Cliques maximales d'un graphe et treillis de Galois

Michel Crampes - Michel Plantié - Julien Bidault

Laboratoire de Génie Informatique et d'Ingénierie de Production (LGI2P)
EMA - Site EERIE, Parc Scientifique Georges Besse
F-30 035 Nîmes cedex 1
{Michel.Crampes, Michel.Plantié, Julien Bidault}@mines-ales.fr

RÉSUMÉ. *L'énumération des cliques maximales d'un graphe non orienté est un sujet récurrent de la recherche sur les graphes et de l'analyse de réseaux (sociaux, biologiques, etc.). Elle est connue pour être un problème NP-Complet. Par ailleurs la visualisation de ces cliques maximales est peu étudiée. Nous montrons qu'il existe une bijection entre les cliques maximales d'un graphe et un sous ensemble des concepts d'un treillis de Galois dont il est dérivé. Les concepts de ce sous ensemble ont pour propriété remarquable que leur intention est égale à leur extension. Partant de ce résultat nous pouvons d'une part calculer les cliques maximales d'un graphe à partir du treillis de Galois obtenu en utilisant des outils de construction de ces treillis et les visualiser ensuite avec leur relations sur un diagramme de Hasse .*

SUMMARY. *The enumeration of maximal cliques of an undirected graph is a recurring theme of graph research and network analysis of (social, biological, etc.). It is known to be NP-complete. Moreover, visualization of these maximal cliques is little studied. We show that there is a bijection between the maximal cliques of a graph and a subset of the concepts of a Galois lattice from which it is derived. The concepts of this subset have a remarkable property: their intent is equal to their extent. From this result we first compute the maximal cliques of a graph from the Galois lattice issued via tools building these lattices and then view their relations in a Hasse diagram.*

MOTS-CLÉS : *Treillis de Galois lattices, graphes, réseaux sociaux, cliques maximum.*

KEYWORDS: *Galois lattices, graphs, social networks, maximum cliques.*

1. Introduction

De nombreux réseaux peuvent être représentés par un graphe non orienté. Il en est ainsi par exemple des réseaux sociaux où la relation entre deux individus peut être représentée par l'arête d'un graphe dont les sommets sont les individus. De même en biologie l'interaction entre gènes, métabolites (Kose et al 2001) ou protéines . (Byong-Hyon Ju et Kyungsook Han, 2003) peut être représentée par des arêtes d'un graphe. Une clique d'un graphe est un sous ensemble de sommets tous en relation deux à deux. Elle définit un sous graphe à n sommets et n(n-1)/2 arêtes. La recherche des cliques d'un graphe représentant un réseau est d'un intérêt majeur puisqu'elle consiste à rechercher les sous-groupes d'individus tous en relation. Plus particulièrement on s'intéresse aux cliques maximales qui contiennent tous les individus en relation les uns avec les autres. Il est donc particulièrement utile de pouvoir énumérer toutes les cliques maximales d'un graphe et en particulier la ou les cliques maximum qui sont les cliques du graphe de cardinalité maximale. Ce problème est un sujet récurrent de la recherche sur les graphes et de l'analyse de réseaux (sociaux, biologiques, etc.). Il est connu pour être NP-Complet. Par ailleurs la visualisation de ces cliques maximales est peu étudiée. Nous montrons qu'il existe une bijection entre les cliques maximales d'un graphe et un sous ensemble des concepts d'un treillis de Galois dont il est dérivé. Les concepts de ce sous ensemble ont pour propriété remarquable que leur intention est égale à leur extension. Partant de ce résultat nous pouvons d'une part calculer les cliques maximales d'un graphe à partir du treillis de Galois dérivé à l'aide d'outil de construction de ces treillis et les visualiser ensuite avec leur relations sur un diagramme de Hasse. La suite de l'article est organisée de la manière suivante. La section 2 situe l'état de l'art en matière de visualisation de cliques maximales d'un graphe. Le treillis de Galois dérivé d'un graphe est défini dans la section 3. Les principaux théorèmes sont posés et démontrés dans la section 4. La visualisation des cliques maximales d'un graphe par treillis de Galois est développée dans la section 5. La section 6 introduit une étude partielle de la complexité comparée entre recherche de cliques maximales et calcul du treillis de Galois dérivé. La conclusion discute les perspectives de recherche.

2. Etat de l'art

Si la visualisation de graphes a fait l'objet de très nombreux travaux de recherche en particulier dans la communauté Graph Drawing (Battista et al.,1999)(Herman et al 2000), la visualisation des cliques maximales d'un graphe a fait l'objet de moins d'attention. L'approche la plus immédiate consiste à réorganiser le graphe pour faire apparaître les cliques. C'est ainsi que (Dongbo Bu et al, 2003) partent d'un graphe d'interactions entre protéines et identifient les cliques par analyse spectrale pour mieux dévoiler visuellement les cliques ou quasi-cliques en jouant sur la disposition et la coloration des sommets. Les autres approches opèrent des

transformations. La représentation en étoiles dont les centres sont les cliques permet de réduire le nombre de sommets. Par exemple (Byong-Hyon Ju et Kyungsook Han, 2003) proposent une disposition en étoile d'un graphe dérivé d'un graphe d'interaction entre protéines. C'est avec la même perspective que (Brandenburg,1998) s'intéresse au cas particulier de cycles de cliques. (Kose et al 2001) explorent des représentations sous forme de tableaux à deux entrées, les sommets du graphe et les cliques ; des réorganisations et des couleurs permettent de repérer les appartenances. La représentation matricielle est aussi utilisée par (Henry Riche N. and Fekete JD 2010) pour visualiser les groupes et les interactions entre groupes dans les réseaux sociaux. Dans la mesure où une clique peut être considérée comme une hyper-arête d'un hypergraphe et plus généralement que l'ensemble des cliques maximales d'un graphe simple peuvent être représentées par un hypergraphe simple (Berge 1973) certains auteurs, en étudiant les représentations d'hypergraphes indirectement proposent des solutions partielles pour les cliques maximales d'un graphe. Par exemple (Henry Riche, 2010) explorent une représentation par boites de diagrammes d'Euler ; mais cette dernière approche trouve sa limite dans le fait que même un hypergraphe avec peu d'hyper-arêtes est en général non planaire et que la représentation devient alors très difficile à lire. Pour sortir de cette difficulté, la représentation d'un hypergraphe par un graphe support est étudiée aussi dans (Brandes et al 2010) dans la perspective de l'utilisation d'une métaphore dite du métro. L'intérêt pratique reste à démontrer. Dans ce même article, une représentation d'un hypergraphe par un diagramme de Hasse est aussi étudiée, mais sa présentation est limitée à la hiérarchie portant sur une seule hyper-arête. De plus si les hyper-arêtes d'un hypergraphe peuvent représenter les cliques maximales d'un graphe, ces modèles perdent l'information de l'interconnexion précise des cliques au travers des arêtes entre les sommets du graphe. Quand le graphe commence à être complexe, il ne reste guère qu'une représentation des cliques maximales par un graphe biparti (V, C, E) où V est l'ensemble des sommets du graphe origine, C l'ensemble de ses cliques maximales et E les arêtes reliant les nœuds des deux ensembles. Le dessin devient cependant vite confus et ne présente pas les relations entre les cliques. Dans le présent article nous chercherons à représenter les cliques maximales d'un graphe simple non orienté à l'aide d'un treillis de Galois dérivé de la topologie du graphe. Contrairement aux articles cités ci-dessus, le diagramme de Hasse qui en découle conserve toute l'information du graphe. Un diagramme réduit est aussi proposé pour simplifier cette vue tout en conservant l'information du graphe initial.

3. Treillis de Galois dérivé d'un graphe non orienté

3.1. Définitions et formalismes

Un graphe G est défini comme une paire (V,E) où V est un ensemble fini de sommets et E un ensemble fini d'arêtes. Les éléments de E sont de la forme $(v_i,v_j) \in$

V^2. Dans cet article nous ne considérons que des graphes simples (il n'existe au plus qu'une arête entre deux sommets et il n'existe pas de boucles, c'est-à-dire de sommets reliés à eux-mêmes) et non orientés : $(v_i,v_j)\equiv(v_j,v_i)$. La matrice d'adjacence d'un graphe G est la matrice carrée A de dimension $|V|^2$ dont les éléments a_{ij} sont tels que $a_{ij}=1$ si $(v_i,v_j) \in E$ et $a_{ij}=0$ si $(v_i,v_j) \notin E$. Si $a_{ij}=1$, v_i et v_j sont dits adjacents. Une clique dans G est un sous-ensemble de sommets de G tous reliés deux à deux : $C = (V_c, E_c)$, $V_c \subseteq V$, $E_c \subseteq E$ et $\forall (v_i,v_j) \in V_c{}^2$, $(v_i,v_j) \in E_c$. Une clique C est maximale s'il n'existe pas d'autre clique qui la contient : C est maximale $\Leftrightarrow \neg\exists$ C', $V_c \subseteq V_{c'}$. Dans la suite nous désignerons par C_{max} une clique maximale de G. Nous désignons par Δ_G l'ensemble des cliques maximales d'un graphe G : $\Delta_G=\{C_{max}\}$. Une clique est maximum (nous la noterons par C_{mum}) si sa cardinalité est un maximum parmi les cardinalités des cliques maximales : $C_{mum} \Leftrightarrow |V_{Cmum}| = \max(|V_{Cmax}|)$.

La définition des treillis de Galois qui suit est extraite de notre article (Crampes et al 2009) dans lequel plusieurs exemples sont présentés auxquels le lecteur pourra se référer. Un treillis de Galois est un ensemble d'objets dotés de propriétés (ou attributs) organisés en un treillis de concepts. Un concept contient l'ensemble des objets qui possèdent un même sous-ensemble de propriétés, en retenant le plus grand des sous-ensembles de propriétés communes entre les objets (Ganter B. & Wille, 1999). Le processus d'organisation débute avec la construction d'un contexte formel (nous le désignerons par T), ou plus simplement contexte, qui est une table avec les objets disposés en lignes et les propriétés en colonnes. Chaque case est marquée (par exemple avec la valeur 1) si l'objet en ligne possède la propriété en colonne; elle n'est pas marquée (évaluée à 0) dans le cas contraire. Formellement, un contexte est un triplé (O,M,B) où O est un ensemble d'objets, M un ensemble de propriétés, et B une relation binaire entre les objets et les propriétés, i.e. $B \subseteq O \times M$.

La construction du treillis de concepts se poursuit avec la recherche des concepts. Un concept est défini par une paire de sous-ensembles : un sous-ensemble d'objets, appelé **l'extension** du concept, et un sous-ensemble de propriétés appelé **l'intension**, qui est le sous-ensemble maximal des propriétés que les objets du concept partagent. Pour un ensemble d'objets $W \subseteq O$ et un ensemble de propriétés $H \subseteq M$, on définit l'ensemble des propriétés communes aux objets de W : $f : 2^O \rightarrow 2^M$, $f(W) = \{h \in H \mid \forall w \in W, (w, h) \in B\}$ et l'ensemble des objets qui ont toutes leurs propriétés dans H : $g : 2^M \rightarrow 2^O$, $g(H) = \{w \in W \mid \forall h \in H, (w, h) \in B\}$. La paire (f, g) est une correspondance de Galois entre $(2^G, \subseteq)$ et $(2^M, \subseteq)$. Un concept formel du contexte (O, M, B) est une paire (W, H) avec $W \subseteq O$, $H \subseteq M$ et $H = f(W)$ et $W = g(H)$. Après l'identification des concepts, le but suivant est de construire le treillis dont les éléments sont les concepts. Nous définissons un ordre partiel dans l'ensemble des concepts, tel que toute paire de concepts a un supremum (le plus petit majorant commun aux deux concepts), et un infimum (le plus grand minorant commun). Formellement, soit L l'ensemble des concepts de (O, M, B) et soit $\leq_L$, l'ordre partiel défini par : $(W_1, H_1) \leq_L (W_2, H_2) \Leftrightarrow H_1 \subseteq H_2 \Leftrightarrow W_2 \subseteq W_1$. La paire $(L, \leq_L)$ est appelée le treillis des concepts de (O,M,R). Un treillis de Galois peut être

représenté par un diagramme de Hasse ; c'est un graphe orienté dont l'orientation est implicite en raison de la disposition de haut en bas. On trouvera des exemples de diagramme de Hasse dans notre article (Crampes et al 2009). C'est un graphe dont les nœuds sont les concepts, ordonnés de haut en bas, selon leur ordre dans le treillis. Chaque concept montre son extension et son intension. Les arêtes du graphe relient les concepts qui sont dans une relation d'ordre directe, sans concept intermédiaire.

3.2. Treillis de Galois dérivé d'un graphe

Soit un graphe simple non orienté G=(V,E) ayant A pour matrice d'adjacence. Nous nous proposons de le transformer en un treillis de Galois G'=(U,F) à U sommets et F arêtes de contexte T pour en révéler les cliques maximales et parmi celles-ci les cliques maximum. Dans ce but nous posons T= A+I avec I la matrice identité de même dimension et de même rang que A. Du point de vue du graphe, cela revient à rajouter une boucle à chaque sommet. Du point de vue du treillis de Galois, on a un contexte dans lequel les propriétés sont aussi les objets. Un objet a un autre objet pour propriété si cet autre objet est lui-même ou si cet autre objet lui est adjacent dans G. Nous appelons dérivation *d* cette transformation. Formellement, en considérant que V' est la recopie de V : *d* : G=(V,E) → G'=(U,F) est le treillis de contexte (O,M,B), O = V, M = V', T =A+I. Le tableau 1 montre un exemple de graphe à 6 sommets ; dans la première case la matrice d'adjacence du graphe, au milieu le dessin du graphe (Figure 1) et à droite le contexte du treillis de Galois dérivé. La figure 2 montre le diagramme de Hasse associé au treillis de Galois de ce contexte dérivé. Il a été construit avec l'outil Lattice Miner (Lahcen and Kwuida, 2010). La Figure 3 montre ce même treillis reconstruit à l'aide du logiciel In-Close pour l'identification des concepts à partir du contexte dérivé, suivi de la construction du treillis et d'un diagramme de Hasse réalisée avec notre grapheur Molage (Crampes et al 2006) qui fait apparaître les cliques maximales comme indiqué dans la section suivante.

	1	2	3	4	5	6
1	0	1	1	1	0	0
2	1	0	1	0	1	0
3	1	1	0	0	0	1
4	1	0	0	0	0	0
5	0	1	0	0	0	0
6	0	0	1	0	0	0

Matrice d'adjacence

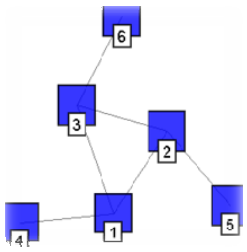

Figure 1 – Graphe de la matrice d'adjacence

	a	b	c	d	e	f
1	1	1	1	1	0	0
2	1	1	1	0	1	0
3	1	1	1	0	0	1
4	1	0	0	1	0	0
5	0	1	0	0	1	0
6	0	0	1	0		1

Matrice du contexte dérivé

Tableau 1. Graphe et contexte du treillis de Galois dérivé

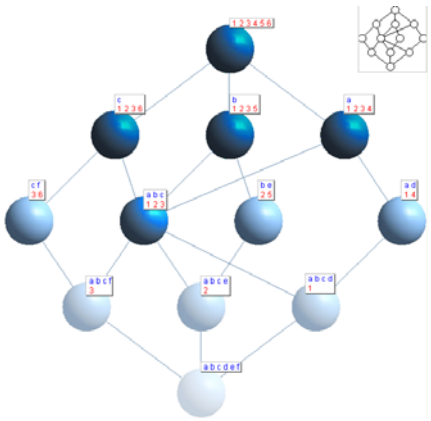

Figure 2 – avec Lattice Miner

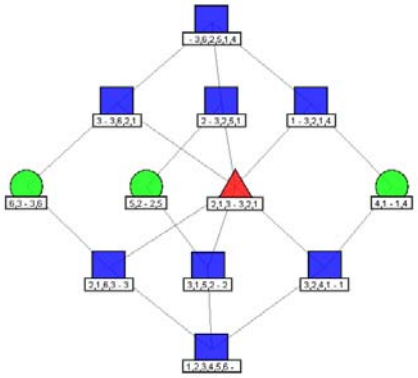

Figure3 - avec Molage, cliques max apparentes

Tableau 2. – Treillis de concepts associé au contexte du Tableau 1

4. Propriétés du treillis de Galois dérivé

Le treillis de Galois ainsi dérivé d'un graphe présente plusieurs propriétés intéressantes qui pourront être mises à profit pour visualiser les cliques maximales de G. Nous appelons concepts miroirs d'un treillis de Galois dérivé d'un Graphe les concepts qui ont la propriété particulière d'avoir leur extension identique à leur intension : (W,H) ∈ U est un concept miroir ⇔ W = H. Soit U_m l'ensemble des concepts miroirs de G' : $U_m = \{(W,H) \in U \mid W = H\}$

4.1 Théorème des concepts miroirs :

Nous énonçons la première propriété sous forme d'un théorème :

Théorème : Il y a une bijection entre les cliques maximales d'un graphe G et les concepts miroirs du treillis de Galois G' dérivé de G (ceux dont l'extension est égale à l'intension). Formellement :

$\forall C_{max} \in \Delta_G \; \exists! \; (W,H) \in U_m \mid W = V_{Cmax}$

et $\forall (W,H) \in U_m \, , \; \exists! \; C_{max} \in \Delta_G \mid W = V_{Cmax} \, ,$

(le symbole ∃! Signifie « il existe en unique exemplaire »).

Par exemple, dans la Figure 2, nous voyons qu'il existe quatre concepts miroirs qui sont disposées sur l'axe de symétrie du treillis (voir plus loin l'axe de symétrie). Ils correspondent aux quatre cliques maximales de G. Pour prouver ce théorème nous prouvons d'abord un lemme qui énonce l'existence d'un concept correspondant à chaque clique maximale, puis un second lemme énonçant l'unicité de ce concept dans U, et par conséquent dans U_m. On montrera ensuite (lemme 3)

qu'à un concept appartenant à G'_m il existe une clique maximale dans G et que celle-ci est unique (lemme 4).

Lemme 1 : $\forall C_{max} \in \Delta_G \ \exists (W,H) \in U_m \mid W = V_{Cmax}$

Preuve: Une clique maximale est un ensemble de sommets adjacents deux à deux. A chaque sommet, par construction de G' il correspond donc un objet dans G ayant tous les autres objets de la clique pour propriétés. De plus par l'adjonction de la boucle sur le sommet il possède lui-même comme propriété. A tous ces objets qui partagent les mêmes propriétés aucune autre propriété ne peut être rajoutée car sinon il lui correspondrait un objet qui n'est pas dans la clique et la clique ne serait pas maximale. Ces objets et leurs propriétés définissent donc un concept dans U. En conséquence à toute clique maximale il correspond donc bien au moins un concept de U pour lequel $W = V_{Cmax}$ dont les objets de l'extension ont pour propriétés ces mêmes objets, c'est-à-dire un concept appartenant à U_m .□

Lemme 2 : $\forall C_{max} \in \Delta_G , (W^1 ,H^1) \in U_m \mid W^1 = V_{Cmax}$
et $(W^2 ,H^2) \in U_m \mid W^2 = V_{Cmax} \Rightarrow (W^1 ,H^1) = (W^2 ,H^2)$

Preuve : Soit une clique maximale $C_{max} \in \Delta_G$, s'il lui correspond deux concepts de U_m, (W^1 ,H^1) et (W^2 ,H^2) alors $V_{Cmax} = W^1 = W^2$. Comme (W^1,H^1) et (W^2,H^2) ont respectivement la propriété $W^1 = H^1$ et $W^2 = H^2$ on a $W^1 = W^2 = H^1 = H^2$ et donc $(W^1 ,H^1) = (W^2 ,H^2)$. □

Lemme 3 : $\forall (W,H) \in U_m , \exists C_{max} \in \Delta_G \mid W = V_{Cmax}$

Preuve : Supposons qu'il existe un concept $(W,H) \in U_m$ pour lequel il ne correspond pas une clique maximale dans ΔG. A ce concept il correspond nécessairement un ensemble de sommets reliés deux à deux dans G puisqu'ils partagent tous les mêmes propriétés qui sont ces sommets adjacents plus eux-mêmes. Ce concept représente donc une clique. Si cette clique n'est pas maximale, elle est nécessairement incluse dans une clique maximale. Les objets possèdent donc en commun d'autres propriétés qui sont les sommets faisant partie de la clique maximale et qui ne font pas partie de la clique qu'ils représentent. Ils ne constituent donc pas un concept au sens des treillis de Galois puisque dans un concept toutes les propriétés partagées par les objets de l'extension sont considérées dans l'intension. Donc il ne peut pas exister $(W,H) \in U_m$ tel qu'il ne lui corresponde pas une clique maximale $C_{max} \in \Delta_G$. Par conséquent à tout concept appartenant à U_m il correspond au moins une clique maximale de G. □

Lemme 4 : $\forall (W,H) \in U_m , \forall C_{max}^1, C_{max}^2 \in \Delta_G \mid W = V_{Cmax}^1$
et $W = V_{Cmax}^2 \Rightarrow C_{max}^1 = C_{max}^2$

Preuve : Soit un concept $(W,H) \in U_m$, s'il lui correspond deux cliques maximales C_{max}^1 et $C_{max}^2 \in \Delta_G$ telles que $W = V_{Cmax}^1$ et $W = V_{Cmax}^2$ alors nécessairement $V_{Cmax}^1 = V_{Cmax}^2 = W$ et par conséquent $C_{max}^1 = C_{max}^2$. □

4.2 Autres propriétés.

Propriété 1 : G' dérivé de G est symétrique, et les concepts miroirs en sont l'axe de symétrie.

Soit *R* la relation d'ordre dans G' :

$\forall \{(W^1,H^1), (W^2,H^2) \in UxU , R ((W^1,H^1),(W^2,H^2)) \Leftrightarrow W^1 \subseteq W^2 \Leftrightarrow H^1 \supseteq H^2$

La propriété s'exprime alors formellement :

$\forall \{(W^1,H^1), (W^2,H^2)\} \in UxU, \exists! \{(W^3,H^3), (W^4,H^4)\} \in UxU \setminus \{(W^1,H^1), (W^2,H^2)\} \mid W^3=H^1 , H^3=W^1 , W^4 = H^2 , H^4 = W^2 , R ((W^1,H^1),(W^2,H^2)) \Leftrightarrow R ((W^3,H^3),(W^4,H^4)) , \neg R ((W^1,H^1),(W^2,H^2)) \Leftrightarrow \neg R ((W^3,H^3),(W^4,H^4))$

Preuve : La matrice d'adjacence A de G est symétrique ; comme le tableau du contexte T de G' est égal à A+I et que I est symétrique, T est symétrique : $T = T^t$ où T^t est la matrice transposée de T. Comme par ailleurs O = M, on a (O,M,B) = (M,O,B) et les treillis construits à partir de ces deux contextes sont isomorphes. En conséquence le demi-treillis inférieur construit à partir d' (O,M,B) est symétrique du demi-treillis supérieur de ce même contexte (O,M,B). □

Propriété 2 : l'ensemble des concepts miroirs est un ensemble stable du treillis de Galois, Formellement : $\forall \{(W^1,H^1), (W^2,H^2)\} \in U_m \, xU_m , \neg R((W^1,H^1),(W^2,H^2))$

Preuve : Supposons qu'il existe une relation entre deux concepts miroirs (W^1,H^1) et (W^2,H^2), c'est-à-dire que par exemple $W^1 \subset W^2$ (respectivement $H^2 \subset H^{1)}$) alors il existerait une clique maximale dans G' qui serait contenue par une autre clique maximale, ce qui est impossible par définition d'une clique maximale. □

5. Visualisation des cliques maximales d'un graphe

5.1 *Treillis de Galois des cliques maximales.*

Le diagramme de Hasse d'un treillis de Galois dérivé d'un graphe permet de mettre en relief les cliques maximales et maximum en jouant sur la symétrie. Dans ce but nous avons élaboré une application qui révèle les concepts miroirs d'un treillis de Galois et nous l'avons interfacée à notre grapheur Molage (Crampes et al 2006). La figure 3 montre le résultat sur le petit graphe décrit dans le tableau 1. Les cliques maximales sont placées au centre entre les nœuds top et bottom du treillis. Les concepts miroirs correspondant aux cliques maximales sont indiqués avec un rond vert, et parmi elles les cliques maximum (une seule dans l'exemple) par un triangle rouge. Dans cet exemple on voit cependant que le treillis de Galois paraît plus complexe que le graphe dont il est dérivé. Cet aspect est encore plus visible sur l'exemple suivant où G est le graphe de Turàn $T_{3,8}$. La figure 4 montre ce graphe qui comporte huit sommets et 21 arêtes. Le diagramme de Hasse dérivé présenté dans la figure 5 montre 100 concepts dont 18 miroirs (ronds rouges) qui représentent les cliques maximales contenant toutes 3 sommets. Il n'y a donc que des cliques maximum parmi les cliques maximales. Pour dessiner ce graphe nous

avons posé la propriété sans la démontrer que les concepts miroirs étaient un axe de symétrie pour le diagramme de Hasse.

L'explosion du nombre de concepts est bien connue de la communauté de chercheurs sur les treillis de Galois. Comme souvent dans des problèmes qui font intervenir les treillis de Galois, il est nécessaire de diminuer le nombre de concepts pour rendre le résultat plus lisible. La sous-section suivante va s'y employer.

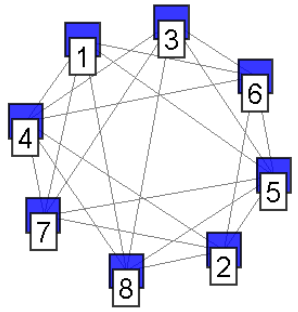

Figure 4 – Le graphe $T_{3,8}$ (dessin Molage)

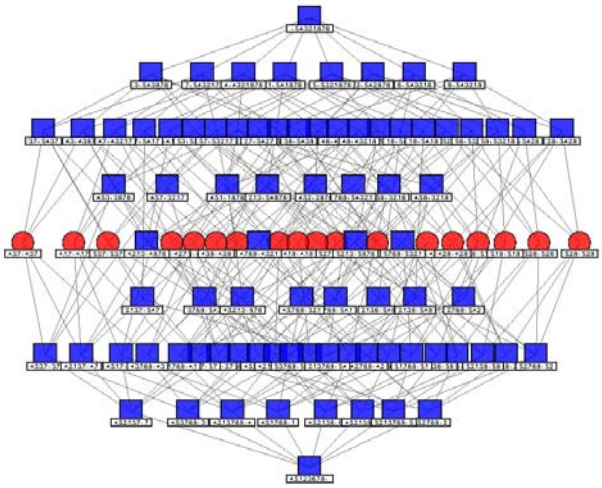

Figure 5 - Diagramme de Hasse intégral du treillis de Galois dérivé (dessin Molage)

5.2 Sous treillis de Galois des cliques maximales.

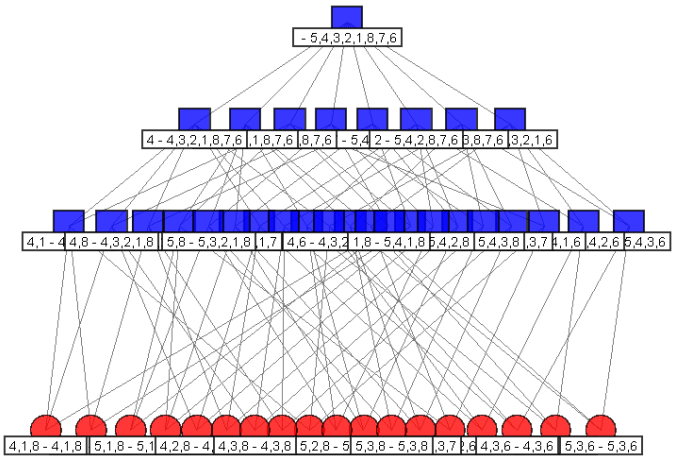

Figure 6 – Diagramme de Hasse du treillis de Galois réduit à sa partie supérieure et connexe aux cliques maximales

Considérons le treillis de Galois G'', sous treillis de G' dont on ne retient que les concepts plus grands ou égaux aux concepts miroir avec les arêtes les reliant :

$G''_{top} = (U'',F'') \subseteq G'=(U,F)$ avec

$U''= \{(W,H) \mid \exists(W^1,H^1) \in U_m, (W^1,H^1) \leq_L (W,H)\}$,

$F''= \{f = ((W^1,H^1),(W^2,H^2)) \in F \mid (W^1,H^1),(W^2,H^2) \in U''\}$

En ne retenant que ces concepts, on obtient la figure 6 qui montre le diagramme de Hasse réduit avec les cliques maximales et les différents niveaux d'intersection des cliques maximales.

6- Conclusion

L'énumération des cliques maximales d'un graphe et de ses cliques maximum a fait l'objet d'une grande littérature étant donné son importance pratique dans beaucoup d'applications et son caractère NP-complet pour le cas général. (Pardalos,et Xue 1994) en proposent une revue et la liste continue de s'allonger (Ostergard, 2002).

Le but du présent article n'est pas de proposer un nouvel algorithme d'énumération, mais d'établir une relation entre cliques maximales et treillis de Galois pour que le premier problème bénéficie des résultats du second. Le passage d'un treillis de Galois aux cliques maximales d'un graphe a déjà été rapporté dans plusieurs articles. En particulier il est montré qu'il est possible d'énumérer les bicliques d'un graphe (les sous-graphes biparti maximaux d'un graphe) en transformant ces bicliques en cliques maximales d'un graphe dérivé du premier (Alexe et al. 2004)(Makino and T. Uno 2004)(Gély et al. 2009). La recherche des bicliques maximales revient alors à appliquer des algorithmes connus de recherche des cliques maximales d'un graphe. Mais cette transformation rajoute de nombreuses arêtes et l'énumération s'en trouve plus longue. Dans le présent article nous avons proposé le processus inverse, à savoir la transformation d'un problème d'énumération des cliques maximales d'un graphe en un problème de recherche et de visualisation de bicliques maximales, c'est-à-dire d'un treillis de Galois. Nous obtenons comme résultat un treillis de Galois et un diagramme de Hasse qui met en relief les cliques maximales et maximum.

La représentation par diagramme de Hasse simplifié où extensions et intensions sont visibles (Fig. 6) présente l'intérêt de donner toute l'information et seulement l'information utile par rapport aux cliques maximales du graphe. Elle est déjà complexe pour un graphe à huit sommets, mais elle est au moins utilisable même pour des hyper-graphes non planaires. Il est difficile d'envisager moins de nœuds en ce qui concerne les cliques maximales puisqu'on a le nombre exact de celles présentes dans le graphe. Pour les autres nœuds, on peut se limiter pour encore plus de simplicité aux concepts immédiatement voisins des concepts miroirs (non présenté ici faute de place). D'autres améliorations sont aussi possibles mais elles ne sont pas traitées ici. Un second intérêt de la transformation est que le raisonnement sur les cliques maximales a fait l'objet de peu d'attention par les chercheurs. A l'inverse le passage au treillis de Galois permet de bénéficier d'un important travail

de recherche d'extraction de connaissance par raisonnement inductif. C'est la piste de réflexion la plus intéressante que nous explorerons dans de futurs articles en relation avec les réseaux sociaux.

Le problème de la complexité pose d'autres questions. En effet il est démontré que le nombre de concepts d'un treillis de Galois est en général une fonction exponentielle par rapport au nombre d'objets et de propriétés, ce que l'on peut observer sur la Figure 5. Si pour énumérer les cliques maximales d'un graphe il faut le transformer en treillis de Galois qui produit un graphe encore plus gros, l'intérêt est mince. Pour répondre en partie à ce problème nous avons amélioré l'énumération des concepts avec des algorithmes récents qui utilisent les propriétés des treillis comme In-Close présenté dans (Andrews 2009). Nous avons pu alors générer les listes des concepts de treillis de Galois dérivés de graphes à faible densité avec quelques centaines d'objets. Mais la taille de la mémoire utilisée croit de manière exponentielle et quand le graphe est plus dense l'application trouve ses limites. On notera que nos essais réalisés avec des outils comme « Lattice Miner » ou encore Galicia ont été encore plus limités avec des débordements de la mémoire. De nombreuses améliorations sont possibles. In-Close permet de ne lister que les concepts au dessus d'un seuil d'extension ou d'intension. Si on veut se limiter à la recherche de cliques maximum, ou du moins de taille intéressante, cas fréquents, cette fonctionnalité améliore fortement les performances, ce que nous avons pu tester avec des graphes plus denses. Pour aller encore plus loin, si on ne s'intéresse qu'aux cliques maximales il n'est pas nécessaire de mémoriser tous les concepts et on doit pouvoir très largement limiter la croissance exponentielle de l'utilisation de la mémoire, ce que nous étudierons par la suite en améliorant l'élagage du treillis (on notera déjà que l'élagage peut porter sur le calcul direct du treillis de la Figure 6 sans passer par celui de la Figure 5). Il reste donc une large marge de manœuvre pour étudier et visualiser les cliques maximales de graphes avec le prisme des treillis de Galois et des diagrammes de Hasse.

8. Bibliographie

Alexe G., Alexe S., Crama Y., Foldes S., Hammer P.L., Simeone B., Consensus algorithms for the generation of all maximal bicliques, *Discrete Appl. Math. 143 (11–21)*. 2004.

Andrews S., In-Close, a fast algorithm for computing formal concepts, *in: International Conference on Conceptual Structures (ICCS), Moscow*, 2009.

Battista G., P. Eadeaus, R. Tamassia, and I. Tollis. *Graph drawing. Algorithms for the visualisation of graphs,* Prentice Hall, 1999.

Berge C., *Graphs and Hypergraphs*. North-Holland, 1973.

Brandenburg Franz J., Graph Clustering I: Cycles of Cliques, *In: Graph Drawing, September 18-20, Rome, Italy , pp. 158-168,* 1997.

Brandes Ulrik, Cornelsen Sabine, Pampel Barbara, and Sallaberry Arnaud. Pathbased supports for hypergraphs. *In Proceedings of the 21th International Workshop on Combinatorial Algorithms, V6460 of L.N. in Computer Science, p20-33. Springer*, 2010.

Byong-Hyon Ju and Kyungsook Han, Complexity Management in visualizing protein, *interaction networks. Vol. 19 Suppl. 1, pages i177–i179,* 2003.

Cook, Stephen A, "The Complexity of Theorem-Proving Procedures", *Proceedings of the Third Annual ACM Symposium on Theory of Computing: 151-158,* 1971.

Crampes M., de Oliveira-Kumar J., Ranwez S., Villerd J., Indexation de photos sociales par propagation sur une hiérarchie de concepts. *In Proc. IC2009. Hammamet,* 2009.

Crampes, M., Ranwez, S., Emery, A., Velickovski, F. Mooney, C., Mille, N., Concept Maps for Designing Knowledge Maps, *Journal of Information Visualisation, Palgrave, V 5, n° 4,* 2006.

Dongbo Bu, Yi Zhao, Lun Cai, Hong Xue, Xiaopeng Zhu, Hongchao Lu, Jingfen Zhang, Shiwei Sun, Lunjiang Ling, Nan Zhang, Guojie Li and Runsheng Chen, Topological structure analysis of the protein-protein interaction network in budding yeast, *Nucleic Acids Research, Vol. 31, No. 9 2443-2450*, 2003.

Karp Richard, "Reducibility Among Combinatorial Problems", *Proceedings of a Symposium on the Complexity of Computer Computations, Plenum Press,* 1972.

Kose Frank, Weckwerth W., T. Linke, O. Fiehn: Visualizing plant metabolomic correlation networks using clique-metabolite matrices. *Bioinformatics 17(12): 1198-1208,* 2001.

GANTER B., WILLE R., *Formal Concept Analysis:Mathematical Foundations,* Springer 1999.

Gély Alain, Nourine Lhouari and Sadi Bachir, Enumeration aspects of maximal cliques and bicliques, *Discrete Applied Mathematics, Vol. 157, Issue 7, 6 April, p1447-*1459, 2009.

Herman, G. Melancon, and S. Marshall, "*Graph visualization and naviga-tion in information* visualization: A survey", IEEE TVCG, Vol. 6, No. 1, pp. 24–43, 2000.

Lahcen Boumedjout and Kwuida Leonard, Lattice Miner: A Tool for Concept Lattice Construction and Exploration. *In Suplementary Proceeding of International Conference on Formal concept analysis (ICFCA'10),* 2010.

Makino K. and Uno T., New algorithms for enumerating all maximal cliques, *Proc. 9th Scand. Worksh. Algorithm Theory, SWAT 2004, LNCS, pp. 260–272*, 2004.

Ostergard, P.R.J. A fast algorithm for the maximum clique problem, *Discrete Applied Mathematics 120 197–207,* 2002.

Pardalos P.M., Xue J., The maximum clique problem, *J. Glob. Optim. 4 () 301–328,* 1994.

Riche Henry N. and Dwyer T. *Untangling Euler Diagrams.* In Proceedings of Infovis 2010.

Riche Henry N. and Fekete JD., Novel Visualizations and Interactions for Social Networks Exploration, *In Handbook of Social Network Technologies and Applications*, *Borko Furht (Ed.),Springer, ISBN 978-1-4419-7141*, 2010.

Ruskey Frank and Weston Mark, A survey of venn diagrams. *The electronic journal of combinatorics*, 2005.