
Utilisation des treillis de Galois pour l'extraction et la visualisation des règles d'association spatiales

R. Marghoubi^{1,2} - A. Boulmakoul¹ - K. Zeitouni³

¹ Faculté des Sciences Technique de Mohammedia (FSTM), B.P. 146 Mohammedia, Maroc.

² Agence Nationale de Réglementation des Télécommunications (ANRT), Complexe d'Affaires aile sud Ar-ryad, Hay riad 10100, BP 2939 Rabat Maroc

³ Laboratoire PRiSM, Université de Versailles, 45 avenue des Etats-Unis, 78035 Versailles Cedex France.

E-mail: {marghoubi@anrt.net.ma; boulmakoul2000@yahoo.fr, Karine.Zeitouni@prism.uvsq.fr }

Jeune chercheur

RÉSUMÉ. Dans ce papier nous présentons des solutions concernant le problème d'extraction de la connaissance spatiale à l'aide des règles d'association spatiales. La première étape est consacrée à l'élaboration du contexte. Une fois ce dernier est défini, nous exploitons le paradigme de Galois permettant ainsi de créer un treillis de concepts dans un contexte spatial. Ce treillis permet d'extraire les règles d'association. Afin de bien visualiser les règles d'association spatiale, nous avons utilisé la plate forme Galicia au contexte spatial. Dans ce travail, nous présentons aussi le prototype développé, il concerne l'application du data mining spatial à la promotion de l'utilisation d'Internet au Maroc ; plus précisément il se focalise sur la gestion des Services à Valeur Ajoutée que nous avons amélioré.

ABSTRACT. In this paper we present some solutions dealing with knowledge discovering problem by using spatial association rules. The first step will be reserved to elaborate a spatial context, once this last is defined; we will use the Galois Lattice paradigms. This lattice makes it possible to extract the closed spatial generators, and to deduce the spatial association rules. In order to visualize extracted spatial association rules, we will extend Galicia platform to spatial context. In this work, we present also the prototype relating to the promotion of the use of Internet in Morocco, and more precisely in the management of the Added value services which we improved.

MOTS-CLÉS : Data Mining Spatial, prédicats spatiaux, treillis de Galois, itemsets fermés fréquents, règles d'association spatiales, Services à Valeur Ajoutée., Galicia.

KEYWORDS: Spatial data mining, spatial predicates, Galois lattice; Frequent closed itemsets, Spatial Association Rules, Added value services, Galicia

1. Motivations et problématique

La découverte de la connaissance devrait être comprise comme découverte de l'information combinée avec la création de la connaissance. La création de la connaissance de l'information peut être favorisée par les représentations appropriées d'information qui rendent la structure logique inhérente de l'information transparente [13]. Puisque les concepts sont les unités de base de la pensée humaine et par conséquent des structures de base de la logique, la structure logique d'information est basée sur des concepts et des systèmes de concept. Par conséquent, les treillis de concept en tant qu'abstraction mathématique des systèmes de concept peuvent soutenir des humains pour découvrir l'information et puis pour déduire la connaissance. Dans ce contexte, nous utilisons les treillis de concept pour l'extraction de la connaissance spatiale, dans l'objectif de générer les règles d'association spatiale.

Le but de l'utilisation des règles d'association, comme méthode de fouille de données est l'identification des relations significatives entre les données de taille très importante. Dans le cas des bases de données transactionnelles, étant donné un ensemble d'articles, le but est de découvrir si l'occurrence de cet ensemble dans une transaction est associée à l'occurrence d'un ensemble d'articles. Par exemple, « 90% des clients qui achètent du thé achètent le sucre » est une règle d'association, sous la forme $\mathfrak{R} : \text{thé} \rightarrow \text{sucre}$, associant l'article thé à l'article sucre.

Dans le cas des bases de données géographiques ou spatiales, la nature et le volume de données dépassent les capacités humaines en terme d'analyse. D'où le besoin de l'utilisation des méthodes simples et faciles à comprendre, parmi ces méthodes on peut citer les règles d'association spatiale. Ces dernières, appliquées dans plusieurs domaines d'activités, présentent des résultats claires, faciles à interpréter. Cependant, deux problèmes majeurs liés à l'utilisation des règles d'association spatiales sont posés, à savoir le problème des temps d'extraction des règles d'association spatiales à partir du jeu de données et le problème de la pertinence et de l'utilité des règles d'association spatiales extraites.

Ce papier répond à deux objectifs essentielles, le premier relatif à l'amélioration du temps de calcul pour l'extraction des règles d'association spatiales, en utilisant le paradigme de Galois basé sur l'analyse formelle [13], permettant ainsi de créer un treillis de concepts dans un contexte spatial. Le deuxième concerne la visualisation des différentes règles d'association spatiale extraites à l'aide du treillis de Galois.

La suite de ce papier est organisée comme suit : dans la section 2, nous présentons un état de l'art sur les algorithmes d'extraction des règles d'association spatiale. Dans la section 3, nous proposons notre approche afin de réduire le temps de calcul et améliorer la visualisation des règles d'association spatiale extraites par l'extension de la plate forme Galicia pour le contexte spatial. La section 4, est consacrée au prototype développé. La section 5 est réservée à la conclusion et aux perspectives de notre travail de recherche.

2. Etat de l'art des algorithmes d'extraction des règles d'association spatiales

Les algorithmes d'extraction des règles d'association spatiales sont une extension des algorithmes des règles d'association, incluant la composante spatiale. A ce titre avant de présenter les travaux de recherches traitants les règles d'association spatiales, nous allons présenter un bref aperçu sur les algorithmes d'extraction des règles d'association.

Plusieurs travaux de recherches [1-6 ; 9 ;12] ont proposé des solutions pour les deux problèmes majeurs de l'extraction des règles d'association, à savoir :

1. le temps de réponse de l'extraction des règles d'association qui dépend principalement des temps d'extraction des itemsets fréquents, car plusieurs balayages du contexte doivent être réalisés. En comptant pour chaque itemset fréquent potentiel le nombre d'objets du contexte dans lesquels il est contenu.
2. l'utilité des règles d'association extraite, afin d'obtenir les règles d'association pertinentes à partir de celles qui sont extraites.

Les algorithmes d'extraction des règles d'association peuvent être classés en trois grandes catégories : l'extraction des fréquents, l'extraction des maximum et l'extraction des fermés. Les algorithmes de la première catégorie, sont basés sur la propriété d'antimonotonie :

1. tout sous ensemble d'un ensemble d'article fréquent est fréquent ;
2. tout sur ensemble d'un ensemble non fréquent est non fréquent.

Parmi les algorithmes de cette catégorie, on peut citer l'algorithme Apriori [1] dans lequel, deux paramètres, support minimum (minsup) et confiance minimale (minconf) sont introduits et deux phases principales résument l'algorithme :

1. la première phase a pour objet la recherche de tous les sous ensembles d'articles fréquents (frequent Itemset) tel que $\text{support}(X) > \text{minsup}$, autrement dit c'est la recherche des règles d'association utiles ;
2. la deuxième a pour objet la recherche de toutes les règles d'association intéressantes et pertinentes à partir des sous ensembles X fréquents déterminés lors de la première phase tel que $\text{confiance}(X) > \text{minconf}$.

Les algorithmes de la deuxième catégorie, extraction des maximums, sont basés, en plus de la propriété d'antimonotonie, sur la définition des itemsets fréquents maximum, c'est à dire les itemsets dont tous les sur-ensembles sont inférieurs, forment une bordure au dessous de laquelle tous les itemsets sont fréquents. Parmi les algorithmes qui utilisent cette approche, nous citons Pincer-Search, Max-Clique, Max-Eclat et MaxMiner. L'avantage de ces algorithmes par rapport à ceux présentés auparavant est qu'ils réalisent deux types de parcours : de bas en haut et de haut en bas pour la découverte des itemsets fréquents maximums lors de chaque itération. Ce nouveau mode de parcours anticipatif ou en avance de haut en bas, « look ahead » permet de réduire le nombre d'itérations et par conséquent de diminuer le nombre de balayage du contexte réalisé.

En ce qui concerne la dernière catégorie, relative à l'extraction des fermés. Elle est basée sur la fermeture de la connexion de Galois [10], les itemsets sont des itemsets fréquents qui sont fermés selon l'opérateur de fermeture de la connexion de Galois. Les itemsets fermés fréquents, selon cet opérateur de fermeture, forment un ensemble générateur non redondant minimal pour tous les itemsets fréquents et leurs supports. Tous les itemsets fréquents et leurs supports, peuvent donc être déduits efficacement, sans accéder au jeu de données. Cette propriété découle du fait que le support d'un itemset fréquent est égal au support de sa fermeture. Les itemsets fermés fréquents forment un treillis dont la taille est bornée par la taille du treillis des itemsets fréquents. Toutefois, en pratique, la taille de ce treillis est en moyenne bien inférieure à la taille du treillis des itemsets (treillis des parties).

Koperski et Han [7] ont prolongé la méthode de l'extraction des règles d'association aux bases de données spatiales. Une règle d'association spatiale est de la forme :

$$P1 \wedge P2 \wedge \dots \wedge Pn \rightarrow Q1 \wedge Q2 \wedge \dots \wedge Qm (s\%, c\%). \quad [1]$$

Où au moins un des prédicats $P1, P2, \dots, Pn$ et $Q1, Q2, \dots, Qm$ est un prédicat spatial, $s\%$ est le support de la règle et $c\%$ est la confiance de la règle.

Plusieurs types de prédicats spatiaux peuvent être utilisés au niveau des règles d'association spatiales. Ils peuvent être utilisés pour représenter des relations topologiques entre objets spatiaux, comme : disjoint, intersects, à l'intérieur de, contient, adjacent à, couvre, égal, ou pour représenter un ordre ou une orientation spatial, comme : à gauche, à droite, nord, sud, ou apporter une information de distance : près de, loin de, etc.

Deux méthodes ont traité la problématique d'extraction des règles d'associations spatiales à savoir la méthode de Koperski [8] et celle de Salleb [11]. Pour la première méthode, proposée par Koperski, elle se base sur des techniques de recherche, avec approfondissement progressif, de haut en bas (top-down), au niveau d'une hiérarchie de concepts. Ces techniques entament d'abord la recherche des patterns fréquents, pour le premier niveau de concept, ensuite, et uniquement pour chaque itemset fréquent, poursuivent la recherche à un niveau de concept inférieur. En ce qui concerne les attributs descriptifs, ils sont organisés sous forme de hiérarchies de concepts, l'extraction des règles d'associations est alors guidée par ces hiérarchies en parcourant chacune d'elles niveau par niveau. Néanmoins, les informations appartenant à différents niveaux des hiérarchies ne peuvent être apprises. De plus, l'utilisation des attributs non hiérarchiques n'est pas mise en évidence dans son algorithme. Salleb a fait une extension du travail de Koperski en considérant les attributs hiérarchiques et non hiérarchiques et en permettant un « mélange » des niveaux entre les différentes hiérarchies. L'algorithme ARGIS (Association Rules in GIS) proposé par Salleb a défini deux catégories de couches thématiques (couche de référence, couche descriptive). Cet algorithme est basé sur la notion de table de liens. Le principe de base de ARGIS est le suivant : étant donné une couche référence C_R , et un ensemble de couches descriptives C_D ; on cherche alors l'ensemble des règles d'associations spatiales décrivant les objets de la couche référence C_R par rapport à chacune des couches descriptives C_D . Autrement dit, les règles d'association découvertes dans ARGIS considèrent à chaque fois deux et seulement deux couches thématiques ; il n'est pas possible d'extraire des règles regroupant plus de deux couches thématiques.

3. Notre contribution

Les deux méthodes de Koperski et celles de Salleb (ARGIS), présentent les limitations suivantes :

1. Utilisation de l'algorithme d'Apriori pour l'extraction des règles d'association spatiales, cet algorithme est très coûteux en temps de calcul pour la première phase : plusieurs balayages de la base de données sont activés, il présente aussi le problème de stockage en mémoire.
2. L'algorithme ARGIS proposé par Salleb relative à la découverte des règles d'association spatiales considère deux et seulement deux couches thématiques à chaque comparaison, Autrement dit, il n'est pas possible d'extraire des règles regroupant plus de deux couches thématiques.

Dans ce qui suit, nous proposons une nouvelle approche pour l'extraction des règles d'association spatiales, cette approche se base sur les fondements mathématiques du concept de Treillis de Galois notamment la fermeture de la connexion de Galois dans un contexte spatial. Elle va nous permettre de réduire le temps de réponse par rapport aux deux méthodes présentés ci-dessus et ce, en

construisant l'ensemble des générateurs non redondant minimum des itemsets fermes fréquents.

Par ailleurs, et afin de résoudre le problème posé par ARGIS, concernant la prise en considération deux et seulement deux couches thématiques à chaque comparaison, nous allons utiliser un système de codification des objets spatiaux permettant de savoir la couche thématique de n'importe quel objet spatial et ainsi de trouver les relations spatiales y afférentes.

Pour la visualisation des règles d'association spatiales extraites. Nous avons utilisé la plateforme Galicia.

3. Prototype

4.1. Architecture du système

Dans cette partie nous allons décrire l'architecture de notre système. D'après la figure 1, quatre modules constituent l'architecture de ce système :

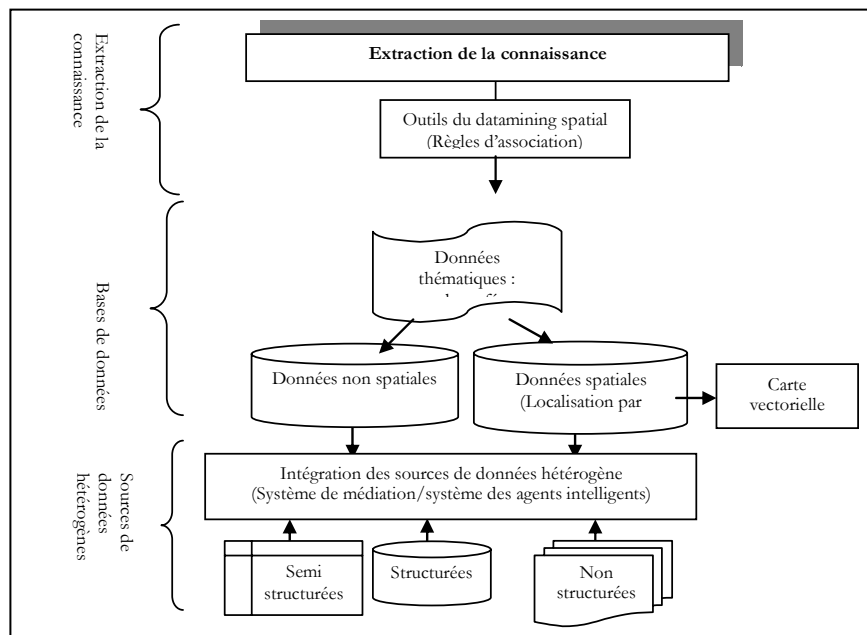


Figure 1. Architecture du système

Module 1 : Sources de données hétérogènes

Pour résoudre le problème de l'hétérogénéité des sources de données (données structurées, semi structurées, et non structurées), on peut utiliser l'un des systèmes basé sur le principe de médiation.

Module 2 : Module des bases de données

Ce module est constitué de différentes couches thématiques qui sont utilisées dans notre prototype et de la carte vectorielle de la ville de Mohammedia (Maroc).

Module 3 : Module d'extraction des connaissances

Une fois le contexte spatial est défini, nous avons implanté un algorithme basé sur l'opérateur de fermeture de Galois. Dans ce prototype, nous nous intéressons à l'extraction des règles d'association spatiales. Par exemple :

$\langle X, \text{cyber} \rangle$ et proche de $\langle Y, \text{maison de jeune} \rangle \rightarrow \langle \text{proche de, école} \rangle$ (40%,60%).

Module 4 : Visualisation des règles d'association extraites

L'objectif de ce module est la visualisation des différentes règles spatiales extraites à l'aide de la plate forme Galicia [14].

4.2. Préparation des données

L'étape la plus importante du processus de l'extraction des règles d'association spatiale est celle de la préparation des données. Cette étape doit être réalisée en prenant en considération l'objectif définit dans la première étape (du processus de l'extraction de la connaissance). Dans notre cas rappelons que l'objectif de cette extraction est la promotion de l'Internet au Maroc.

Selon cet objectif cinq couches thématiques (sources de données) sont traitées:

- source de données pour la gestion des services à valeur ajoutée, dont le service Internet (Cyber,...) ;
- source de données pour la gestion des établissements d'enseignement (Ecoles...)
- source de données pour la gestion des hôtels ;
- source de données pour la gestion des centres de culture et loisir (maisons de jeunes, cinéma, ...);
- source de données pour la gestion des services téléphoniques (téléboutiques, ...).

Chaque objet de ces sources de données possède des attributs spatiaux et des attributs non spatiaux. Le géocodage des adresses postales des différents objets a pris en charge par le SIG MapInfo (figure. 2)

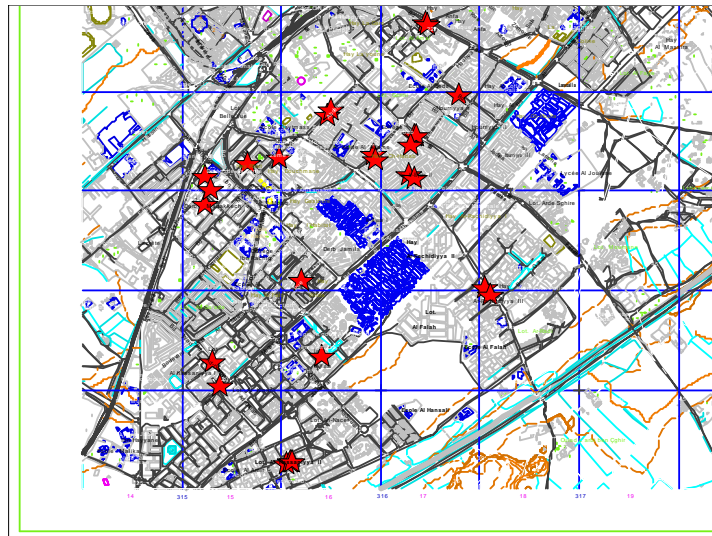


Figure 2. Répartition des objets spatiaux dans la ville de Mohammedia

La figure 3 présente le scénario pour l'élaboration du contexte spatial de notre prototype. Les tables physiques représentant les couches thématiques sont des tables extraites des bases de données de production.

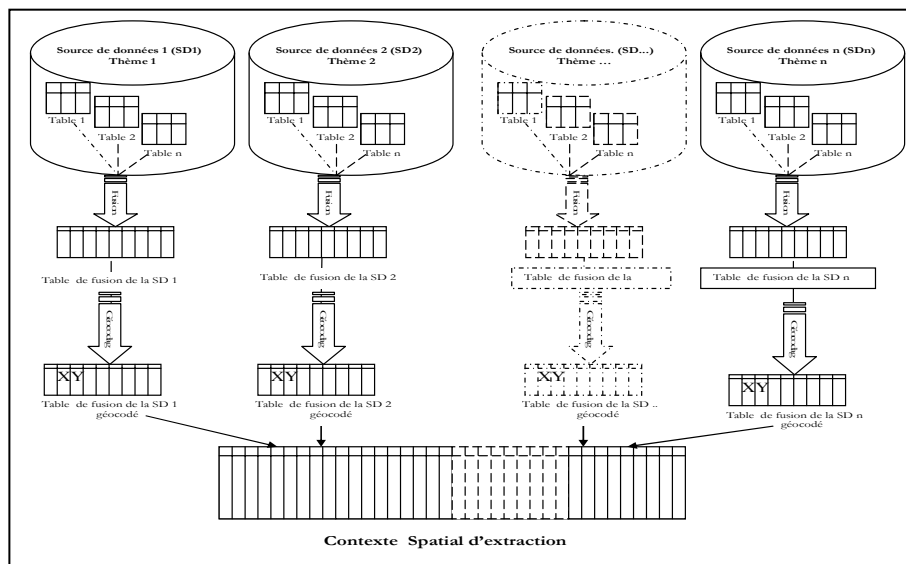


Figure 3. Elaboration du contexte spatiale

Le contexte spatial d'extraction représente

1. le résultat de calcul des relations spatiales de chaque objet spatial, de chaque couche thématique, avec les autres objets spatiaux des différentes sources de données (couches thématiques).
2. l'ajout des autres attributs non spatiaux correspondant à l'objet en question.

Exemple 1:

Table de fusion de la source de données 1 : Soit T_c la table qui représente la fusion de la source de données de la gestion des services à valeur ajoutée et plus précisément, le service Internet (catégorie Cyber).

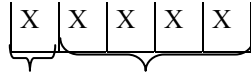
Table de fusion de la source de données 2 : soit T_e la table qui représente la fusion de la source de données des établissements d'enseignement et plus précisément les écoles.

Table de fusion de la source de données 3 : soit T_h la table qui représente la fusion de la source de données des centres médicaux et plus précisément les hôpitaux.

Table de fusion de la source de données 4 : soit T_{cl} la table qui représente la fusion de la source de données des centres de cultures et loisirs.

Table de fusion de la source de données 5 : soit T_t la table qui représente la fusion de la source de données du service téléphonique (catégorie téléboutique).

Les objets des tables de fusion sont codés de la manière suivante :



Code du thème Code de l'objet

Le premier caractère est réservé au code de la couche thématique, les autres sont réservés au code de l'objet. La codification des différentes couches thématiques est représentée dans la table 1.

Afin d'élaborer le contexte spatiale d'extraction, on doit d'abord déterminer les prédicats spatiaux (table 2), selon le choix de l'utilisateur, ensuite on doit calculer les distances de chaque objet spatial avec tous les autres objets spatiaux des différentes couches thématiques.

id_theme	libelle_theme
1	Cyber
2	Ecole
3	Hôpital
4	Culture et Loisir
5	Téléboutique

Table 1. Table de codification de thèmes

id_predicat	libelle_predicat
1	Proche de
2	Contient
3	Loin de

Table 2. Table de codification des prédicats spatiaux

4.3. Détermination des prédicats spatiaux

La détermination des prédicats spatiaux pour les différentes couches thématiques se fait de la manière suivante : en donnant la possibilité à l'utilisateur de choisir les distances entre les différentes couches thématiques et d'affecter le prédicat spatial correspondant.

Exemple 1 :

Si la distance entre les objets spatiaux de la couche thématique « Cyber » et les objets spatiaux de la couche thématique « Ecole » égale 20 m alors le « Cyber » est Proche de «Ecole ».

La table 3 représente quelques exemple de règles pour la détermination des relations spatiales entres les différentes couches thématiques.

id_theme1	id_theme2	distance	id_predicat	id_operateur
1	2	20	1	<=
1	5	100	1	<=
2	4	100	1	<=
3	4	0	2	=
4	5	250	1	<=

Table 3. Détermination des prédicats

Si on a « m » couche thématique est « k » prédicat spatial alors le nombre de combinaison des différentes couches thématiques avec les différents prédicats spatiaux est :

$$k \times \left[\sum_{i=1}^m (i) \right] - m = k \times \left[(m \times (m + 1) / 2) \right] - m \quad [2]$$

Dans notre cas et puisqu'on a 5 couches thématiques et 3 prédicats spatiaux, alors on a 40 relations spatiales entre les 5 couches thématiques.

4.4. Déterminations des relations métriques (basé sur la distance)

Le calcul des différentes distances se fait de la manière suivante : pour chaque objet spatial on va calculer sa distance avec les autres objets, dont les indices sont supérieure à l'indice de l'objet choisi. L'indice représente le numéro d'ordre de l'objet dans la table regroupant les différents objets des différentes couches thématiques.

Distance (id_objet[i], id_objet[i+1]) ; ou i est l'indice de l'objet varie de 1 jusqu'au n

Le nombre de relations entre les différents objets spatiaux de la table 4 est :

$$\sum_{i=1}^n (i) = n \times (n + 1) / 2 \quad [3]$$

Exemple 2 :

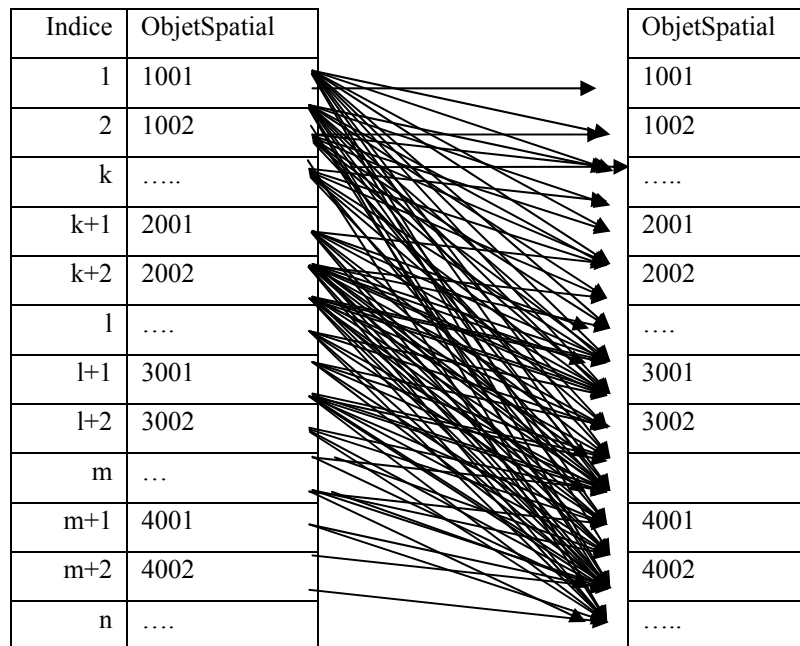


Table 4. Calcul des relations spatiales entre les différentes objets

Une fois les distances sont calculées, on va associer à chaque objets spatiales, dans un premier temps, les prédicats correspondants et ce, en faisant un parcours de la table contenant les relations des différentes couches thématiques (table 3). Ensuite on va choisir à partir des différentes sources de données les attributs non spatiaux répondant à l'objectif défini par le décideur. (Première étape du processus de la fouille de données).

Exemple 3 :

Si la distance entre Cyber 1013 et Ecole 2054 égale 20m alors le Cyber 1013 est Proche de l'Ecole 2054

Si la distance entre Cyber 1013 et la maison de jeune 4093 égale à 10 m alors le cyber 1013 est à coté de la maison de jeune 4093.

Et si on suppose que le Cyber 1013 possède les propriétés suivantes :

- Type d'accès : ADSL ;
- Tarifs appliqués au usagers : 5 DH /h
- Nombre de postes connectés : 10 postes.

Les propriétés de l'objet Cyber 1013 dans le contexte spatiales sont comme suit :

Id_objet 1	Id_prédicat	Id_objet2	Type d'accès	Tarifs
1013	1	2054	ADSL	5

Table 5. *Les propriétés spatiales et non spatiales de l'objet cyber 1013*

L'exemple de l'enregistrement de la table 5 signifie que le cyber 13 qui a comme type d'accès ADSL et les tarifs appliqués aux usagers 5DH est proche de l'école 54.

Cette approche doit être faite pour chaque objet de chaque couche thématique.

Dans notre prototype, nous avons 173 objets spatiaux, soit 15051 relations spatiales entres les objets spatiaux

4.5. Diagramme de class

La figure 4 montre les différentes classes principales qui sont implémentées dans notre prototype pour l'extraction des règles d'association spatiales.

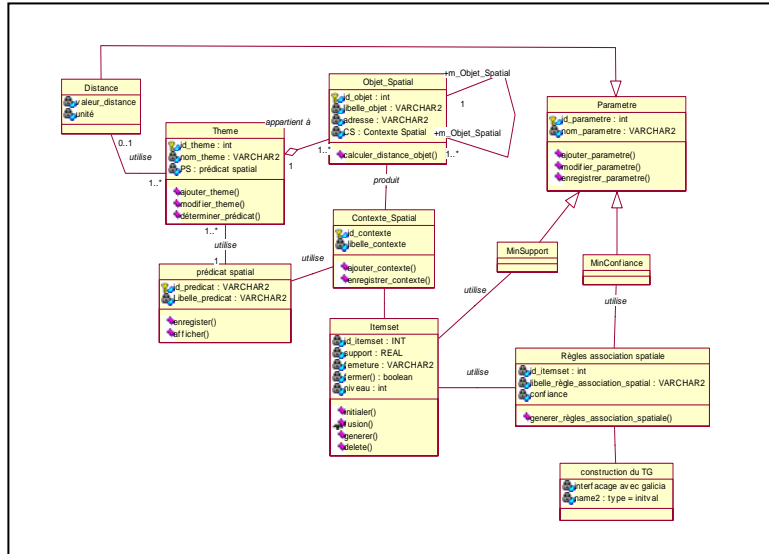


Figure 4. Diagramme de classe

4.6. Algorithme d'extraction des règles d'association spatiales

L'algorithme proposé dans notre approche peut être résumé en quatre étapes :

1ère étape : préparation du contexte spatial (Cs)

- détermination des prédicats spatiaux ;
- calcul des relations spatiales entre les différents objets spatiaux ;
- affectation de chaque prédicat spatial à la relation spatiale calculée à l'étape précédente

2ème étape : génération des Générateur Fermés spatiales GFs (Cs, support minimal)

3ème étape : extraction des règles d'association spatiale (GFs, confiance minimale)

4ème étape : visualisation de l'ensemble des générateurs fermés spatiale à l'aide de la plateforme galicia

code_generateur	support	fermeture	niveau_generateur
a	0,34	aci	1
af	0,33	acfi	2
acf	0,33	afc	3
c	0,45	ci	1
cf	0,44	cfi	2
f	0,44	fi	1
i	0,46	i	1
l	0,31	l	1
lq	0,3	lq	2
q	0,3	q	1

Table 6. Extrait de quelques générateurs fermés spatial (support = 30%)

Où les éléments [a,c,f,i,l q] représentent respectivement les relations spatiales suivantes : (cyber,C) \wedge (proche du Cyber,C') ;(cyber,C) \wedge (proche de, Ecole E) ; (cyber, C) \wedge (contient téléboutique, T) ; (cyber,C) \wedge (type acces,ADSL) ; (école, E) \wedge (proche de , Ecole,E') ; (école, E) \wedge (proche de , centre culture et loisir) ;

code_generateur	fermeture	support	Code_règle	libelle	confiance
a	aci	0,34	0	a-->ci	0,76
c	ci	0,45	1	c-->i	0,98
f	fi	0,44	2	f-->i	0,96
af	acfi	0,33	6	af-->ci	0,73
cf	cfi	0,44	7	cf-->i	0,96

Table 7. Extrait de quelques règles spatiales (confiance = 70%)

La table 7 représente un extrait de quelques règles d'association spatiales avec une confiance de 70%.

La règle (cf $\rightarrow_{(44\%,96\%)}$ i) signifie que 96% des cybers qui sont proche des écoles ont comme type d'accès ADSL.

4.7. Visualisation des générateurs fermés spatiaux à l'aide du Treillis de Galois

La figure 5 montre une représentation visuelle de quelques résultats de notre prototype, quant à la figure 6, elle montre d'une manière trop détaillée des relations qui existent entre les objets spatiaux et les prédicats spatiaux et non spatiaux..

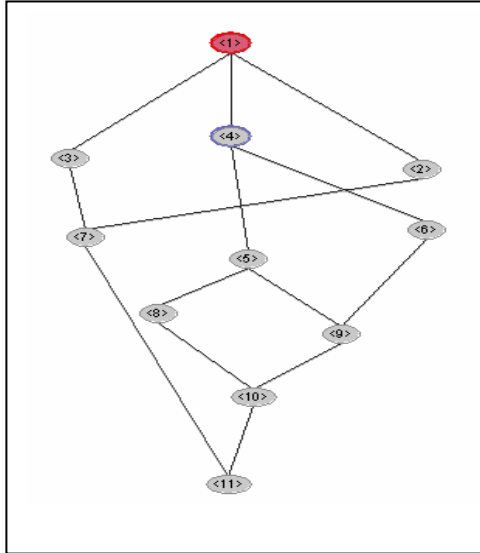


Figure 5. Visualisation de quelques générateurs spatiaux

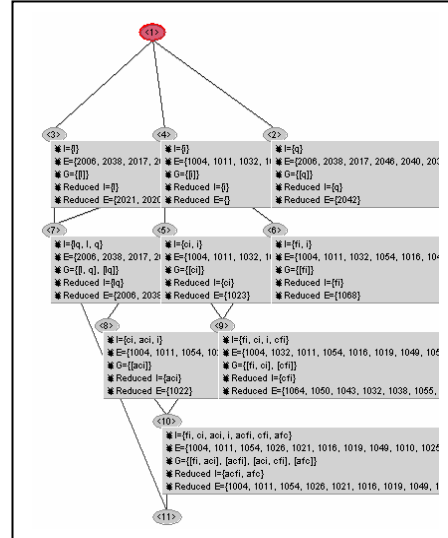


Figure 6. Visualisation détaillée de quelques générateurs spatiaux

D'après les tests de performance, nous pouvons dire que, lorsqu'il s'agit d'un jeu de données faiblement corrélées, l'algorithme Apriori donne un temps d'exécution relativement faible, mais lorsqu'il s'agit d'un jeu de données denses et corrélées, les performances de l'algorithme Apriori se dégradent considérablement, l'utilisation de l'algorithme close, basé sur la fermeture de la connexion de Galois est une alternative très importante dans le cas d'un jeu de données denses et corrélés, En plus cet algorithme a permet de réduire considérablement l'espace mémoire.

4. Conclusion et perspectives

Dans ce papier, nous avons essayé de résoudre quelques problèmes relatives à l'extraction des règles d'association spatiale et ce, en utilisant le paradigme de Galois et plus précisément, la fermeture de la connexion de Galois permettant ainsi de créer un treillis de concepts dans un contexte spatial. Ce concept nous a permis de réduire le temps de calcul et l'espace mémoire. Nous avons aussi présenté les résultats de notre prototype sous forme de règles d'association spatiale et sous forme d'un treillis aidant, ainsi les décideurs de bien visualiser les résultats obtenus.

Comme perspective de ce travail de recherche, et afin de sélectionner les règles d'associations spatiales pertinente répondant à l'objectif défini par le décideur, nous proposons d'intégrer les techniques de l'analyse multicritère. Ces techniques vont permettre au décideur d'associer à son objectif d'extraction de la connaissance des sous critères.

5. Bibliographies

- [1] Agrawal R. and al “*Mining Association Items in Large Databases*”. In Proc. 1993 ACM-SIGMOD Int. Conf. Management of Data, Washington, DC, 1993, pp. 207-216.
- [2] Agrawal R. and al “*Fast Algorithms for Mining Association Rules*”. In Proc. 20th VLDB Conference. Santiago, Chile, 1994, pp. 487-499.
- [3] Fu Y. and Han J. “*Meta-rule-guided Mining of Association Rules in Relational Databases*”. In Proc. 1st Intl. Workshop on Integration of Knowledge Discovery with Deductive and Object-Oriented Databases (KDOD'95), Singapore, 1995, pp. 39-46.
- [4] Han J. and Fu Y. “*Discovery of Multiple-level Association Rules from Large Databases*”. In Proc. 1995 Int. Conf. Very Large Data Bases. Zurich, Switzerland. 1995. pp. 420-431.
- [5] Kamber, M. and al “*Meta rule-Guided Mining of Multi-Dimensional Association Rules Using Data Cubes*”. In Proceeding of the 3rd Int. Conf, on Knowledge Discovery and Data Mining. Newport Beach, California. 1997, pp. 207-210.
- [6] Klemettinen M and al “*Finding Interesting Rules from Large Sets of Discovered Association Rules*”. In Proc. 3rd Int. Conf. Information and Knowledge Management. Gaithersburg. Maryland. 1994, pp. 401-408.
- [7] Koperski K. «*A progressive refinement approach to spatial data mining*». M.Sc, Warsaw University of Technology. 1988
- [8] Koperski K. and Han J. «*Discovery of Spatial Association Rules in Geographic Information Databases*». In Proc. 4th Int'l Symp. on Large Spatial Databases (SSD'95), pp. 47--66, Portland, Maine, August 1995
- [9] Ng R. T and al “*Exploratory Mining and Pruning Optimizations of Constrained Associations Rules*”. In Proc. of 1998 ACM-SIGMOD Conf. on Management of Data. Seattle, WA, 1998, pp. 13-23.
- [10] Pasquier N. “*Extraction de bases pour les règles d'association à partir des itemsets fermés fréquents*” Proceedings of the INFORSID'2000 conference, pp 56-77, may 2000.
- [11] Salleb A. et Vrain C.. «*An application of association rule Discovery to geographic information systems*» . In : PKDD'2000, 4th European Conference on Principles and Practice of Knowledge Discovery in Databases. Lyon, France, Septembre 2000.
- [12] Tsur D. and al. “*Query Flocks : A Generalization of Association-Rule Mining* “ In Proc. of 1998, ACM-SIGMOD Conf. on Management of Data, Seattle, Washington. 1998, pp. 1-12.
- [13] Wille D. “*Why Can Concept Lattices Support Knowledge Discovery in Databases?*” ICCS'01 Int'l. Workshop on Concept Lattices-based KDD 2001
- [14] <http://www.iro.umontreal.ca/~galicia/>