



HAL
open science

Le contrôle de la forme des réseaux par leurs membres

Alexandre Delanoë, Bernard Conein

► **To cite this version:**

Alexandre Delanoë, Bernard Conein. Le contrôle de la forme des réseaux par leurs membres : Le fil de discussion comme réseau d'interaction. 2013. hal-00979580

HAL Id: hal-00979580

<https://hal.science/hal-00979580>

Preprint submitted on 16 Apr 2014

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.

Le contrôle de la forme des réseaux par leurs membres: le fil de discussion comme réseau d'interaction

Auteurs :

Alexandre Delanoë (*) et Bernard Conein

Affiliations :

Alexandre Delanoë est chercheur au CNRS/EHESS, laboratoire CAMS (Centre d'Analyse et de Mathématiques Sociales, 190-198 avenue de France, 75244 Paris Cedex 13)

et associé au à Mines ParisTech, laboratoire CSI (Centre de Sociologie de l'Innovation), membre de l'Institut des Systèmes Complexes.

Bernard Conein est professeur à l'Université Nice Sophia Antipolis, laboratoire GREDEG, 250 rue Albert Einstein, 06560 Valbonne

(*) Auteur correspondant :

Alexandre Delanoë.

Mail : alexandre@delanoe.org .

Tel : 07 81 12 53 49.

Fax : 09 58 58 25 08

Le contrôle de la forme des réseaux par leurs membres: le fil de discussion comme réseau d'interaction

Abstract:

Cet article enquête sur les modes de coordination sur les listes de discussions électroniques en montrant qu'un fil de discussion peut être identifié comme un réseau de répliques s'étendant à partir de la dynamique des processus d'interaction. L'observation sur une durée de 10 ans et l'analyse de 20 ans d'archives des listes de diffusion d'un projet de logiciel libre permettent d'appréhender le phénomène aussi bien au niveau des interactions locales que sur la longue durée. L'analyse des processus d'interaction ne permet pas seulement de visualiser des coordinations que l'on ne voyait pas avant mais aussi et surtout de comprendre le processus génératif d'élaboration des connaissances à partir des alignements entre les contributeurs.

Mots clefs : interaction, réseau, contrôle, identité, logiciel libre, listes de discussion électroniques

Les archives de listes de diffusion se présentent comme un dispositif d'enregistrement de traces d'interaction restituables sous forme de séquences d'actions au sein d'un fil diachronique. Ces dispositifs possèdent des propriétés dynamiques qui peuvent être étudiées pour elles-mêmes (Lazer *et al.* 2009) et que l'on retrouve sur une longue durée¹ grâce au fonctionnement collaboratif des projets de logiciels libres (Ducheneaut, 2005). Des observations peuvent en effet être construites en s'appuyant sur des enregistrements d'interactions numérisées (dans notre cas sous forme de fils de discussion, *threads*) et non sur des liens rapportés ou déclarés (*reported data*). Plus précisément, les technologies des listes de discussion (*mailing-list*), les encyclopédies collaboratives (*wiki*) et les journaux en ligne (*blogs*) ont des effets sur la façon dont les interactions à distance sont représentées ; le plus souvent de façon statique et non évolutive. D'autre part, ces technologies transforment non seulement l'enregistrement des interactions sociales en préservant la dynamique temporelle mais de façon plus

1 Le projet GNU vient de fêter ses 30 ans et le projet Debian ses 20 ans.

générale elles rendent manifeste des échelles de coordination à plusieurs niveaux et donc l'évolution de la forme des collectifs du local au global.

Le premier point a été noté par le manifeste paru dans *Science* par Lazer et alii en 2009 qui prônait la formation d'une science sociale dite « computationnelle »:

« Un champ de recherche émerge offrant la possibilité de collecter et d'analyser des données à une échelle qui révèle des patterns de comportements individuels et collectifs. (...) Les nouvelles technologies rendent une image processuelle des interactions sur une période de temps étendue, fournissant des informations à la fois sur la structure et le contenu des relations » (Lazer *et al.*, 2009).

Le second point a été moins souligné. La nature processuelle des collectifs devient représentable à plusieurs niveaux, notamment par les membres de la liste eux-mêmes lorsqu'ils travaillent de façon collaborative pour développer une discussion. Le point que nous voulons souligner est le suivant. Les contributeurs du logiciel du libre tendent à contrôler la façon dont leurs interactions évoluent en choisissant le placement de leurs messages dans un fil de discussion. C'est ainsi que leur intervention a pu être caractérisée comme un marqueur d'identité (White, 2011) car les listes tendent à être des dispositifs permettant d'étendre la capacité de contrôle des contributeurs².

1. La spécificité du fil de discussion comme réseau de messages

Plusieurs caractéristiques des listes de diffusion offrent des opportunités de contrôle par les contributeurs. Les listes de discussion rendent possibles des échanges collaboratifs de messages dans l'objectif de résoudre des problèmes informatiques qui

² Selon White (White, 2011), le contrôle se mesure à la capacité d'action d'un agent par rapport à un autre.

réclament une approche exploratoire du système d'exploitation et de son intégration au réseau. La multitude des messages y est cartographiée grâce à un protocole relativement standardisé qui en facilite l'analyse au moment où un utilisateur intervient. Chaque message dispose d'un système de coordonnées multidimensionnel permettant de positionner le message dans une suite d'autres messages: le sujet (titre du message), le temps (la date d'expédition du message), l'index du message directement antérieur (ou non dans le cadre de l'ouverture d'un fil) et la suite des messages précédents auxquels le message fait référence³.

La position d'un message dans une série l'inscrit dans une séquence dont l'unité relationnelle de base est une interaction dyadique questionneur/répondant. Mais les listes ne constituent pas seulement une ressource pour observer des messages, elles permettent de capter des propriétés d'interaction. Des traces d'interaction sont en effet enregistrées et rendues visibles pour tout contributeur au moyen de logiciels dédiés révélant la forme d'un fil de discussion autour d'un centre d'intérêt. La forme temporelle ordonnée de cet ensemble de messages rend manifeste une structure représentable sous la forme d'un graphe évolutif. Cette structure est strictement un arbre hiérarchique orienté: chaque message est directement relié au message antécédent. Le défi est de restituer cette dynamique temporelle inscrite partiellement dans les données elles-mêmes.

L'attente implicite des participants au fil de discussion est la résolution d'un problème dont le cadre est d'abord publiquement formulé à l'ouverture du fil, c'est-à-dire un message initial qui est souvent reformulé et/ou précisé par la suite dans le fil. À la résolution du problème initial, le fil se clôt (virtuellement puisque le lien n'est pas matérialisé), la structure se stabilise, l'arbre est bouclé, le réseau s'actualise comme une suite de messages.

Ainsi, un réseau de fil se présente à son lecteur comme une structure agencée sous la forme de relations de répliques dont les nœuds sont des messages. En d'autres

3 Champs « References: » des headers.

termes, la version statique du fil est le réseau, la version dynamique du fil est le processus d'interaction entre les auteurs (Livet et Nef, 2009). Pouvoir compter sur le bouclage d'un fil (sa résolution et donc sur sa clôture) semble être la forme la plus élémentaire du processus d'interactions et de la façon dont le réseau se structure. Cependant, la robustesse d'un réseau se mesure à son pouvoir de réparer des pannes (Livet et Nef, 2009), c'est-à-dire de pallier l'absence de l'un des membres qui pourrait répondre à la question posée au cours du processus de discussion. De fait, la substituabilité des membres devient nécessaire. Or la forme des fils facilite la coordination en permettant la mise en suspension de l'un de ses membres. En examinant ces alignements entre un contributeur et son destinataire, un collectif en formation sur une liste de discussion (*mailing list*) peut ainsi être repéré à deux échelles: au niveau local et à un niveau global. Nous proposons l'analyse des contributions dans les fils de discussion pour montrer l'importance du processus de contrôle exprimé par la forme des réseaux. Nous suivons ici l'idée d'Harrison White (1992: 92) selon laquelle, le contrôle est exprimé par la forme du réseau et la nature des actions sont rendues manifestes par les liens construits et par le degré de connectivité entre les noeuds :

« Les influences et les actions sont actualisées à travers des liens. En outre, la façon dont elles se transforment en actions significatives est fortement influencée par la forme et la connectivité dans le réseau tissé par ces liens. Chercher et observer le contrôle requiert une compréhension de la manière dont le réseau est actualisé. » (White, 1992)

Dans ce processus de contrôle, la connaissance par les membres du problème traité joue un rôle central. Le fil de discussion se développe en fonction de la façon dont les connaissances et les informations se distribuent selon les alignements épistémiques des

contributeurs (Lazega, 2011). Les positionnements de contributeurs révèlent des engagements divers vis-à-vis de la connaissance selon la nature des réponses mais aussi des questions qui sont postées. Les réponses comme les questions expriment un degré d'accès à un territoire d'information ou à un domaine de connaissance (Heritage, 2012). La prise en compte du niveau local du fil de discussion est donc nécessaire car on peut y observer les micro-structures de co-orientation qui se tissent entre deux porteurs du message selon le destinataire de la réplique. Un contributeur B se co-orienté vers un message antécédent d'un contributeur A soit en lui répondant s'il pose une question, soit en commentant son message de réplique. Dans les deux cas, il exprime un alignement épistémique défini par un degré d'ignorance ou de connaissance. Le contrôle se présente comme l'extension d'un domaine de connaissance. La question exprime en général un aveu d'ignorance, la réponse une prétention à la connaissance. L'alignement épistémique d'un contributeur sur un autre exprime donc une distribution sociale de la connaissance entre deux agents avec des avantages de position selon la façon dont le questionneur exprime son degré d'ignorance et la façon dont les répondants expriment leur degré de connaissance (Heritage et Raymond, 2005):

« On doit tenir compte que l'accès relatif à un domaine ou territoire d'information est stratifié entre les interactants car chacun occupe des positions différentes sur un gradient épistémique (plus connaissant C+ ou moins connaissant C-) qui varie entre faiblement ou bien étayé » (Heritage, 2012, p.32).

De l'analyse locale à l'inférence de liens au niveau global, une question se pose. Dans quelle mesure ces observations sur les micro-structures des fils sont généralisables à un collectif et/ou à des individus ? Au niveau de l'ensemble des fils, il a déjà été montré que plus un fil tend vers une forme caractéristique, filiforme, plus le nombre de

contributeurs différents tend à diminuer (Conein et Latapy, 2008). Dans cet article, nous interprétons ce résultat comme étant la conséquence du fait qu'un fil tend vers une résolution d'un problème et donc tend vers sa clôture. Dans la continuité des recherches antérieures, on se propose désormais d'approfondir l'analyse des interactions au niveau de la structure locale des fils de façon à pouvoir traiter l'évolution des structures réticulaires sur la longue durée.

2. La signification du fil pour les contributeurs eux-mêmes

Une façon d'analyser les fils de discussion comme des réseaux est de caractériser les mécanismes de coordination au niveau local pour, ensuite, élargir et comprendre les logiques de participation et de production de messages au niveau global c'est-à-dire au sein des listes de discussion.

2.1. Espaces ouverts mais contrôlés

Bien que les listes de discussion des projets de logiciel libre soient des espaces de travail nécessitant une expertise avancée, elles se présentent néanmoins comme des espaces ouverts: elles sont « non modérées *a priori* ». Chacun est libre de poster un message et de placer son message à un certain moment du développement du fil. Pour autant, après plusieurs années d'existence, cette ouverture de l'espace de discussion ne prend pas pour autant la forme d'un « bazar », comme le suggérait (Raymond, 1998). Les listes et les forums sont des espaces régulés et organisés (Akrich et Méadel, 2009) par les processus de contrôle mis en œuvre dans les liens de répliques. Ainsi, c'est parce qu'ils sont ouverts et non modérés qu'ils offrent des ressources diverses de contrôle disponibles aux contributeurs. Quelles sont ces ressources et sous quelles formes s'exerce le contrôle? Nous nous inspirons de la définition du contrôle d'H. White où le contrôle est lié au positionnement des identités au sein d'un réseau:

« Les identités, qui forment les nœuds du réseau, se déclenchent à partir des luttes de contrôle en recherchant des

appuis entre elles; elles évoluent conjointement avec les réseaux dans un domaine concret d'activité donné (White, 2006, p36). »

Un premier sens du contrôle provient du placement de l'auteur dans une séquence d'interactions. Un second sens du contrôle provient des ressources fournies par des instruments numériques qui accentuent la visibilité des liens au sein du fil.

On peut expliquer le premier sens de contrôle de la manière suivante. Le placement d'un message au sein d'un fil de discussion acquiert une signification pour les contributeurs des listes à un projet de logiciel libre. Le placement des messages dans les fils expriment un alignement des contributeurs avec des avantages de position⁴ selon la place qu'ils prennent par rapport à la question posée. Ainsi, la distribution de la connaissance et des statuts épistémiques (degré d'expertise dans un domaine) peut prendre la forme d'une activité de discussion si les réponses sont commentées. Dans ce cas, le principe de production des messages est de maintenir un processus d'interaction collaboratif dans la durée de façon à obtenir un contenu de qualité selon deux modes d'apprentissage collectif, soit sous la forme d'une demande de conseil, soit sous la forme d'un commentaire de conseil. Chacun des deux modes exprime un alignement épistémique différent. Ainsi, c'est en observant le placement des messages dans les fils, qu'on peut identifier certaines actions de stabilisation locale de la structure des interactions entre répondants. Par ailleurs, les contributeurs ont développé des outils numériques pour se placer dans les fils, le contrôle est dans ce contexte appuyé par des logiciels *ad-hoc* et conduit au second sens du contrôle défini précédemment.

L'argumentation repose en effet sur l'idée que la forme du fil a une signification pour les contributeurs eux-mêmes – et pas seulement pour le sociologue. La forme du fil de discussion, comme réseau d'interactions, ne semble pas être un pur artefact sociologique puisque les contributeurs l'utilisent pour se coordonner eux-mêmes et

4 Un avantage de position peut provenir du fait qu'un contributeur détient l'information que le questionneur réclame (Williams, 2006).

donc la contrôler.

2.2. Mécanismes de contrôle et forme du réseau

Un contributeur co-construit la micro-structure du fil dans lequel il intervient. L'étude des mécanismes de contrôle de la forme du réseau par les contributeurs passe alors par une analyse locale de la structure relationnelle du fil de discussion compris comme un réseau d'interactions.

Le fil de discussion s'étend par l'ajout de nouveaux messages à la suite d'un premier message. L'étude et la pratique des logiciels configurés par les membres construisent le fil de discussion comme un graphe orienté ayant pour nœud un message électronique et pour relation le sens de l'interaction qui est une relation orientée exprimant un lien de réplique entre les participants au fil. Chaque contributeur peut se relier à un des nœuds, c'est-à-dire à un message précédent en créant un nouveau lien. Le contrôle du réseau passe par le contrôle de sa forme mais aussi de sa taille et donc de sa capacité à créer de nouveaux liens par de nouvelles interventions.

Nous montrerons ici que les fils de discussion existent pour les contributeurs dans la mesure où leur forme est perçue d'abord, interprétée ensuite. En effet, l'ensemble des discussions est rendu manifeste et traduit le plus souvent par l'intermédiaire d'un logiciel de lecture des messages. Ensuite, les fils sont interprétés car identifiés selon des schémas de structures (*thread patterns*) par les contributeurs eux-mêmes. Enfin, l'analyse du fil fait apparaître des séquences typiques qui aboutissent à des formes élémentaires, en éventail ou filiforme, qui permettent par transformation des liens de passer de l'interaction au réseau.

2.2.1 Le logiciel comme expression d'une activité collaborative

La fréquence des messages et la quantité des listes électroniques auxquels participent les principaux contributeurs nécessitent une certaine organisation du travail de

contribution (Ducheneaut, 2003, Auray, 2004) pour la gestion de leurs ressources⁵. Les contributeurs ont en effet développé des méthodes de tri et de lecture des messages qui leur permettent de suivre et d’agir face aux flux de messages conséquents auxquels ils sont soumis quotidiennement. Ces méthodes sont notamment traduites au travers des logiciels qu’ils ont développés et/ou configurés⁶.

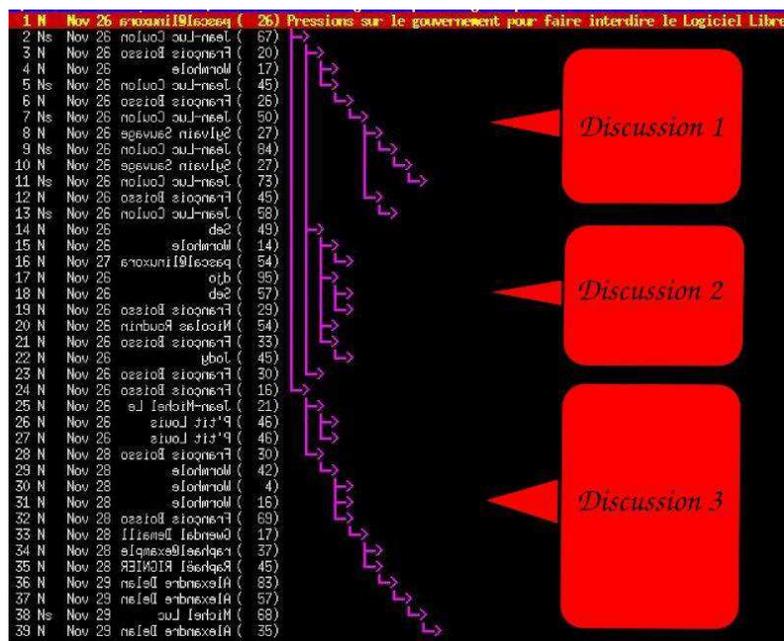


Figure 1: Capture d’écran du logiciel de messagerie électronique Mutt. De gauche à droite, les colonnes indiquent: 1) la numérotation du message, 2) la date 3) L’auteur du message 4) La taille du message 5) le graphe du fil. (Source: mutt d’un utilisateur lisant un fil de la liste Debian user-french)

Parmi les logiciels utilisés par les contributeurs, le logiciel Mutt est le plus utilisé

5 La spécialisation des listes de discussion n’est pas forcément spécifique aux contributeurs Debian, l’observation étant déjà faite par ailleurs (Butler, 2001).

6 Les résultats présentés ont été obtenus par une immersion d’un des auteurs dans la communauté Debian sur une durée de 10 ans. Le fil de discussion électronique n’existe pas en soi, il ne peut être abordé qu’à travers une médiation technique au travers de logiciels spécifiques.

comme en témoignent les statistiques des logiciels utilisés sur les mailing-lists⁷. Cet outil a été développé à partir de 1995, au moment même où le projet Debian prenait son essor. Il participe toujours à la coordination et l'organisation du projet. Le lecteur de mail est un objet technique qui exprime une activité collective (Simondon, 1958) au travers de ses pratiques que nous rendons ici manifestes (Latour, 2012. p. 212).

Ce logiciel est avant tout un logiciel de lecture des messages électroniques. Un contributeur expert concentre son attention non seulement sur un message précédent auquel il peut répliquer mais surtout sur une visualisation panoramique des messages électroniques organisés en graphe. Ce logiciel reconstruit les séquences de messages automatiquement à partir de paramètres configurables par l'utilisateur.

Par exemple, certains experts configurent (*hackent*) leur logiciel de manière à filtrer les messages selon au moins trois manières différentes. La première consiste à adopter une règle de filtrage à la récupération du message électronique. Par exemple, une règle de filtrage réalisée peut être du type:

```
:0
* « Nom du contributeur X »
/dev/null
```

Cette règle indique que tous les messages du contributeur X sont directement envoyés dans la poubelle. Lorsque cette règle est explicitement formulée à l'interlocuteur, elle permet d'indiquer que tous ses messages seront tout simplement ignorés. En changeant l'adresse d'envoi « /dev/null » par un dossier de liste électronique, les utilisateurs organisent leurs messages dès la réception avec des règles

7 Une analyse des messages échangés sur les listes Debian à la fois celles des développeurs et des utilisateurs anglophones et francophones, soit une étude sur un échantillon de près d'un demi million de mails, permet d'entrevoir quels sont les logiciels majoritairement utilisés. 51% des mails échangés ont utilisé une version standard de Mutt (les versions personnalisées ne sont pas prises en compte), 15% des messages ont été expédiés avec Gnus/Emacs, 19% avec Mozilla Thunderbird ou Icedove qui sont plus récents, 7% Kmail et 8% restants pour les autres logiciels. La présente analyse se concentre sur le logiciel permettant la production de la majeure partie des messages postés sur les listes étudiées: Mutt.

de filtrage assez sophistiquées⁸.

Une fois le message accepté une deuxième règle de filtrage intervient. Il s'agit du « *scoring* ». Son principe est d'accorder un certain nombre de points à un message en fonction de règles simples comme nom de l'auteur, les termes du sujet. Ces scores attribués à chaque message permettant d'organiser différemment les messages à la lecture: les messages avec un score élevés sont mis en haut de la pile des mails. De manière transitive, les fils de discussions comportant des messages à scores élevés sont également proposés à la lecture avant les autres.

Enfin, la dernière règle de filtrage⁹ met en forme les fils de discussion selon leur positionnement dans la discussion.

2.2.2 La forme du fil comme indice pour la construction d'un lien

Comment l'outil est-il utilisé par les contributeurs pour se coordonner ? Concernant les contributeurs les plus actifs, nous formulons l'hypothèse que la forme du fil de discussion est un indicateur de qualité des contenus échangés. Ensuite, la forme du fil, en tant qu'indicateur du contenu de la discussion en ligne, permet la coordination des contributeurs si bien que l'apprentissage des manières de s'orienter dans un fil de discussion fait partie du travail d'intégration au collectif.

Lorsque le nouveau venu dans le monde du logiciel libre installe pour la première fois le système d'exploitation, en fait le compte-rendu sur la liste et coupe malgré lui le fil, un contributeur plus expérimenté intervient pour lui indiquer:

« Profites-en pour installer quand même un client de mail moderne, un qui ne coupe pas les fils de discussion. » (Un

8 Même si le logiciel Mutt ne comporte pas toutes les fonctionnalités sophistiquées du logiciel scientifique de cartographie des conversations en ligne #, Mutt comporte néanmoins des fonctions avancées de filtrage des fils de discussion (*threads*) et des messages qui permettent la lecture quotidienne d'un grand nombre de conversations. Ce logiciel présente en effet visuellement les messages sous la forme de fils et s'inspire initialement des fils de discussion propres aux logiciels d'accès au protocole usenet.

9 Nous omettons volontairement le filtre Bayésien permettant de détecter les messages non désirés. Cette méthode est connue mais plus difficilement contrôlable.

contributeur sur la liste Debian User French) ¹⁰

Comment procède un « grand » de la cité du logiciel libre (Boltanski et Thévenot, 1991) ? À la manière du prestidigitateur expliquant le *modus operandus* de son tour de magie, le développeur Debian explique dans cet exemple sur son blog comment il peut suivre plusieurs centaines de messages électroniques par jour (Figure 2). Ce témoignage révèle selon nous la manière dont est évaluée la pertinence des fils de discussion à partir de leurs formes respectives, évidemment à l'aide du logiciel Mutt présenté précédemment. Dans son blog, le développeur annonce l'effet pour ensuite préciser l'explication:

«Voici comment souscrire aux listes de discussion et suivre plus de 2000 messages par jour et y survivre. Tout est affaire de reconnaissance de formes (Joey Hess, 30 octobre 2005¹¹).»

Certaines des interprétations des « formes » proposées par J. Hess semblent discutables néanmoins elles sont le résultat d'une activité régulière de suivi des messages. Ces formes permettent de subsumer le contenu des messages à partir de la structure des graphes des discussions en ligne. Joey Hess décrit une collection de formes. Nous proposons de commenter quatre exemples de fil : *think before* (i), *obvious answer* (k), *too late* (n) et *worthwhile* (n).

```
bob    +-> Foo
bob    +->
```

Figure 2: Capture d'écran du schéma du fil intitulé « *think before* » expliqué sur le blog de Joey Hess, développeur Debian. L'auteur est indiqué dans la colonne de

10 Source: Liste de discussion Debian User French, <http://lists.debian.org/debian-user-french/2013/05/msg00118.html>

11 Source: http://joeyh.name/blog/entry/thread_patterns

gauche (bob), ses messages sont indiqués par le signe «+», il envoie un premier message titré «Foo» qui signifie «divers». Enfin le signe «->» indique le sens de lecture du fil (qui répond à qui).

Le premier fil ayant pour légende «*think before*» fait référence à un fil à deux messages dans lequel l’auteur du premier message est aussi l’auteur du deuxième et dernier message. Le questionneur se répond à lui-même en ayant apparemment résolu sa question initiale. Sur son blog, J. Hess indique que ce contributeur aurait pu réfléchir plus longuement avant de poster sa question, d’où le «*think before*»! Cet exemple illustre la position réflexive des utilisateurs vis-à-vis de messages postés.

```
bob    +-> Foo?  
fred   +->  
alice  +->  
george +->  
nancy  +->  
steven +->  
nick   +->
```

Figure 3: Schéma du fil intitulé «*obvious answer*» expliqué sur le blog de Joey Hess, développeur Debian. Dans ce fil, fred, alice, george, nancy, steven et nick répondent tous directement à bob au sujet de son message intitulé «*Foo?*».

Le fil «*obvious answer*» fait référence à un fil en éventail avec plusieurs auteurs répondants une seule fois chacun. Plusieurs réponses à une seule question indiquent selon J. Hess que le problème avancé est simple à résoudre; l’argument étant que la séquence d’un si grand nombre de réponses consécutives ne sont pas elles-mêmes commentées. Cette structure est liée à au moins trois paramètres indépendants: la facilité de la question initiale, la disponibilité du nombre de contributeurs au moment de l’envoi de la question mais aussi la latence du réseau (les contributeurs ne perçoivent pas les réponses des autres au moment où ils écrivent et envoient leur réponse). C’est pour cette raison que certains membres mettent en place des stratégies

pour retarder leur contribution et faire émerger la forme des fils avant leur propre intervention.

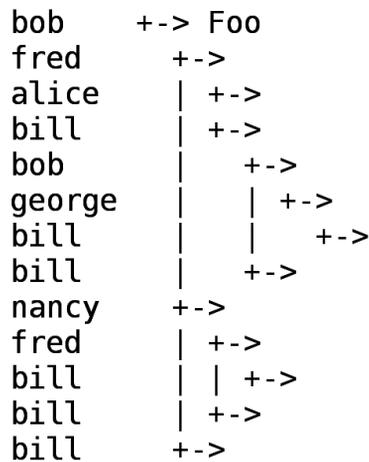


Figure 4: Schéma du fil intitulé « *too late* » expliqué sur le blog de Joey Hess, développeur Debian. Sur ce fil, trois personnes (Fred, Nancy et Bill) répondent directement à bob qui a posté le premier message. Ensuite, les sous-fils s’étendent en discussions à l’intérieur de l’ensemble du fil.

Le fil de discussion « *too late* » indique un fil de discussion potentiellement intéressant mais déjà trop entamé selon J.-Hess. Trois discussions émergent à partir de la première question. Ainsi, rattraper le cours du développement de ce fil lui demanderait un investissement trop coûteux.

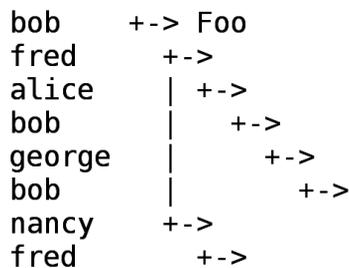


Figure 5: Schéma du fil intitulé « *worthwhile* » expliqué sur le

blog de Joey Hess, développeur Debian.

Contrairement au fil *too late*, celui intitulé *worthwhile* semble digne d'intérêt pour J.- Hess. En effet, ce fil de discussion attire son attention étant donné son état d'avancement en termes de forme de graphe.

En opérant de cette manière, le développeur révèle un certain nombre de formes typiques qui, par la vision panoramique de la connaissance échangée qu'ils donnent à voir, permettent d'optimiser son temps de travail limité entre la production, c'est-à-dire le développement informatique, et la transmission des connaissances, c'est-à-dire l'assistance ou l'apprentissage. Aussi, depuis que Joey Hess a repéré ces formes, celui-ci affirme s'efforcer d'en préserver les formes les plus idéales en contribuant à leurs branches les plus fécondes. En quelques sortes, ses réponses contribuent à un *design* des fils, à la manière d'un esthète de l'élaboration de la connaissance collaborative comme le révèle son entretien avec le développeur Debian Raphaël Hertzog qui l'a publié sur son blog.

- « Raphaël Hertzog: Tu contribues régulièrement aux listes Debian. À ce jour, je n'ai pas noté de message agressif ou frustré de ta part. Comment tu t'y prends ? Évites-tu les discussions polémiques ?
- Joey Hess: Mince, on dirait que mes magnifiques interventions enflammées ont été oubliées! Sérieusement, depuis que j'ai repéré les formes de fils, j'évite de participer dans des mauvaises formes (patterns). Aujourd'hui je me limite à l'expression d'une seule opinion et il m'importe peu de savoir qui a le dernier mot. Si les gens ne peuvent être persuadés, il est venu le temps de trouver une autre approche. De toutes les façons, le code parle de lui-

même.¹²»

Il apparaît donc que la forme du fil de discussion agit comme une partie du code car il permet au contributeur de limiter ses interventions pour éviter de perdre du temps.

2.3. Modalités de la participation sur les mailing-lists

Si la nature des problèmes informatiques posés dans les questions (requêtes) semblent souvent hétérogènes, et donc les fils de discussion deviennent contingents dans leur enchaînement, une certaine régularité apparaît néanmoins. Pour tester l'intuition que les listes et les contributeurs tendent à privilégier certaines formes d'interactions, il convient alors de définir un référentiel de manière à décrire les développements et les variations de la morphologie des fils de discussion.

Ce processus apparaît sous forme d'un réseau d'interactions (Dorat *et al.*, 2007) dont on peut décrire le fil temporel, « *moment by moment* ». Lorsqu'un fil de discussion se développe, il prend la forme de liens d'échanges tissés entre des individus et se développe un processus d'interaction restitué par le fil lui-même. De l'interaction au réseau, le fil de discussion apparaît comme le résultat d'un processus séquentiel où un message définit une place pour un message virtuel suivant et ainsi de suite à chaque tour de contribution (Livet & Nef, 2009).

12 Entretien de Joey Hess par R. Hertzog, développeur Debian français, 11/11/2010
<http://raphaelhertzog.com/2010/11/11/people-behind-debian-joey-hess-of-debhelper-fame/>

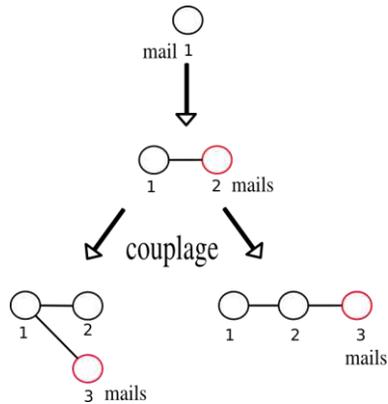


Figure 6: Couplage entre forme filiforme et/ou forme en éventail
 Le nœud gris indique le dernier message posté en date.

C'est à partir de 3 messages que la structure d'un fil de discussion commence à apparaître comme un réseau qui s'étend. Ce processus est issu d'un couplage (Livet et Nef, 2009) (Figure 6) entre la poursuite d'une séquence d'interaction continue, représentée par une forme filiforme, et/ou un embranchement rompant la séquence initiale. Le fil adopte alors une forme qui devient la combinaison de deux formes élémentaires idéal-typiques, soit une forme en éventail ou une forme filiforme (Conein et Latapy, 2008) (Figure 7). Le déploiement du fil de discussion fait ainsi apparaître un graphe orienté sur un axe temporel qui peut être exprimé par ces deux dimensions¹³.

13 Néanmoins, ces deux structures filaires ne rendent pas compte d'un troisième cas intermédiaire: le saut de tour où le message ne réplique pas au message directement précédent mais à un message antécédent éloigné. (Watson *et al.* 2014, Rapport ANR)

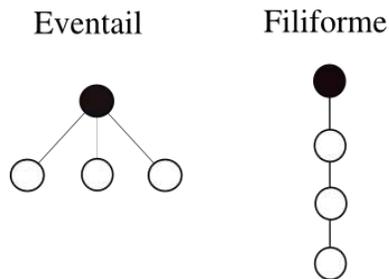


Figure 7: Schéma d'un fil en éventail et filiforme

Le nœud noir indique le premier message posté dans le fil.

Si le respect du fil de discussion revient à respecter le contenu de la discussion et donc si la structure exprime la qualité de son contenu alors se pose la question de l'évolution de l'ensemble des structures locales des fils au niveau global c'est-à-dire liste par liste et sur la longue durée. Une approche quantitative et statistique nous permettra alors d'en étudier la portée.

3. Big data et sous-structure : variations locales sur la longue durée

Dans quelle mesure l'interprétation locale de la forme des fils de discussion, et le travail sur leur structure, est-elle généralisable à l'ensemble des contributeurs des listes de discussion électroniques étudiées ? N'existe-t-il pas un effet de loupe inhérent à la focale de l'observation de l'interaction locale ? De plus, si toutefois l'inférence était raisonnable, n'existe-t-il pas des spécificités collectives ou individuelles ?

Pour déterminer dans quelle mesure nos observations deviennent généralisables, et méritent éventuellement d'être nuancées, une étude comparative à la fois quantitative et statistique peut être utile. Sont alors étudiées les variations des structures locales des fils de discussion à un niveau plus « global », c'est-à-dire sur la longue durée. Cette partie de l'enquête visent à valider l'hypothèse que la forme du réseau est un indice de contrôle et devient un marqueur d'identité (Lorrain et White, 1971). Dans cette

perspective, trois questions se posent:

1. Tout d'abord, comment la structure des fils de discussion évolue-t-elle dans le temps ? Ses variations globales reflètent-elles des variations constatées au niveau local ?
2. Ensuite, la structure des fils de discussion se différencie-t-elle selon les espaces de discussion ? En d'autres termes la structure des fil se différencie-t-elle en fonction des listes, c'est-à-dire des niveaux d'expertise (utilisateurs et développeurs) ou des langues d'usage (anglais ou français) ?
3. Enfin, la structure des fils de discussion varie-t-elle significativement en fonction des sujets de discussion et/ou permet-elle de typifier les contributions et ainsi caractériser des profils de contributeurs ?

Trois principaux problèmes d'ordre méthodologique apparaissent immédiatement. Tout d'abord, le premier problème concerne la constitution du corpus qui doit correspondre à ce que « voient » les contributeurs sans opérer une traduction supplémentaire sur la structure des fils.

Ensuite, nous allons progressivement reconstruire le fil pour en décrire quantitativement les formes et variations élémentaires. Le travail sur un grand nombre de données nécessite en effet un certain nombre de concepts mathématiques propres à une analyse des fils de discussion comme réseaux : quantité/densité, longueur/hauteur, moyenne/barycentre.

Enfin, dans quelle mesure les variations constatées sur la longue durée traduisent-elles l'organisation des séquences dans les structures locales ? Si les pratiques ne sont pas généralisables alors une évolution aléatoire des formes locales des fils de discussion devraient être observée à la fois liste par liste et sur la longue durée. Si toutefois l'évolution des formes n'étaient pas aléatoires, et qu'elles s'orientaient tendanciellement vers une forme particulière, la question devient alors: dans quelle mesure cette forme particulière correspond avec celle constatée dans la description locale présentée jusqu'à présent ?

3.1. Quantité et densité

La forme des processus d'interactions dans les listes de discussion permet-elle de retrouver au niveau du collectif de contributeurs la persistance des formes observable au niveau local ? C'est au moyen d'une étude exhaustive des listes Debian et Ubuntu (comme échantillon témoin), utilisateurs et développeurs, anglophones et francophones fut recherchée une réponse à cette question en étudiant les variations locales sur la longue durée. Avant l'analyse proprement dite, un corpus a été construit s'appuyant sur des archives d'une durée de 20 ans pour les listes les plus anciennes (Debian¹⁴ et Ubuntu¹⁵). Nous avons obtenu un corpus initial de près de 1,5 millions de messages électroniques.

En nous concentrant sur les listes principales en termes de quantité de mails publiés, c'est-à-dire les listes des utilisateurs du système d'exploitation Debian, il est possible de constater que la quantité de messages électroniques postés sur les listes créées au début des années 1990 diminue tendanciellement depuis 2002. Si la quantité de mails par an diminue, est-ce pour autant le signe qu'il existe moins de discussions sur les

14 L'intérêt scientifique de ces projets de logiciels libres est notamment de rendre un libre accès aux archives des listes de discussion électronique. 2 Corpus Debian ont été constitué de manière à pouvoir étudier finement la structure des fils. Le corpus Debian en format HTML se compose initialement de l'ensemble des messages postés depuis leur constitution respective, 1994 pour la première liste officielle du projet Debian. Liste « user » (646 005 messages, 17 years, 104 messages/jour en moyenne): liste des utilisateurs anglophones; liste « user-french » (178 844 messages, 10 years, 49 messages/jour): liste des utilisateurs francophones; liste « devel » (305 106 messages, 17 years, 49 messages/jour): liste des développeurs (anglophone); liste « vote » (17 266 messages, 12 years, 4 messages/jour): liste spécifiques aux votes du projet Debian; liste « mentors » (58 778 messages, 13 years, 12 messages/jour): liste spécifique à l'intégration des nouveaux contributeurs. Le format dit « mbox » concerne les mêmes messages sur un historique de 10 ans avec en revanche un formatage permettant une analyse fine des structures de fils.

15 Ubuntu est un projet plus récent, qui prolonge le projet Debian avec quelques différences notables dans son organisation qui salarie certains de ses développeurs contrairement à Debian. La première liste du projet Ubuntu est référencée à partir de 2004: Liste « user » (246 233 messages, 7 years, 96 messages/jour en moyenne; liste des utilisateurs anglophones; liste « fr » (37 073 messages, 7 years, 15 messages/jour): liste des utilisateurs francophones; liste « devel » (33 134 messages, 7 years, 13 messages/jour): liste des développeurs (anglophone); liste « devel-discuss » (12 676 messages, 5 years, 7 messages/jour): liste de discussion des problèmes généraux.

listes ? La quantité de mails par liste apparaît comme un indicateur faible pour observer les fils de discussion. Il convient donc d'étudier l'évolution des fils de discussion dans le temps. Pour réaliser cet objectif, il convient de rassembler les messages par fil de discussion. Mais rassembler les mails par fils, suppose une agrégation. Nous avons utilisé le logiciel Mutt présenté précédemment pour reconstruire les fils de discussion de manière à produire un corpus de l'intervention des contributeurs sans reconstruire les fils de discussion à l'aide de recodages particuliers. Cette étude analyse les fils tels qu'ils sont produits et non tels qu'ils sont reconstruit selon des hypothèses sociologiques sous-jacentes.

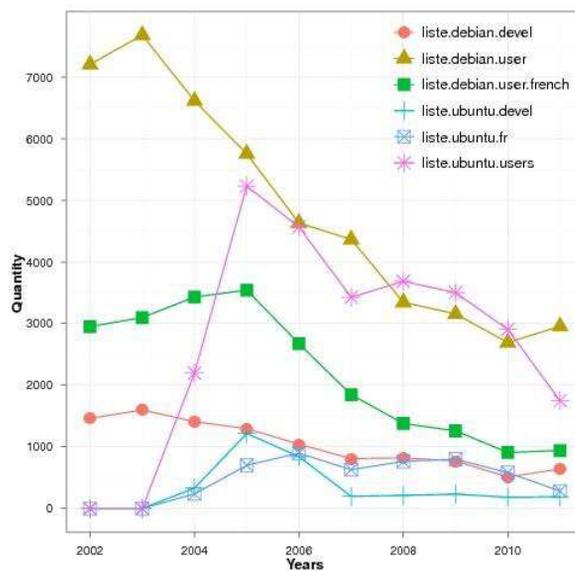


Figure 8: Evolution annuelle de la quantité de fils de discussion (nombre de mails > 1) par année et par liste

Après une croissance continue, la quantité de mails par an décroît progressivement et le nombre de fils diminue corrélativement (Figure 8). Mais est-ce pour autant un indice qu'il y existe progressivement moins de discussions? Le nombre de sujets différents décroît mais, pour autant, le nombre messages par fil de discussion augmente année après année (Figure 9). Comment expliquer ce paradoxe ?

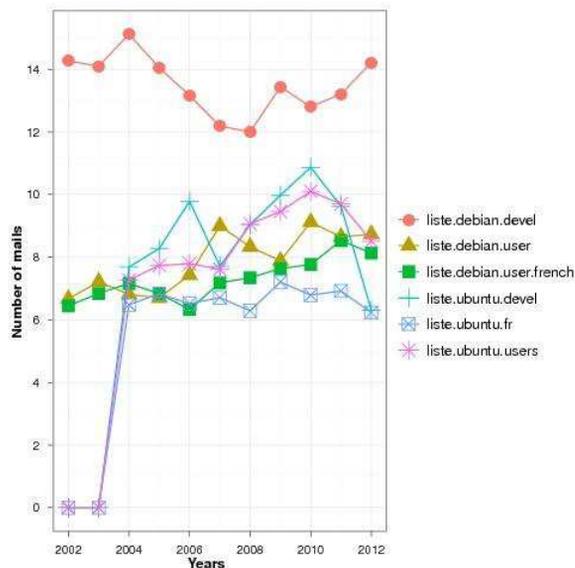


Figure 9: Evolution annuelle de la quantité de mails par fil de discussion et par liste.

Nous faisons l’hypothèse que le phénomène n’est pas seulement statistique mais relationnel, c’est-à-dire propre à la forme du fil de discussion; forme dont il convient d’estimer son évolution à partir de la structure des interactions. On peut faire par exemple l’hypothèse que l’augmentation du nombre de fils filiformes exprime une augmentation du contrôle par sélection des messages par les contributeurs les plus compétents.

De nouveaux indicateurs doivent être créés pour mieux analyser l’évolution des fils de discussion. Pour créer ces indicateurs, une meilleure compréhension du fil comme processus doit être mis en œuvre.

3.2. Longueur et hauteur des fils

Comment étudier l’évolution d’un collectif à partir de la structure de ses interactions ? Cette question suppose que l’on s’interroge sur la nature de cette structure sous au

moins plusieurs angles différents:

1. Deux collectifs peuvent ils être spécifiés par une différence entre la structure de leurs interactions ?
2. Comment s'exprime l'évolution diachronique de la structure de ses interactions au cours du temps ?
3. Est ce que l'évolution de la morphologie de la structure des interactions dépend des individus eux-mêmes ou des sujets abordés?

Ces trois autres questions conduisent à voir la dynamique d'extension d'un réseau au travers de ses interactions locales, sous forme de réseaux locaux qui augmentent leur taille, l'évolution de la taille pouvant agir sur la forme du réseau lui-même et sur l'extension de sa taille.

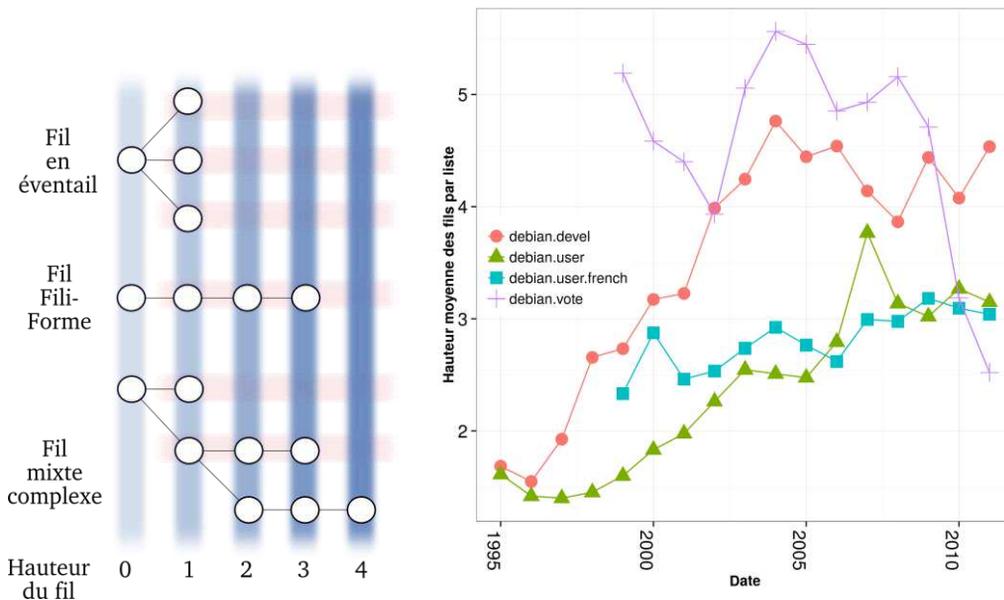


Figure 10: évolution annuelle par liste et par année moyenne des hauteurs moyennes.

Dans un premier temps, la hauteur moyenne d'un fil de discussion est évaluée de façon

diachronique (Figure 10). L'évaluation de la hauteur moyenne de tous les fils de discussion et de toutes les listes par année indique une stabilité dans le temps et par liste de discussion.

La liste politique, Debian Vote, accueille toutes les discussions liées à l'élection annuelle du leader, la taille moyenne de ses fils suit une évolution conjoncturelle qui dépend des débats internes et des sujets évoqués; ceux-ci peuvent en effet fournir un support intéressant pour des entretiens d'explicitation avec les contributeurs. Par contre les listes de développeurs et d'utilisateurs suivent une évolution similaire: une lente croissance progressive et une stabilisation sur les dernières années. On peut interpréter la croissance progressive comme un apprentissage collectif des formes des fils et une stabilisation car au-delà d'une certaine taille, le fil devient illisible. Il est intéressant de noter que ces évolutions, en particulier pour les listes d'utilisateurs, sont indépendantes de la langue. Que la liste soit anglophone ou francophone la hauteur moyenne des fils par année est la même. Aussi, la hauteur moyenne des listes des experts est plus grande que la hauteur moyenne des listes des utilisateurs. Donc la hauteur moyenne des fils semble correspondre à différents niveaux d'expertise.

Le contrôle de la forme des interactions permet d'établir le mode d'existence et de reconnaissance d'un fil, dans le sens où la hauteur moyenne évolue dans le temps et qu'elle différencie les espaces de discussion. On peut effectivement remarquer que les participants à chacune des listes ne sont pas les mêmes, en d'autres termes les listes sont cloisonnées. Ces différences répondent à notre questionnement initial mais pas complètement. Les différentes listes de discussion sont spécialisées et leurs contributeurs également. Il devient alors difficile de qualifier ce qui est en commun puisque les différences sont surtout mises en valeur. L'évolution diachronique de la hauteur conduit à reformuler notre problème de recherche. Si la hauteur traduit une augmentation du taux d'interaction (rapport message/ hauteur cf. (Conein et Latapy, 2008), permet-elle pour autant d'être un indice du niveau de contrôle exercé par les membres ?

3.3. Moyenne et barycentre de la structure des interactions

Comment décrire l'évolution des fils et des réseaux du fait de l'évolution diachronique de la structure de ses interactions ? Et dans quel sens vont ces évolutions ? La hauteur moyenne d'un fil n'est pas suffisante pour appréhender la stricte structure du réseau puisque la moyenne lisse les différences structurales. Un nouvel indicateur est donc créé.

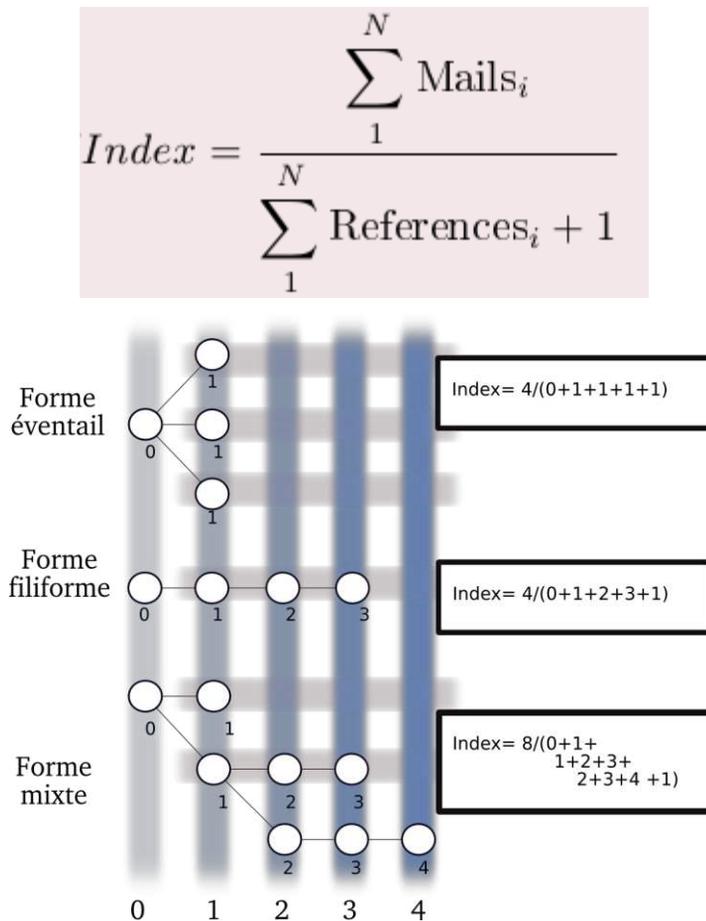


Figure 11: Formule et exemples du calcul de l'index

Cet indicateur calcule le « barycentre » (ou les points de forces) d'un fil de discussion en réseau. Son mode de calcul permet de caractériser la structure d'un fil: si

le fil est plutôt de forme filiforme ou plutôt de forme en éventail. Il s'agit d'un index puisque l'indicateur est respectivement compris entre 0 et 1. Ce « barycentre » tend vers 1 pour les fils plutôt en éventail tandis qu'il tend vers 0 pour les fils plutôt filiformes.

En étudiant l'évolution annuelle de cet indicateur par liste de discussion, il est possible d'observer une évolution commune, une décroissance continue et progressive vers 0. Ainsi la participation dans les fils de discussion différents ont ceci de commun que les membres tendent à privilégier une forme particulière d'interaction, vers des fils filiformes.

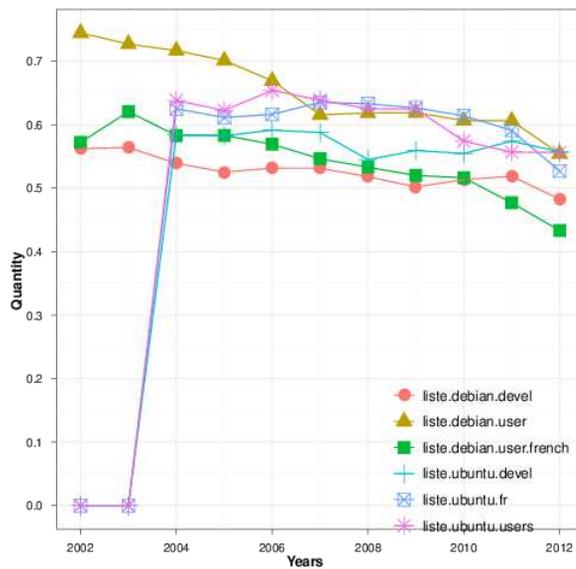


Figure 12: évolution annuelle des indices par liste de discussion (moyenne des barycentres).

Une étude statistique spécifique (Delanoë, 2013) permet de discriminer l'effet du sujet du fil, le nombre de participants mais aussi les positions structurales liées à chacun des individus. Ainsi, l'évolution filiforme des fils de discussions ne signifie pas

uniquement une augmentation des relations dyadiques sur des sujets spécifiques¹⁶. L'analyse de l'orientation tendancielle vers la structure filiforme nécessite d'être poursuivie de manière compréhensive du point de vue des acteurs. Au niveau local, l'étude de l'usage des fils a relevé comment un contributeur utilise la forme du fil de discussion. La forme visuelle des échanges facilite la coordination et donc la qualité des contributions qui non seulement font la réputation de la liste. Cette manière de procéder lui permet d'optimiser la répartition de son temps: entre temps de production de connaissances (le développement) et temps de transmission des connaissances (sur les listes électroniques). De fait, l'absence de contrôle sur le cours des fils de discussion est problématique pour la survie du projet. Ce point est tellement critique, et structurant, qu'il devient objet à controverse sur les listes politiques spécialement conçues pour les développeurs Debian. En 2012, le leader du projet doit alors recadrer la discussion pour que les attaques *ad hominem* (flames) ou les glissements hors sujet « trolls » ne polluent pas l'espace de communication.

Sur la liste des développeurs et dans un fil de discussion qui devient controversé, le leader du projet Debian intervient en rappelant 3 règles d'usage des mailing-lists¹⁷ :

1. « N'alimentez pas le troll »
2. « Reportez l'abus aux responsables des listes concernées »
3. « Changement du caractère public des listes électroniques qui deviendraient alors modérées »

Nous pouvons effectivement noter sur l'illustration 12 qu'à partir de 2011, les fils tendent à devenir plus en éventail et moins filiformes sur la liste des développeurs; en suivant le référentiel de comparaison avec l'évolution du barycentre moyen de la liste des utilisateurs francophones. Nous observons ainsi comment un espace *a priori*

16 Ce qui laisserait à penser des problèmes de recrutement pour les collectifs observés. On pourrait en effet penser que seuls des contributeurs affirmés restent fidèles aux listes de discussion étudiées.

17 Stefano Zacchiroli le 3 mai 2012. <http://lists.debian.org/debian-devel/2012/05/msg00114.html>

ouvert, c'est-à-dire public, peut se fermer, c'est-à-dire devenir modéré.

Mais pourquoi la forme des fils, comme indicateur, de qualité du contenu est-elle associée à une augmentation du contrôle au sens de capacité d'action ? Nous pensons que la spécificité du collectif du logiciel libre est de favoriser la mise en place de dispositifs qui substituent l'autorité d'expertise à l'autorité sociale en valorisant la qualité de la connaissance. Dans cette perspective, la critique de la règle de la majorité fait partie intégrante des modes d'existence du collectif dont les traits se traduisent par la morphologie en éventail de ses interactions dans l'élaboration de ses connaissances. Car en effet, la somme individuelle des préférences n'est pas transitive (Guilbaud, 1952), et donc un fil de discussion n'est pas *de facto* un sondage d'opinion mais plutôt un espace critique d'élaboration et de commentaire des connaissances. Les interventions d'opinion en éventail qui ont une visée normative de par leur agrégation majoritaire sont immédiatement nuancées par les contributeurs eux-mêmes; ce n'est donc pas étonnant si le leader du projet Debian est élu annuellement par un vote électronique sécurisé qui prend en compte le paradoxe de Condorcet¹⁸.

4. Conclusion

Cet article expose d'abord comment et pourquoi un fil de discussion en ligne dans une *mailing-list* peut être caractérisé en fonction de critères de formes et de structures des interactions entre ses contributeurs : l'interprétation des fils en termes de réseau ne permet pas seulement de représenter des coordinations qu'on ne voyait pas avant mais de comprendre le processus génératif d'élaboration des connaissances à partir des alignements qu'il dessine.

En effet, nous avons souligné que les fils de discussion des listes électroniques peuvent être traités comme des séquences d'interactions, processus que l'on peut interpréter par des formes typiques en réseau. Sur la longue durée, on a pu constater

18

Pour une présentation des conséquences sociales de ce paradoxe, voir (Guilbaud, 1952).

que les réseaux de fils évoluent tendanciellement vers une certaine forme (filiforme). La prise en compte de l'activité des contributeurs experts, est centrale pour la compréhension de l'évolution des structures de réseau, elle s'est faite en tenant compte de la façon dont les experts construisent leurs répliques au niveau local. Pour autant, l'approche morphologique suivie dans cet article appréhende les collectifs réticulaires tels qu'ils se donnent à voir mais aussi tels qu'ils sont vus par les participants à la discussion lorsqu'ils utilisent des outils de visualisation que les membres ont configurés pour leur propre utilisation. Cet usage n'est pas indépendant de l'augmentation des processus de contrôle par les membres qui se présente comme une capacité d'usage des fils, elle-même relative à la maîtrise de la connaissance des logiciels et des codes.

Les contributeurs experts contrôlent la structure des réseaux dans lesquels ils interviennent dans la mesure où le fil de discussion offre une vue panoramique des interactions leur permettant à leur tour de faciliter leur coordination. Cet usage du fil de discussion peut être considéré comme un indice identitaire: les contributeurs deviennent membres de *Debian*. Dans cette perspective, les membres du projet se caractérisent non seulement par leurs contributions à un projet collectif mais surtout par l'activité de contrôle de la structure des discussions en ligne, contrôle que les participants mènent pour sélectionner les messages ou leurs partenaires de discussion.

Remerciements

Remerciements à l'ISC.

Références

Akrich, Madeleine. Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse. Rap. tech., Interdisciplinary Institute for Innovation, Centre de sociologie de l'innovation Mines ParisTech, 2012.

Akrich, Madeleine et Méadel, Cécile, 2009. Les échanges entre patients sur l'internet. *La presse médicale*, 38 (10):1484–1490.

Auray, Nicolas, 2004. La régulation de la connaissance: arbitrage sur la taille et gestion aux frontières dans la communauté debian. *Revue d'économie politique*, (113):161–182.

Boltanski, L. et Thévenot, L., 1991. De la justification. Les économies de la grandeur. Gallimard.

Butler, B., 2001. Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research*, 4 (12):346–362.

Conein, B. et Latapy, M., 2008. Les usages épistémiques des réseaux de communication électronique : Le cas de l'open-source. *Sociologie du travail*.

Delanoë, A., 2013. Open source community watch towards a mailing-list sociometer. *Intelligence Journal*, (2).

Dorat, R., Latapy, M., Conein, B. et Auray, N., 2007. Multi-level analysis of an interaction network between individuals on a mailing list. *Analyse des trafics et traces d'usages*, *Annals of Telecommunication*.

Ducheneaut, N., 2003. The reproduction of Open Source Software programming communities. Thèse de doctorat, University of California, Berkeley.

Ducheneaut, N., 2005. Socialization in an open source software community: A socio-technical analysis. *Computer Supported Cooperative Work*, 14 (4):323–368.

Gómez, Vicenç, Kaltenbrunner, Andreas et López, Vicente. Statistical analysis of the social network and discussion threads in slashdot. In *Proceedings of the 17th international conference on World Wide Web*, p. 645–654. ACM, 2008.

Guilbaud, Georges-Théodule, 1952. Les théories de l'intérêt général et le problème de la logique de l'agrégation. *Economie appliquée*, 15:509–511.

Heritage, John, 2012. Epistemic engine. *Research on Language and social interaction*, 45 (1):30–52.

Heritage, John et Raymond, Geoffrey, 2005. The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly*, 68 (1):pp. 15–38. ISSN 01902725. URL <https://www-jstor-org.acces-distant.sciences-po.fr/stable/4148779>.

Latour, B., 2012. *Enquête sur les modes d'existence*. La Découverte, Paris.

Lazega, E., 2011. Pertinence et structure. *Swiss Journal of Sociology*, 37 (1):127–149.

Lazega, Emmanuel, 1992. *The micropolitics of knowledge: communication and indirect control in workgroups*. Transaction Publishers.

Lazer, David, Pentland, Alex, Adamic, Lada, Aral, Sinan, Barabasi, Albert-Laszlo, Brewer, Devon, Christakis, Nicholas, Contractor, Noshir, Fowler, James, Gutmann, Myron, Jebara, Tony, King, Gary, Macy, Michael, Roy, Deb et Alstyne, Marshall Van, February 2009. Social science: Computational social science. *Science*, 323(5915):721–723. URL <http://dx.doi.org/10.1126/science.1167742>.

Livet, P. et Nef, F., 2009. *Les êtres sociaux. Processus et virtualité*. Hermann Philosophie, Paris.

Lorrain, François et White, Harrison C., 1971. Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology*, 1 (1):49 – 80. ISSN 0022250X. URL <https://search-ebshost-com.acces-distant.sciences-po.fr/login.aspx?direct=true&db=sih&AN=14316820&site=ehost-live>.

Raymond, E. S., 1998. *La cathédrale et le bazar*. Linux France. Trad. Blondeel, S.

Sack, Warren, 2000. Conversation map: An interface for very large-scale conversations. *Journal of Management Information Systems*, 17 (3):73 – 92. ISSN 07421222.

Simondon, G., 1958. *Du mode d'existence des objets techniques*. Aubier.

Watson, R., Conein, B. et Delanoë, A., 2014. Skip tying on mailing-lists. (unpublished yet).

White, H. 1992. Agency as Control in Formal Networks. In : R.G. Eccles and Nohria (Ed). Networks and Organizations : Structure, Form and Action. Cambridge: Harvard Business School Press.

White, H., 2011. Identité et contrôle : Une théorie de l'émergence des formations sociales. Editions de l'Ecole des Hautes Etudes en Sciences Sociales. Trad. M. Grossetti et F. Godart.

Williams, Bernard, 2006. Vérité et véracité, Paris, Gallimard, Essais.

Yeh, J.-Y. et Harnly, A.. Email thread reassembly using similarity matching. In CEAS - Third Conference on Email and Anti-Spam, July 27-28 2006.