



HAL
open science

Des réseaux de photos aux réseaux sociaux. Création et utilisation d'un réseau social à partir de photos

Michel Plantié, Michel Crampes

► To cite this version:

Michel Plantié, Michel Crampes. Des réseaux de photos aux réseaux sociaux. Création et utilisation d'un réseau social à partir de photos. MARAMI, conférence sur les Modèles et l'Analyse des réseaux: Approches Mathématiques et Informatique, Oct 2010, TOULOUSE, France. pp.100-110. hal-00808040

HAL Id: hal-00808040

<https://hal.science/hal-00808040>

Submitted on 4 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Des réseaux de photos aux réseaux sociaux. Création et utilisation d'un réseau social à partir de photos

Michel Plantié, Michel Crampes

*EMA-LGI2P, Parc Scientifique Georges Besse, 30035 Nîmes Cedex, France
{michel.plantie,michel.crampes}@ema.fr; {http ://www.lgi2p.ema.fr}*

RÉSUMÉ. Avec l'avènement du monde numérique pour la photographie, les photos prises lors d'événements familiaux ou autres réunions sociales font apparaître des groupes cohérents de personnes intéressants dans l'optique des réseaux sociaux. Dans cet article, nous montrons les différentes analyses et conclusions que l'on peut tirer de ces photos sociales. Ces résultats peuvent contribuer à donner les réponses les plus appropriées pour différents objectifs sur les réseaux ou groupes de personnes identifiés. L'organisation du réseau de photos utilise des méthodes d'analyse formelle de concepts.

ABSTRACT. With the new digital photography world, photos taken at family events or other social gatherings reveal coherent groups of people having an interest for social networks concerns. In this paper, we layout the different analysis and conclusions that can be drawn from these social photos. These results may provide the most appropriate answers for different objectives on networks or identified groups of people. The network organization of photos uses methods of concepts formal analysis.

MOTS-CLÉS : Photos sociales , Réseaux sociaux, Analyse formelle de concepts, Diagrammes de Hasse. Techniques et outils de conception, Interfaces utilisateur, Environnements interactifs, Théorie des graphes, Algorithmes de Graphes.

KEYWORDS: Social Photos, Social Networks, Formal Concept Analysis, Hasse Diagrams. Design Tools and Techniques, User interface, Interactive environments, Graph Theory, Graph algorithms.

1. Introduction

Avec l'arrivée des photographies digitales et la disponibilité récente des appareils photos dans les téléphones mobiles, il est maintenant possible pour une seule personne de prendre des milliers de photos en une seule année. Actuellement, chacun fait face à des problèmes soulevés par l'abondance : comment organiser, visualiser, rechercher et partager des milliers de photos.

Dans ce papier, nous considérons uniquement les photos "sociales" indexées manuellement. Nous nous référons avec ce terme aux photos qui sont prises durant des événements familiaux ou lors de soirées entre amis. Ces photos représentent des individus ou des groupes de personnes. En considérant le rôle important que les photos numériques semblent prendre pendant des événements "sociaux", il est intéressant d'explorer leur capacité à agir en tant que nouveau langage social reflétant des relations sociales. Il est aussi pertinent d'explorer de nouvelles applications dérivées de ce nouveau langage permettant de renforcer la socialisation. Des applications telles que "Facebook" ou Flickr manipulent déjà des photos dans des réseaux sociaux. Cependant, dans ces cas les réseaux sont construits en utilisant les personnes. Les photos sont uniquement des objets passifs qui sont volontairement partagés. Elles ne jouent pas un rôle actif dans la construction des réseaux sociaux. Pour considérer les photos comme des marqueurs sociaux, il est nécessaire d'interagir avec elles en les considérant comme des objets actifs qui par leur seule présence témoigne d'une relation sociale. Cette exigence doit être appliquée à différents niveaux dans le cycle de vie du réseau de photo et du réseau social. Chaque niveau est développé dans les sections ci-dessous. La section 2 décrit les techniques pour organiser les photos sociales de manière à révéler les relations sociales entre personnes. Pour remplir cet objectif, nous utilisons les méthodes d'analyse formelle de concepts pour construire les réseaux de photos expressifs, en particulier les hiérarchies de concepts et les diagrammes de Hasse. La section 3 présente plusieurs solutions pour extraire des réseaux sociaux de collections de photos organisées. La section 4 présente les différentes méthodes de formation des tribus. La section 5 présente une application pratique de constitution de réseaux. La section 6 présente la constitution d'albums personnalisés à partir des réseaux construits.

2. Etat de l'art

Organiser socialement des photos et élaborer des réseaux sociaux à partir du contenu des photos est depuis peu un domaine très actif de recherche. La méthode la plus proche de nos travaux est présentée en [GOL 08]. L'auteur justifie le fait que les photos représentant des personnes peuvent être utilisées pour élaborer des réseaux sociaux. La méthode d'évaluation est bien fondée, et les résultats sont proches des nôtres. Cependant, l'auteur ne présente pas de réseau social déduit du réseau de photos. Par conséquent, il n'y a pas d'aide pour la construction et le partage d'albums personnalisés, ainsi la méthode d'évaluation présentée par l'auteur n'est pas très pertinente pour comparer la construction de réseaux sociaux. Il n'y a pas de réelle justification pour l'élaboration du réseau social à partir des photos. [CRA 09] présente une méthode différente incluant ce couplage entre deux outils que nous étendons dans ce papier. L'utilisation des hiérarchies de Galois pour gérer des ensembles de photos, est décrite dans plusieurs applications telles que dans [EKL 06, FER 07]. L'analyse formelle de

Concepts (FCA, Formal Concept Analysis) et les diagrammes de Hasse ont été introduits, il y a longtemps pour représenter les réseaux sociaux [FRE 93] et restent un domaine fructueux de recherche comme dans [ROT 06]. Nous introduisons une caractéristique principale de notre système : la capacité à organiser les photos sociales dans un réseau particulier. Un diagramme de Hasse est utilisé pour révéler le contenu social et élaborer le réseau social sous-jacent. Un diagramme de Hasse est une représentation bien connue d'un treillis de concepts (ou treillis de Galois) comme présenté dans [PLA 10]. Nous utilisons ce formalisme pour générer notre réseau de photos. La théorie des réseaux sociaux [STR 05] s'intéresse à la morphologie des réseaux afin de trouver des "super-nodes" des éléments référents dans le réseau qui permettent la diffusion fluide de l'information à travers le réseau. Cependant, notre objectif est différent car nous cherchons à constituer des réseaux à partir de la connaissance, déduite des photos, que représente la présence conjointe de personnes et leurs fréquences dans un événement. Cette connaissance analysée est suffisante pour induire des liens sociaux entre ces personnes. Les techniques bien connues de clustering sont fondées sur des distances entre éléments. Notre objectif est différent et se situe en amont de ces techniques car il est à la fois de définir des distances liées aux liens sociaux et de résonner à partir de ces liens qui permettent de conserver la sémantique du modèle.

3. Groupes et tribus

3.1. Des réseaux de photos aux réseaux sociaux

Chaque concept montre un groupe particulier de personnes (l'intension du concept) et contient une collection de photos (l'extension du concept) avec ce même groupe de personnes. A partir de ce diagramme, nous pouvons déduire des réseaux sociaux par le calcul de distance, l'inverse de la proximité entre toute paire de personne. Différentes formules de proximité sont possibles et nous introduisons quelques unes d'entre elles. Nous définissons les variables suivantes : x_i : un individu parmi les personnes trouvées dans la collection de photos, N : le nombre de personnes de la collection de photos, c : un concept dans le diagramme de Hasse, C : le nombre de concepts dans le diagramme de Hasse, n_c : le nombre de personnes dans un concept c , $[c/x_i]$: l'ensemble des concepts qui contiennent x_i , $[c/[x_i, x_j]]$: l'ensemble des concepts qui contiennent uniquement x_i et x_j , $[c/x_i, x_j]$: l'ensemble des concepts qui contiennent au moins x_i et x_j , $[c/x_i, x_j, x_k]$: l'ensemble des concepts qui contiennent x_i et x_j et x_k , $[c/x_i \vee x_j]$: l'ensemble des concepts qui contiennent au moins x_i ou x_j , $Car[.]$: la cardinalité d'un ensemble particulier, $[c/x_i \oplus x_j]$: l'ensemble des concepts qui contiennent x_i ou x_j mais non les deux, $[c/\oplus(x_i, x_j)]$: l'ensemble des concepts ne contenant ni x_i ni x_j .

3.2. Le concept de tribu

L'objectif de notre travail est de créer des albums personnalisés selon les liens entre les personnes présentes lors d'un événement et permettre le partage de photos entre les personnes concernées. Plus formellement de trouver les individus x_i (les individus dans les photos) qui sont concernés par l'attribution d'un ensemble de photos. Nous devons pour cela introduire une nouvelle notion. Nous appelons Tribu un sous-ensemble de personnes x_i concernées par un album personnalisé. Tous les individus

qui appartiennent à la même tribu possèdent un réseau social commun qui est mis en évidence par l'apparition de ces individus sur différentes photos sous des formes différentes. Nous résonnerons ici sur un couple d'individus x, y , plus précisément sur le lien entre x et y . Notre ensemble de photos et de personnes est un hypergraphe [MIK 05] dont les sommets sont des personnes x_i et les arêtes représentent les tribus. Un hypergraphe est la généralisation d'un graphe, et les liens entre les personnes x_i seront toujours représentés par les photos dans l'hypergraphe.

3.3. Modèles de Forces de couples et tribus

Une tribu est définie par un sous ensemble d'individus obéissant à des règles. Il est possible de définir plusieurs règles d'appartenance. Ci-dessous nous définissons différentes lois qui constitueront des tribus différentes et nous comparerons ensuite ces différents découpages. Dans ce modèle nous considérons les occurrences de personnes en relation dans les concepts. Un concept est représenté par une photo qui caractérise l'intension d'un concept, c.a.d. le groupe de personnes. La notion de tribu et les notions de forces définies ci-après sont plus générales et vont au delà de la simple co-occurrence entre personnes. Ces notions prennent en compte des mesures sur l'ensemble des concepts et non pas uniquement la présence conjointe de deux personnes.

3.3.1. La force simple d'un couple

Ici nous définissons la "fréquence" d'un couple (deux individus dans une certaine forme de relation) comme le nombre d'occurrences de ce couple parmi les concepts divisé par le nombre de concepts :

$$FORCESIMPLE(x_i, x_j) = \frac{Car[c/x_i, x_j]}{C} \quad (1)$$

Cette métrique est intéressante car elle représente la fréquence d'apparition d'un couple capturé par le photographe dans différentes situations sociales. Si un couple est vu dans de nombreuses situations alors la relation entre ces deux personnes est plutôt stable. Par contre cette distance n'exprime pas la force de la liaison entre deux personnes.

3.3.2. La proximité ou force pondérée d'un couple

En conformité avec [GOL 08] nous considérons que plus il y a de personnes dans une photo, moins les liens sont forts entre ces personnes. En conséquence pour exprimer la force d'un couple, le nombre de personnes qui sont présentes dans les photos avec ce couple doit être pris en compte. La proximité d'une tribu (respectivement d'un couple) est la somme des inverses des nombres d'individus présents dans les concepts qui contiennent la tribu, divisée par le nombre total de concepts. Donc, plus le couple est "dilué" dans un grand nombre d'individus pour un concept, moins la contribution de ce concept à la proximité est forte. Quand une tribu (un couple) n'apparaît dans aucun concept, la force est nulle. Deux personnes apparaissant dans un seul concept, donne la proximité maximum de $2/2/1 = 1$.

$$PROXIMITE(x_i, x_j) = \frac{(\sum_{k, (k/(x_i, x_j \in c_k))} \frac{2}{Car[c_k/x_i, x_j]})}{C} \quad (2)$$

3.3.3. Cohésion d'un couple

Intuitivement, on peut comprendre que le modèle social ci-dessus ne capture pas certaines observations qui peuvent être extraites automatiquement à partir du diagramme de Hasse. Parmi elles, il en existe une qui est particulièrement intéressante. Plus deux personnes sont vues séparément, plus elles sont indépendantes l'une de l'autre. Inversement quand deux personnes sont observées toujours ensemble, cela renforce l'hypothèse de liens particuliers entre eux. Nous l'appelons "cohésion". Le couple (x_i, x_j) a une forte cohésion lorsque x_i et x_j sont souvent représentés dans le même concept par rapport au nombre de fois où ils apparaissent dans un concept. On a donc :

$$COHESION(x_i, x_j) = \frac{Car[c/x_i, x_j]}{Car[c/x_i \vee x_j]} \quad (3)$$

4. Méthodes de formation des tribus

A partir des forces définies ci-avant, nous sommes en mesure de déterminer des liaisons entre personnes et donc de construire des graphes de personnes, c'est à dire un réseau social. Nous construisons un graphe valué en considérant les personnes comme les sommets et en insérant un arc entre chaque sommet dès que la force entre deux personnes est non nulle. Chaque arc comportera un poids qui est égal à la force de liaison entre les deux personnes reliées par cet arc. De nombreuses possibilités s'offrent à nous selon les graphes générés et leurs utilisations. La manière de diffuser les photos découlera directement de la méthode de calculs des groupes de personnes que nous nommons tribus. La seule application de diffusion de photos ne nécessite pas obligatoirement la détermination des tribus. Nous nous focalisons tout de même sur le calcul de tribus, car ces groupes pourront être utiles pour d'autres applications, en particulier la détermination d'intérêts communs. Dans ce chapitre, nous proposons différentes politiques et nous appliqueront l'une d'entre elle sur une application pratique réelle.

4.1. Calculs Introductifs : Densité et densité pondérée d'un sous graphe

Dans un graphe non valué (i.e. les arêtes n'ont pas de poids), la théorie des graphes définit la densité d'un graphe comme le rapport du nombre d'arêtes sur le nombre d'arêtes du graphes complet (clique) correspondant. Le nombre d'arêtes d'un graphe complet est le nombre maximal d'arêtes dans un graphe et est égal à $\frac{p \times (p-1)}{2}$, p étant le nombre de sommets. Donc la densité est $\frac{2 \times nb_arêtes}{p \times (p-1)}$. Dans un graphe valué (i.e. les arêtes ont un poids), nous définissons la densité pondérée d'un graphe comme le rapport de la somme des poids des arêtes du graphes sur le produit du nombre d'arêtes du graphes complet (clique) correspondant :

$$D_w = \frac{2 \times \sum_{i=1}^{nb_arêtes} poids_arête(i)}{(nb_sommets \times (nb_sommets - 1))} \quad (4)$$

4.2. Première méthode : extraction de tribus sur le graphe complet

Cette méthode procède en plusieurs étapes : tracer le graphe à partir des concepts et l'une des forces présentées au paragraphe précédent, contenant toutes les arêtes non nulles, énumérer tous les sous graphes possibles, retenir les tribus intéressantes : les sous graphes qui ont la densité pondérée la plus forte.

4.3. Deuxième méthode : extraction des tribus en considérant chaque personne

Dans cette méthode, si le graphe était non valué, nous considérerions qu'une personne peut avoir des liens intéressants avec d'autres personnes si le nombre d'arcs qui les séparent est inférieur ou égal à 2. Au delà, nous considérons que ces deux personnes se connaissent beaucoup moins en général. Dans notre cas où les graphes valués, en plus de la considération précédente nous considérons qu'une personne peut avoir des liens intéressants avec d'autres personnes, si la somme des poids des arcs nécessaires pour la liaison entre deux personnes est au minimum égal à 2 fois le poids minimum des arêtes partant de ce sommet. Cela signifie qu'une personne peut être liée à une autre si elle la connaît directement et en ayant une valeur de poids minimum, ou bien si elle la connaît par l'intermédiaire d'une autre personne qui la connaît avec au minimum la même force de liaison que cette personne intermédiaire. Cette méthode demande un calcul par personne. Pour chaque personne : tracer le graphe selon les consignes précédentes à partir du sommet représentant la personne considérée, retenir les tribus intéressantes : générer les sous-graphes et retenir ceux qui ont la densité pondérée la plus forte, Comparer les tribus obtenues, et fusionner les tribus identiques.

4.4. Troisième méthode : extraction des tribus en procédant par filtrage sur les composantes connexes

Dans cette méthode, le principe est de classer les arêtes du graphe de concepts par distances décroissantes. Puis à chaque valeur seuil de distance en débutant par la valeur de distance la plus faible, nous conservons les arêtes dont la distance est inférieure ou égale au seuil considéré ainsi que les sommets qui leurs sont reliés. Nous obtenons ainsi un graphe constitué des arêtes et sommets dont le seuil de distance est supérieur au seuil fixé. Pour chaque valeur seuil un graphe est construit. Les graphes sont de plus en plus étoffés au fur et à mesure que le seuil augmente. Cette méthode procède en plusieurs étapes : classer les arêtes du graphe de concepts par distances décroissantes, Pour toutes les valeurs croissantes des seuils de distance à partir de la distance la plus faible, tracer les graphes, choisir une valeur seuil, qui sélectionne un graphe parmi tous les graphes construits, dans ce dernier graphe, deux alternatives possibles : soit nous limitons le graphe construit à 2 fois le poids minimum des arêtes partant de ce sommet, soit nous générons tous les sous graphes possibles à partir du graphe obtenu par seuillage, et retenons les tribus intéressantes : les sous graphes qui ont la densité pondérée la plus forte.

4.5. Quatrième méthode : extension des calculs de forces aux groupes

Dans cette méthode, nous étendons les concepts de forces appliqués à des couples tels que présentés précédemment à des groupes. Nous calculons alors pour chaque groupe possible la force de ce groupe au sens de la force choisie. Cette méthode procède en plusieurs étapes : énumérer tous les sous graphes possibles, Calculer la mesure de force choisie pour chaque sous graphe, retenir les tribus intéressantes : les sous graphes qui ont la mesure de force la plus forte.

4.6. Cinquième méthode : combinaison des deux méthodes précédentes (3 et 4)

Dans cette méthode, nous procédons comme dans la troisième méthode ci dessus, mais en utilisant l'extension des calculs de forces de couples aux tribus. Cette méthode procède en plusieurs étapes : classer les arêtes du graphe de concepts par distances décroissantes, pour toutes les valeurs croissantes des seuils de distance à partir de la distance la plus faible, tracer les graphes, choisir une valeur seuil, qui sélectionne un graphe parmi tous les graphes construits, dans ce dernier graphe, retenir les tribus intéressantes : les sous graphes qui ont la mesure de force de tribu la plus forte.

4.7. Choix des tribus intéressantes

Le nombre de tribus possibles en fonction de la méthode choisie peut être important. Il est donc important de limiter les tribus à traiter.

Nous pouvons considérer quelques méthodes de sélection : classer les tribus par mesures de force ou de cohésion des plus fortes au plus faibles, puis choisir une valeur seuil en fonction : d'un nombre maximum de tribus, d'une valeur seuil sur la cohésion ou la force associée à la tribu, un compromis des deux précédentes propositions. Cette proposition nécessite donc de trouver une valeur seuil. La problématique se reporte alors sur le choix de la bonne valeur de seuil. soit considérer les tribus ayant un nombre de personnes maximum : les tribus trop importantes ne regroupent pas des personnes très proches, soit un compromis des deux précédentes propositions.

Plusieurs combinaisons des méthodes ci-dessus sont possibles. Nous avons déjà évoqué une combinaison dans la cinquième méthode. Chacune des méthodes ci-dessus pourra être utilisée en choisissant un type de force différent à chaque fois. Cela donnera donc au moins 15 méthodes différentes. Il est même possible de calculer les graphes avec une combinaison des forces, ou également de pondérer une force comme la proximité par la force de cohésion.

4.8. Discussion : avantages et inconvénients de chaque méthode

Nous avons proposé ci-dessus plusieurs méthodes pour déterminer les tribus intéressantes. Chacune a ses avantages et ses inconvénients. La méthode 1 a l'avantage de pouvoir comparer tous les graphes possibles. Cependant l'inconvénient majeur est le nombre de graphes possibles, il est au nombre de 2^N , N étant le nombre de personnes. Pour une application réelle de 40 personnes comme l'application que nous présentons ci-dessous, le nombre de tribus est déjà très important. La méthode 2 a l'avantage de limiter, à un facteur près, le nombre de tribus générées au nombre de personnes. Elle est également très adaptée à l'application de diffusion de photos. L'inconvénient est bien sûr de calculer des tribus différentes pour chaque personnes. Dans cette méthode, on ne détermine pas de tribus pour l'ensemble des personnes et l'utilisation que l'on pourra faire des tribus ainsi trouvées sera plus limitée. La méthode 3 a l'avantage de ne pas proposer un calcul exponentiel mais un calcul linéaire. Cependant la deuxième alternative peut générer un nombre de tribus potentiellement important. Il sera cependant limité au graphe généré. L'inconvénient est que nous devons choisir un seuil qui déterminera le graphe construit. Il faudra donc trouver des critères pertinents pour le choix de ce seuil. L'inconvénient majeur de la méthode 4 est le nombre de graphes possibles, il est au nombre de 2^N , N étant le nombres de personnes. Son avantage est d'étendre le concept de force à des graphes. La méthode 5 combine les avantages

des méthodes trois et quatre tout en éliminant le problème du nombre exponentiel de graphes. Son avantage est d'étendre le concept de force à des graphes. L'inconvénient est que nous devons choisir un seuil qui déterminera le graphe construit. Il faudra donc trouver des critères pertinents pour le choix de ce seuil.

5. Expérimentation pratique sur la troisième méthode

Nous avons considéré un événement réel de type mariage comprenant : 27 personnes et 127 photos. Conformément à la règle des treillis de concepts, nous considérons comme hypothèse que deux photos comprenant les mêmes personnes représentent le même concept. Nous calculons alors la matrice de présence 27×127 des personnes dans les concepts. Il y a en moyenne 3 personnes par photos réparties selon le tableau 1 .

nombre de personnes sur les photos	1	2	3	4	5	6	7	11	12
nombre de photos	21	32	26	18	16	7	5	1	1

Tableau 1. répartition du nombre de personnes par photos

Puis pour chaque force nous calculons les différents graphes associés aux niveaux de seuil sur les forces de liaisons entre sommets. Cela produit des graphes successifs de plus en plus complets au fur et à mesure que le seuil baisse. Afin de pouvoir valider les tribus obtenues par nos calculs, nous représentons ci-après (cf. figure 5) le réseau familial comprenant les personnes présentes sur les photos. Dans ce graphique, un certain nombre de personnes ne sont en relation avec aucun membre : ce sont les amis et autres relations.

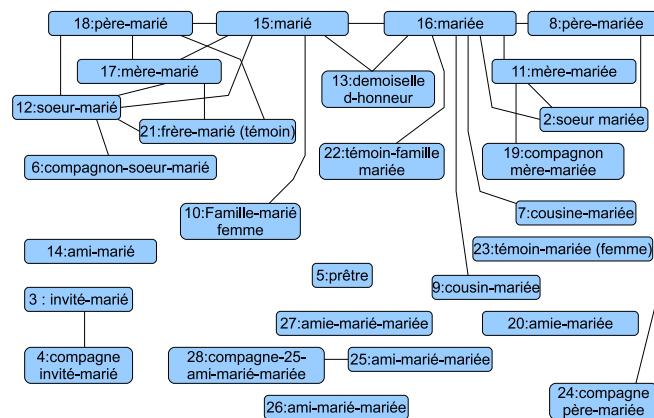


Figure 1. réseau familial réel

5.1. Graphes des personnes selon les différentes forces

La force simple produit un graphe évolutif et toujours connexe (cf figure 2). Cette force fait apparaître une seule tribu pour obtenir davantage de tribus nous devons générer les sous-graphes de la composante connexe sélectionnée. La force de cohésion produit des tribus séparées qui deviennent de plus en plus connexes (cf figure

3). Cette force fait apparaître rapidement des sous groupes intéressants. La force de proximité produit un graphe évolutif de plus en plus connexe (cf figure 4). Cette force fait apparaître deux tribus.

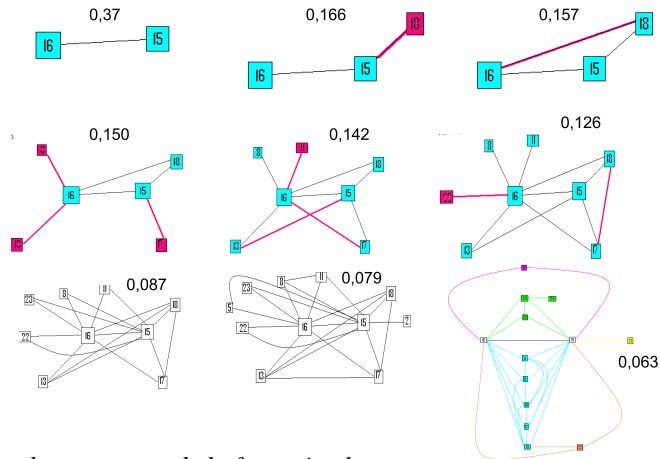


Figure 2. graphes connexes de la force simple

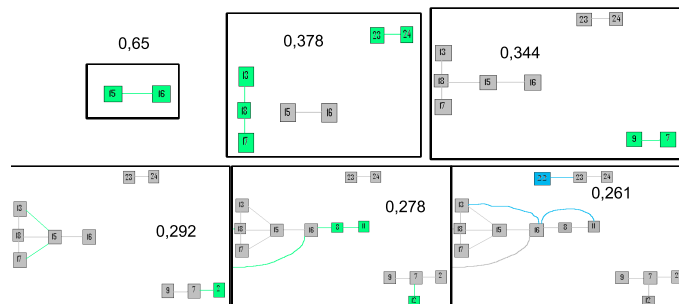


Figure 3. graphes connexes de la force de cohésion

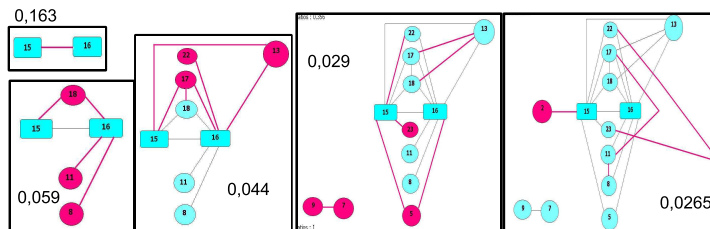


Figure 4. graphes connexes de la force de proximité

5.2. Choix du seuil des composantes connexes et sous-graphes

Afin de pouvoir comparer les différents effets des forces, nous choisissons de limiter le nombre de personnes à 13, soit une personne de plus que le nombre de personnes

maximum sur une photo. Cette limite nous permet de fixer un seuil pour chaque force. Il est de 0,062 pour la force simple, de 0,277 pour la force de cohésion et 0,0265 pour la force de proximité. Nous choisissons la deuxième alternative proposée dans la méthode à savoir générer toutes les sous-graphes possibles à partir de la composante connexe sélectionnée par le seuil.

5.3. Sélection des sous-graphes intéressants de la force simple

Le tableau 2 montre les tribus générées et leur densité pondérée (4.1) pour la force simple. Nous avons choisi de citer uniquement les sous-graphes dont la connexité brute a une valeur élevée. Et bien sûr nous avons volontairement éliminé les sous-graphes constitués d'un seul couple de personnes et de trois personnes. la sélection se fait selon le degré de connexité en observant le graphe, car le nombre de sous-graphes est déjà important (2^{13}). Le dernier sous-graphe présenté correspond à la composante connexe complète. On remarquera que le calcul de densité pondérée est de plus en plus faible au fur et à mesure que le nombre de personnes augmente.

n° tribu	tribu	densité	densité pondérée	description
1	15-16-11-8	1	0,16	mariés, parents-mariée
2	15-16-22-23	1	0,14	mariés, témoins-mariée
3	15-16-11-8-18	0,9	0,1354	mariés, père marié, parents-mariée
4	15-16-11-8-17	0,9	0,1323	mariés, mère marié, parents-mariée
5	15-16-11-8-18-17	0,93	0,0945	mariés, parents-mariés
6	15-16-11-8-18-17-13	0,714	0,0937	mariés, parents-mariés demoiselle d'honneur
7	15-16-22-23-24	0,7	0,08	mariés, témoins-mariée..
8	15-16-11-8-18-17 13-22-23-24-5-2-21	0,423		composante connexe

Tableau 2. densité des tribus de la force simple

5.4. Sélection des sous-graphes intéressants de la force de cohésion

La force de cohésion pour une composante connexe comprenant 13 sommets (personnes), fait apparaître naturellement plusieurs sous-graphes non connectés (voir tableau 3). Nous limiterons les calculs à ces sous-graphes "naturels".

n° tribu	tribu	densité	densité pondérée	description
1	23-24	1	0,444	témoin-mariée et compagne-père-mariée
2	9-7-2-12	0,5	0,154	cousin-cousine-mariée soeur-marié-soeur-mariée
3	13-18-17-15-16-8-11	0,6	0,139	mariés, parents-mariés demoiselle d'honneur

Tableau 3. densité des tribus de la force de cohésion

5.5. Sélection des sous-graphes intéressants de la force de proximité

La force de proximité pour une composante connexe comprenant 13 sommets (personnes), fait apparaître naturellement deux graphes non connectés (voir tableau 4).

Nous générons également les sous-graphes du graphe le plus grand dont la connexité brute a une valeur élevée. On remarquera que le calcul de densité pondérée n'est pas systématiquement proportionnel au nombre de personnes de la tribu.

n° tribu	tribu	densité	densité pondérée	description
1	15-16-22-23	0,66	0,0473	mariés, témoins-mariée
2	8-11-15-16-18-17	0,6	0,0368	mariés, parents-mariés
3	8-11-15-16-18-17-13	0,571	0,0346	mariés, parents-mariés demoiselle d'honneur
4	9-7	1	0,03	cousin-cousine-mariée
5	8-11-15-16-18-17-13 2-5-23-22	0,381	0,0176	mariés, parents-mariés demoiselle d'honneur prêtre, témoins mariée

Tableau 4. densité des tribus de la force de proximité

6. Politique de diffusion de photos, constitution d'albums personnalisés

Afin de déterminer les albums à construire, nous calculons la distance de chaque photo à une tribu. Nous utilisons la distance de Jaccard [KAU 90] qui mesure la dissimilarité entre deux ensembles. Elle est obtenue en divisant la différence de taille de l'union et de l'intersection de deux ensembles : $J_{\delta}(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$. Les deux ensembles sont les individus présents dans une photo et les individus présents dans une tribu. Nous pondérons alors la distance d'une photo à une tribu par la densité des tribus associée à chaque force. Nous obtenons pour chaque force un classement décroissant des couples photos-tribus en fonction de leur distance pondérée.

6.1. Diffusion de photos

Cinq alternatives de diffusion sont possibles de la plus restrictive à la plus sociale :

- 1) Diffusion des photos aux seules personnes présentes sur les photos
- 2) Diffusion des photos aux tribus dont la distance photo-tribu est inférieure à une valeur donnée. Fixons comme règle pour diffuser une photo à tous les membres d'une tribu, le fait que la distance de Jaccard photo-tribu soit inférieure à 0,5, qui indique qu'une photo montre au moins une majorité des individus de la tribu.
- 3) Diffusion des photos aux tribus dont la distance photo-tribu est inférieure à une valeur donnée. Cette distance est déterminée par un critère économique c'est à dire le nombre d'exemplaire de photos à diffuser, ici 1000 par exemple.
- 4) Nous Classons les couples "photo-tribu" par ordre décroissant selon leur distance et nous diffusons les photos en déterminant un seuil à partir duquel on ne diffuse plus les photos aux tribus désignées. Ce seuil peut revêtir un critère plus sémantique comme le fait que dès qu'une personne de la tribu est présente sur la photo elle est diffusée à toute la tribu.. Cette solution est une variante sémantique de la solution 2.
- 5) Diffusion de toutes les photos à toutes les personnes.

La solution 1 représente 415 photos et la solution 5 représente 3429 photos et sont déjà abondamment pratiquées. La solution 2 a été évaluée avec les différentes forces. Pour la force simple le nombre de photos est de 197 en considérant les tribus 6 et 7 qui couvrent toutes les personnes de la composante connexe avec des scores élevés de densité, pour la force de cohésion le nombre de photos est de 175 en prenant les

tribus “naturelles” 1,2 et 3, et pour la force de proximité le nombre de photos est de 215 en prenant les tribus 1, 3 et 4 qui couvrent toutes les personnes de la composante connexe. Avec cette méthode on obtient moins de photos que la solution 1. Ce qui est normal car les tribus ne concernent pas toutes les personnes de l'événement. Les stratégies évaluées permettent globalement de diffuser un nombre de photos plus restreint tout en augmentant la pertinence et l'intérêt des photos pour chaque participant. La diffusion des photos est habituellement un casse-tête pour les personnes en charge de cette diffusion, qui ne souhaitent pas envoyer toutes les photos mais plutôt les plus intéressantes, trouve ici une solution élégante et efficace.

7. Conclusion

Nous avons présenté des méthodes d'extraction de réseaux sociaux à partir de photos pour leur diffusion, dans l'intention d'utiliser également ces réseaux pour d'autres objectifs. Les différentes stratégies présentées sont modifiables et montrent la richesse du modèle présenté. Nous avons évalué quelques stratégies sur un cas réel assez important. Les modèles peuvent être étendus pour prendre en compte plus efficacement la sémantique des relations et utiliser les réseaux extraits à des fins plus sémantiques. Nos travaux futurs vont explorer ces modèles sur des cas réels plus volumineux où le problème de passage à l'échelle devra être pris en compte.

8. Bibliographie

- [CRA 09] CRAMPES M., DE OLIVEIRA-KUMAR J., RANWEZ S., VILLERD J., « Indexation de photos sociales par propagation sur une hiérarchie de concepts », *Actes de la conférence IC 2009, Hammamet Tunisie*, , 2009, p. 13-24.
- [EKL 06] EKLUND P., DUCROU J., WILSON T., « An Intelligent User Interface for Browsing and Search MPEG-7 Images using Concept Lattices », *Proc of the 4th International Conference on Concept Lattices and Their Applications, LNAI, Springer-Verlag*, , 2006.
- [FER 07] FERRÉ S., « CAMELIS : Organizing and Browsing a Personal Photo Collection with a Logical Information System », *Proc. of the 5th International. Conference on. Concept Lattices and Their Applications*, , 2007, p. 112-123.
- [FRE 93] FREEMAN L., WHITE D., « Using Galois Lattices to Represent Network Data », *Sociological Methodology*, vol. 23, 1993, p. 127-146.
- [GOL 08] GOLDER S. A., « Measuring Social Networks with Digital Photograph Collections », *ACM Conference on Hypertext and Hypermedia. June 19-21. Pittsburgh, Pennsylvania*, , 2008.
- [KAU 90] KAUFMAN L., ROUSSEEUW P., P. J. B., *Finding groups in data : An introduction to cluster analysis*, WILEYInterscience, 1990.
- [MIK 05] MIKA P., « Ontologies are us : A unified model of social networks and semantics », *In International Semantic Web Conference*, , 2005, p. 522-536.
- [PLA 10] PLANTIÉ M., CRAMPES M., « Création, interaction et visualisation d'un réseau social obtenu avec des photos », *REISO 2010, Atelier REcherche et REcommandation d'information dans les REseaux sOciaux, Marseille*, , 2010, p. 1-13.
- [ROT 06] ROTH C., BOURGINE P., « Lattice-based dynamic and overlapping taxonomies : The case of epistemic communities », *Scientometrics (impact factor : 1.74)*, vol. 69(2), 2006, p. 429-447.
- [STR 05] STRAHILEVITZ L. J., « A Social Networks Theory of Privacy », *American Law & Economics, Association Annual Meetings*, , 2005.