



HAL
open science

De l'influence de l'enrichissement de profil utilisateur sur la propagation de buzz dans les médias sociaux. Expérimentations sur Delicious

Manel Mezghani, André Péninou, Florence Sèdes, Sirinya On-At, Arnaud Quirin, Marie-Françoise Canut

► To cite this version:

Manel Mezghani, André Péninou, Florence Sèdes, Sirinya On-At, Arnaud Quirin, et al.. De l'influence de l'enrichissement de profil utilisateur sur la propagation de buzz dans les médias sociaux. Expérimentations sur Delicious. *Revue des Sciences et Technologies de l'Information - Série ISI: Ingénierie des Systèmes d'Information*, 2016, vol. 21 (n° 4), pp. 67-81. 10.3166/isi.21.4.67-81 . hal-01429680

HAL Id: hal-01429680

<https://hal.science/hal-01429680>

Submitted on 9 Jan 2017

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.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 17252

To link to this article :

URL : <http://isi.revuesonline.com/article.jsp?articleId=36935>

To cite this version : Mezghani, Manel and Péninou, André and Sèdes, Florence and On-At, Sirinya and Quirin, Arnaud and Canut, Marie-Françoise *De l'influence de l'enrichissement de profil utilisateur sur la propagation de buzz dans les médias sociaux. Expérimentations sur Delicious.* (2016) Ingénierie des Systèmes d'Information, vol. 21 (n° 4). pp. 67-81. ISSN 1633-1311

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

De l'influence de l'enrichissement de profil utilisateur sur la propagation de buzz dans les médias sociaux

Expérimentations sur Delicious

Manel Mezghani^{1,2}, André Péninou², Florence Sèdes²
Sirinya On-At², Arnaud Quirin², Marie-Françoise Canut²

1. Université de Sfax, laboratoire MIRACL, Sfax, Tunisie

2. IRIT, Université de Toulouse, CNRS, INPT, UPS, UTI, UT2J, France

{mezghani.manel, aquirin}@gmail.com

{andre.peninou, florence.sedes, sirinya.on-at}@irit.fr

marie-francoise.canut@univ-tlse2.fr

RÉSUMÉ. L'utilisateur est la source principale de l'information diffusée dans les médias sociaux et mais il est, en même temps, influencé par les informations partagées sur les réseaux. Le phénomène de buzz, c'est-à-dire faire « du bruit » autour d'une information (fait ou rumeur) pour que plusieurs utilisateurs soient intéressés par celle-ci simultanément, peut être défini comme une information populaire dans un temps spécifique. Notre étude porte sur l'influence de l'enrichissement dynamique de profil utilisateur (Mezghani et al., 2014) sur la propagation des buzz avec application au réseau social Delicious. Delicious contient des annotations sociales (tags) fournies par les utilisateurs et qui contribuent à influencer les autres utilisateurs afin de suivre certaines informations ou de les utiliser. Notre étude suit la méthodologie suivante : 1) nous analysons la propagation des tags considérés comme des buzz dans le temps, 2) nous appliquons l'enrichissement dynamique de profil utilisateur et nous étudions l'influence de cet enrichissement dans la propagation de buzz, 3) nous analysons si l'approche d'enrichissement anticipe la propagation de buzz. Nous montrons dans cet article l'intérêt, lors de l'enrichissement de profil, de filtrer les informations afin de proposer des résultats pertinents à l'utilisateur et éviter de « mauvaises » recommandations.

ABSTRACT. The user is the main contributor for creating information in social media and is influenced by the information shared through such networks. There are so-called "buzz", which is a technique to cause a stir around a piece of information (fact or rumour) so that several users will be interested in simultaneously, can be defined as a popular piece of information in a specific time. We are interested in studying the influence of the dynamic user profile enrichment (Mezghani et al., 2014) on the buzz propagation and we experiment it to the social network Delicious. Delicious contains social annotations (tags) provided by users and that contribute to influence other users to follow some information or to use them. Our study is grounded on the following methodology: 1) we analyse the propagation of tags considered as buzz through time 2) we apply the dynamic user profile enrichment and we

analyse the influence of this enrichment in the buzz propagation, 3) we analyse if the enrichment approach anticipates the buzz propagation. Thus, we show interest, during profile enrichment, of filtering the information in order to propose relevant results to the user and avoid “bad” recommendations.

MOTS-CLÉS : profil utilisateur, enrichissement, tag, ressource, buzz, temps.

KEYWORDS: user profile, enrichment, tag, resource, buzz, time.

1. Introduction

Dans un contexte social, le « buzz » est une technique pour créer et entretenir du bruit autour d'un événement : un buzz est une information « populaire » sur une durée donnée. De fait, cela sous entend que plusieurs utilisateurs à la fois seront intéressés par cet événement. Un buzz peut être un fait (information) ou une rumeur (information fausse). L'origine de ce mécanisme repose sur le fait que les utilisateurs des médias sociaux sont influencés par les informations partagées dans le réseau.

(Rosnow et Kimmel, 2000) définissent une rumeur comme une proposition non vérifiée à prétention de vérité activement diffusée. Selon (Hashimoto *et al.*, 2011), une rumeur est caractérisée par sa diffusion rapide. Cependant, la détection de rumeur est un problème crucial puisqu'il exige une connaissance supplémentaire qui vérifie des informations/propositions.

Dans ce papier, nous proposons d'étudier un buzz comme étant une rumeur potentielle. Nous nous sommes intéressés à la propagation dans le temps du buzz dans le réseau social *Delicious*¹ (plus précisément un ensemble de données de *Delicious* (Cantador *et al.*, 2011)). Nous nous sommes aussi intéressés à l'étude de l'influence d'enrichissement dynamique de profils utilisateurs, proposée dans (Mezghani *et al.*, 2014), sur cette propagation.

Notre approche d'enrichissement dynamique considère la dynamique temporelle du réseau social : l'enrichissement du profil utilisateur est effectué à chaque période de temps (période prédéterminée), et non par accumulation d'enrichissements antérieurs issus des périodes précédentes. Cette approche d'enrichissement prend en considération la popularité de l'information (du tag), la fraîcheur de l'information (du tag) et la similarité des utilisateurs annotant la même ressource dans une période de temps donnée.

Dans ce papier, nous étudions quelques informations du réseau social *Delicious*. *Delicious* contient des annotations sociales (tags) posées sur des documents et

1. www.delicious.com

fournies par les utilisateurs. Ces tags contribuent à influencer d'autres utilisateurs pour « suivre » ces informations ou pour les utiliser.

Cette étude est effectuée selon la méthodologie suivante :

1) nous nous concentrons sur des tags considérés comme des buzz (des tags populaires) et analysons leur propagation au cours du temps ;

2) nous considérons qu'un profil utilisateur est constitué de l'ensemble des tags fournis par l'utilisateur lui-même. Nous utilisons le résultat de nos travaux sur l'enrichissement dynamique de profil utilisateur (Mezghani *et al.*, 2014) pour analyser l'influence de cet enrichissement sur la propagation de buzz ;

3) nous analysons chaque profil utilisateur enrichi pour vérifier si l'approche d'enrichissement anticipe ou non la propagation de buzz. Ainsi, nous étudions l'intérêt de filtrer les informations pour éviter d'intégrer dans les profils des rumeurs potentielles et, ainsi, proposer des résultats pertinents (appropriés) à l'utilisateur et éviter, par exemple, de « mauvaises » recommandations.

Ce papier est structuré comme suit. En premier lieu, nous donnons une vue d'ensemble de l'approche d'enrichissement dynamique. Ensuite, nous détaillons l'ensemble de données utilisées dans *Delicious*, en étudiant quelques cas de propagation de buzz dans le temps avec et sans l'approche d'enrichissement. Puis nous analysons si l'approche d'enrichissement anticipe la propagation de buzz. Enfin, nous concluons et donnons quelques perspectives.

2. Aperçu de l'approche d'enrichissement dynamique

Dans cette section, nous donnons un aperçu de l'approche validée dans (Mezghani *et al.*, 2014) qui permet d'enrichir les profils des utilisateurs. L'évolution dynamique du profil de l'utilisateur est traitée en enrichissant ses intérêts avec des tags jugés pertinents pour chaque période de temps. Dans un environnement social, l'utilisateur consulte les ressources stockées dans le réseau, communique et interagit avec d'autres utilisateurs afin de trouver l'information dont il a besoin. L'enrichissement dans ce contexte se fait par analyse de l'environnement de l'utilisateur pour détecter les intérêts concernés (tags pertinents).

L'environnement social de chaque utilisateur est considéré, dans ce contexte, comme étant constitué des informations : tags, ressources (nous considérons que les ressources sont des données semi-structurées) et amis. En effet, selon (Meo *et al.*, 2014), l'utilisation de tags dénote implicitement les intérêts de l'utilisateur. Analyser les tags de l'utilisateur constitue donc un puissant outil de gestion des connaissances. Les ressources peuvent être notées (évaluées) et ceci reflète le degré d'intérêt de l'utilisateur vis-à-vis de ces ressources (Kim *et al.*, 2011 ; Joly *et al.*, 2010). L'information associée aux amis a déjà montré son efficacité pour détecter les intérêts d'un utilisateur (Tchunte *et al.*, 2013 ; Zheng et Li, 2011).

La pertinence d'un intérêt est généralement calculée à partir de la fréquence d'utilisation du tag sur une période. La fréquence varie périodiquement. Nous avons

traité cette représentation en s'inspirant du travail de (Manzat *et al.*, 2010), à travers le concept de « température ». Cette notion est intéressante car elle modélise la variation de popularité d'un terme au fil du temps.

Le profil utilisateur est construit de manière implicite en utilisant la liste des tags fournie par l'utilisateur. Le profil utilisateur est enrichi avec des tags (considérés comme ses intérêts) dans chaque période de temps afin de refléter les intérêts actuels de l'utilisateur.

La première étape du processus d'enrichissement consiste à diviser une base de données selon chaque période de temps Δt (voir figure 1). Le choix de cette période est stratégique pour détecter l'évolution des intérêts des utilisateurs entre deux périodes successives. Cette période devrait être compatible avec la quantité de données présente dans le réseau social. En divisant une base de données, nous obtenons des informations temporelles sur l'activité de l'utilisateur dans chaque période, comme ses voisins, ses tags et les ressources annotées. La durée d'une période de temps Δt sera fixée lors des expérimentations.

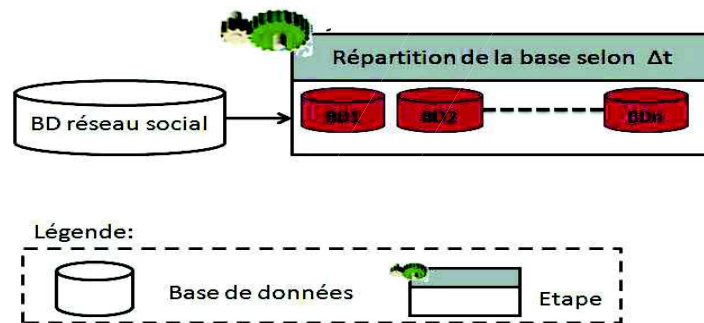


Figure 1. Préparation des données

La deuxième étape consiste à calculer la température de chaque ressource sur une période donnée (issue de l'étape précédente). Pour calculer cette métadonnée, nous proposons la formule (1), qui prend en compte plusieurs paramètres : la fraîcheur d'un tag associé à la ressource (r), la similarité des utilisateurs qui ont annoté la même ressource dans la même période de temps et le nombre de tags associés à la ressource (popularité).

$$\text{Temperature } \Delta t (r) = \alpha * \text{fraîcheurTag} + \beta * \text{similaritéUtilisateurs} + \gamma * \text{popularitéTag} \quad (1)$$

α , β , γ sont des constantes ayant une valeur positive. Ces constantes reflètent le degré de l'influence de chaque paramètre (fraîcheur du tag, les utilisateurs similaires et la popularité des ressources). La température de la ressource varie au cours du temps. Elle peut augmenter ou diminuer. Nous considérons qu'une ressource est intéressante si sa température augmente entre deux périodes de temps.

La troisième étape consiste à détecter les ressources dont la température augmente : après avoir calculé la température de chaque ressource, nous ne considérons que celles dont la température augmente entre deux périodes de temps successives, reflétant l'intérêt croissant des utilisateurs pour cette ressource. Cependant, dans les réseaux sociaux qui se caractérisent par la grande quantité des ressources présentes, beaucoup de ressources peuvent voir leur température augmenter. Donc, le traitement de ces ressources peut être complexe. Afin de surmonter un tel problème, nous proposons de ne considérer que les ressources les plus pertinentes pour l'utilisateur. À cette fin, nous analysons le contenu des ressources et plus précisément leurs métadonnées. Nous utilisons donc les métadonnées qui décrivent le contenu de la ressource, afin de filtrer les ressources les plus pertinentes, en attribuant un poids aux tags associés aux ressources. Ce poids est calculé en fonction du degré de correspondance des tags avec les métadonnées de la ressource associée selon la formule (2) proposée par (Joly *et al.*, 2010) :

$$W(\text{tag}, r) = \alpha' * |\text{tag} \in T_r| + \beta' * |\text{tag} \in K_r| + \gamma' * |\text{tag} \in D_r| \quad (2)$$

Cette fonction compte le nombre d'occurrences de chaque tag sur une ressource r en appliquant les coefficients α' , β' et γ' selon l'emplacement du tag dans les métadonnées de la ressource. $|\text{tag} \in T_r|$ est le nombre d'occurrences du tag dans l'élément de titre (*< title >*) de la ressource, $|\text{tag} \in K_r|$ est le nombre d'occurrences du tag dans l'élément de mots-clés (*< keywords >*) de la ressource, $|\text{tag} \in D_r|$ est le nombre d'occurrences du tag dans l'élément de description (*< description >*) de la ressource.

La quatrième étape consiste à enrichir le profil de l'utilisateur avec les mots-clés associés aux ressources. Après avoir calculé le poids des tags associés aux ressources les plus pertinentes, nous enrichissons, dans cette étape, le profil de l'utilisateur avec les tags qui reflètent au mieux ses intérêts. Plus le tag a un poids élevé, mieux il reflète le contenu de la ressource, et donc mieux il reflète les intérêts de l'utilisateur. Nous choisissons à partir du résultat de l'étape précédente, les tags qui sont les plus pertinents pour l'utilisateur. Un tag est indiqué comme un intérêt potentiel si son poids est supérieur à un seuil donné (défini lors des expérimentations).

À l'issue de ce processus, un profil enrichi est disponible pour chaque période de temps. La figure 2 récapitule les étapes expliquées ci-dessus (hormis la division de la base de données déjà expliquée à travers la figure 1).

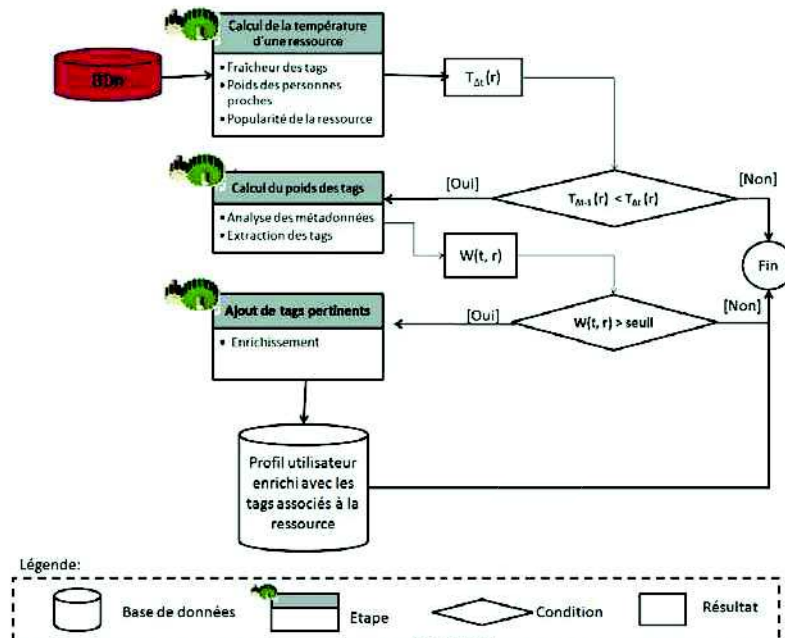


Figure 2. Les étapes d'enrichissement

3. Étude de cas sur un ensemble de données de *Delicious* sur la propagation de buzz

Dans cette section, nous présentons d'abord l'ensemble de données utilisées dans notre expérimentation. Nous analysons ensuite l'évolution du top-10 buzz (tags populaires) à travers le temps. Puis, nous analysons l'influence de l'approche d'enrichissement sur la propagation des top-10 buzz. Enfin, nous analysons si l'enrichissement du profil utilisateur anticipe la propagation du buzz.

3.1. Présentation de la base de données *Delicious*

La base de données *Delicious* contient le réseau égo-centrique² de chaque utilisateur, des marques-pages et des tags pour chaque utilisateur. Les utilisateurs U sont décrits par leur ID (Identifiant) par exemple : ID utilisateur=8. Les ressources R sont décrites par leur ID, leur titre et leur URL par exemple : 1 IFLA – le site web officiel des Internationaux fédération d'associations de bibliothèque et institutions <http://www.ifla.org/>. Les tags T sont décrits par leur ID et leur valeur par exemple : 1 développement. La base de test contient :

2. L'ensemble des individus avec qui l'utilisateur est en relation explicite directe.

- 1 867 utilisateurs,
- 7 668 relations bidirectionnelles et une moyenne de 8,236 relations par utilisateur. Ces relations sont les relations explicites d’amitié,
- 69 226 URL dont 38 581 URL principales,
- 53 388 tags, 437 593 *tag assignments* (tas), sous forme de tuples [user, tag, URL], et une moyenne de 234,383 tas par URL et une moyenne de 6,321 tas par tags,
- 104 799 bookmarks (considérés comme les ressources), une moyenne de 56,132 URL annotées par utilisateur et une moyenne de 1,514 utilisateur annotant une URL.

Le comportement d’annotation est fourni en fonction du temps. Cette information permet de connaître quel tag a été utilisé dans quelle période de temps. Un exemple de comportement d’annotation temporel est présenté dans le tableau 1.

Tableau 1. Un exemple de comportement d’annotation temporel

ID utilisateur	ID bookmark	ID tag	jour	mois	année	heure	min	seconde
8	1	1	8	11	2010	23	29	22

3.2. Suivi de l’évolution du buzz

Dans cette section, nous présentons l’évolution des tags considérés comme buzz dans le réseau social *Delicious* entre 2003 et 2010. Dans ce travail, nous considérons le top-10 des tags les plus populaires comme étant des buzz sur l’ensemble des données étudiées.

Tableau 2. Les top-10 tags les plus populaires dans Delicious entre 2003 et 2010

Tag	Design	Tools	Vidéo	Education	Webdesign	Web	Inspiration	Art	Web20	Google
Popularité	4060	2929	2236	2041	1907	1733	1723	1691	1653	1648

L’évolution de chaque tag est présentée comme un graphique de sa popularité (nombre d’utilisations) sur un axe temporel. Dans cette étude, nous utilisons la granularité de mois pour étudier l’évolution de chaque buzz. Le graphique de visualisation est présenté dans la figure 3.

Nous pouvons voir que la plupart des tags étudiés représentent la caractéristique du buzz : leur popularité augmente légèrement au début puis explose à une période de temps et diminue par la suite. Nous observons que la période la plus populaire de tous les tags étudiés est autour de septembre-octobre 2010.

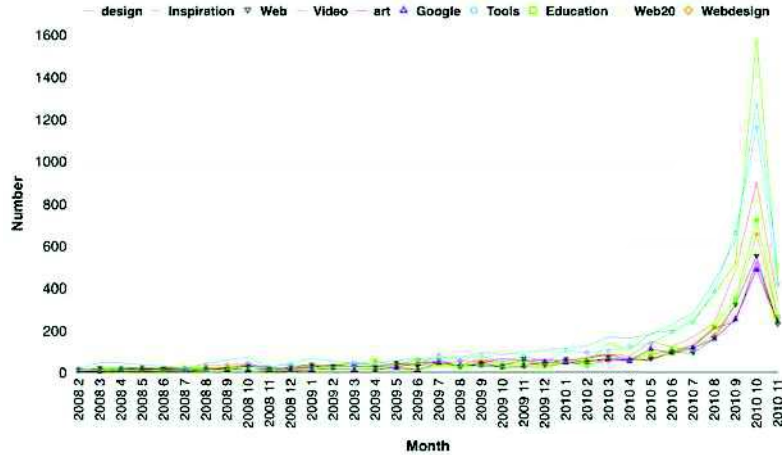
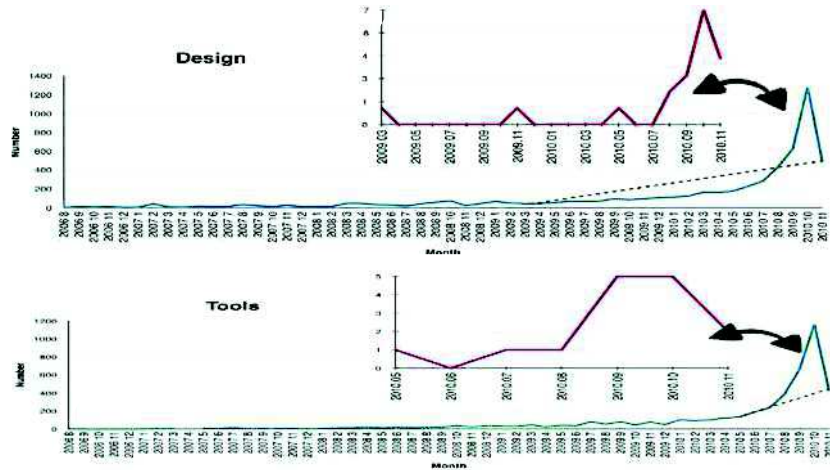


Figure 3. L'évolution des top-10 tags populaires dans Delicious entre 2003 et 2010

3.3. Analyse de l'influence de l'enrichissement du profil de l'utilisateur sur la propagation de buzz

Pour analyser l'impact de l'enrichissement du profil de l'utilisateur sur la propagation de buzz, nous nous sommes intéressés à la corrélation entre la propagation de buzz dans le jeu de données et la propagation de buzz dans les profils utilisateurs enrichis. En fait, nous n'analysons que le résultat de l'approche d'enrichissement (pas la totalité du profil enrichi) avec la propagation de buzz.



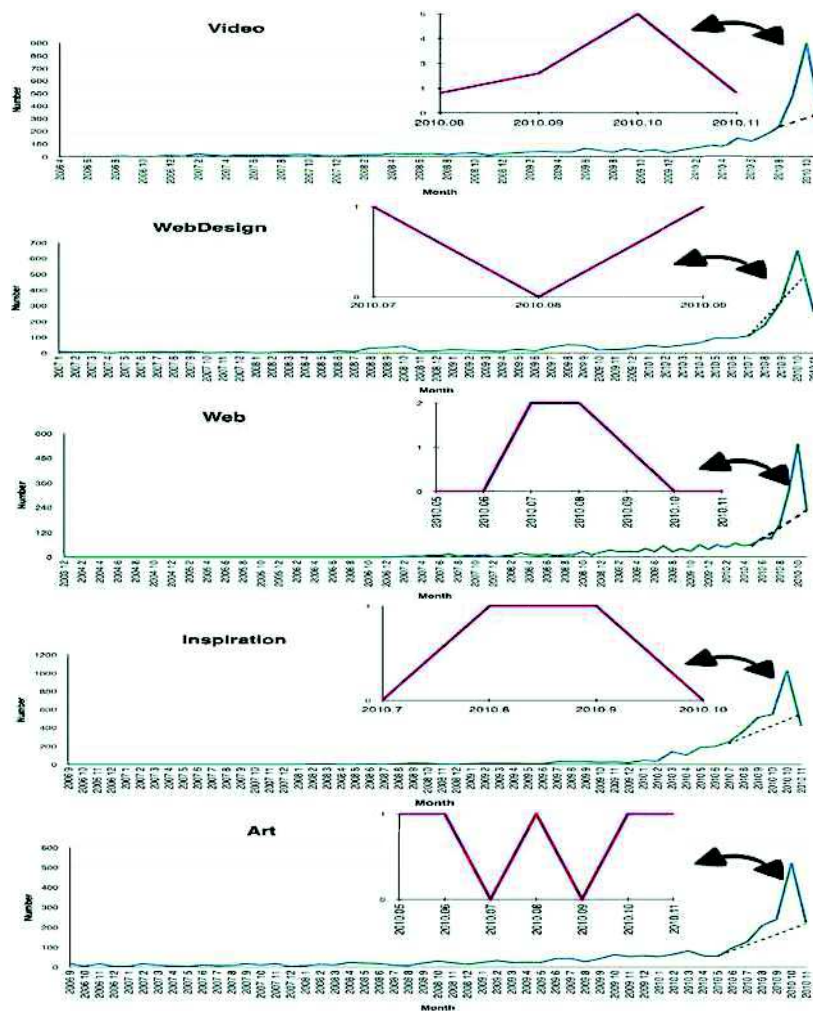


Figure 4. (Bleu) La propagation du buzz dans la base de données
(Rouge) la propagation du buzz dans les profils enrichis

Parmi les top-10 tags étudiés dans la section précédente, nous ne trouvons que 8 tags dans les résultats d'enrichissement (pour tous les utilisateurs dans l'ensemble de données). Les graphiques de visualisation de ces 8 tags sont présentés dans la figure 4. Celle-ci représente, pour chaque tag, sa popularité (nombre d'utilisations) dans *Delicious* et son utilisation dans l'enrichissement de profils (nombre de fois où il est utilisé pour enrichir un profil).

Les graphiques (figure 4) ci-dessus montrent que les tags étudiés sont principalement retenus dans le profil de l'utilisateur après le processus

d'enrichissement dans la période au cours de laquelle ils deviennent populaires. Par exemple, le tag *Google* est retenu dans le profil de l'utilisateur entre septembre-novembre 2010, la période pendant laquelle le tag est le plus utilisé par l'ensemble des utilisateurs de ce réseau social.

Cette analyse illustre que la popularité des tags peut être un facteur important dans le processus d'enrichissement de profil d'utilisateur. Si le tag est un buzz pendant une période, il a plus de chance d'être extrait dans le processus d'enrichissement de profil d'utilisateur pour cette période. Donc, le processus d'enrichissement de profil utilisateur peut contribuer à la propagation des buzz (comme les rumeurs potentielles) dans les réseaux sociaux.

3.4. L'enrichissement du profil utilisateur peut-il anticiper la propagation du buzz ?

Tout au long de l'analyse précédente, nous avons analysé la propagation de buzz dans l'ensemble du réseau, indépendamment du profil de l'utilisateur. Dans cette section, nous analysons chaque profil utilisateur enrichi afin de montrer si l'approche d'enrichissement peut anticiper la propagation de buzz.

Tableau 3. Analyse du tag « Design » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1^{re} utilisation après enrichissement
1094	8/10/2010	9	1	03/11/2010
1113	27/08/2010	11	51	29/08/2010
1113	29/08/2010	14	48	30/08/2010
16915	30/03/2009	1	6	16/04/2009
24802	10/11/2009	9	4	18/11/2009
8315	31/05/2010	1	2	24/06/2010
62070	20/09/2010	10	0	-
9960	28/09/2010	0	0	-
51543	30/09/2010	9	3	09/10/2010
2032	01/10/2010	6	0	-
8691	12/10/2010	32	38	13/10/2010
3233	21/10/2010	12	1	22/10/2010
1296	26/10/2010	0	0	-
11699	09/10/2010	0	3	01/11/2010
1701	09/10/2010	10	2	13/10/2010
15728	09/10/2010	29	17	11/10/2010
13222	03/11/2010	3	0	-
8452	04/11/2010	5	0	-
6067	04/11/2010	7	0	-

Pour chaque buzz trouvé dans le résultat de l'enrichissement, nous détaillons l'ID utilisateur associé, la date d'enrichissement, le nombre d'occurrences du tag (buzz) pour l'utilisateur avant la date d'enrichissement, le nombre d'occurrences du tag (buzz) pour l'utilisateur après la date de l'enrichissement et la date de la première utilisation après enrichissement par l'utilisateur. Ces résultats sont détaillés dans les tableaux 3, 4, 5, 6, 7, 8, 9 et 10 suivants.

Tableau 4. Analyse du tag « Tools » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1^{re} utilisation après enrichissement
8315	31/05/2010	1	20	01/06/2010
6120	31/07/2010	1	5	24/08/2010
46715	29/08/2010	1	6	12/10/2010
35745	16/09/2010	18	30	17/09/2010
11699	24/09/2010	2	34	26/09/2010
1328	29/09/2014	0	0	--
7396	19/10/2010	2	5	20/10/2010
2315	21/10/2010	7	2	27/10/2010
8554	22/10/2010	1	13	25/10/2010
70894	27/10/2010	12	6	01/11/2010
1505	29/10/2010	11	0	29/10/2010
13102	05/11/2010	8	2	06/11/2010
23135	06/11/2010	16	0	06/11/2010

Tableau 5. Analyse du tag « Video » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1^{re} utilisation après enrichissement
74708	16/08/2010	7	9	23/08/2010
13084	14/09/2010	6	8	15/09/2010
4742	30/09/2010	2	4	06/10/2010
6796	12/10/2010	16	1	21/10/2010
8452	20/10/2010	4	0	--
11690	20/10/2010	7	0	--
8775	21/10/2010	2	1	21/10/2010
1701	21/10/2010	12	2	04/11/2010
12847	02/11/2010	2	0	--

Tableau 6. Analyse du tag « Webdesign » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1 ^{re} utilisation après enrichissement
6120	23/07/2010	2	0	--
9660	28/09/2010	1	3	01/10/2010

Tableau 7. Analyse du tag « Web » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1 ^{re} utilisation après enrichissement
13973	16/07/2010	2	12	02/08/2010
12506	05/07/2010	1	0	--
1113	27/08/2010	1	17	29/08/2010
1113	29/08/2010	3	15	31/08/2010
13084	14/09/2010	1	0	--

Tableau 8. Analyse du tag « Inspiration » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1 ^{re} utilisation après enrichissement
1113	27/08/2010	7	22	29/08/2010
51543	30/09/2010	7	2	9/10/2010

Tableau 9. Analyse du tag « Art » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1 ^{re} utilisation après enrichissement
31272	25/05/2010	8	5	09/06/2010
11962	30/06/2010	2	55	02/07/2010
10567	26/08/2010	1	1	09/11/2010
1701	27/10/2010	6	1	06/11/2010
8452	04/11/2010	4	0	--

Tableau 10. Analyse du tag « Google » selon chaque utilisateur

ID utilisateur	Date enrichissement	Occurrence avant la date enrichissement	Occurrence après la date enrichissement	Date de la 1 ^{re} utilisation après enrichissement
1505	05/09/2010	2	1	13/09/2010
11853	05/11/2010	7	0	--

De l'analyse de ces tableaux nous constatons que :

1) En ce qui concerne l'occurrence avant/après la date d'enrichissement : elle varie selon les différents cas. En fait, nous remarquons que les utilisateurs qui n'ont utilisé le tag qu'avant l'enrichissement sont environ 23,63 %, les utilisateurs qui ne l'ont utilisé qu'après l'enrichissement sont environ 1,81 %, les utilisateurs qui l'ont utilisé avant et après l'enrichissement sont environ 69,09 % et les utilisateurs qui n'ont jamais utilisé le tag et dont nous avons enrichi le profil avec celui-ci sont environ 5,45 %. Nous pouvons conclure que l'approche d'enrichissement est en quelque sorte liée avec l'activité précédente d'un utilisateur. Toutefois, le nombre de ces buzz trouvés dans les résultats de l'enrichissement est relativement faible par rapport à leur popularité dans l'ensemble de données initial.

2) En ce qui concerne la date de la première utilisation d'un tag après enrichissement : cette date vise à montrer la capacité de l'approche d'enrichissement à anticiper le buzz. Plus l'intervalle entre la date de l'enrichissement et la première utilisation du buzz est grand, plus le buzz est anticipé. Selon ces tableaux, nous constatons que la valeur minimale de l'anticipation est de 0 jour (nous enrichissons le jour même d'une activité en cours) et concerne 4,8 % de tous les cas. La valeur maximale de l'anticipation est de 75 jours (associée à l'ID utilisateur = 10567 dans le tableau 9). La valeur moyenne de l'anticipation est de 9 jours et la valeur médiane de l'anticipation est de 5 jours.

4. Conclusion

Dans cet article, nous avons fait une étude de cas concernant la propagation de buzz sur des informations du réseau social *Delicious*. Ce dernier contient des annotations sociales (tags) qui sont fournies par les utilisateurs. Ces tags contribuent à influencer les autres utilisateurs à suivre ces informations ou à les utiliser.

Cette étude repose sur la méthodologie suivante :

1) Nous nous sommes concentrés sur les tags considérés comme buzz et nous avons analysé leur propagation dans le temps. Dans cette analyse, nous avons remarqué que le nombre d'utilisateurs dans le réseau influence la propagation. Plus un utilisateur est actif dans des périodes spécifiques, plus le buzz est présent dans ces périodes.

2) Nous avons considéré un profil utilisateur comme l'ensemble des tags fournis par l'utilisateur lui-même. Nous avons utilisé le résultat de nos travaux sur l'enrichissement temporel du profil utilisateur afin d'analyser l'influence de cet enrichissement dans la propagation du buzz. Nous avons remarqué que le processus d'enrichissement contribue à propager le buzz dans presque tous les cas (8 tags sur 10 ont été trouvés dans le résultat de l'enrichissement). Ainsi, l'enrichissement contribue à propager le buzz dans le réseau social.

3) Nous avons également analysé chaque profil utilisateur enrichi afin de montrer que l'approche d'enrichissement permet d'anticiper la propagation de buzz.

Ainsi, nous pouvons voir l'intérêt de filtrer les informations afin d'éviter d'intégrer dans les profils des rumeurs potentielles et, ensuite, de proposer des résultats pertinents pour l'utilisateur en évitant, par exemple, de « mauvaises » recommandations. Nous avons trouvé que l'approche d'enrichissement est en quelque sorte liée avec l'activité précédente d'un utilisateur. C'est à dire, la majorité des utilisateurs dont leur profil a été enrichi, ont utilisé un tag considéré comme buzz, avant et après l'enrichissement. En outre, le nombre de buzz trouvés dans les résultats de l'enrichissement est relativement faible par rapport à leur popularité dans l'ensemble de données initial. L'anticipation varie de 0 à 75 jours. La moyenne d'anticipation est de 9 jours.

Afin de réduire la propagation des buzz qui peuvent être des rumeurs potentielles, nous envisageons de prendre en compte un processus de filtrage de buzz avant d'appliquer notre approche d'enrichissement. En outre, nous prévoyons d'élargir cette étude de cas à plus de 10 tags. Ainsi, nous pourrions étudier l'évolution d'autres buzz ainsi que l'influence de l'approche d'enrichissement sur la propagation de buzz.

Remerciements

Ce travail a été soutenu financièrement par le programme PHC Utique du ministère français des Affaires étrangères et le ministère de l'Enseignement supérieur et la Recherche et le ministère tunisien de l'Enseignement supérieur et la Recherche scientifique sous le numéro de projet CMCU 30540XX.

Bibliographie

- Cantador I., Brusilovsky P., Kuflik T. (2011). 2nd Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec 2011), *Proceedings of the 5th ACM conference on Recommender systems (RecSys'11)*, ACM, New York, NY, USA.
- Hashimoto T., Kuboyama T., Shirota Y. (2011) Rumor analysis framework in social media, *TENCON 2011-2011 IEEE Region 10 Conference*, p. 133-137.
- Joly A., Maret P., Daigremont J. (2010). Contextual recommendation of social updates, a tag-based framework. *In Proceedings of the 6th International Conference on Active Media Technology, AMT'10*, p. 436-447, Berlin, Heidelberg. Springer-Verlag.
- Kim H.-N., Alkhalidi A., El Saddik A., Jo G.-S. (2011). Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications*. vol. 38, n° 7, p. 8488-8496.
- Manzat A., Grigoras R., Sèdes F. (2010). Towards a User-aware Enrichment of Multimedia Metadata, *Workshop on Semantic Multimedia Database Technologies (SMDT'10)*, Saarbrücken, Germany, vol. 680, CEUR Workshop Proceedings, p. 30-41.
- Meo P. D., Ferrara E., Abel F., Aroyo L., Houben G.-J. (2014). Analyzing user behavior across social sharing environments. *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol 5, n° 1, p. 14, 2013, ACM.

- Mezghani M., Zayani C-A., Amous I., Péninou A., Sèdes F. (2014). Dynamic Enrichment of Social Users' Interests. *IEEE Eighth International Conference on Research Challenges in Information Science (RCIS'14)*. IEEE, p. 1-11.
- Rosnow R.L., Kimmel A.J. (2000). Rumor. *Encyclopedia of Psychology*, vol. 7, edited by A.E. Kazdin. New York: Oxford University Press, p. 122-123.
- Tchunte D., Canut M-F., Jessel N., Peninou A., Sedes F. (2013). A community-based algorithm for deriving users' profiles from egocentrics networks : experiment on facebook and DBLP. *Social Network Analysis and Mining*, vol. 3, n° 3, p. 667-683.
- Zheng N., Li Q. (2011). A recommender system based on tag and time information for social tagging systems. *Expert Syst. Appl.*, vol. 38, n° 4, p. 4575-4587.