souvent volumineux. Les SIG doivent donc également gérer et traiter de type de données, souvent volumineuses.

Le dessin, l'édition cartographique, la cartographie automatique 3D : Comme la plupart des systèmes de gestion de données, les SIG ont pour but l'affichage des données résultant d'une requête. Cette édition est souvent graphique puisque le traitement concerne souvent les données fortement liées à la localisation. Les modules de cartographie automatique à partir des données administrées par le système sont donc fondamentaux pour l'utilisateur. Avec l'évolution des systèmes, la troisième dimension est intégrée en permettant l'édition de données en perspective. Mais la saisie et la maintenance de la troisième dimension restent une tâche complexe malgré le développement dans les techniques d'imagerie.

Internet et accessibilité distante : L'Internet permet depuis longtemps de nouvelles perspectives d'accès distant aux données. Les SIG doivent donc être capables d'exploiter ce privilège, en gérant la complexité de structure de l'information localisée, de façon à offrir aux utilisateurs des méthodes simples de consultation et de cartographie via Internet.

1.5 Composantes d'un SIG

Selon Burrough [4], un SIG est constitué en trois composantes principales :

Hardware : Cette partie "Hardware" se compose d'un ordinateur ou unité centrale de traitement (CPU : Central Processing Unit) qui est connectée à un disque permettant l'enregistrement des données et des programmes, d'un digitaliseur ou tout autre dispositif permettant la conversion analogique / numérique (un scanner numérique ou autre), ainsi que des moyens de sortie (un écran graphique, un traceur, une imprimante, etc.). Aussi, une unité de bandes magnétiques peut se révéler nécessaire pour des opérations de stockage ou de communication avec d'autres systèmes.

Software : la partie "Software" d'un SIG comprend cinq modules et les outils de base qui assurent les opérations suivantes.

1. Saisie et vérification des données ;
2. Gestion et sauvegarde de la base de données ;
3. Affichage ou représentation des données ;
4. Transformation et conversion des données ;
5. Interaction avec l'utilisateur (IHM :interface homme-machine).

Aspect organisationnel : les deux composantes Hardware et Software, même ultra puissante, ne garantissent pas l'efficacité du SIG qu'elles composent. Pour garantir la rentabilité d'un tel système, les différentes composantes, y compris les différents acteurs (opérateurs, utilisateurs, gestionnaires, etc.) doivent être mis dans un

contexte organisationnel approprié. Pour ce faire, une étape d'étude et de conception est indispensable.

La figure 1.4 illustre une représentation schématique d'un SIG.

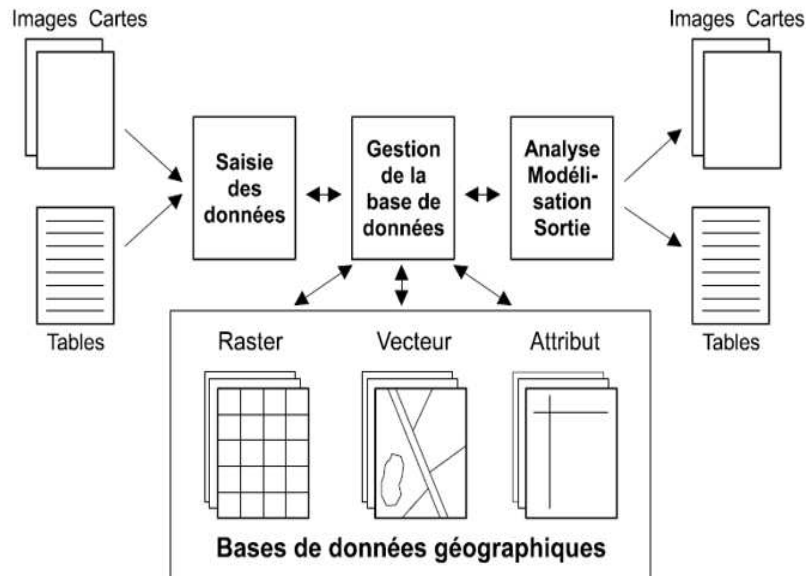


FIG. 1.4 – Schéma représentatif d'un SIG.

Bien que cette structure paraît ancienne, mais on peut dire que la structure proposée par Burrough (1986) représente les composantes minimum qui doit contenir un SIG. On note aussi que l'évolution de hardware et de software concerne le nombre et l'hétérogénéité des ressources matérielles, logicielles et humaines [9]. Des nouveaux défis de d'administration et de gestion sont ouverts. En effet, les systèmes distribués, les grilles et les clouds ouvrent des nouveaux perspectives et assurent l'extensibilité et l'évolution des SIG.

1.6 Acquisition des données Géographiques

Il existe plusieurs méthodes d'acquisition des données géographique. Ces méthodes sont caractérisées par le matériel utilisé, la précision et la référence des coordonnées des objets à localiser.

Levés de terrain : les trois caractéristiques de cette méthode sont les suivantes.

1. Matériel : station totale, carnet de terrain électronique (voir Fig 1.5).
2. Précision : dépend du matériel utilisé et des erreurs des mesures.
3. Coordonnées : Les coordonnées sont des coordonnées cartésiennes de deux ou trois dimensions, rectangulaires, géodésiques et ou astronomiques.

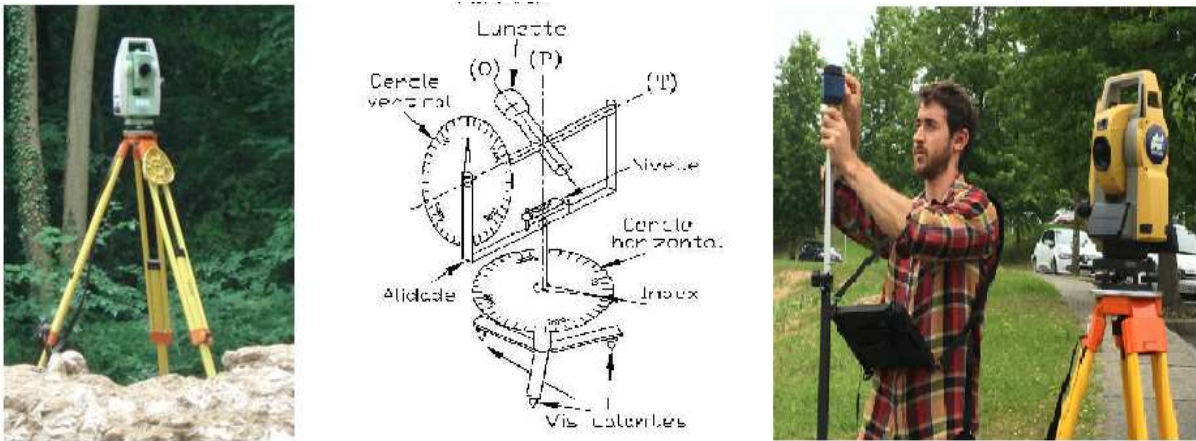


FIG. 1.5 – Matériel du levé du terrain.

Mesures de G.P.S (Globale Positioning System) : un système globale de localisation qui emploie un satellite et un terminale mobile ou fixe pour l’affichage des positions (voir Fig 1.6).

1. Matériel : station GPS.
2. Précision : dépend du matériel utilisé et des mesures absolues et les données géodésiques.
3. Coordonnées : Les coordonnées sont des coordonnées cartésiennes de deux ou trois dimensions, rectangulaires et géodésiques.



FIG. 1.6 – Matériel GPS.

Photogrammétrie analytique ou digitale : Un appareil qui permet d’agrandir et d’éclaircir les cartes afin de mieux visualiser et d’analyser les petits détails sur les cartes.

1. Matériel : restituteur analytique / digital (voir Fig 1.7).
2. Précision : dépend du matériel, orientations, échelle des clichés et les erreurs de mesure.

3. Coordonnées : Les coordonnées sont des coordonnées cartésiennes de deux ou trois dimensions, rectangulaires et ou géodésiques.



FIG. 1.7 – Matériel Photogrammétrie digitale.

Téledétection : La téledétection est basée sur l'utilisation d'un appareil mobile (avion, voiture) est un scanner embarqué pour la détection d'un objet ou d'une collection d'objet.

1. Matériel : un scanner embarqué (voir Fig 1.8).
2. Précision : dépend de la résolution, modèle de corrections géométriques.
3. Coordonnées : les coordonnées sont des coordonnées cartésiennes de deux dimensions, rectangulaires et image.



FIG. 1.8 – Matériel de téledétection.

L'acquisition des données géographique peut aussi être classifiée selon la façon par laquelle les données sont numérisées en trois catégorie.

Mode de saisie : dans ce mode, on utilise le clavier et les instruments du levé du terrain.

Mode sémi-automatique : dans ce mode, on utilise un digitaliseur, il capture l'information point par point ou en mode continue.

Mode automatique : Un ensemble des outils est déployé pour assurer l'automatisation d'acquisition des données.

1. Scanner du bureau pour la saisie des documents papier(cartes, etc.).
2. Enregistreur binaire en niveau de gris et en couleurs
3. Scanner embarqué pour la télédétection.
4. enregistreur de l'énergie réfléchie par le sol

1.7 Conclusion

Les Systèmes d'information géographique sont actuellement employés dans plusieurs domaines (le cadastre, l'agriculture, la marine..etc), de ce fait, un nombre important des disciplines et axes de recherche intervient dans la réalisation et le déploiement de ces systèmes. Dans nos jours, les systèmes d'information géographique peuvent être considérés comme des systèmes distribués qui prennent en charge le domaine spatial. Dans ce chapitre, nous avons présenté les définitions, les structures, les architectures et aussi une brève description des fonctionnalités qui doivent être chargés et relative aux SIG. Dans les chapitres suivants, nous détaillons les composants et les fonctions de base accordés aux systèmes d'information et aussi les méthodes de modélisation et de structuration des données dans les SIG.

Questions

1. Donnez les 05 fonctionnalités applicatives de base d'un SIG, dites les cinq "A".
2. Donnez les fonctionnalités métier de base d'un SIG.
3. Donnez les différentes façons comment localiser un objet.
4. Il existe deux modélisations importantes pour implémenter un SIG. Citez-les.
5. Donnez les deux formats de l'information géographique

Réponses

1. Les cinq "A" d'un système d'information géographique :
 - Abstraction : Modélisation de l'information.
 - Acquisition : récupération des informations existantes.
 - Archivage : stocker les données de façon à les retrouver et les interroger facilement.
 - Analyse : réponse aux requêtes.
 - Affichage : restitution graphique.
2. Les fonctionnalités métier d'un système d'information géographique
 - (a) Où : Localisation géographique d'un ou de plusieurs objets
 - (b) Quoi : Recherches des objets présents à un endroit ou dans une zone donnée
 - (c) Comment : Relations entre objets/phénomènes
Réalité sous-jacente révélée par leur répartition géographique ?
 - (d) Quand : analyse temporelle : évolution d'un objet ou phénomène.
 - (e) Si : simulation, c'est-à-dire l'étude des conséquences d'une décision du fait de la localisation des objets/phénomènes concernés
3. Les différentes façons pour localiser des objets sont :
 - Indirecte :
 - Référence à un objet déjà géolocalisé
 - Exemples : code postal, code communal ou zone de référence
 - Lien avec un autre objet de la base
 - Directe :
 - Position sur l'ellipsoïde géodésique modélisant la Terre : $O_{i,j,H}$
 - Coordonnée géographique (longitude, latitude) : i, j
 - Altitude H
 - Représentation : paire $[i, j]$ voire triplet $[i, j, H]$
4. Les deux modélisations importantes à considérer pour implémenter un SIG sont la modélisation dans les BD et la modélisation cartographique.

2

Structures des données spatiales

Sommaire

2.1	Introduction	23
2.2	Base de données géographique	23
2.2.1	les avantages d'utilisation des bases de données	24
2.2.2	Les modèles de base de données	25
2.3	Les modes de présentation des données géographiques	28
2.4	Structure des données dans le mode vectoriel	30
2.4.1	Structures non topologiques	30
2.4.2	Structures topologiques	31
2.5	Structure spécifique à l'information altimétrique	37
2.5.1	Structure en réseaux triangulaires irréguliers : RTI (en anglais, Triangulated Irregular Network : TIN)	37
2.5.2	Structure topologique pour les courbes de niveau	39
2.6	Mode raser	43
2.7	Conclusion	48

2.1 Introduction

Historiquement, les données spatiales sont représentées sous forme des cartes. Elle comporte les zones et peu des données sémantiques qui décrivent brièvement les localités concernées. Avec l'avancée des systèmes matériels et logiciels de traitement des données, le linkage entre les données sémantiques et spatiales est fortifié. Cependant, l'acquisition, le filtrage et la mise à jour des données restent une tâche qui ne cesse pas à se développer. Les ressources des données sont souvent complexes et porte des imprécisions. Les données doivent être donc réorganisées et être restructurées de façon à ce qu'ils soient faciles à les réutiliser facilement. Généralement, les structures données sont conçues afin d'éviter la redondance et faciliter les opérations de traitement, autrement dit, l'espace mémoire alloué et aussi et surtout le temps de calcul sont à optimiser. En effet, les données complexes portent des imprécisions qui sont du à cause de

1. redondance d'information : l'information redondante peut-être due d'une redondance directe par exemple en trouve la même information sur l'adresse d'une localité dans plusieurs fichiers ou par induction. Par exemple, on trouve les coordonnées des objets et aussi les distances, alors que les distances peuvent être calculées par les coordonnées
2. sémantique floue : par exemple si les calculs de distance sur une carte se font à base d'échelle mal adoptée.
3. support de mauvaise qualité, tel que les textes mal scannés.

2.2 Base de données géographique

C'est une collection des données spatiales et non spatiales structurées et organisées d'une façon à être interrogeables et analysables de manière interactive ou automatique. Une base de données géographique vise souvent une zone définie. Elle a les caractéristiques suivantes.

1. Elle est gérée et administrée par un logiciel SIG.
2. Elle regroupe les données elles-mêmes ainsi que leurs métadonnées.
3. Une base de données est un ensemble de représentations de la réalité sous forme de données inter-reliées, aussi cohérentes que possible, mémorisées et structurées de manière à faciliter leur exploitation pour satisfaire une grande variété d'applications.

Au début, les données spatiales sont représentées sous forme des données graphiques. Les cartes (cartes topographiques, cartes thématiques, plans cadastraux, etc.) fournissent un grand nombre d'informations à référence spatiale qui sont ouverte pour des applications très diverses. Ces documents cartographiques composent bien des bases de données graphiques. Mais, il est facile de déduire et de comprendre les limites de ces bases graphiques et qui sont principalement

1. la perte d'information,
2. Le recours indispensable aux spécialistes du graphisme,

3. Le caractère statique de la carte qui rend difficile toute déduction sur la dynamique du phénomène à représenter,
4. le coût de l'impression, etc.

Avec l'évolution des systèmes distribués et les systèmes d'information en général, les données graphiques sont numérisées et structurées d'une façon à pouvoir :

1. Conservation et sauvegarde des données originales : les données sont enregistrées indépendamment de la représentation graphique,
2. Gestion dynamique : il ouvre la possibilité de mises à jour rapides et continues dans le temps ;
3. Facilité de traitement et de manipulation : cette facilité est assurée par le changement d'échelle ou de projection, combinaison de thèmes, etc.,
4. Richesse de traitement graphique : le mode numérique réduit considérablement le coût des réalisations cartographiques.

Comme il est indiqué dans le premier chapitre, les systèmes d'information géographique doivent être conçus pour qu'ils soient capables de gérer les bases de données standard (sémantique) et spatiale ainsi les dépendances fonctionnelles entre les caractéristiques spatiales et sémantiques. Pratiquement, la réalisation des SIG exploite les systèmes de gestion de bases de données existants (SGBD) par trois approches.

1. Extension du SGBD : Les données sont complètement accessibles à travers le système de gestion, et doivent structurer selon les modèles conçus par le réalisateur du SGBD. Donc, les structures de données existantes sont adoptées pour gérer convenablement le caractère spatial.
2. Solution mixte : La solution mixte est basée sur la coexistence de deux modèles pour les données spatiales et thématiques. Les données sémantiques sont gérées par le biais du système (souvent les tableaux des attributs et aussi les relations) parce qu'elles agissent sur les modèles du SGBD. Les données qui modélisent l'information spatiale (les données de géométrie et de localisation) sont traitées par le système spatial parce qu'elles ne reflètent pas les modèles des SGBD.
3. Intégration du domaine spatial dans le SGBD : Mise au point de SGBD incluant les fonctionnalités nécessaires pour gérer les entités spatiales. C'est le cas des SGBD orientés - objets qui sont en cours de réalisation.

2.2.1 les avantages d'utilisation des bases de données

Nous citons quelques avantages d'utilisation des bases de données et les approches correspondant aux systèmes dédiés au gestion et l'administration. Ces avantages sont associés aux caractéristiques fonctionnelles.

1. minimisation des données dupliquées (la redondance) : les modèles conceptuels des données ont pour objectif de minimiser le maximum les données afin d'optimiser l'utilisation d'espace mémoire. L'expérience acquise et le nombre de recherches important dans la modélisation dans le domaine contribuent clairement à une meilleure structuration des données.

2. Assurance de l'intégrité des données et amélioration de la qualité : les modèles structurels assurent les liens d'association et donnent une visibilité simple et compréhensible.
3. Préservation des définitions sémantiques avec les données (métadonnées) : bien que les données ne soient pas souvent répliquées, les caractéristiques sémantiques et les liens fonctionnels et logiques sont préservés.
4. Réduction du coût de développement des logiciels (il exist beaucoup d'opérations fondamentales qui sont prises en charge par le SGBD) et éviter la modification les logiciels d'application à chaque fois qu'on modifie la structure de la base de données ;
5. Minimisation de l'accès aux données et conservation de la sécurité des informations.

Les intérêts cités concernent principalement les caractéristiques fonctionnelles fournies par les bases de données mais les caractéristiques non fonctionnelles telles que la disponibilité, fiabilité, sécurité et d'autres sont aussi assurées. En bref, l'évolution des systèmes d'information et les systèmes distribués conduisent à une meilleure formalisation de structures de données.

2.2.2 Les modèles de base de données

Différents modèles de base de données peuvent être exploités pour construire une base de données géographique. Pratiquement, certains modèles conduisent parfois aux mêmes structures de données tels que le modèle réseau, entité association, et relationnel mais, les visions conceptuelles diffèrent et aussi l'évolutivité du modèle définit le modèle à choisir. Les types de modèles de bases de données pratiquement utilisés sont les suivants.

1. Le modèle tabulaire : les données sont enregistrées sous forme d'un simple tableau de lignes et de colonnes, comme dans un logiciel de tableur.
2. Hiérarchique : les données sont sauvegardées sous forme d'une hiérarchie. Un fils ne peut avoir qu'un seul père et un père peut avoir plusieurs fils. Pour identifier les relations entre les noeuds, la hiérarchie doit être respectée. L'exemple dans la figure 2.1 illustre la simplicité d'utiliser ce modèle pour modéliser l'information géographique d'une paye. La carte dans la figure 2.2 présente une localité composée de deux zones. La zone I est composée de quatre sommets (1, 2, 3 et 4) et quatre lignes. La zone II est aussi composée de quatre sommets (4,3,5 et 6) et quatre lignes. La ligne de deux sommets (3, 4) est une ligne frontière entre les deux zones. La figure 2.3 illustre le modèle hiérarchique associé à la carte M. Au sommet, le nom de la localité « M » est montré. Puis la carte M a deux fils : la zone I et II. Les zones ont des segments (ligne) comme des fils. Les sommets définissent les fils des segments. Les coordonnées des sommets peuvent être ajoutés comme des fils des sommets.
3. modèle en réseau : les données sont organisées pour construire des réseaux entre les entités représentées. Une entité peut avoir plusieurs pères et plusieurs fils à la fois. En effet, le modèle réseau beaucoup plus flexible que le modèle hiérarchique et il supporte les relations spatiales complexes. D'autre part, le modèle réseau minimise beaucoup plus la duplication des données. Dans la figure ??, le modèle réseau est appliqué. La ligne « c » a deux parents : la zone I et la zone II. Donc, il est facile

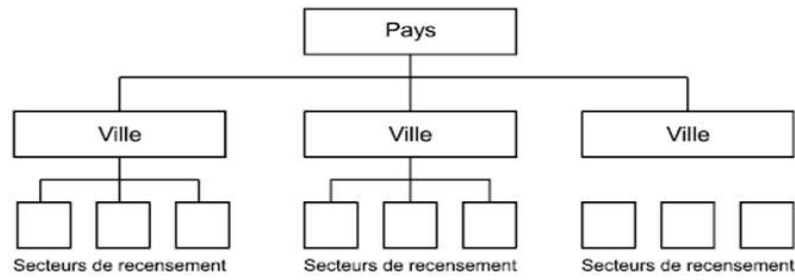


FIG. 2.1 – Le modèle hiérarchique d'une paye.

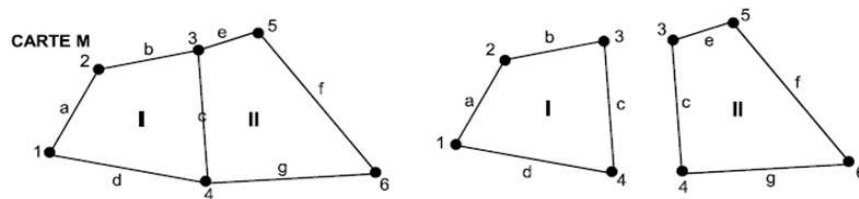


FIG. 2.2 – Une carte M composée de deux régions.

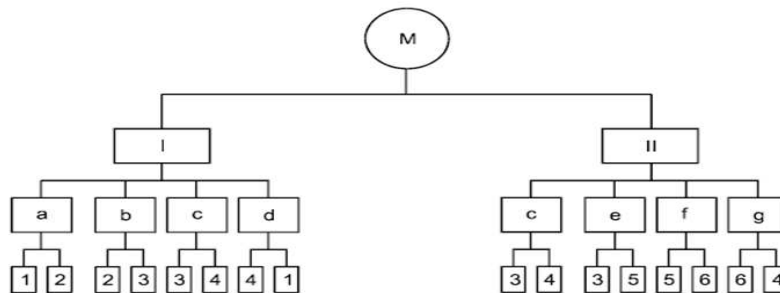


FIG. 2.3 – Le modèle hiérarchique de la carte M.

d'extraire les segments en commun entre les zones. Contrairement au modèle hiérarchique où on est obligé de parcourir une grande partie de la structure pour extraire les segments en commun. D'autre part, le segment c est répliqué dans le modèle hiérarchique.

4. Relationnel : les données sont reliées par des relations qui établissent la correspondance entre des informations distribuées dans plusieurs tableaux. Le modèle relationnel a été défini pour la première fois par E. F. Codd, chercheur chez IBM, en 1970. Ce modèle fournit une approche flexible des liens entre les enregistrements, ce qui permet une modélisation plus satisfaisante des relations spatiales complexes pouvant exister entre les objets représentés dans la base de données. Les données sont arrangées dans des enregistrements simples contenant une collection ordonnée de valeurs d'attributs. Ces collections sont rassemblées dans des tables bidimensionnelles. Chaque table est enregistrée souvent dans un fichier séparé. Les relations

entre les tables sont assurées par les clés ou chaque enregistrement est identifié par une clé. Les trois critères suivants doivent être respectés.

- (a) Deux enregistrements (lignes) ne sont jamais identiques.
- (b) L'ordre des lignes n'est pas important.
- (c) L'ordre des colonnes doit être respecté.

La figure 2.4 représente le modèle relationnel qui correspond à la carte « M ».

ID_zone	Nom du zone
I	Région est
II	Région ouest

Id_sommet	Coordonnée X	Coordonnée y
1	10	10
2	15	30
3	46	35
4	50	2
5	65	40
6	80	6

ID_seg	sommet_1	Sommet_2	Zone_gauche	Zone_droite
a	1	2	-----	I
b	2	3	-----	I
c	3	4	I	II
d	1	4	-----	I
e	3	5	-----	II
f	5	6	II	-----
g	4	6	-----	II

FIG. 2.4 – Le modèle relationnel de la carte M.

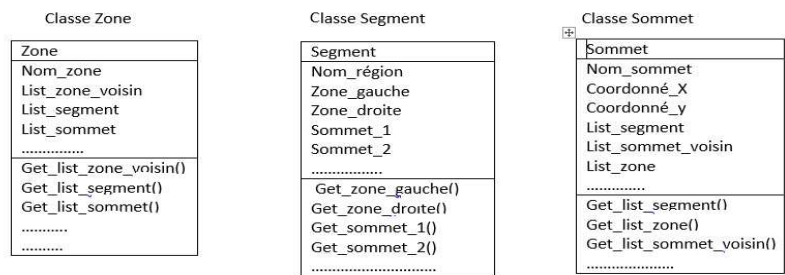


FIG. 2.5 – Les classes du modèle orienté objet de la carte M.

5. Le modèle orienté-objet : les structures données sont modélisées autour du concept d'objets qui peuvent être assimilés aux entités du monde réel. Ils sont peu réponsus du fait que les relations par des pointeurs entre les objets sont difficiles d'être sauvegardées dans des mémoires persistantes. La puissance de ces modèles en terme de conception, traitement rapide et modularité évoque ces derniers à jouer un rôle important dans les SIG. Le besoin des nouveaux concepts et outil qui dépasse les limites

des modèles relationnels et l'apparition des langages de programmation orientés-objet représente des facteurs supplémentaires qui poussent les concepteurs à utiliser ces modèles. Les données et le traitement ne sont plus séparables. Chaque objet possède ces propres fonctions. Les relations entre les objets peuvent être réalisées par plusieurs façons (ensemble, liste .etc.). L'indexation aux objets peuvent se faire direction sans passé par les opérations de jointure. La figure 2.5 représente les objets associés aux composants de la carte : la région, les zones, les segments (les lignes) et les points. Les sommets composant une zone sont faciles à extraire à partir de la fonction « `get_list_sommet` » de la classe « `zone` ». Alors que dans le modèle relationnel, une recherche dans les trois tables relationnelles est nécessaire.

2.3 Les modes de présentation des données géographiques

Les données géographiques sont représentées par deux modes : le mode vectoriel et le mode raster [12] [15] [4] [1]. L'unité logique de base dans le modèle vectoriel est la ligne (limite d'une zone, rivière, courbe de niveau, etc.). En effet, une série des points (qui définissent les localisations par les coordonnées, X, Y) sont sauvegardés le long de la ligne. Un point peut être défini comme une ligne de longueur nulle. En mode raster, l'unité de base est le pixel (element picture en anglais) : une cellule qui représente une partie de la surface. Donc, les zones sont subdivisées à des cellules adjacentes de la même taille. La forme de ces cellules est souvent rectangulaire bien que d'autres formes peuvent être utilisées (hexagonal ou triangulaire). En mode raster, les approches et les méthodes utilisées sont issues du domaine de traitement d'image. Pratiquement, les systèmes d'information géographique utilisent les deux modes à la fois pour exploiter les avantages de chaque mode. En effet, chaque mode complète l'autre.

Pour concevoir l'information géographique localisée, trois niveaux d'information sont principalement utilisés : le niveau géométrique qui concerne la géométrie des objets géographiques, le niveau sémantique dédié à la nature des objets géographiques et le niveau topologique, qui représente les relations connexité et de liaison entre les objets géométriques [12][10].

1. Les données géométriques sont constituées par de points, de lignes et de polygones. Une ligne peut être modélisée par un ensemble de segments (un segment est une droite composée de 2 points), un ensemble de courbes plus complexes (cubiques ...) ou les deux à la fois, c'est-à-dire un ensemble de segments et de courbes. Alors qu'un polygone est composé par une ligne fermée. Tout polygone peut être à l'intérieur d'une zone décrit aussi par d'autres polygones (un parc dans une ville, une île sur un lac, une cour à l'intérieur d'une maison ...).
2. Les données sémantiques représentent les caractéristiques non géométriques des données géographiques. Elles peuvent décrire des entités simples (carrefour, bâtiment, portion de route) ou des entités plus complexes (route, échangeur, ensemble de bâtiments ...etc.). Effectivement, les caractéristiques sémantiques d'un objet telles que le nom, l'adresse, largeur peuvent être associé à l'objet concerné.

2.4 Structure des données dans le mode vectoriel

Afin de mieux concevoir les modèles de stockage des données spatiales dans le mode vectoriel, les structures de données sont catégorisées en deux grandes classes : structures topologiques et non topologiques [14] [15] [4] [1]. Les modèles non topologiques ne sauvegardent pas les relations de connectivité entre les objets géométriques, par conséquent, ces relations sont déduites par voie logicielle. Alors que dans les modèles topologiques, les relations topologiques sont enregistrées explicitement.

2.4.1 Structures non topologiques

Dans les modèles les plus simples, seulement les notions de point et de ligne sont enregistrées. Un dictionnaire de données est utilisé pour stocker les successions de points. Deux points successifs représentent une ligne droite ou curviligne. Seulement la taille et la forme des objets spatiaux sont facilement déduites aux moyennes des successions des points. Ces modèles conduisent à ce qui connut sous le nom « structure spaghetti ». La « structure séquentielles simple » est considérée comme une structure non topologique. La base de données est sauvegardée sous forme d'un fichier des séries des paires des coordonnées des points : X, Y. l'organisation de la structure est la suivante.

1. Une succession de deux paires de points représente une ligne.
2. Une entité spatiale polygone est décrite par une série de paires.
3. Un polygone est indépendant des autres, autrement dit, chaque série des paires est dédiée pour un seul polygone.
4. Une paire de coordonnées spatiales est dédié pour séparer les séries (polygone qui correspond à une zone). fermeture des polygones est implicite en répétant la paire des coordonnées du point initial

La figure 2.7 montre un exemple d'utilisation du modèle séquentiel pour enregistrer les zones de la carte M_1 . La région de la carte M_1 est composée de 4 zones fermées. Dans le fichier séquentiel, les zones sont séparées par -1. La paire des coordonnées de point initial est répétée (pour la zone z_1 , c'est x1, y1). Les coordonnées en commun entre les zones sont répétées pour chaque zone. Pour alléger l'opération de double saisie du point, un dictionnaire de point est utilisé (voir Fig. 2.8).

Les structures séquentielles sont faciles à concevoir, mais elles surmontent certaines incontinences qui sont principalement :

1. l'espace de stockage alloué : les données répliquées nécessitent un espace mémoire supplémentaire.
2. le cout de développement : pour implémenter les fonctions et les procédures de base telles que la recherche des segments en commun entre les zones, des efforts importants doivent être fournis.
3. le temps de traitement dans la phase d'exploitation est aussi important : pour obtenir les segments entre deux zones, par exemple, le fichier doit être parcouru pour identifier les zones concernées. Puis, il faut déterminer les segments par la succession des deux paires dans les deux séries qui représentent les deux zones. A la fin, une comparaison est entamée.

X	Y	X	Y	X	Y
x1	y1	x12	y12	x19	y19
x2	y2	x13	y13	x1	y1
x3	y3	x14	y14	-1	-1
x4	y4	x15	y15	x19	y19
x5	y5	x16	y16	x15	y15
x6	y6	x6	y6	x14	y14
x7	y7	x5	y5	x13	y13
x8	y8	-1	-1	x12	y12
x1	y1	x1	y1	x11	y11
-1	-1	x8	y8	x17	y17
x5	y5	x7	y7	x18	y18
x9	y9	x6	y6	x19	y19
x10	y10	x16	y16	-1	-1
x11	y11	x15	y15		

FIG. 2.7 – Structure séquentielle de la carte M_1 .

Fichier des zones

Pointeurs			
p1	p11	p11	
p2	p12	p17	
p3	p13	p18	
p4	p14	p19	
p5	p15	-1	
p6	p16	p1	
p7	p6	p8	
p8	p5	p7	
p1	-1	p6	
-1	p19	p16	
p5	p15	p15	
p9	p14	p19	
p10	p13	p1	
	p12	-1	

Dictionnaire

Etiquette	X	Y
p1	x1	y1
p2	x2	y2
p3	x3	y3
p4	x4	y4
p17	x17	y17
p18	x18	y18
p19	x19	y19

(a) (b)

FIG. 2.8 – Structure séquentielle avec dictionnaire de points de la carte M_1 .

2.4.2 Structures topologiques

Des structures plus évoluées sont désignées d'être des structures topologiques du fait que la majorité des relations topologiques sont explicitement sauvegardées. Ces struc-

tures sont conçues comme un graphe, par conséquent, un ensemble des règles issues des caractéristiques des graphes sont appliquées pour modéliser convenablement les relations topologiques. Ces règles sont énumérées comme suit.

1. Les nœuds ou sommets du graphe modélisent les points du fond de la carte,
2. Les arcs ou les arrêtes connectant deux sommets du graphe décrivent les segments.
3. Les zones sont représentées par une séquence d'arcs et apparaissent donc comme des circuits dans le graphe.
4. Des fichiers séparés sont créés pour distinguer entre les éléments de graphe.
5. Les pointeurs entre les fichiers décrivent les relations topologiques entre les éléments des graphes qui forment les entités principales de la carte (les zones, les lignes et les points).

Les points de fond qui décrivent sur la carte l'intersection entre deux lignes (les routes, les rivières, les frontières entre deux zones) sont représentés par les sommets de graphe. Les lignes sont décrites par des arcs ou arrêtes et sont aussi enregistrées dans un autre fichier. Les circuits dans le graphe modélisent les zones. Pour chaque entité (points, lignes, zones), un fichier est souvent dédié pour le stockage. Dans la plupart des structures, la relation topologique zone-ligne (la relation topologique entre les lignes et les zones) est représentée par des pointeurs entre le fichier des lignes et le fichier des zones. La relation ligne-point est aussi représentée par des pointeurs. La relation zone-point est déduite par voie logicielle. La relation point-point est assurée par ordre dans les séquences des sommets. Les autres relations topologiques (« zone-zone », « ligne-ligne ») dépendent de la structure utilisée (elles sont sous forme des pointeurs ou déduite par voie logicielle).

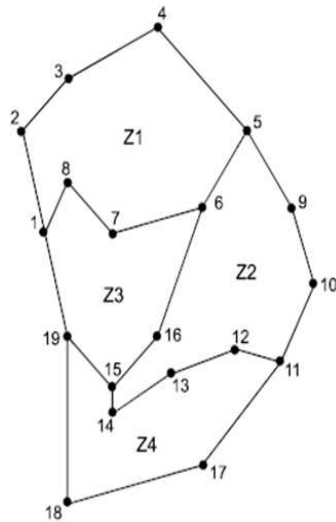
Structure DIME : Dual independent Map encoding

DIME est considérée comme étant la première et le plus célèbre système opérationnel qui permet de modéliser une structure de données spatiale. Ce système a été conçu pour la visualisation et la gestion d'une base de données présentant les réseaux urbains de voirie et les unités statistiques intra-urbaines utilisées pour les recensements américains dès 1970. Le segment associé à une face d'un îlot urbain est l'élément de base de cette structure. Les lignes complexes telles que les lignes de chemin de fer et les rues sont décrites par une série de segments successifs. La disposition des arrêtes identifie les îlots, ou blocs à base desquels les données de recensement sont rassemblées. Un fichier d'étiquettes d'arrêtes et un fichier dictionnaire de nœuds sont utilisés pour définir les zones.

En effet, la base de données est constituée d'un fichier de coordonnées, enregistrant les sommets de graphe, et un fichier d'arrête. Le fichier des sommets se compose des enregistrements. Chaque enregistrement contient l'étiquette d'un nœud et les deux coordonnées du sommet : X et Y . Par conséquent, deux nœuds n'ont pas la même étiquette et les mêmes coordonnées. Chaque enregistrement du second fichier est associé à un arrête. Il contient une étiquette, un pointeur du nœud de départ, un pointeur du nœud de destination et deux codes identifiant les deux blocs de part et d'autre de l'arrête. D'autres informations sémantiques (attributaire), telles que le nom de la rue, peuvent être jointes.

Afin de mieux contrôler l'étape d'acquisition des données, les critères suivants doivent être respectés.

1. Deux arcs n'ont pas la même étiquette.
2. Le pointeur du nœud de départ n'est pas vide et n'est pas nul.
3. Le pointeur du nœud de destination n'est pas vide et n'est pas nul.
4. Les pointeurs de nœud de départ et de nœuds de destination ne sont pas les mêmes et ne pointent pas sur les mêmes arcs (dans la plupart des systèmes).
5. Au moins l'un des codes des deux voisins n'est pas vide.
6. Les codes, indiquant les zones voisin d'un arc, ne sont pas les mêmes.
7. Les sommets représentent des chevauchements entre les arcs, autrement dit, deux arcs n'ont pas une intersection.



(a)

Etiquette	X	Y
N1	X_{N1}	Y_{N1}
N2	X_{N2}	Y_{N2}
N3	X_{N3}	Y_{N3}
...
N17	X_{N17}	Y_{N17}
N18	X_{N18}	Y_{N18}
N19	X_{N19}	Y_{N19}

(b)

FIG. 2.9 – Nœuds de la carte M_1 selon la structure DIME. (a) la carte M_1 ; (b) fichier des nœuds.

Conceptuellement, ces critères sont respectés, car ce modèle est basé sur les graphes qui sont soutenus par une base mathématique solide (théorie de graphe). Cependant dans l'étape d'acquisition des données (saisie, acquisition par radar...etc.), il suffit qu'il existe une faute de saisie de coordonnées d'un nœud, les critères cités dessus seront violés. La vérification de ces critères est faite par des méthodes non automatiques telles que l'affichage de la carte à partir de la structure. Un SIG au niveau d'un pays nécessite des bases de données distribués et un nombre important des outils d'acquisition (y compris les agents humains). Il est important de mettre en place des outils(logicielles) pour la vérification des structures.

Les deux figures 2.9 et 2.10 illustrent la modélisation d'une carte par l'utilisation du modèle DIME. La carte M_1 est composée de 19 nœuds et 22 arcs. Le fichier des nœuds

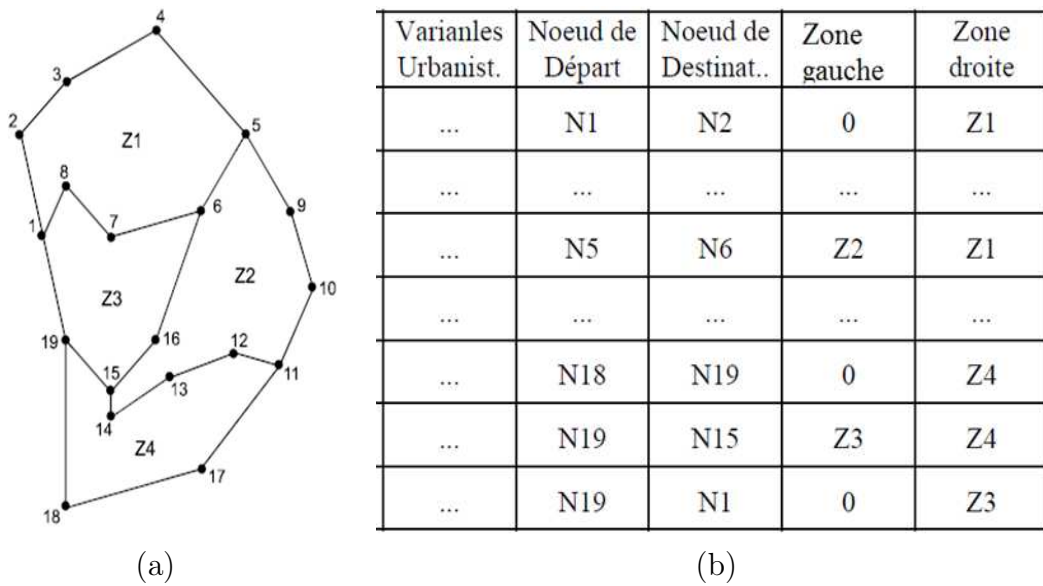


FIG. 2.10 – Arrêtes de la carte M_1 selon la structure DIME. (a) la carte M_1 ; (b) fichier des arrêtes.

(voir Fig. 2.9 (b)) est composé de trois attributs : étiquette de nœud, coordonnée X et coordonnée Y. Conventionnellement, les attributs du fichier d'arrête (voir Fig. 2.10 (b)) sont ordonnés comme : Etiquette, nœud de départ, nœud de destination, Zone gauche, Zone droite.

La structure POLYVRT (POLYgone conVerTer)

Cette structure emploie une modélisation plus détaillée des composants spatiaux qui constituent le fond de la carte. Cette structure est considérée comme référence à plusieurs systèmes conçus par la suite. La ligne est aussi un élément de base sur laquelle elle est fondée. Une chaîne est définie pour qu'elle soit constituée d'un ou de plusieurs segments successifs. Elle peut seule représenter la frontière comme une île. Les segments sont reliés par des points intermédiaires. Une chaîne est limitée par deux nœuds, alors qu'un segment se trouve entre deux points intermédiaires ou un point intermédiaire et un nœud s'il est sur l'extrémité d'une chaîne. Par conséquent :

1. un point intermédiaire n'appartient qu'une seule chaîne ;
2. les nœuds représentent les points de rencontre entre les chaînes, et de ce fait, ils peuvent appartenir à plusieurs chaînes.
3. une chaîne doit être appartenie au moins à une zone et au plus à deux zones adjacentes.
4. les zones sont définies par une succession des chaînes formant un circuit fermé.
5. une chaîne doit contenir, au moins et au plus, deux nœuds sur leurs extrémités.

La structure POLYVERT est composée de quatre fichiers : un fichier de zones, un fichier chaînes, un fichier pour les points intermédiaires et un fichier pour les nœuds. Si

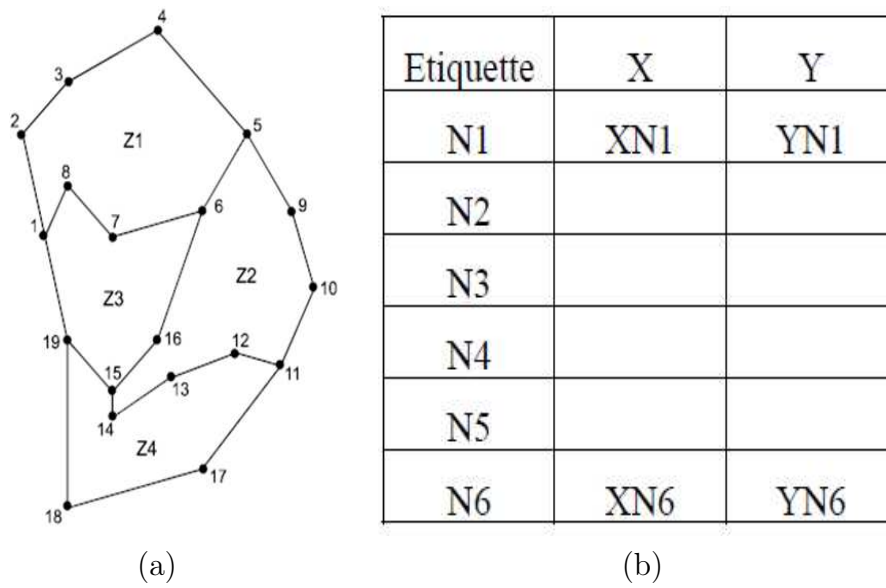


FIG. 2.11 – Nœuds de la carte M_1 selon la structure POLYVRT. (a) La carte M_1 ; (b) Le fichier des nœuds.

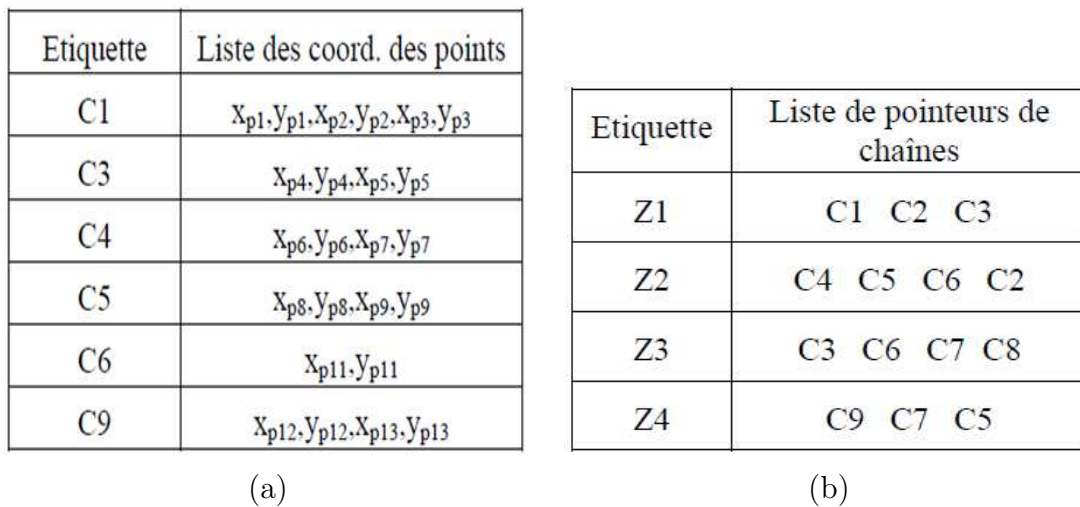


FIG. 2.12 – Fichier des points intermédiaires et fichier des zones de la structure POLYVRT de la carte M_1 . (a) le fichier des points intermédiaires; (b) le fichier des zones.

l'on exclut les informations sémantiques, les deux fichiers de zones et de chaînes ne sont constitués que des pointeurs. Les enregistrements de chaque fichier sont organisés comme suit.

1. Fichier des nœuds : chaque enregistrement de fichier des nœuds sauvegarde une étiquette qui désigne le nœud concerné. Un champ est consacré pour le coordonnées X et un autre pour le coordonné Y .
2. Fichier point intermédiaire : l'étiquette de la chaîne concernée est stockée après que

Etiquette	Nombre de points	Noeud de départ	Noeud de destination.	Zone gauche	Zone droite
C1	3	N1	N2	0	Z1
C2	0	N2	N3	Z2	Z1
C3	2	N3	N1	Z3	Z1
C4	2	N2	N4	0	Z2
C5	3	N4	N5	Z4	Z2
C6	1	N5	N3	Z3	Z2
C7	0	N5	N6	Z4	Z3
C8	0	N6	N1	0	Z3
C9	2	N4	N6	0	Z4

FIG. 2.13 – Fichier chaîne de la structure POLYVRT de la carte M_1 .

la liste des coordonnées des points intermédiaires soit enregistrée. Les coordonnées sont séparées par des virgules ou un autre caractère non numérique.

3. Fichier chaîne : un enregistrement de ce fichier contient les informations suivantes par ordre.
 - (a) pointeur vers l'étiquette représentant les listes de point intermédiaire de la chaîne concernée
 - (b) le nombre de points intermédiaires.
 - (c) pointeur vers le nœud de départ.
 - (d) pointeur vers le nœud de destination.
 - (e) pointeur vers la zone gauche de la chaîne (ou haut).
 - (f) pointeur vers la zone droite (ou en bas).
4. Fichier des zones : l'étiquette de la zone, suivie par la liste des chaînes constituant par ordre les zones, est sauvegardée.

Les figures 2.11, 2.12 et 2.13 illustrent les fichiers utilisés pour la présentation de la carte M_1 selon le modèle POLYVRT. Le nombre des nœuds dans la carte est 6, les chaînes sont 9, les points intermédiaires sont 13 alors que les zones sont 4. Les mêmes étiquettes, qui sont utilisées dans le fichier des points intermédiaire pour désigner une chaîne, sont employées comme des pointeurs dans le fichier « zone » et le fichier « chaîne ». De la même manière, les étiquettes des nœuds dans le fichier « nœuds » sont utilisées comme des pointeurs dans le fichier « chaîne ».

Deux remarques qui attirent l'attention dans la structure POLYVERT. Les nœuds sont sauvegardés dans un fichier distingué pour éviter la redondance. Les coordonnées des points intermédiaires sont listées directement dans le fichier des points intermédiaire, car un point intermédiaire n'appartient qu'une seule chaîne. Cependant, un fichier est consacré

aux zones bien que les zones peuvent être déduites du fichier de chaîne. L'intérêt de sauvegarder explicitement les relations topologiques est l'augmentation de la performance de système et aussi la diminution du coût de développement.

2.5 Structure spécifique à l'information altimétrique

L'information altimétrique concerne les données associées à la hauteur des unités spatiales. Pour modéliser les reliefs tels que les montagnes, les plateaux, les profondeurs des océans), des structures spécifiques sont conçues telles que le réseau triangulaire irrégulier et les courbes de niveau.

2.5.1 Structure en réseaux triangulaires irréguliers : RTI (en anglais, Triangulated Irregular Network : TIN)

Cette structure a été proposée la première fois en 1970 [15]. Après l'étape d'acquisition des points d'appui (de fonds), une opération de sélection des points les plus significatifs selon le changement d'altitude du point à un autre. Les points sont liés en formant des triangles. Trois points A, B, C sont liés pour former un triangle, si l'on peut dessiner un cercle « C » passant par ces trois points sans qu'il existe un autre point dans le cercle « C ». La figure 2.14 montre le processus de choix des triangles.

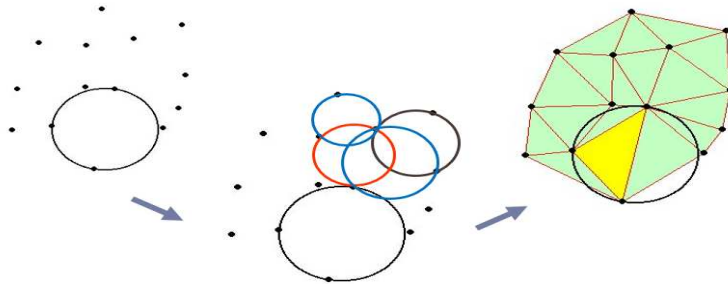


FIG. 2.14 – Construction des triangles à partir des points.

Deux méthodes peuvent être employées pour encoder les triangles : par triangle ou par sommet.

1. Présentation de RTI par triangle : La base de données est sous forme d'un seul fichier. En associant à chaque triangle un enregistrement dans le fichier contenant
 - (a) le numéro de référence de triangle concerné,
 - (b) les coordonnées des trois sommets constituant le triangle,
 - (c) et les numéros de référence des triangles voisins.
2. Présentation de RTI par sommet : les enregistrements de cette structure contiennent
 - (a) le numéro d'identification du sommet,
 - (b) les coordonnées du sommet concerné,

(c) et les pointeurs vers les sommets voisins.

La figure 2.15 montre un ensemble de points en réseaux triangulaires irréguliers. La figure 2.16 montre un exemple de structuration d'un ensemble des points d'altitude selon la présentation par triangle.

La figure 2.17 illustre la présentation par sommet de l'exemple précédent.

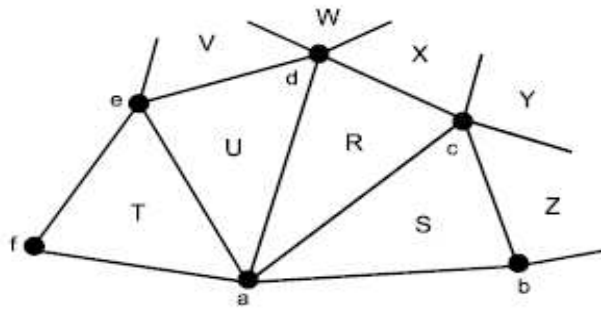


FIG. 2.15 – Structure en réseaux triangulaires irréguliers.

Triangle	Sommets	Triangles adjacents
R	a(x,y,z), c(x,y,z), d(x,y,z)	S, X, U
S	a(x,y,z), c(x,y,z), b(x,y,z)	Z, R
etc.		

FIG. 2.16 – Sauvegarde des données par triangles.

Sommet	Coordonnées	Sommets connectés
a	x, y, z	f, e, d, c, b, ...
b	x, y, z	a, c, ...
c	x, y, z	a, d, b, ...
etc.		

FIG. 2.17 – Sauvegarde des données par sommets.

Pour l'édition (affichage) des données, une projection sur le plan des deux coordonnées « x » et « y » est effectuée. Une coloration dégradée des couleurs est utilisée pour servir davantage de l'inspection visuelle (voir Fig. 2.18).

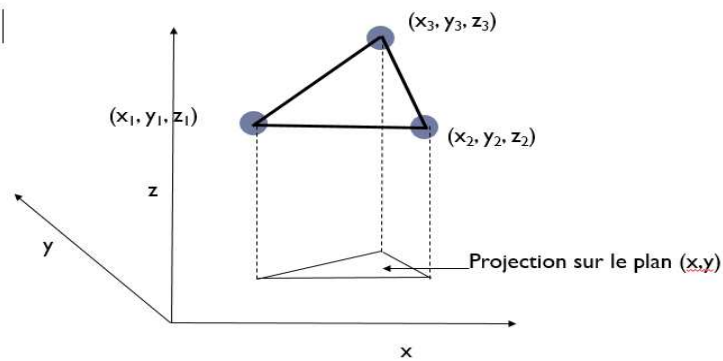


FIG. 2.18 – Projection des points sur le plan (XY).

2.5.2 Structure topologique pour les courbes de niveau

Les structurations des données altimétriques sont fondées sur les courbes de niveau pour répondre aux besoins de traitement complexe de l'information altimétrique. L'idée de base de cette structure est d'identifier et présenter les points voisins de la même altitude dans la même ligne, ce qui assure une organisation efficace des données de base. En effet, la relation topologique entre les points voisins est assurée par une relation d'ordre dans le fichier de point. Les fichiers de voisins supérieurs et inférieurs permettent l'accès direct aux courbes immédiatement supérieures ou inférieures permettant de modéliser convenablement la relation " ligne-ligne " pour un accès facile et efficace à l'information topologique. La figure 2.19 représente un exemple d'une carte des courbes de niveau.

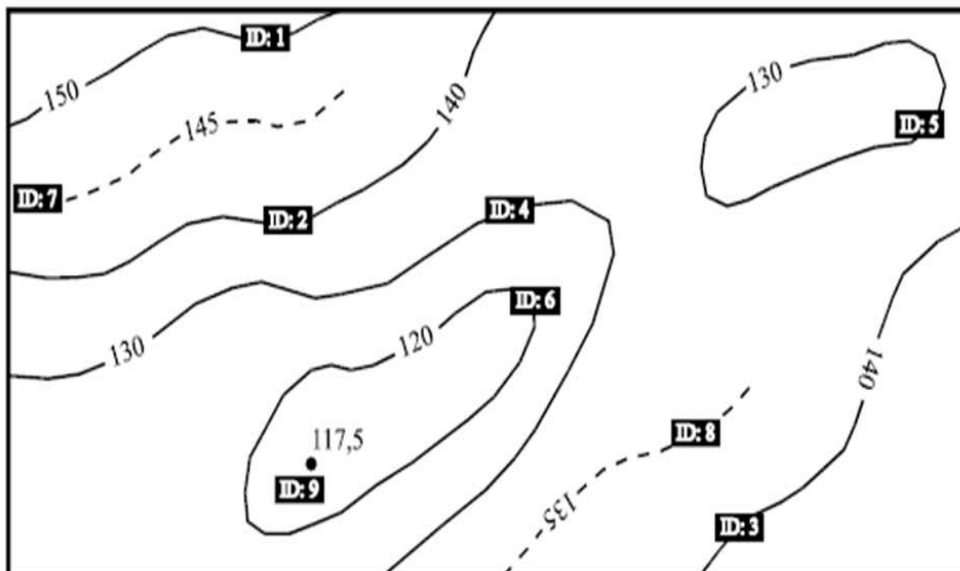


FIG. 2.19 – Courbe de niveau.

Cette structure est développée afin de répondre aux besoins de traitements complexes de l'information altimétrique. La modélisation de la structure topologique des courbes de niveau est réalisée en quatre fichiers et qui sont

1. fichier des entités (ligne ou courbe de niveau),
2. fichier des coordonnées des points de fonds,
3. fichier des courbes des voisins inférieurs,
4. fichier des courbes des voisins supérieurs.

Fichier des entités

C'est le fichier de base de la structure. Chaque courbe est caractérisée dans ce fichier par son identificateur, son altitude, sa forme et les pointeurs vers les autres fichiers permettant de déterminer les points composites et les courbes voisins. Ce fichier fait le lien avec les autres fichiers et donc les relations topologiques entre les entités spatiales. Une succession des enregistrements est sauvegardée dans le fichier des courbes. Chaque enregistrement est dédié pour une courbe de niveau. Les informations d'un enregistrement sont organisées comme suit.

1. Identifiant(ID) : Un numéro associé à chaque courbe serve comme identifiant et donc deux courbes ne peuvent pas avoir le même numéro. Les différentes caractéristiques de la courbe concernée sont accessibles par ce numéro.
2. Altitude (ALT) : il représente la hauteur de la courbe du niveau. Du fait que les points constituant la courbe de niveau ont la même hauteur (la troisième coordonnée), cette information est stockée dans l'entité courbe pour éviter la redondance des données.
3. Nombre de points (NP) : C'est le nombre de points constituant une courbe. Effectivement, le nombre de points est déterminé par l'exactitude du matériel d'acquisition des données, la capacité de stockage et de traitement de systèmes et aussi par la forme des différentes parties de la courbe. Une ligne, approximativement droite, peut être décrite seulement par deux points. Ce numéro vaut 1 dans le cas des points de côté qui représente (les profondeurs ou sommet des montagne, océan...etc.).
4. Adresse du point de départ (APD) : c'est une référence vers l'adresse du premier point de la courbe dans le fichier des points. La liste des points constituant la courbe est déduite facilement du fichier de point grâce à l'information contenue dans NP.
5. Forme (FOR) : cette rubrique est destinée pour l'allure de la ligne représentant la courbe. « For » peut avoir l'une des trois valeurs suivantes,
 - (a) F indiquant que la courbe est fermée, autrement dit, une ligne existe entre le premier et le dernier point.
 - (b) O montrant que la courbe est ouverte, il n'existe pas une ligne entre le premier et le dernier point
 - (c) P représentant un point de côté, la ligne se réduit à un point.
6. continuité(CON) : cet attribut montre si la courbe est continue, indiqué par « C », ou interrompue (discontinue), indiqué par « I ». Le point de côté est mentionné par « P ».

7. Longueur (LON) : est la longueur exprimée généralement en mètre de la courbe. Ce paramètre est employé d'une part pour l'identification de voisinage et, d'autre part, pour l'estimation des autres paramètres comme la pente générale du terrain. Les points de côté ont une valeur de longueur 0.
8. Adresse du voisinage supérieur(AVS) : c'est un pointeur qui permet d'accéder à la liste des voisins supérieurs dans le fichier des voisinages supérieurs. Deux courbes sont considérées comme des voisins si l'on peut dessiner une ligne droite entre les deux courbes sans faire l'intersection avec une autre courbe dans la région à étudier.
9. Nombre d'entités voisines supérieures (NVS) : c'est le nombre d'entités composant le voisinage supérieur d'une courbe de niveau ou d'un point coté.
10. Adresse du voisinage inférieur : il représente un pointeur vers le fichier des voisinages inférieurs y compris les courbes ou les points cotés.
11. Nombre d'entités voisines inférieures(NVI) : le nombre d'entités constituant la liste des voisinages inférieurs d'une courbe ou d'un point coté.

La figure 2.20 illustre le fichier « entité »de la carte précédente.

<i>ID</i>	<i>Alt</i>	<i>NP</i>	<i>APD</i>	<i>LON</i>	<i>FOR</i>	<i>CON</i>	<i>NVI</i>	<i>AVI</i>	<i>NVS</i>	<i>AVS</i>
1	150.0	8	1	2749.3	O	C	2	1	0	0
2	140.0	15	9	4754.0	O	C	3	2	3	1
3	140.0	9	24	2862.0	O	C	3	3	1	2
4	130.0	21	33	6893.6	O	C	1	4	4	3
5	130.0	12	54	3505.2	F	C	0	0	4	4
6	120.0	13	66	4852.2	F	C	1	5	1	5
7	145.0	7	79	2563.4	O	I	1	6	1	6
8	135.0	7	86	2571.6	O	I	2	7	2	7
9	117.5	1	93	0.0	P	P	0	0	1	8

FIG. 2.20 – Le fichier « entité »de la carte précédente.

Fichier des coordonnées

Le fichier des coordonnées enregistre tous les points constituant les courbes de niveau et les points cotés. L'ordre de point est important où chaque succession de deux points dessine un segment de ligne qui représente une portion de courbe de niveau. Il faut d'abord identifier les points constituant une courbe de niveau. Pour cela, on utilise l'adresse de point de départ dans le fichier des courbes de niveau et le nombre de points. L'adresse

de point de départ représente l'ordre d'un point dans ce fichier. Il faut aussi noter qu'il n'existe pas un segment entre le point de départ et le point juste avant. En effet, si la courbe est fermée d'après le paramètre " forme ", dans le fichier des courbes de niveau, un segment doit relier le premier point (identifié par l'adresse point de départ) et le dernier point de la courbe de niveau. Du fait qu'un point ne peut appartenir qu'une et une seule entité de niveaux, les données dans ce fichier ne sont pas redondantes.

Le fichier des points de la carte précédente est montré dans la figure 2.21.

Adresse	X	Y	Adresse	X	Y
1	x1 ₁	y1 ₁	-	-	-
2	x1 ₂	y1 ₂	-	-	-
3	x1 ₃	y1 ₃	88	x8 ₃	y8 ₃
4	x1 ₄	y1 ₄	89	x8 ₄	y8 ₄
5	x1 ₅	y1 ₅	90	x8 ₅	y8 ₅
6	x1 ₆	y1 ₆	91	x8 ₆	y8 ₆
-	-	-	92	x8 ₇	y8 ₇
-	-	-	93	x9	y9

FIG. 2.21 – Le fichier des coordonnées.

Fichier relatif aux voisinages inférieurs

Ce fichier contient les listes des identificateurs voisins inférieurs d'une courbe de niveau ou d'un point coté. Les voisins inférieurs représentent comme il est indiqué précédemment les courbes de niveau ou les points cotés qui ont une altitude inférieure à l'altitude de la courbe ou de du point coté correspondant à la liste. Généralement quand deux entités ont la même liste de voisinage inférieur, la liste est sauvegardée une nouvelle fois avec une nouvelle adresse, pour assurer le caractère d'extensibilité de la structure de base de données. Autrement dit, il sera plus facile dans la future d'ajouter de nouvelles courbes sans entamer des modifications profondes dans le fichier des voisinages. Dans la figure 2.22, le fichier de voisinage inférieur est illustré.

Fichier de voisinage supérieur

Le fichier de voisinage supérieur est modélisé pour supporter les mêmes données que le fichier des voisins inférieurs, mais il est associé aux courbes de niveau qui ont une altitude plus grande que la courbe ou le point coté concerné. La figure 2.23 montre les listes des voisins supérieurs des courbes de niveau de fichier « entité ».

Adresse	ID des Voisins inférieurs
1	7, 2
2	4, 5, 8
3	4, 5, 8
4	6
5	9
6	2
7	4, 5

FIG. 2.22 – Le fichier de voisinage inférieur.

Adresse	ID des Voisins supérieur
1	1, 7, 3
2	2
3	2, 5, 3, 8
4	2, 3, 4, 8
5	4
6	1
7	2, 3
8	6

FIG. 2.23 – Le fichier de voisinage supérieur.

2.6 Mode raser

Le mode « raser » consiste à une subdivision de la zone étudiée par un maillage régulier. L'élément de base de ce mode est une unité de surface. Généralement, trois types de maillage sont utilisés.

1. Le maillage carré.
2. Le maillage triangulaire.
3. Le maillage hexagonal.

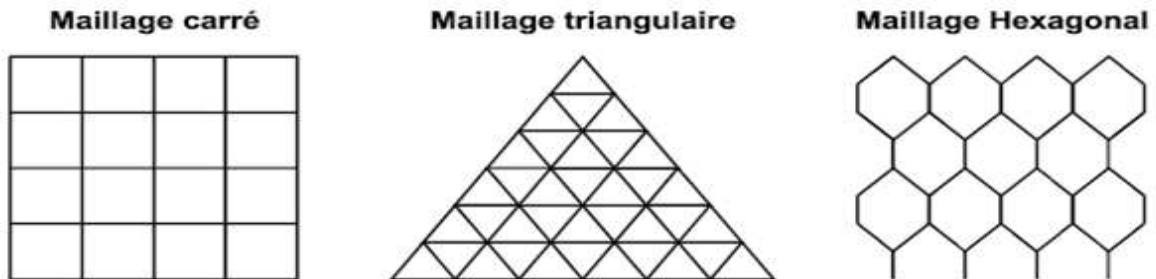


FIG. 2.24 – Les trois types de maillage fréquemment utilisés. De gauche à droite : Maillage carré, Maillage triangulaire et Maillage hexagonal.

La figure 2.24 montre les trois maillages fréquemment utilisés pour la division de l'espace à des unités de la même taille. Le maillage carré est le maillage le plus utilisé pour les deux raisons suivantes.

1. L'utilisation facile par les calculateurs.
2. Compatibilité avec les dispositifs d'acquisition et d'affichage de données spatiales.

La zone d'étude est divisée sous forme d'une grille où chaque carré de la grille correspond à un pixel. Le pixel est référencé par un numéro d'ordre de la ligne et de la colonne auxquelles il appartient. La forme de la grille est donc présentée sous forme de matrice. Chaque pixel de la grille est référencé par un numéro d'ordre de la ligne et de la colonne auxquelles il appartient au sein de la matrice. De ce fait, un pixel est associé à trois attributs : l'ordre de ligne, l'ordre de colonne et un pointeur vers une table des valeurs mémorisant les niveaux de gris des couleurs. L'enregistrement direct de la matrice de la grille nécessite un large volume de stockage. Des techniques de compression sont employées pour réduire considérablement les exigences énormes de stockage. Dans les structures « raser », une zone géographique est présentée de la façon suivante.

1. un point est représenté par un pixel (une cellule de la grille).
2. une ligne par un ensemble des pixels adjacents selon une direction.
3. Une zone par une agrégation des lignes adjacentes.

Une zone géographique est considérée comme étant une surface cartésienne et plane alors qu'un pixel est associé à une parcelle de terrain de forme carrée. Dans ce cas, la résolution est définie par la relation entre la taille pixel dans la base de données et la taille de surface sur la terre dont il est associé. Un pixel ne porte qu'une seule valeur. De ce fait, une structure de trois couches est fondée pour supporter les trois valeurs associées aux couleurs. Plusieurs méthodes sont employées pour le stockage des données dans le mémoire. Certaines méthodes sont purement destinées pour le compactage des données alors que d'autres forment une structuration complète des données. Les plus utilisées parmi ces méthodes sont

Ligne	0	1	2	3	4	5	6	7	8	9
0	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	3	3	3	3	3	3
3	2	2	2	2	3	3	3	3	3	3
4	2	2	2	2	3	3	3	3	3	3
5	2	2	2	2	3	3	3	3	3	3
6	2	2	2	3	3	4	4	4	4	4
7	2	2	2	3	4	4	4	4	4	4
8	2	2	2	4	4	4	4	4	4	4
9	2	2	2	4	4	4	4	4	4	4

Ligne	Valeur	Long.
0	1	10
1	1	10
2	1	4
2	3	6
3	2	4
3	3	6
4	2	4
4	3	6
5	2	4
5	3	6
6	2	3
6	3	2
6	4	5
7	2	3
7	3	1
7	4	6
8	2	3
8	4	7
9	2	3
9	4	7

FIG. 2.25 – La méthode RLC (Run Length Coing). A gauche, une image en format BitMap. A droite, l'image en format RLC.

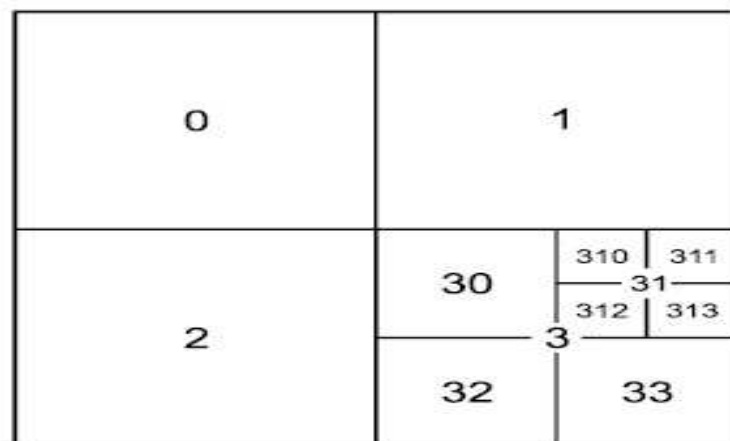


FIG. 2.26 – La méthode des blocs carrés appliquée sur la matrice de l'image précédente.

1. Bitmap,
2. RLC,
3. et Blocs carrés tétra-arbres.

Bitmap est une structure séquentielle simple. Une image est sauvegardée ligne par ligne, du pixel supérieur gauche au pixel inférieur droit. La taille de l'image sauvegardée est égale à « L x C x T » tel que « L » représente le nombre de lignes, « C » le nombre des colonnes et « T » le nombre d'octets selon le type des valeurs quantifiant un pixel (entier, réel ..etc.).

Run Length Coing (RLC) est un format compressé lors de l'enregistrement du fichier d'image. Les séquences des valeurs identiques dans l'image sont exploitées. Chaque ligne de l'image est codée en deux valeurs :

1. la valeur du pixel,
2. le nombre d'occurrences successives de cette valeur dans la ligne.

Le coefficient de compression est variable selon le nombre d'occurrences des pixels successifs ayant les mêmes valeurs. Cette compression est efficace si le nombre d'occurrences successives est supérieur à 2.

La figure 2.27 montre un exemple du codage d'une matrice d'image selon la méthode RLC.

Blocs carrés est une généralisation du concept valeur-pixel du codage utilisé dans RLC mais avec deux dimensions. Les blocs carrés de pixels ayant les mêmes valeurs sont exploités. La plus petite valeur du bloc est égale un pixel. Chaque bloc est identifié par trois attribut ;

1. les coordonnées du premier pixel (ligne et colonne) dans l'image
2. la longueur (en nombre de pixels) du bloc.
3. la valeur (qui est uniforme) du pixel du bloc.

Le compactage de la méthode du bloc carré est efficace si la longueur moyenne du bloc est supérieure à 2 pixels. La figure 2.26 illustre un exemple de la présentation d'une image sous forme des blocs carrés.

Tétra-arbres est une structure hiérarchique en arbre quatre-nœud. La méthode est basée sur une subdivision successive de l'image en 2^n partie, jusqu'à l'obtention des valeurs des pixels identiques. L'image est structurée en doublet.

1. L'ordre de la branche qui correspond à la partie d'image.
2. La valeur des pixels de la partie d'image.

L'ordre de la branche incluant une série du numéro dont les valeurs entre 0 et 3 de tous les nœuds de la branche. Un exemple expliquant l'utilisation de cette méthode est montré dans la figure 2.27. Les structures du mode raser sont employé pour les disciplines qui nécessitent des informations plus détaillée sur les attributs et leurs distribution spatiale. Elle est aussi utilisée dans le cas où des combinaisons complexes des attributs sont requises (analyse spatiale et modélisation). Les structures vectorielles sont plus destinées vers les

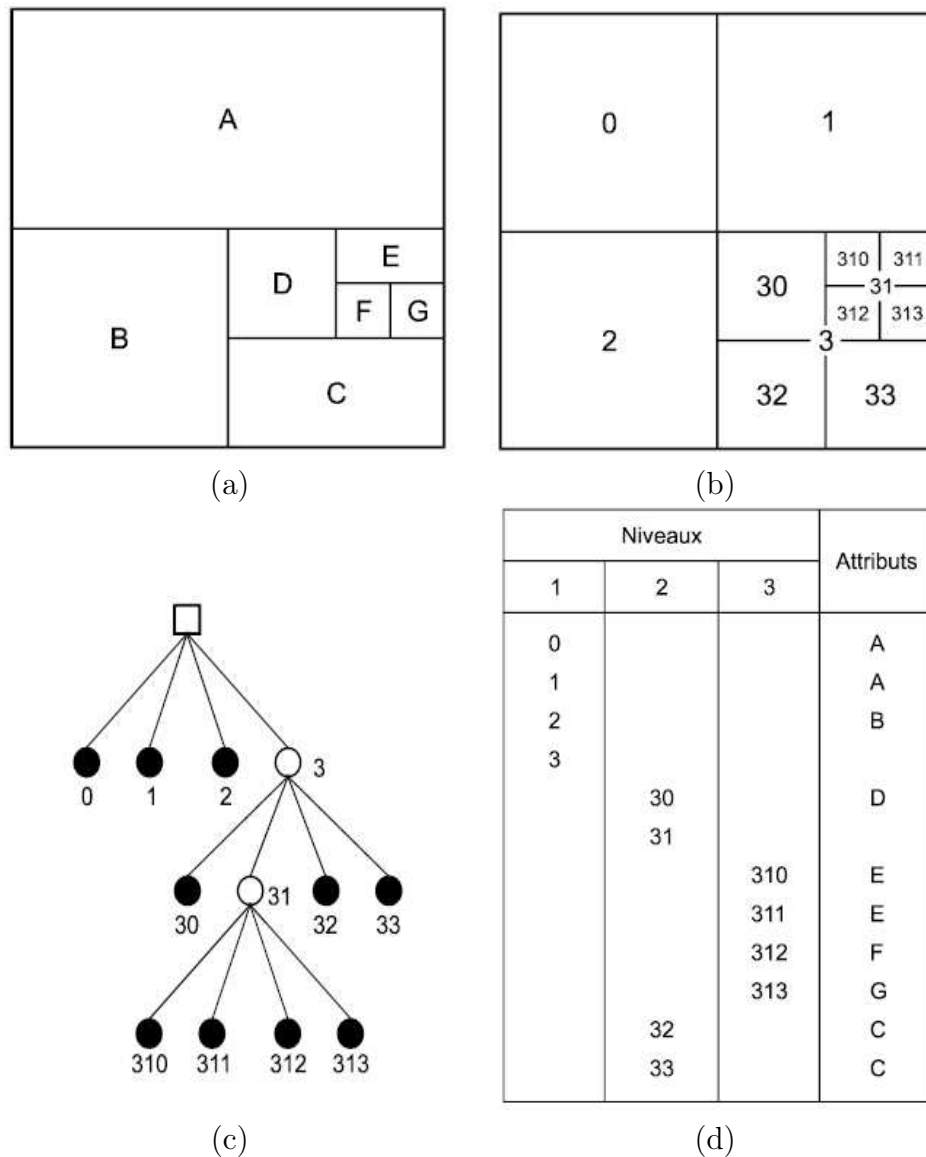


FIG. 2.27 – La méthode de Tétra-arbre. (a) La matrice des valeurs des bloc. (b) Codage des blocs. (c) représentation des blocs sous forme d'une arbre (la hiérarchie). (d) L'enregistrement de Tétra-arbre dans le mémoire.

domaines qui ont seulement un besoin d'une localisation spatial précise tel que le cadastre et la cartographie topologique. Pratiquement dans nos jours et avec le développement des applications, les deux modes sont souvent combinés. De ce fait, des structures multi couches est adopté incluant des structures à la fois raser et vectorielles.

2.7 Conclusion

Pour répondre aux besoins multidisciplinaires des systèmes spatiaux, plusieurs approches de plusieurs disciplines interviennent telles que l'imagerie, théories de graphe, la géométrie ..etc. De ce fait, l'acquisition, le stockage, l'archivage, l'analyse et l'édition des données spatiales sont conçus, d'une façon, à ce que le caractère de modularité de chaque partie de système est préservé le maximum. D'autre part, la plupart des concepteurs des SIG tirent profit d'autres systèmes afin de réduire le coût de développement alors que les recherches sur des techniques spécialisés pour le domaine spatial se continuent à nos jours.

Exercice 1

Soit la carte C suivante :

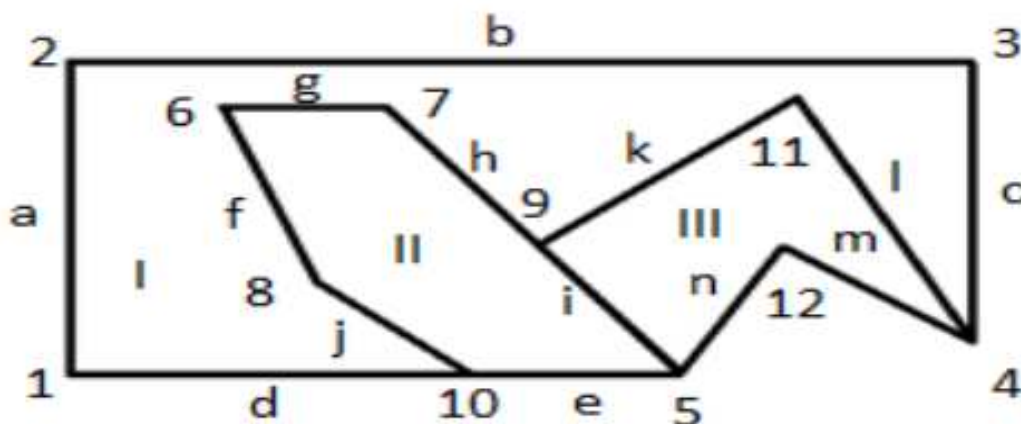


FIG. 2.28 – Carte C

- Donnez le model hiérarchique de la carte C .
- Donnez le model réseau de la carte C .
- Donnez le model relationnel de la carte C .

On supposant que chaque point possède des coordonnées en $2D$, modifiez le model relationnel en intégrant, pour chaque point, ses coordonnées.

Exercice 2

Soit la carte C de l'exercice 1.

- Donnez la structure séquentielle simple de cette carte.
- Quel est l'inconvénient majeur de cette approche ?
- Donnez une solution afin de remédier à cet inconvénient.

Solution

Exercice 1

Le modèle hiérarchique de la carte C est présenté dans la figure 2.29.

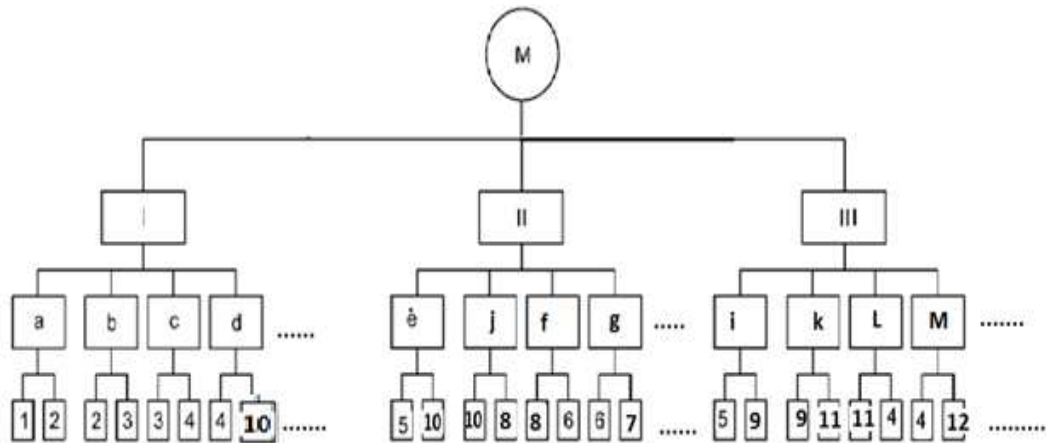


FIG. 2.29 – Le modèle hiérarchique de la carte C

Le modèle réseaux de la carte C est montré dans la figure 2.31.

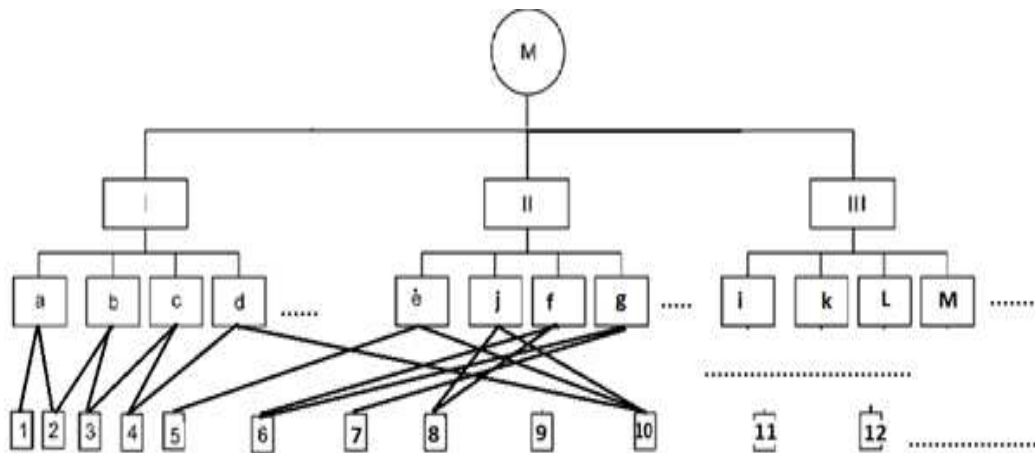


FIG. 2.30 – Le modèle hiérarchique de la carte C

Le modèle relationnel est constitué de trois tables.

1. table-Zone(id-zone, id-ligne);
2. table-ligne (id-ligne, point-1, point-2);
3. table-point(id-point, X, Y);

Cette dernière table permet de répondre à la deuxième partie de la question.

Exercice 2

La structure séquentielle simple de la carte.

X	Y
X1	Y1
X2	Y2
X3	Y3
X4	Y4
X11	Y11
X9	Y9
X7	Y7
X8	Y8
X10	Y10
X1	Y1
-1	-1
X5	Y5

X9	Y9
X7	Y7
X6	Y6
X8	Y8
X10	Y10
X5	Y5
-1	-1
X5	Y5
X12	Y12
X4	Y4
X11	Y11
X9	Y9
X5	Y5
-1	-1

FIG. 2.31 – La structure séquentielle simple de la carte C

L'inconvénient major de cette structure est la redondance de l'information. La solution est d'utiliser des structures topologiques telles que la structure POLYVRT.

Exercice 3

Soit la carte vectorielle C de la ville de Chlef décomposée en 05 zones :

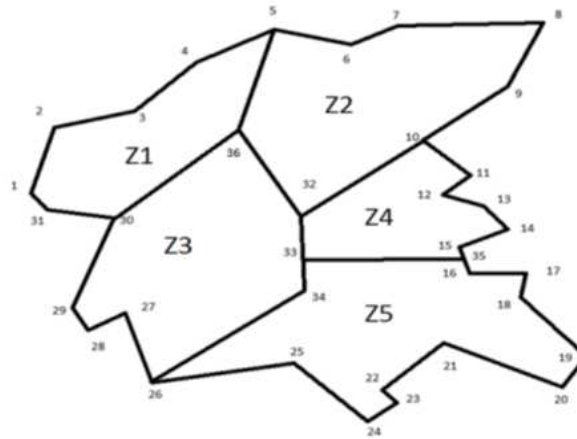


FIG. 2.32 – la carte C

1. Donnez la structure DIME de la carte C en donnant tous les fichiers nécessaires. (Notez que les points importants).
2. Donnez la structure POLYVRT de la carte C en donnant tous les fichiers nécessaires. (Notez que les points importants).

Exercice 4

1. Donnez la description par triangles et par sommets de la carte 2.33 :

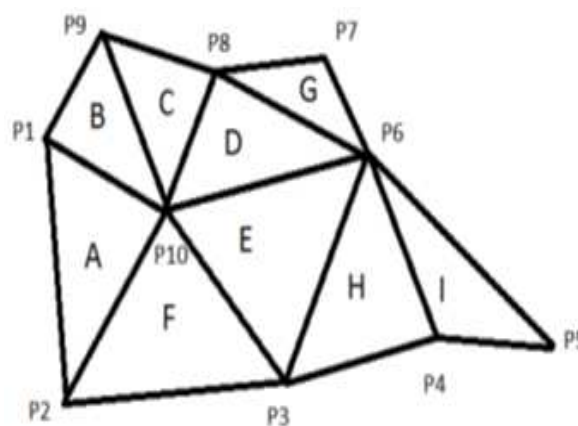


FIG. 2.33 – la carte C-2

2. Comment les triangles sont-ils construits ?

Solution de l'exercice 3

1. Les fichiers de la structure DIME de la carte C.

Etiquète Nœud	X	Y
N1	X1	Y1
N2	X2	Y2
N3	X3	Y3
.....
.....
N36	X 36	Y36

FIG. 2.34 – Fichier des noeuds de la carte C

Etiquète arrête	Neoud de départ	Nœud de destinati	Zone gauche	Zone droite
N(1-2)	N1	N2	-	Z1
N(2-3)	N2	N3	-	Z1
N(3-4)	N3	N4	-	Z1
.....
N(36-30)	N30	N36	Z1	Z3
.....
N(26-34)	N26	N34	Z3	Z5

FIG. 2.35 – Fichier des arrêtes de la carte C

2. Les fichiers de structure POLYVRT de la carte C.

Etiquète_zone	Liste des pointeurs de chaines
Z1	C(30 ,5)C(5,36)C(36,30)
Z2	C(5,10)C(10,32)C(32,36)
Z3	C(36,30) C(32,36) C(32,33) C(33,26) C(26,30)
Z4	C(32,33) C(33,35) C(35,26) C(35,10) C(35,10)
Z5	C(35,26) C(33,26) C(33,35)

FIG. 2.36 – Fichier des zones de la carte C.

Etiquète_Nœud	X	Y
N30	X30	Y30
N5	X5	Y5
N36	X36	Y36
N10	X10	Y10
N32	X32	Y32
N33	X33	Y33
N26	X26	Y26
N35	X35	Y35

FIG. 2.37 – Fichier des noeuds de la carte C

Etiquète chaine	Points intermédiaires
C(30,5)	X31,Y31,X1,Y1,X2,Y2,X3,Y3,X4,Y4
C(5,10)	X6,Y6,X7,Y7,X8,Y8,X9,Y9
C(10,35)	X11,Y11,X12,Y12,X13,Y13,X14,Y14,X15,Y15
C(35,26)	X16,Y16,X17,Y17,X18,Y18,X19,Y19,X20,Y20, X21,Y21,X22,Y22,X23,Y23,X24,Y24,X25,Y25
C(33,26)	X34,Y34
C(26,30)	X27,Y27,X28,Y28,X29,Y29

FIG. 2.38 – Fichier des points intermédiaires de la carte C

Chaîne	Nbre_point_int	Nœud_départ	Nœud_arrivé	Zone_gauche	Zone_droite
C(30,5)	5	N30	N5	-	Z1
C(5,36)	0	N5	N36	Z1	Z2
C(36,30)	0	N36	N30	Z3	Z2
C(5,10)	4	N5	N10	Z2	-
C(10,32)	0	N10	N32	Z3	Z4
C(32,36)	0	N32	N36	Z3	Z2
C(32,33)	0	N32	N33	Z3	Z4
C(33,26)	1	N33	N26	Z3	Z5
C(26,30)	3	N26	N30	-	Z3
C(10,35)	5	N10	N35	Z4	-
C(33,35)	0	N33	N35	Z4	Z5
C(35,26)	10	N35	N26	Z5	-

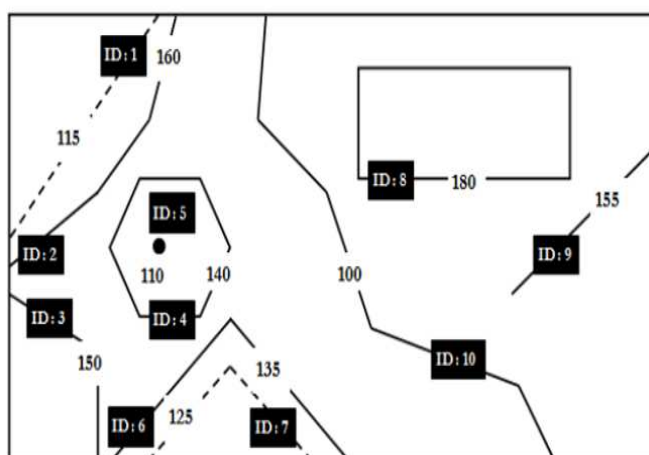
FIG. 2.39 – Fichier des chaines de la carte C

Exercice sur les structures altimétriques

Exercice 4

Soit la carte suivante en mode vecteur.

- Donnez la description complète des fichiers de la structure topologique pour les courbes de niveau. (Spécifiez que les points importants)



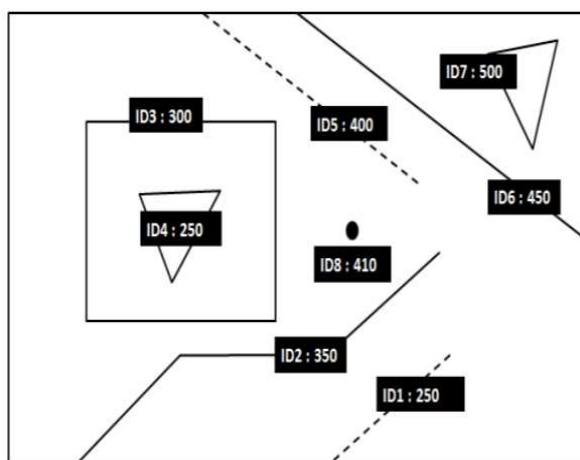
-Extrait d'une carte topographique 1/50.000 (équidistance : 10 m)

ID	Longueur (m)
1	3780
2	4521
3	2805
4	4320
5	0
6	5106
7	3860
8	9870
9	2087
10	10322

FIG. 2.40 – Carte des courbes de niveaux C-3

Exercice 5

Soit la carte suivante en mode vecteur.



- Extrait d'une carte topographique 1/50.000 (équidistance : 20 m) -

ID	Longueur (m)
1	1520
2	4570
3	6280
4	1930
5	1870
6	2310
7	1930
8	0

FIG. 2.41 – Carte C-4 des courbes de niveaux.

- Donnez la description complète des fichiers de la structure topologique pour les courbes de niveau. (Spécifiez que les points importants).

Solution de l'exercice 4.

Les fichiers de la structure POLYVRT de la carte C-3 sont les suivants.

ID	ALT	NP	APD	LON	FOR	CON	NVI	AVI	NVS	AVS
1	115	2	1	3580	O	I	----	----	1	1
2	160	4	3	4521	O	C	5	1	--	-----
3	150	3	7	2805	O	C	3	2	1	1
4	140	6	10	4320	F	C	3	3	2	2
5	110	1	16	0	P	P	----	----	1	3
6	135	3	17	5106	O	C	2	4	3	4
7	125	3	20	3860	O	I	-----	-----	1	5
8	180	4	23	9870	F	C	2	5	-----	-----
9	155	2	27	2087	O	C	1	6	1	6
10	100	6	29	10321	O	C	-----	-----	6	7

FIG. 2.42 – Carte des courbes de niveaux C-2.

Adresse	Liste	Adresse	Liste	Adresse	X	Y
1	1, 3, 4, 6, 10	1	2	1	X1	Y1
2	4, 6, 10	2	2, 3	.	.	.
3	6, 10, 5	3	4	.	.	.
4	7, 10	4	3, 4, 2	.	.	.
5	10, 9	5	6	34	X34	Y34
6	10	6	8			
		7	8, 9, 2, 3, 4, 6			

FIG. 2.43 – Les trois fichiers de la sstructure des courbes de niveaux. (a) Fichier des voisins inférieurs. (b) Fichier des voisins supérieurs. (c) Fichier des points.

On suit les mêmes étapes pour résoudre l'exercice 5.

3

Analyse des données géographiques

Sommaire

3.1	Introduction	58
3.2	Notion d'analyse des données spatiales	58
3.3	Types des données géographiques	59
3.4	Implémentation des SIG	63
3.5	Exemple d'application	66
	3.5.1 Les requêtes attributaires	68
	3.5.2 Les requêtes spatiales	68
3.6	Conclusion	69

3.1 Introduction

Les systèmes géographiques sont conçus de tel sort que ces systèmes soient capables de répondre aux requêtes relatives aux fonctions métier. Un SIG peut être considéré comme étant une extension des systèmes d'information afin de supporter le domaine spatial. Autrement dit, une donnée d'un SIG est modélisée en deux parties : une partie attributive et une partie spatiale. La partie attributive représente l'information sémantique et la partie spatiale formalise l'information géométrique. Entre les deux parties, des relations de couplage sont conçues afin de garder toute relation topologique entre les objets. Les fonctionnalités appliquées dans le mode spatial que ce soit en mode raster ou en mode vectorielle sont importantes. Dans ce chapitre, une introduction sur les principes de couplage entre les données attributaires et spatiales sera présentés.

3.2 Notion d'analyse des données spatiales

Une bonne partie d'effort de réalisation des SIG est fournie sur la réalisation des outils nécessaires pour effectuer une analyse spatiale. De ce fait, plusieurs définitions ont été proposées dans la littérature pour bien encadrer et préciser les objectifs et les notions de l'analyse spatiale. Les définitions suivantes sont adoptées.

Définition de Champoux et Bédard (1992) [5] : « c'est un processus cognitif et itératif qui utilise différents opérateurs, dont un spatial au moins, pour déduire les caractéristiques descriptives ou spatiales d'un phénomène isolé, réel ou simulé dans l'espace ».

Définition de Charre (1995) [6] : « rechercher, dans les caractères d'unités spatiales, ce qui relève de leur position géographique, en particulier relative, ce qui oblige à modéliser la structure spatiale ».

Définition de Brunet et al. (2009) [2] : « ensemble de méthodes mathématiques et statistiques visant à préciser la nature, la qualité, la quantité attachées aux lieux et aux relations qu'ils entretiennent -l'ensemble constituant l'espace- en étudiant simultanément attributs et localisations ».

Définition de Pumain et Saint-Julien (2010) [13] : « étude formalisée de la configuration et des propriétés de l'espace produit et vécu par les sociétés humaines ».

Un SIG doit permettre de faire une analyse spatiale sur les données et leurs attributs contenus dans la base donnée. Dans le tableau 3.1, une typologie générale sur les catégories des requêtes de SIG est présentée.

La conception de la base de données doit impérativement considérer les requêtes des utilisateurs. Nous avons déjà présenté dans le chapitre précédent les trois façons qui sont utilisées pour la réalisation d'un SIG. Majoritairement, les concepteurs profitent des SGBD actuels, soit en adopte une extension de ses SGBD, soit que les SGBD et les systèmes géométriques sont couplés pour réaliser un SIG. Les systèmes géographiques, qui sont complètement conceptuellement nouveaux, eux-mêmes tirent avantage des SGBD. De ce fait, une requête dans les SIG est souvent basée sur la logique d'algèbre relationnelle. Une requête SQL a la forme suivante

TAB. 3.1 – Typologie des requêtes d'analyse spatiale.

Exemple de requête	Traitement	Usage
Qui a-t-il à tel endroit ?	Localisation	Inventaire localisé
Où trouve-t-on tel genre de phénomène ?	Distribution	Analyse thématique
Qu'est-ce qui a changé depuis.. ?	Evolution	Analyse temporelle
Que se produirait-il si .. ?	Modélisation	Simulation des processus
Quelle est la meilleure façon de.. ?	Optimisation	Aide à la décision

SELECT < nom (s) de l'attribut(s) > **FROM** < tableau1,..ect > **WHERE** < condition de recherche >

Trois catégories des opérateurs de SQL peuvent être appliquées sur les attributs :

1. les opérateurs de comparaisons : <, >, =, ≥, ≤, <>.
2. les opérateurs arithmétiques : +, -, *, / (appliquer seulement sur les champs numériques).
3. les opérateurs booléens : and, or, xor, nor.

Soit l'exemple des deux tables relationnelles de la figure 3.1. La requête « Quel est le nom, l'adresse du propriétaire et la superficie de la parcelle 103 ? » est exprimée en SQL par :

```
SELECT  PARCELLE#, SUPERFICIE, NOM_PROP, ADRESSE
FROM   PARCELLES, PROPRIETAIRE
WHERE  PARCELLE = '103'
AND   PARCELLES.PARCELLE#=PROPRIETAIRES.PARCELLE#
```

Les paramètres après **SELECT** sont des attributs ou des fonctions appliquées sur les attributs tels que `sum(< nom de l'attribut >)` où les valeurs des attributs sont numériques, `mean(< nom d'attribut >)`, `max(< nom d'attribut >)`, `min(< nom d'attribut >)` ect. D'autres fonctions sont appliquées sur les attributs nominaux ou ordinaires telles que `count(< nom d'attribut >)`.

3.3 Types des données géographiques

Les données géographiques sont composées de deux parties fondamentales. Les données attributaires qui sont les données standards des systèmes des gestions des bases de données et les données spatiales sous les deux formats vectoriels et raster. Les bases de données spatiales représentent une extension de la base de données standard qui nécessite impérativement une extension structurelle et extension fonctionnelle. La figure 3.2 illustre un exemple des composants de l'information géographique. Tenant compte que la majorité des SGBD actuels supportent les données géographiques, les données standard non spatiales sont structurées sous forme des tables.

Parcelle#	Superficie (en m ²)	Périmètre (en m)	Parcelle#	Nom_prop	Adresse
101	10453.4	489.0	101	Mansour	Kénitra
102	4105.9	284.7	102	Benmalek	Casablanca
103	11943.8	625.3	103	Moukhtar	Casablanca
104	5457.4	361.0	104	Ijlal	Rabat
105	14456.1	584.9	105	Farah	Temara

Table parcelle

Parcelle#	Nom_prop	Adresse
101	Mansour	Kénitra
102	Benmalek	Casablanca
103	Moukhtar	Casablanca
104	Ijlal	Rabat
105	Farah	Temara

Table propriétaire

FIG. 3.1 – Exemple d’une base de données composée de deux tables : la table parcelle et la table propriétaire.

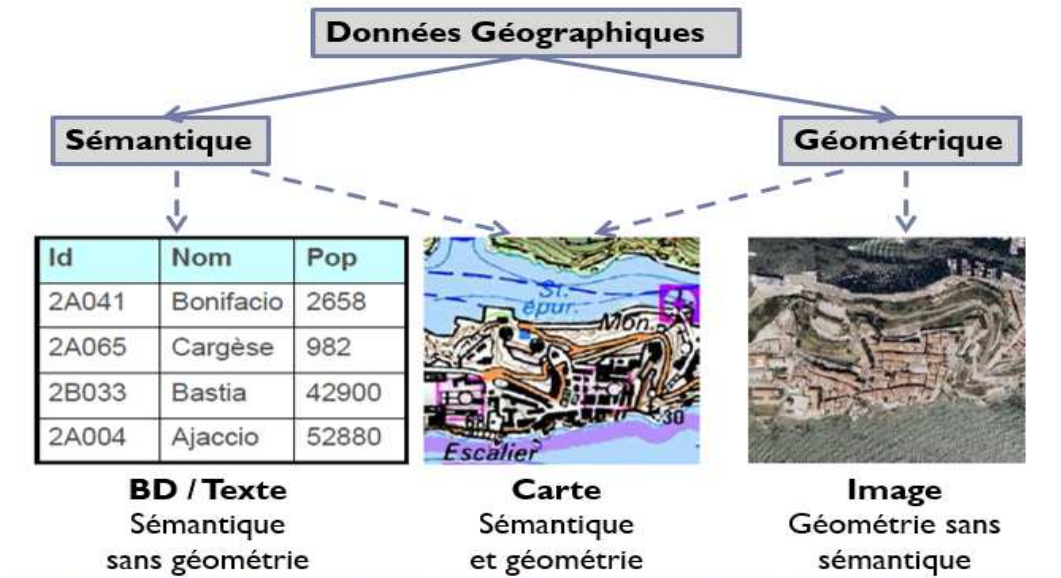


FIG. 3.2 – Les composantes de l’information géographique.

Les données spatiales raser sous forme des images et enregistrées sous forme des matrices multidimensionnelles. Les données vectorielles comme il est montré dans le chapitre 2 sont sous forme des fichiers et selon la structure utilisée, elle peut aussi être sauvegardée selon le modèle relationnel. L’information sémantique donne une description des objets

et aussi les relations qui existent entre les objets. Par conséquent, l'information sémantique peut être décrite par deux parties : la partie attributaire des objets et la partie des relations entre les objets. L'organigramme de la figure 3.3 illustre les deux parties de l'information géométrique.

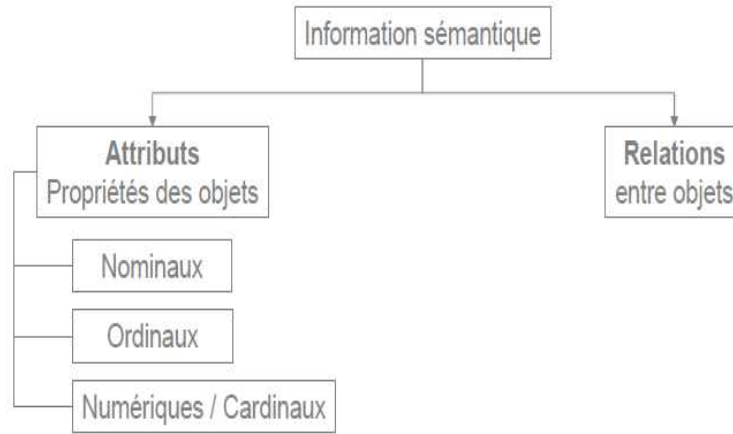


FIG. 3.3 – Les deux parties de l'information sémantique.

1. Les données attributaires peuvent être classées de trois types.

Toponyme : c'est les données nominales. Les zones géographiques sont reconnues par leurs noms souvent le cas par exemple Chlef, Oran, Istanbul etc. Parfois dans de grandes zones, des noms sémantiques sont affectés tels que « zone A93 ».

Donnée ordinaire : les données ordinaires sont utilisées pour décrire l'importance des zones tel que grand, petit, noire etc.

Données numériques : montre la quantification numérique de certaines caractéristiques telles que surface, population, densité des zones par des valeurs numériques, intervalles, ratioetc.

2. les relations sémantiques modélisent les liens entre les objets tels que « intersect avec », « est chef-lieu de ».

En effet, une base de données géographique est constituée d'un ensemble des tables avec des données attributaires ou relatives associées aux objets géoréférencés. Souvent, les données géométriques sont moins volumineuses que les données sémantiques. La figure 3.4 montre un exemple sur les deux types des données. Dans les SGBD étendus aux données spatiales, le modèle conceptuel de données (MCD) est basé sur le modèle entité-association.

Dans le cas où les données vectorielles sont constituées des zones ou poly lignes avec des segments droits, le modèle relationnel est constitué de quatre tables suivantes.

1. La table « segment » contient au moins un identificateur de l'arc.
2. La table point est composé au moins d'un identificateur de la point et ces coordonnées.

id	Nom	Superficie km2
2	Chlef	4 975

id	Longitude	latitude
2	1° 20' 12" E	36° 10' 26" N

FIG. 3.4 – deux tables représentant l'information sémantique et l'information géométrique du même endroit.

3. La table « poly-ligne » : une poly lignes est un ensemble des segments sans intersections. Les poly lignes ont la forme d'une zone ouverte. Une poly linge doit posséder au moins un identificateur.
4. Table polygone : un ensemble des lignes qui représente une zone fermée. Un identificateur d'un polygone est nécessaire.

Trois relations entre les trois entités sont nécessaires pour modéliser les relations topologiques entre les quatre unités spatiales.

« **est-début|fin** » : permet de modéliser la relation topologique entre un segment qui représente une ligne droite est les deux points constituant les limites du segment concernée. Un segment est constitué de deux et seulement deux points limites. Un point peut limiter un ou plusieurs segments à la fois. De ce fait, les cardinalités associées du côté de la table segment est « 2,n » et du côté de la table point : « 1,n ».

« **est-Elément** » : représente la relation composition entre un segment et une poly ligne. Une poly ligne est composée d'un ou de plusieurs segments droits.

« **est-Frontière** » : est une relation topologique qui modélise les frontières d'un polygone. Cette relation est associée à un segment et un polygone. Effectivement, cette relation peut être associé à un polygone et des poly lignes. Cependant, il est plus performant en terme de temps de calculs pour les requêtes géométriques (calcul de périmètre, surface ...etc.) de garder la première relation. En outre, les modèles standard tels que les structures POLYVRT et DIME sont conçues par la même relation ce qui facilite l'extensibilité et le couplage des systèmes spatiaux avec les SGBD.

La figure 3.5 montre le modèle relationnel correspondant aux segments droits. le modèle ci-dessous représente un modèle de base inspiré du modèle standard et théorie de graphe. En effet, les segments ne suivent pas toujours des lignes droites. Une amélioration du modèle précédent est effectuée. En effet, un segment est limité par deux points d'appui et une fonction associée à chaque segment pour représenter l'obliquité des segments. La figure 3.6 montre le modèle relationnel associé.

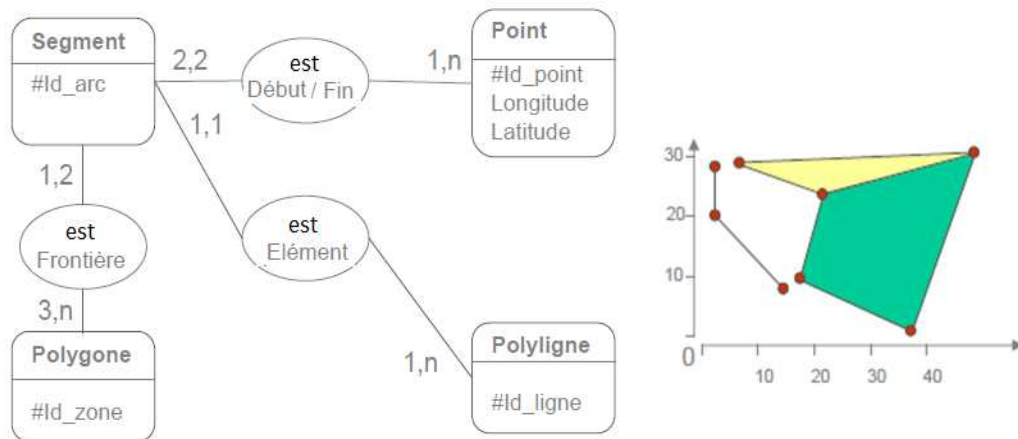


FIG. 3.5 – le modèle entité-association correspond aux segments droits.

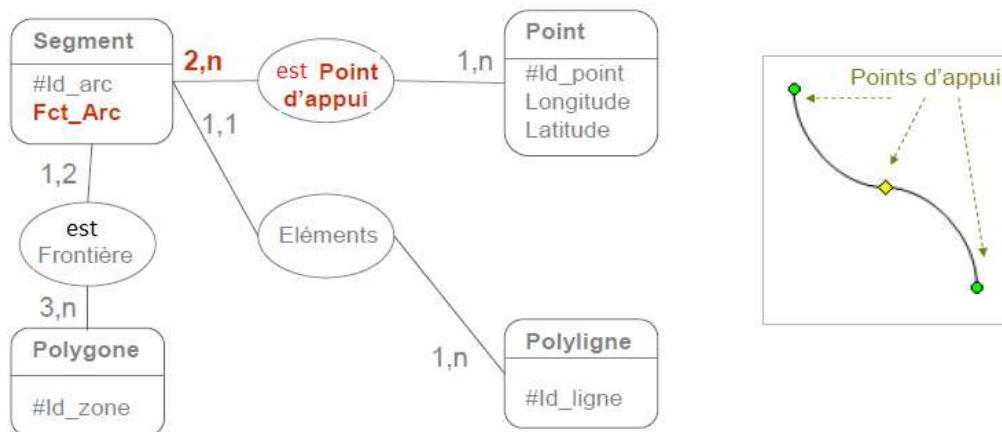


FIG. 3.6 – le modèle-entité association correspondant aux segments obliques.

3.4 Implémentation des SIG

Généralement, comme il est déjà cité dans le chapitre précédent, un SIG est réalisé par l'une des trois méthodologies suivantes.

1. Extension des systèmes de gestion de base de données (SGBD) : Les SGBD sont largement exploités dans cette méthodologie. Autrement dit, tous les efforts concernant la modélisation, la conception et la réalisation des outils spécifiques pour le traitement des données attributaire sont déjà fournis. De ce fait, il reste à concevoir et à implémenter la partie spatiale. En effet, le langage SQL est développé (PL-SQL) pour supporter de nouvelles fonctionnalités permettant d'étendre les SGBD attributaires pour supporter le domaine spatial.
2. Couplage entre les SGBD et les systèmes spatiaux : cette méthodologie est basée sur le couplage entre les SGBD et les systèmes géométriques. La partie attributaire

est traitée par les SGBD est traité par les SGBD alors que la partie géométrique est déléguée au système spatial. Le couplage entre les deux systèmes est fait via des intergiciels qui assure d'une façon transparente le déroulement des requêtes.

3. Conception de nouveaux systèmes : cette méthodologie conçoit complètement des nouveaux systèmes, autrement dit, les modèles de stockage et de traitement des données sont complètement nouveaux. L'exemple le plus connu sur ces modèles est le SQL3 qui est en cours de développement.

Pour les deux premières méthodologies, un SGBD-R garde l'information sémantique et les liens vers l'information géographique alors que le système d'information spatiale garde l'information géométrique des données géographique. Le modèle de la figure 3.7 illustre un exemple de couplage entre les deux parties. Le tronçon dans la figure qui représente une partie de la route est représenté géométriquement par une poly ligne. Pratiquement, un champ nommé « Obj » est utilisé pour accéder aux fonctions géométriques. Il représente un pointeur vers la structure géométrique et les fonctions associées.

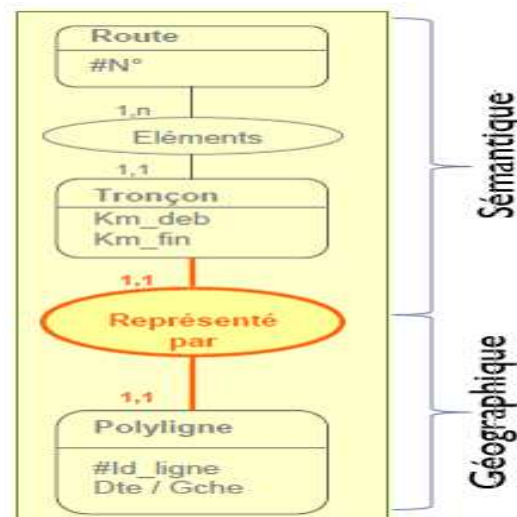


FIG. 3.7 – Exemple de couplage entre les SGBD-R et SI spatial.

Les fonctions spatiales peuvent être catégorisées en deux grands types.

1. Les fonctions en mode vectorielles : Les fonctions de base de cette catégorie s'appliquent sur les unités spatiales de base, à savoir le point, la ligne (mix ligne), les polygones et aussi sur les relations topologiques relatives à savoir,
 - (a) la relation point-point,
 - (b) les relation segment-segment, point-segment
 - (c) les relation polygon-point, ploygon-segment, polygon-polygon
2. Les fonctions en mode raser : Les fonctions de ce mode sont les fonctions issues du domaine de traitement d'image et la reconnaissance de forme. Les fonctions de l'imagerie les plus utilisées incluent « Fonction Locale », « Fonction Focale », et « Fonction Zonale ».

En plus de ces fonctions, d'autres fonctions spécifiques à un domaine bien défini peuvent être introduites telles que les courants maritimes ou les réseaux hydrauliques. Un réseau hydraulique est identifié via un mode raser (reconnaissance des formes), puis il est sauvegardé en mode vectorielle. Un réseau hydraulique verse dans un bassin. Le bassin peut être considéré comme une surface. Considérons les fonctions de base suivantes qui peuvent être injectées directement dans une requête spatiale.

1. Les fonctions « contenir (contains) ».
 - (a) **PointInZone(<zone>.obj, <point>.obj)** : c'est une fonction Booléen, c'est-à-dire qui comme sortie booléenne. Deux paramètres d'entrées, un pointeur d'objet d'une zone et pointeur d'un objet de type point. La fonction a comme sorti « true », si le point est à l'intérieur de la zone concernée.
 - (b) **LineInZone(<zone>.obj, <Line>.obj)** : cette fonction donne comme sortie la valeur booléenne « true » si la ligne concernée est entièrement dans la zone concernée.
 - (c) **ZoneInZone(<zone A>.obj, <zone B>.obj)** : le résultat de cette fonction est vrai « true », si la zone « B » est à l'intérieur de la zone « A ».
2. Les fonctions d'intersection
 - (a) **ZoneInZone(<zone A>.obj, <zone B>.obj)** : cette fonction donne comme résultat la valeur booléenne « true » si la zone « A » inclut la zone « B ».
 - (b) **IntersectZone(<zone A>.obj, <zone B>.obj)** : S'il existe une intersection entre les deux zones « A » et « B » alors la valeur résultante est « true ».
 - (c) **IntersectLineZone(<line>.obj, <zone>.obj)** : la valeur « true » est donnée par cette fonction s'il existe une intersection entre la ligne et la zone.
 - (d) **Adjacent(<zone A>.obj, <zone B>.obj)** : deux zones sont considérées comme adjacentes s'il partage la même frontière. La fonction « adjacent » peut aussi être définie entre une ligne et une zone. Les notions d'intersection, adjonction et contenir doivent être bien précisées. D'autres fonctions peuvent être facilement déduites par les fonctions définies ci-dessous. Par exemple deux zones, ligne, ou zone avec ligne sont disjointes si la fonction d'intersection donne « false ». Deux zones se touchent (l'intersection est un point seulement) si l'intersection est « true » et l'adjonction est « false ». Si on utilise la fonction « contains », on peut déduire si l'une des zones est à l'intérieur de l'autre zone.
3. Les fonctions de périmètre, surface, distance.
 - (a) **Surface (<zone>.obj)** : cette fonction retourne la surface d'une zone.
 - (b) **Intersection(<zone A>.obj, <zone B>.obj)** : cette fonction retourne une zone correspondant à la partie en commun entre les deux zones. Réellement, la zone retournée ne correspond pas toujours à une zone fermée, mais à une structure topologique contenant, selon le cas, des zones fermées, des lignes ou des points.
 - (c) **notIntersect(<zone A>.obj, <zone B>.obj)** : cette fonction retourne une zone (une structure topologique). Contenant les parties de la zone « A » qui

ne sont pas incluses dans la zone B. En résumé, la notion d'intersection peut être généralisée sur toutes les unités spatiales et les structures topologiques. De ce fait, dans les langages orientés-objet (et même dans les modèles relationnels), une classe père qui modélise structurellement et fonctionnellement la « structure topologique » est conçue pour généraliser la notion d'intersection.

- (d) **Distance(<point>.obj, <point>.obj)** : la fonction de la distance entre deux points dans les SIG considère la distance euclidienne. Bien que la notion de la distance est basée sur deux points, la distance entre une ligne et un point correspond à la distance calculée via le point de la droite le plus proche au point concerné, idem sur la distance entre deux lignes, deux zones, ou même deux structures topologiques différents.

3.5 Exemple d'application

Pour mieux comprendre l'utilisation des SGBD dans les SIG, considérant la base de données formée de quatre tables : Commune, Route, Bâtiment, Wilaya (voir les figures 3.8, 3.9, 3.10 et 3.11). Le modèle relationnel doit respecter au moins la première forme normale. Autrement dit, les attributs sont atomiques, l'ordre n'est pas important et pas de redondance. La table commune est constituée d'un ensemble des lignes ou chaque ligne représente une commune. Les attributs d'une commune sont les suivantes.

1. Id : l'« Id » est un attribut qui est considéré comme étant la « clé » des attributs de la table. Cette clé est atomique. Un clé d'une ligne correspond à une et une seule commune. Le modèle administratif adopté généralement pour la gouvernance d'un payé est hiérarchique. De ce fait, l'« Id » correspond dans notre exemple au code administratif de la Wilaya suivi du code de la commune.
2. Nom : Il correspond au nom de la commune.
3. Wilaya : cet attribut est le code de la wilaya dont la commune appartient.
4. Popu : il représente la population de la commune.
5. Obj : comme il est indiqué précédemment, c'est un pointeur vers la structure topologique qui représente géométriquement les caractéristiques spatiales de la commune. Ce pointeur permet aussi d'utiliser les fonctions spatiales associées à la structure topologique concernée. La commune est présentée par une zone (polygone).

id	Nom	Popu	Obj
6	Ouled fares	500 000	*
7	Aougrout	100 000	*

FIG. 3.8 – La table commune.

La table route sauvegarde les routes de la payé. En effet, toutes les routes de l'Algérie sont sauvegardées dans cette table. Cette table est constituée de 5 attributs.

1. Id : c'est l'identificateur de la route. L'identificateur des routes nationales commence par « N », des routes de la wilaya par « W » et communale par « C ».
2. Long : c'est la longueur en kilomètre de la route.
3. Type : il existe trois types des routes. Les routes nationales qui traverse aux moins trois Wilayas. Les Routes Wilaya qui traversent au moins trois communes de la wilaya et les routes communales qui se trouve au sein de la même commune. Trafic : il correspond au nombre moyen des véhicules qui circulent dans une unité de distance par unité de temps.
4. Obj : une route correspond géométriquement à une poly ligne. Donc, l'attribut obj pointe vers une poly ligne.

id	Long	Type	Trafic	Obj
N14	180	Nationale	4500	*

FIG. 3.9 – La table Route.

La table Bâtiment est associée aux bâtiments. Dans cet exemple, tous les bâtiments du payé sont sauvegardés dans cette table. Les attributs constituant la table Bâtiment sont les suivantes.

1. Id : il correspond à l'identificateur du bâtiment.
2. Type : il représente le type de bâtiment (école, parking ..ect).
3. Superficie : c'est la superficie du bâtiment.
4. Obj : c'est le pointeur vers la structure topologique d'un bâtiment. Un bâtiment peut correspondre à une zone ou un point.

id	Type	Superficie	Obj
17	Université	800	*
18	Bibliothèque	400	*

FIG. 3.10 – La table Bâtiment.

La table Wilaya représente toutes les Wilayas du payé.

1. Id : c'est l'identificateur de la Wilaya.
2. Nom : c'est le nom de la Wilaya.
3. Superficie : correspond à la superficie de la Wilaya.
4. Densité : le nombre des individus de la population dans le km^2 .
5. Nombre-daira : le nombre de Daira de la Wilaya.
6. Obj : le pointeur vers la zone correspondant à la wilaya. La zone dans ce cas, il représente un polygone (agrégation des polygones qui correspondent aux communes de la Wilaya).

Nom	Superficie	Densite	Nombre_Daira	Obj
Chlef	4 791	209	13	*
Laghouat	25 057	18	10	*

FIG. 3.11 – La table Wilaya.

3.5.1 Les requêtes attributaires

Les requêtes attributaires concernent les requêtes standard d'un SGBD. Les relations sémantiques entre les tables et les opérations standard sont utilisées pour filtrer des réponses exactes aux requêtes. Par exemple

1. Quel est la longueur de la route nationale N4 ?
Requête : `SELECT Long FROM routes WHERE id='N14'` ;
2. Quelle est le Trafic moyen des routes nationales ? Requête : `SELECT AVG(trafic) FROM Routes WHERE type = " nationale"`.
3. Quelle est la commune la plus peuplée de la Wilaya de Chlef ?
Requête : `SELECT Nom FROM communes WHERE Popu IN (SELECT Max(Pop) FROM communes WHERE ville = '02') AND ville= '02'` ;
4. Quelle est la population moyenne des communes dans chaque Wilaya d'Algérie ?
`SELECT Ville, AVG(Pop) FROM Communes WHERE GROUP BY ville` ;

3.5.2 Les requêtes spatiales

Quand il n'existe pas un lien relationnel entre les entités, les fonctions spatiales sont invoquées pour répondre aux requêtes. On parle dans ce cas sur les requêtes spatiales. Souvent, les requêtes spatiales sont plus couteuses en terme de temps de calcul. De ce fait, il est plus préférable parfois de sauvegarder certaines informations spatiales au lieu de les extraire via des calculs(par voie logicielle). Par exemple, il est préférable de sauvegarder les superficies des zones au lieu de les calculer chaque fois quand on utilise. Les requêtes suivantes sont des exemples de requêtes spatiales.

1. Quelles sont les routes qui croisent la route nationale « N4 » ?
Requête : `SELECT Id FROM Route1 AS Route, Route2 As Route WHERE Route1.Id= "N4" AND Route2.Id <> "N4" AND IntersectLineLine (Route1.obj , Route2.obj)`
Si la fonction booléenne d'intersection n'inclut pas la ligne avec elle-même, on peut supprimer la condition « `Route2 <> "N4"` ».
2. Quelle commune dispose d'une banque ?
Requête : `SELECT Commune.Nom FROM Commune , Bâtiment WHERE Contains (Commune.obj , Bâtiment.obj) and Bâtiment.type = "banque"`
3. Quelles sont les communes frontières de la Wilaya de Chlef ?
Requête : `SELECT Commune FROM Commune, Wilaya1 AS Wilaya, Wilaya2 AS`

Wilaya WHERE Wilayas1.nom = "Chlef" AND ZoneInZone (Wilaya1.obj , commune.obj) AND Adjacent(wilaya2.Obj , Commune.Obj) AND Wilaya1.Id <> Wilaya2.Id

4. Quelle sont les universités qui dispose d'une bibliothèque et d'un parking tel que le parking est loin de la bibliothèque de moins de 100 mètre ?
 Requête : SELECT Bâtiment.Id FROM Bâtiment1 AS Bâtiment, Bâtiment2 AS Bâtiment, Bâtiment3 AS Bâtiment WHERE Bâtiment1.type = "bibliothèque" AND Bâtiment2.type = "parking" AND Bâtiment3.type = "université" AND ZoneInZone(Bâtiment1.Obj , Bâtiment3.Obj) and ZoneInZone(Bâtiment2.Obj , Bâtiment3.Obj) and Distance(Bâtiment1.Obj , Bâtiment2.Obj) < 100.

3.6 Conclusion

Les systèmes spatiaux font intervenir les deux modes, le mode raster et le mode vectoriel. La modélisation des données géographique en mode vectoriel est basée sur les théories de graphe et la géométrie alors que le mode raster est basé dans sa totalité sur le traitement d'image. Dans ce chapitre, une introduction sur le couplage des données attributaires et spatiales est présentée. Bien que la troisième catégorie de la méthodologie de développement des systèmes géographiques est établie sur des notions considérées comme nouvelles, l'utilisation des concepts relative à la géométrie et à l'imagerie reste indispensable.

Exercice

Soit les tables d'une base de données SIG suivantes :

id	Type	Superficie	Obj
17	Université	800	*
18	Bibliothèque	400	*

FIG. 3.12 – Table Bâtiment

id	Long	Type	Trafic	Obj
N14	180	Nationale	4500	*

FIG. 3.13 – Table Route

id	Nom	Popu	Obj
6	Ouled fares	500 000	*
7	Aougrouit	100 00A0	*

FIG. 3.14 – Table Daira

Nom	Superficie	Densite	Nombre_Daira	Obj
Chlef	4 791	209	13	*
Laghouat	25 057	18	10	*

FIG. 3.15 – Table Wilaya

1. A quoi sert la colonne **Obj**?
2. Ecrire les requêtes suivantes (SQL Etendu) et dites s'il s'agit d'une requête attributive ou spatiale :
 - Quelles sont les Dairas de la Wilaya de Chlef
 - Quelle est la population totale de la Wilaya de Chlef?
 - Combien est le total des Dairas de l'Algérie?
 - Combien de routes nationales traversent la Wilaya de Chlef?
 - Quelle est la longueur de la Nationale 14?
 - Combien est la population de la Wilaya de Chlef?
 - Quelles Wilayas disposent d'une université?
 - Quelles sont les Dairas de la Wilaya de Chlef?

Solution

1. La colonne objet est un pointeur vers des fonctions qui sont exécutables par la partie spatiale du système et qui ne peut pas être exécutées par le SGBD traditionnel.

2. Les requêtes SQL3

(a) Les daïras de la Wilaya de chlef est une requête spatiale.

```
SELELCT Daira.Nom
```

```
FROM Daira, Wilaya
```

```
WHERE Wilaya.Nom = "Chlef" AND ZoneInZone(Wilaya.Obj, Daira.Obj);
```

(b) Nombre de population de la Wilaya de chlef est une requête spatiale.

```
SELECT SUM (Daira.Population)
```

```
FROM Daira, Wilaya
```

```
WHERE Wilaya.Nom = "Chlef" AND ZoneInZone(Wilaya.Obj, Daira.Obj);
```

(c) Le nombre total des Daira de l'Algérie est une requête attributaire.

```
SELELCT SUM (Wilaya.Nombre_Daira)
```

```
FROM Wilaya
```

(d) Les routes nationales traversant la Wilaya de chlef est une requête attributaire.

```
SELECT COUNT(route.id)
```

```
FROM Route, Wilaya
```

```
WHERE Route.type = "National" AND Wilaya.Nom = "CHLEF" AND IntersectLineZone(Route.Obj, Wilaya.obj);
```

(e) La longueur du national 14 est une requête attributaire.

```
SELELCT Rout.Long
```

```
FROM Route
```

```
WHERE Route.id = "N14";
```

(f) La population de la Wilaya de Chlef est une requête spatiale.

```
SELECT SUM(Daira.Popu)
```

```
FROM Wilaya, Daira
```

```
WHERE Wilaya.Nom = "Chlef" AND ZoneInZone(Daira.obj, Wilaya.Obj);
```

(g) Les Wilayas disposant d'une université est une requête spatiale.

```
SELECT Wilaya.Nom
```

```
FROM Wilaya, Batiment
```

```
WHERE Batiment.Type = "universite" AND ZoneINzone(Wilaya.Obj, Batiment.Obj);
```

Conclusion générale

Les systèmes d'information géographique jouent un rôle très important dans le développement des entreprises et dans la vie quotidienne des individus. Donc, il est utilisé davantage par le grand public. Pour répondre au besoin croissant des utilisateurs, les SIG ne cessent pas de se développer. Par conséquent, les développeurs se trouvent face à de nouveaux défis. Parmi ces défis, on cite

1. l'hétérogénéité des ressources humaines et matérielles : les SIG sont utilisés par des ingénieurs d'aménagement de territoire comme par transporteurs de marchandise dans leurs véhicules. D'un autre côté, un SIG ou un terminal d'accès à l'information géographique sont installés actuellement sur terminaux mobiles, tablette, véhicule, avions...etc.
2. les objectifs pour lesquels les SIG sont conçus : comme il est indiqué dans le premier chapitre, les objectifs d'un SIG sont nombreux et différents. Un SIG permet de déduire des statiques sur le développement urbain et économique d'un territoire comme il permet de faire la simulation sur l'emplacement d'un échangeur routier. Autrement dit, plusieurs disciplines sont intervenues pour la réalisation des SIG.
3. différents modèles de stockage, de présentation, d'analyse et d'implémentation : les SIG sont considérés comme des systèmes évolués. Donc, les modèles de conception, d'implémentation et de traitement de l'information géographique sont différents. Cette diversité nécessite des coûts et des efforts supplémentaires.

Par conséquent, nous avons insisté dans ce polycopié sur la structuration de l'information géographique et sur le mode de présentation vectorielle en donnant des aperçus sur le mode raster et sur l'emploi des autres disciplines dans les SIG.

Bibliographie

- [1] S. Aronoff. Ottawa, WDL Publications, 1989.
- [2] R. Brunet, R. Ferras, and H. Théry. *Les mots de la géographie : dictionnaire critique. Troisième édition.* Reclus, 2009.
- [3] L. M. Bugayevskiy and J. P. Snyder. *Map projections : a reference manual.* London : Taylor Francis, 1995.
- [4] P. A. Burrough and R. A. McDonnell. *Principles of Geographical Information Systems.* New York : Oxford University Press, 1998.
- [5] P. Champoux and Y. Bédard. Notions fondamentales danalyse spatiale et dopérateurs spatiaux. *Revue des sciences de linformation géographique et de lanalyse spatiale*, (2) :187–208, 1992.
- [6] J. Charre. *Statistique et territoire.* Reclus, 1995.
- [7] R. J. Clark. Habiline handexes and paranthropine pedigree at sterkfontein. *World archiology*, 1988.
- [8] M. Didier. *Utilité et valeur de linformation géographique.* Economica. Paris, 1990.
- [9] ESRI. Tout savoir sur les systèmes dinformation géographique. Technical report, Éditeur de logiciel de système d’information géographique (SIG), <https://www.esrifrance.fr/> (Accessed : 5 Mars 2018), 2018.
- [10] B. Huang. *Comprehensive Geographic Information Systems.* Elsevier, 2018.
- [11] G. Konecny. *Geoinformation : remote sensing, photogrammetry, and geographic information systems. Deuxième édition.* Taylor Francis, 2014.
- [12] K. Merry, P. Bettinger, M. Crosby, and K. Boston. *Geographic Information System Skills for Foresters and Natural Resource Managers.* Elsevier, 2023.
- [13] D. Pumain and T. Saint-Julien. *Analyse spatiale : Les localisations. Deuxième édition.* Armand Colin, 2010.
- [14] Y. Yamagata and H. Seya. *Spatial Analysis Using Big Data Methods and Urban Applications.* Elsevier, 2020.
- [15] Q. Zhou, B. Less, and G. Tang. *Advances in Digital Terrain Analysis.* Springer, 08.