



Création, extraction et visualisation d'un réseau social obtenu avec des photos

Michel Plantié, Michel Crampes

► To cite this version:

Michel Plantié, Michel Crampes. Création, extraction et visualisation d'un réseau social obtenu avec des photos. Atelier REcherche et REcommandation d'information dans les REseaux sOciaux, Conférence Inforsid, May 2010, Marseille, France. pp.1-12. <hal-00812177>

HAL Id: hal-00812177

<https://hal.science/hal-00812177v1>

Submitted on 11 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.



HAL Authorization

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.crampes, michel.plantie}@ema.fr, <http://www.lgi2p.ema.fr>

Résumé. Avec les nouvelles possibilités apparues dans les domaines des communications et du management, les réseaux sociaux et les photos font l'objet d'un intérêt grandissant dans le monde numérique. Dans cet article, nous montrons comment les photos sociales, prises lors d'événements familiaux ou lors de soirées entre amis, représentant des individus ou des groupes d'individus, peuvent être considérées comme génératrice d'un réseau social révélant des attributs sociaux. A partir de ce réseau de photos, nous constituons des sous-groupes de personnes dans l'objectif de diffuser des albums photos personnalisés. L'organisation du réseau de photos utilise des méthodes d'analyse formelle de concepts.

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. Dans le passé, une personne prenait en général un petit nombre de photos au cours d'un événement. Ceci était en particulier dû au coût de développement des photos qui générait clairement une limitation du photographe amateur. Aujourd'hui, le problème est très différent avec l'avènement des technologies numériques. Actuellement, chacun fait face à d'autres 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 semble 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 Flick'r 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 acteurs sociaux, il est nécessaire d'interagir avec elles en les considérant comme des objets

Des réseaux de photos aux réseaux sociaux.

actifs. 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. Les albums photos personnalisés peuvent être déduits à partir du réseau résultant et visuellement lié aux personnes choisies afin de vérifier qui peut recevoir quelles photos. La section 4 présente une application pratique de constitution de réseaux. La section 5 présente la constitution d'albums personnalisés à partir des réseaux construits.

2 Photos sociales et réseaux sociaux

Organiser socialement des photos et élaborer des réseaux sociaux à partir du contenu des photos est depuis peu un domaine très actif de recherche, cependant nous ne connaissons pas d'article qui intègre l'organisation de photos et l'extraction de réseaux sociaux. La méthode la plus proche de nos travaux est présentée en Golder (2008). 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, bien que notre méthode d'évaluation ne soit pas aussi précise que celle présentée dans le papier cité. 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. Les photos sont également indexées dans un outil séparé qui n'utilise pas leur contenu social disponible. Il n'y a pas de réelle justification pour l'élaboration du réseau social à partir des photos. Crampes et al. (2009) 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 Eklund et al. (2006); Ferré (2007). 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 Freeman et White (1993) et restent un domaine fructueux de recherche comme dans Roth et Bourguin (2006).

2.1 Le treillis de concepts de photos indexées

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). Dans l'Analyse Formelle de Concepts Ganter et Wille (1999), un ensemble d'objets possédant des propriétés (ou attributs) peut être organisé dans un treillis de concepts. Ces concepts contiennent des objets selon leur propriétés communes. Dans notre cas, nous considérons les photos comme des objets et les noms des personnages sur les photos comme des propriétés. Le processus d'organisation débute avec le contexte formel : un tableau avec les objets sur les lignes et les propriétés sur les colonnes. Chaque case est cochée (i.e. valeur =1) si l'objet correspondant possède la propriété correspondante, et reste vide (i.e. valeur =0) si l'objet ne possède pas la

propriété correspondante. Formellement, un concept formel est un triplé (G, M, I) où G est un ensemble d'objets, M un ensemble d'attributs, I est une relation binaire entre les objets et les attributs, i.e. $I \subseteq G \times M$.

L'étape suivante dans la construction du treillis de concepts est de définir les concepts. Un concept est défini par une paire de sous-ensemble : un sous-ensemble d'objets (appelé l'extension) et un sous-ensemble de propriétés (l'intention) que les objets partagent. Dans notre cas, un concept est un sous-ensemble de photos qui partagent le même sous-ensemble de personnes. Formellement, pour un ensemble $O \subseteq G$ d'objets et un ensemble $A \subseteq M$ d'attributs, nous définissons un ensemble d'attributs commun aux objets de O par

$$f : 2^G \rightarrow 2^M, f(O) = \{a \in A \mid \forall o \in O, (o, a) \in I\}$$

et l'ensemble des objets qui a tous ses attributs dans A , par :

$$g : 2^M \rightarrow 2^G, g(A) = \{o \in O \mid \forall a \in A, (o, a) \in I\}$$

La paire $\{f, g\}$ est une connexion de Galois entre $(2^G, \subseteq)$ et $(2^M, \subseteq)$.

Un concept formel du contexte (G, M, I) est une paire (O, A) avec $O \subseteq G$, $A \subseteq M$, $A = f(O)$ et $O = g(A)$.

L'étape qui suit le processus d'identification des concepts, est la construction d'un treillis dont les éléments sont les concepts. Nous définissons un ordre partiel sur l'ensemble des concepts, où chaque paire de concepts a un supremum unique (la borne supérieure la plus petite d'un concept ; appelée son joint) et un infimum (la borne inférieure la plus grande ; appelée sa réunion). Formellement, soit L l'ensemble des concepts de (G, M, I) ; l'ordre partiel est défini comme suit : $(o1, a1) \leq_L (o2, a2) \Leftrightarrow a1 \subseteq a2 \Leftrightarrow o2 \subseteq o1$. L'ensemble des concepts L est complété si nécessaire par un concept "top" qui contient tous les objets, et un concept "bottom" qui contient tous les attributs. La paire (L, \leq_L) est appelée treillis de concepts de (G, M, I) .

2.2 La sous hiérarchie de Galois

Pour plus de simplicité, il est possible de représenter les extensions réduites avec les "intentions" réduites de concepts. Une extension réduite d'un concept (O, A) est un ensemble d'objets qui appartiennent à O et n'appartiennent à aucun concept de rang inférieur, c.a.d. les objets qui n'ont aucun autre attribut que ceux contenus dans A . De façon duale, l'intension réduite d'un concept (O, A) est l'ensemble d'attributs qui appartiennent à A et n'appartiennent à aucun autre concept de rang supérieur. Nous souhaitons positionner les photos dans le graphe avec les seuls concepts qui contiennent les photos, ainsi nous considérons une réduction du graphe où seul les concepts objets sont considérés. Cela signifie que nous conservons les concepts qui contiennent des objets (photos) à l'endroit où ces objets n'apparaissent qu'une seule fois à leur niveau le plus bas dans la hiérarchie (l'extension réduite). Cet acte de nettoyage est celui proposé en Godin et al. (1995) sous le nom de PCL/X. Ceci est illustré dans notre exemple dans la figure 1. C'est maintenant une sous-hiérarchie de Galois comme expliqué dans Godin et Chau (1999). Pour être plus explicite nous l'appelons une sous-hiérarchie Objet de Galois (SHOG). C'est une visualisation bien plus légère des données où l'on se focalise sur les objets (dans notre cas les photos).

Des réseaux de photos aux réseaux sociaux.

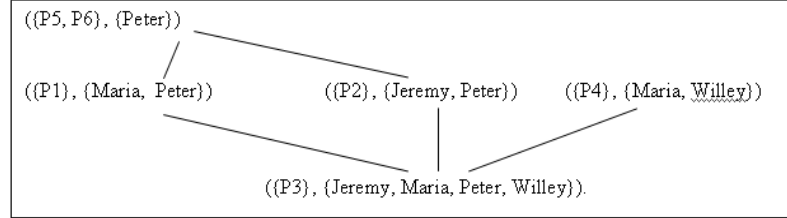


FIG. 1 – Sous hiérarchie objet de Galois

3 Des réseaux de photos aux réseaux sociaux

Le diagramme de la figure 1 révèle les différentes co-occurrences entre les personnes dans les photos. 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 l'ensemble des personnes trouvées dans la collection de photos

N : est le nombre de personnes de la collection de photos

c : est un concept dans le diagramme de Hasse

C : est le nombre de concepts dans le diagramme de Hasse

(Nota : pour la construction du réseau à partir des photos nous considérerons uniquement les concepts dont l'intension contient au moins deux personnes, C est donc le nombre de concepts dont l'intension est de cardinalité supérieure ou égale à 2)

n_c : est le nombre de personnes dans un concept c

$[c/x_i]$: est l'ensemble des concepts qui contiennent x_i

$[c/[x_i, x_j]]$: est l'ensemble des concepts qui contiennent uniquement x_i et x_j

$[c/x_i, x_j]$: est l'ensemble des concepts qui contiennent au moins x_i et x_j

$[c/x_i, x_j, x_k]$: est l'ensemble des concepts qui contiennent x_i et x_j et x_k

$[c/x_i \vee x_j]$: est l'ensemble des concepts qui contiennent au moins x_i ou x_j

$Car[.]$: est la cardinalité d'un ensemble particulier

$[c/x_i \oplus x_j]$: est l'ensemble des concepts qui contiennent x_i ou x_j mais non les deux

$[c/\oplus(x_i, x_j)]$: est l'ensemble des concepts qui ne contiennent ni x_i ni x_j

3.1 Un réseau social est un hypergraphe

La plupart des auteurs considèrent qu'un réseau social est un graphe dont les sommets sont les personnes et les arêtes les liens entre ces personnes. Cependant si l'on analyse plus finement la situation, un réseau social est un hypergraphe comme le suggère également Mika (2005). Les hypergraphes sont des objets mathématiques généralisant la notion des graphes, dans le sens où les arêtes ne relient plus un ou deux sommets, mais un nombre quelconque de sommets.

Un hypergraphe H est un couple (V, E) où $V = v_1, v_2, \dots, v_n$ est un ensemble non vide (généralement fini) et $E = e_1, e_2, \dots, e_m$ est une famille de parties non vides de V . Les éléments de V sont les sommets de H . Les éléments de E sont les arêtes de H . Les hypergraphes correspondent précisément aux matrices de dimension $n \times m$ à coefficients 0 ou 1 (dont chaque colonne a au moins un 1). Tout hypergraphe H correspond de manière univoque à la matrice telle que : $\forall a_{i,j} \in A, a_{i,j} = 1$ si $v_i \in e_j$ et $= 0$ sinon. Un hypergraphe est donc décrit par une matrice que l'on nomme matrice d'incidence.

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 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 Règles de cohésion des tribus ou modèle de force

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.

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.

Des réseaux de photos aux réseaux sociaux.

3.3.2 La proximité ou force pondérée d'un couple

En conformité avec Golder (2008) 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. Avec ce point en tête, l'auteur du Golder (2008) définit la force d'une paire d'individus par la formule suivante :

$$STR(P_a, P_b) = \sum_{i=1}^m \frac{1}{\sqrt{(n_i - 1)}} \quad (2)$$

Où P_k est une photo montrant k , $STR(P_a, P_b)$ est la force du lien entre a et b , n_i est le nombre de personnes dans la photo i et m est le nombre de photos (montrant a et b). Nous voyons trois écueils dans cette définition. Tout d'abord elle est définie par le sens commun mais sans justification formelle. Ensuite, la valeur n'est pas bornée (par exemple à 1). Enfin, notre principal intérêt est d'appliquer cette loi à toutes les photos. En conséquence, si un nombre important de photos des mêmes personnes ont été prises à la suite, elles vont renforcer les liens entre ces personnes de façon incorrecte. Nous préférons considérer les concepts et donc toutes les photos qui représentent le même groupe d'individus ne sont comptées qu'une seule fois. Nous estimons que ce qui différencie un couple dans un groupe de personnes est le fait qu'ils sont plus ou moins socialisés, c.a.d. qu'ils apparaissent plus ou moins dans différents groupes de personnes dans différents concepts comme cela est défini dans l'équation 1. S'ils apparaissent une seule fois et seul, leur force devrait être égale à 1 car le couple peut être identifié clairement. Rigoureusement, la force de toute paire d'individus doit être observée dans les concepts et dans un concept particulier, elle doit être proportionnelle à l'existence de ce couple dans un groupe de personnes qui sont présentes dans le concept. Une formule plus évoluée permet de dire qu'un couple qui apparaît souvent dans des petits groupes d'individus, il semble logique qu'il y ait plus de chance pour que ce couple soit plus proche que s'il apparaissait dans de grands groupes d'individus. 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 $2/2/1 = 1$.

$$Proximite(x_i, x_j) = \frac{(\sum_k \frac{2}{Car[c_k/x_i, x_j]})}{C} \quad (3)$$

\sum_k avec k tel que : $x_i, x_j \in c_k$.

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 :

$$Cohesion(x_i, x_j) = \frac{Car[c/x_i, x_j]}{Car[c/x_i \vee x_j]} \quad (4)$$

3.4 Formation des tribus

Dans cette étape, nous construisons une matrice $B_{n,n}$ qui dépend des règles de formation des tribus. Les règles ci-dessus constituent autant de fonctions tribales. Une fonction tribale révèle le comportement d'un couple d'individus (x_i, x_j) avec $i, j \in [1, n]$. Nous fixons ensuite un seuil ϵ tel que : $\forall b_{i,j} \in B, b_{i,j} = 1$ si la fonction tribale est $> \epsilon$ et 0 sinon. On note que $\forall i \in [1, n] b_{i,i} = 1$.

3.4.1 Co-occurrence et force de tribus

La notion de tribu et les notions de forces définies ci-avant sont plus générales que 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.

3.5 Matrice d'incidence de l'hypergraphe

L'étape précédente nous permet d'obtenir, après avoir choisi un seuil ϵ , une matrice B associée à l'hypergraphe qui met en évidence les relations entre chaque individu présent à un événement. A partir de cette matrice, nous devons donc trouver les tribus de ces individus. Ainsi, nous cherchons la matrice d'incidence de l'hypergraphe H tel que $H = T(G)$, avec $T(G)$ matrice triangulaire supérieure de la matrice associée. La i -ème ligne de H représente une tribu. La j -ème colonne de H représente un individu. Pour un i fixé, si $h_{i,j} = 1$ alors l'individu j est dans la tribu i .

4 Application pratique

Nous supposons un événement où cinq personnes Alain (a), Bernard (b), Céline (c), Daniel (d), Emilie (e) ont été pris en photo. Ces photos sont au nombre de 9. Le tableau suivant relate quels sont les individus présents sur chaque photo.

concept	1	2	3	4	5	6	7	8	9
personnes	a	a,c	d,c,e	d,c	a,c,d,e	b,c,d	b	a,d,e	d,e

TAB. 1 – *photos, personnes et concepts*

4.1 Matrices associées à chaque règle de tribus

Nous présentons dans les tableaux 2, les matrices générées par les différentes règles de constitution de tribus, du paragraphe 3.3.

Des réseaux de photos aux réseaux sociaux.

FORCESIMPLE	A.	B.	C.	D.	E.
Alain	1	0	0,22	0,22	0,22
Bernard		1	0,11	0,11	0
Céline			1	0,44	0,22
Daniel				1	0,44
Emilie					1

PROXIMITÉ	A.	B.	C.	D.	E.
Alain	1	0	0,17	0,13	0,13
Bernard		1	0,07	0,07	0
Céline			1	0,31	0,13
Daniel				1	0,31
Emilie					1

COHESION	A.	B.	C.	D.	E.
Alain	1	0	0,29	0,29	0,33
Bernard		1	0,17	0,14	0
Céline			1	0,57	0,29
Daniel				1	0,67
Emilie					1

TAB. 2 – *forces et couples*

4.2 Forces et tribus

le nombre de tribus possibles est en théorie égal à 2^N mais dans la pratique il ne peut pas dépasser le nombre de photos au carré. Les tribus à considérer sont celles obtenues lorsque les forces sont non nulles. Nous en déduisons donc 14 tribus sans compter les individus uniques qui constituent des tribus particulières dont les scores sont 1 pour toutes les forces. Les forces calculées précédemment (voir leur valeurs dans le tableau 3) permettent d’apporter un coefficient de pondération que nous nommerons “CONNIVENCE” d’une tribu. La connivence peut avoir plusieurs significations, si nous adoptons :

1. La valeur “max” des forces pour chaque tribu, exprime le fait que les individus voient leur empathie déterminée par l’évènement le plus heureux qui les rassemble. Dans le cas des photos, un individu x_i se souviendra d’un individu x_j plus par le fait qu’ils ont discuté en tête à tête (témoigné par une photo) que par le fait qu’ils ont été à un moment donné dans un groupe plus large.
2. La valeur “moyenne” des forces pour chaque tribu, à l’inverse, indique que la connivence est une impression générale sur toute la soirée.
3. La valeur “minimum” des forces pour chaque tribu, indique que les personnes d’une tribu se souviendraient des moments de la soirée où elles sont avec d’autres personnes avec lesquelles elles ont moins d’affinités.

Nous voyons ainsi que le choix de la mesure “CONNIVENCE” peut varier selon l’intention que l’on veut donner à une tribu. Prendre une valeur élevée de connivence c’est favoriser les tribus aux liens sociaux étroits. Prendre une connivence faible correspond à l’observation de tribus dont les membres sont faiblement liés et peu soucieux de partager leurs photos (selon le point de vue socio-photographique que nous adoptons dans cet article). Nous adopterons la valeur max du tableau 3 pour pondérer la distance d’une photo à une tribu. dans le tableau 3

force	FORCE SIMPLE	PROXIMITÉ	COHÉSION	max
tribu (a,c)	0,22	0,17	0,29	0,29
tribu (a,d)	0,22	0,13	0,25	0,25
tribu (a,e)	0,22	0,13	0,33	0,33
tribu (b,c)	0,11	0,07	0,17	0,17
tribu (b,d)	0,11	0,07	0,14	0,14
tribu (c,d)	0,44	0,31	0,57	0,57
tribu (c,e)	0,22	0,13	0,29	0,29
tribu (d,e)	0,44	0,31	0,66	0,66
tribu (a,c,d)	0,11	0,06	0,13	0,13
tribu (a,c,e)	0,11	0,06	0,13	0,13
tribu (a,d,e)	0,22	0,13	0,25	0,25
tribu (b,c,d)	0,11	0,07	0,13	0,13
tribu (c,d,e)	0,22	0,13	0,33	0,33
tribu (a,c,d,e)	0,11	0,06	0,13	0,13

TAB. 3 – *forces et tribus*

5 Politique de diffusion de photos, constitution d’albums personnalisés

La constitution des albums personnalisés découle d’une règle simple : toute photo contenant un membre d’une tribu est diffusée à toute la tribu. Nous devons donc déterminer la règle la plus intéressante pour choisir les tribus.

5.1 Distance brute et pondérée d’une photo à une tribu

Afin de déterminer les albums à construire, nous calculons la distance de chaque photo à une tribu. Nous utilisons la distance de jaccard (voir KAUFMAN et al. (1990)) 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 : $J_\delta(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$. Ici nos deux ensembles sont les individus présents dans une photo et les individus présents dans une tribu. Le tableau 4 montre les distances (plus exactement : le complément à 1 des distances) des concepts aux tribus en fonction des différentes tribus. Plus la mesure indiquée est grande plus la distance entre les éléments est faible et leur proximité importante. Ce tableau indique également que plus une tribu est grande plus la somme de ses distances est grandes. On remarque également que certaines photos obtiennent un score somme important qui n’est pas lié uniquement au nombre de personnes présentes dans les photos. Nous utilisons ensuite la valeur “max” du tableau 3 des forces pour pondérer la distance d’une photo à une tribu. Le tableau des distances des concepts aux tribus est donc modifié sur le tableau 5.

5.2 Diffusion de photos

Le critère de choix entre différentes alternatives peut être fondé sur un critère économique : le nombre total d’exemplaires de photos diffusés. Pour diffuser les photos nous pouvons choisir entre cinq alternatives qui vont de la plus restrictive à la plus sociale :

1. Diffusion des photos aux seules personnes présentes sur les photos

Des réseaux de photos aux réseaux sociaux.

concepts	1	2	3	4	5	6	7	8	9	Total
personnes	a	a,c	d,c,e	d,c	a,c,d,e	b,c,d	b	a,d,e	d,e	
tribu (a,c)	0,5	1	0,25	0,25	0,5	0,25	0	0,25	0	3
tribu (a,d)	0,5	0,33	0,25	0,33	0,5	0,33	0	0,66	0,66	3,58
tribu (a,e)	0,5	0,66	0,5	0	0,5	0	0	0,66	0,66	3,5
tribu (b,c)	0,33	0,33	0,25	0,33	0,20	0,66	0,5	0	0	2,61
tribu (c,d)	0	0,5	0,66	1	0,5	0,66	0	0,25	0,66	4,25
tribu (c,e)	0	0,33	0,66	0,33	0,5	0,25	0	0,25	0,66	3
tribu (d,e)	0	0	0,66	0,66	0,5	0,25	0	0,66	1	3,75
tribu (a,c,d)	0,33	0,66	0,5	0,66	0,75	0,40	0	0,5	0,25	4,07
tribu (a,c,e)	0,33	0,66	0,5	0,25	0,75	0,20	0	0,5	0,25	3,45
tribu (a,d,e)	0,33	0,25	0,5	0,25	0,75	0,20	0	1	0,66	3,95
tribu (b,c,d)	0	0,66	0,5	0,66	0,4	1	0,33	0,2	0,25	4,02
tribu (c,d,e)	0	0,33	1	0,66	0,75	0,75	0	0,75	0,66	4,92
tribu (a,c,d,e)	0,25	0,5	0,75	0,5	1	0,6	0	0,75	0,5	4,85

TAB. 4 – *distances brutes : concepts - tribus*

concepts	1	2	3	4	5	6	7	8	9	Total
personnes	a	a,c	d,c,e	d,c	a,c,d,e	b,c,d	b	a,d,e	d,e	
tribu (a,c)	0,15	0,29	0,07	0,07	0,15	0,07	0	0,07	0	0,87
tribu (a,d)	0,13	0,08	0,06	0,08	0,13	0,08	0	0,17	0,17	0,9
tribu (a,e)	0,17	0,22	0,17	0	0,17	0	0	0,22	0,22	1,16
tribu (b,c)	0,06	0,06	0,04	0,06	0,03	0,11	0,09	0	0	0,44
tribu (c,d)	0	0,07	0,09	0,14	0,07	0,09	0	0,04	0,09	0,6
tribu (c,e)	0	0,19	0,38	0,19	0,29	0,14	0	0,14	0,38	1,71
tribu (d,e)	0	0	0,19	0,19	0,15	0,07	0	0,19	0,29	1,09
tribu (a,c,d)	0,22	0,44	0,33	0,44	0,5	0,27	0	0,33	0,17	2,71
tribu (a,c,e)	0,04	0,09	0,07	0,03	0,1	0,03	0	0,07	0,03	0,45
tribu (a,d,e)	0,04	0,03	0,07	0,03	0,10	0,03	0	0,13	0,09	0,51
tribu (b,c,d)	0	0,17	0,13	0,17	0,1	0,25	0,03	0,05	0,06	1
tribu (c,d,e)	0	0,04	0,13	0,09	0,10	0,40	0	0,10	0,09	0,64
tribu (a,c,d,e)	0,08	0,17	0,25	0,17	0,33	0,2	0	0,25	0,17	1,62

TAB. 5 – *distances pondérées : concepts - tribus*

- Diffusion des photos aux tribus dont le seuil dans le tableau 4 dépasse une valeur donnée pour la tribu. Fixons comme règle pour diffuser une photo à une tribu (c'est à dire à tous ses membres) le fait que la distance d'une photo à cette tribu soit supérieure à 0,5 au sens de la distance de Jaccard, qui indique qu'une photo montre au moins une majorité des individus de la tribu.
- Diffusion des photos aux tribus dont le seuil dans le tableau 5 dépasse une valeur donnée pour la tribu. Ce seuil est déterminé par un critère économique c'est à dire le nombre de photos à diffuser. Prenons par exemple le nombre de 35 photos maximum, le seuil est alors de 0,15.
- Nous Classons les couples "photo-tribu" par ordre décroissant selon la mesure du tableau 5 et nous diffusons les photos en déterminant un seuil à partir duquel on ne diffuse plus les photos aux tribus désignées. Prenons comme seuil les 10 photos les plus importantes

au sens des couples “photo-tribu”. Le classement pour les 10 premiers couples photo-tribu est montré dans le tableau 6.

5. Diffusion de toutes les photos à toutes les personnes.

Les alternatives 1 et 5 sont déjà pratiquées. Nous allons donc comparer toutes ces alternatives et plus particulièrement les solutions 2, 3 et 4. Prenons la solution 2 nous constatons que toutes

priorité	1	2	3	4	5	6	7	8	9	10
score	0,5	0,44	0,44	0,38	0,38	0,33	0,33	0,33	0,29	0,29
photo	(a,c,d,e)	(a,c)	(d,c)	(c,d,e)	(d,e)	(a,d,e)	(c,d,e)	(c,d,e)	(a,c,d,e)	(d,e)
tribu	(a,c,d,e)	(a,c,d)	(a,c,d)	(c,e)	(c,e)	(a,c,d)	(a,c,d)	(a,c,d)	(c,e)	(d,e)

TAB. 6 – Classement des 10 premiers couples photos-tribus par ordre décroissant de distance

les tribus sont sélectionnées pour au moins une photo. Prenons maintenant la solution 3, le tableau 7 montre les photos diffusées. Plusieurs tribus ne sont sélectionnées dans aucune photo : (b,c),(c,d),(a,c,e),(a,d,e),(c,d,e). Selon la solution 4, plusieurs tribus ne sont sélectionnées pour aucune photo : (a,c),(a,d),(b,c),(c,d),(a,c,e),(a,d,e),(b,c,d),(c,d,e). La solution 1 représente 21

concepts	1	2	3	4	5	6	7	8	9
	a	a,c	d,c,e	d,c	a,c,d,e	b,c,d	b	a,d,e	d,e
Alain	X	X	X	X	X	X		X	X
Bernard		X		X		X			
Céline	X	X	X	X	X	X		X	X
Daniel	X	X	X	X	X	X		X	X
Émilie	X	X	X	X	X	X		X	X

TAB. 7 – diffusion des photos pour chaque personne selon la distance brute de Jaccard

photos, la solution 2 représente 38 photos, la solution 3 représente 35 photos, la solution 4 représente 27 photos et la solution complète 5 représente 45 photos. Chacune des solutions présente une stratégie de sélection des photos différentes : soit un critère “anti-social”, soit un critère “semi-social”, soit un critère économique, soit un critère sémantique, soit un critère totalement social. Nous disposons maintenant d’un outil de diffusion d’albums personnalisés qui peut procéder soit de manière automatique, soit sous forme d’assistance à la diffusion manuelle. Les différentes stratégies présentées sont modifiables et montre la richesse du modèle de sélection présenté. De nombreuses voies sont encore à explorer. Nous constatons dans notre exemple, que la “proximité” et la cohésion évoluent dans le même sens. Cependant elles expriment des réalités différentes et des exemples plus grands montreront leur différence.

6 Conclusion

De nombreuses améliorations sont possibles, et les modèles peuvent être étendus en particulier la prise en compte de la sémantique pour l’extraction du réseau social. C’est à dire la prise en compte de certains biais comme celui du photographe, qui n’est pas forcément objectif ou de la personne qui produit les photos ou celle qui diffuse les photos. Nos travaux futurs vont

Des réseaux de photos aux réseaux sociaux.

explorer ces modèles, avec des tests plus approfondis. Le passage à l'échelle sera également étudié par l'utilisation de centaines de photos pour révéler l'évolution des réseaux sociaux dans le temps avec la fusion de plusieurs événements sociaux. Nous exploiterons également les principes énoncés ici pour explorer une autre voie : la construction de réseaux sociaux à partir de l'analyse sémantique de documents partagés entre personnes.

Références

- Crampes, M., J. de Oliveira-Kumar, S. Ranwez, et J. Villerd (2009). Indexation de photos sociales par propagation sur une hiérarchie de concepts. *Actes de la conférence IC 2009, Hammamet Tunisie*, 13–24.
- Eklund, P., J. Ducrou, et T. Wilson (2006). 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*.
- Ferré, S. (2007). 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*, 112–123.
- Freeman, L. et D. White (1993). Using galois lattices to represent network data. *Sociological Methodology* 23, 127–146.
- Ganter, B. et R. Wille (1999). Formal concept analysis. *Mathematical Foundations, Springer*.
- Godin, R. et T. Chau (1999). Comparaison d'algorithmes de construction de hiérarchies de classes. *L objet* 5(3/4).
- Godin, R., G. Mineau, et R. Missaoui (1995). Incremental structuring of knowledge bases. *Proceedings of the International Knowledge Retrieval, Use, and Storage for Efficiency Symposium (KRUSE'95), Santa Cruz* 179-198.
- Golder, S. A. (2008). Measuring social networks with digital photograph collections. *ACM Conference on Hypertext and Hypermedia. June 19-21. Pittsburgh, Pennsylvania*.
- KAUFMAN, L., P. ROUSSEAU, et J. B. P. (1990). *Finding groups in data : An introduction to cluster analysis*. WILEY-Interscience.
- Mika, P. (2005). Ontologies are us : A unified model of social networks and semantics. *In International Semantic Web Conference*, 522–536.
- Roth, C. et P. Bourguin (2006). Lattice-based dynamic and overlapping taxonomies : The case of epistemic communities. *Scientometrics (impact factor : 1.74)* 69(2), 429–447.

Summary

With the new possibilities in communication and information management, social networks and photos have received plenty of attention in the digital age. In this paper, we show how social photos, captured during family events, representing individuals or groups, can be visualized as a network that reveals social attributes. From this photo network, a social network is extracted that can help to build personalized albums. The photo network organization makes use of Formal Concept Analysis methods and a Hasse Diagram representation.