BUSINESS INFORMATION SYSTEMS
WORKING PAPER SERIES

ISSN 1738-1744

What's in a Group? Identification of group types for Enterprise Social Network Analytics using SWOOP data

Kai Riemer [+], Laurence Lock Lee [^], Cai Kjaer [^], Annika Haeffner [*]

*[+] The University of Sydney, ^ SWOOP Inc, * Ulm University*

BIS WP2018-01

What’s in a Group? Identification of group types for Enterprise Social Network Analytics using SWOOP data

Abstract

We report on research, carried out in collaboration with SWOOP Analytics, to identify metrics that allow distinguishing groups in Enterprise Social Networks (ESN) according to their activity patterns. The emerging field of ESN Analytics has made inroads into providing metrics and models to measure 1) the health and structural properties of enterprise social networks, as well as 2) the activity pattern and distinct behavioural roles of individual users. What is lacking so far is ESN Analytics at the group level. Yet, groups play an important role in ESNs for organising communication and collaboration activity. In this study we carry out explorative research employing cluster analysis to identify metrics that best distinguish a sample of 350 ESN groups from three organisations into distinct types. We identify three metrics as most useful: 1) the Gini coefficient, measuring (un)evenness of user participation, 2) density, measuring the extent to which users interact with each other, and 3) reciprocity, measuring the response rate to messages within the group. The resulting typology of four groups, broadcast streams, information forums, communities of practice and project teams, will be useful for network managers and group leaders to check how well their group is tracking against intended group activity pattern.

Keywords: Enterprise Social Networks, Groups, ESN Analytics, Typology, Cluster Analysis

Introduction

Enterprise Social Networks (ESN), online services that allow employees to connect and converse with each other in a secure space, have made fast inroads into organisations, with the promise to foster collaboration and enable new work practices (P.M. Leonardi, 2015). According to a report by McKinsey (2012)), effective use of such services can result in a 20-25 percent improvement in the productivity of knowledge workers. In the case of one large company Forrester Research found a return on investment of 365 percent on their ESN investment over three years (Dodd, 2011). With more organisations adopting ESN, questions about how to measure benefits and success abound (Bughin, 2015).

The emerging sub field of ESN analytics (e.g. Schwade & Schubert, 2017) aims to develop metrics and models to examine ESN log file and content data to gain a better understanding of actual ESN usage pattern. This comprises both metrics for measuring the health and structural properties of the overall network, as well as metrics to characterise and classify individual ESN user behaviour and user roles (Hacker, Bernsmann, & Riemer, 2017).

What is missing from existing work so far are metrics and models for measuring activity at the intermediate, or group level of analysis. Groups play an important role within ESN in organising communication and collaboration practices. In this paper we are concerned with developing metrics to distinguish between ESN group types. We ask the following research question:

Which metrics are best suited to distinguish groups in ESN networks into distinct types?

We utilise an ESN activity meta-data set provided by Australian analytics company SWOOP Analytics, which was sourced from the Yammer networks of three organisations. We engage in explorative research, using cluster analysis, employing a range of general social network and ESN-specific activity metrics to see which ones divide a sample of 350 groups into distinct group types.

A combination of three metrics divided our group sample into clusters that are not only well-interpretable, but relate to, and extend the classification used by Yammer itself. The three metrics are 1) evenness of user participation in the group, 2) the degree to which messages elicit responses from others (reciprocity), and 3) network density, the extent to which users interact with all other users in a group. Clustering with these metrics resulted in four distinct ESN group types:

1) <u>Broadcast Streams</u>: Large groups with largely one-way communication and little participation by the broader membership. Used to broadcast information and announcements.
2) <u>Information Forums</u>: Groups with even participation by the membership but little interaction between members. Used to share information with others.
3) <u>Communities of Practice</u>: Groups that elicit interaction between members but revolve around a core group of active users. Where learning and discussion happen.
4) <u>Project Teams</u>: Small groups with even participation and lots of interaction. Where work gets done.

Our findings have the potential to help ESN network managers and group leaders in understanding the discrepancies between aspiration and actual activity in ESN groups. Utilising metrics to classify groups will help group leaders understand how their group is tracking against the patterns of a particular group type that they envision their group to embody. Our research represents a first step towards making available group-level metrics and typologies for businesses through analytics platforms.

Enterprise Social Network Analytics

Enterprise Social Networks (ESN) are services, accessed through a web browser or mobile app, that allow people to (1) communicate with their co-workers or broadcast messages to everyone within the organisation; (2) explicitly indicate or implicitly reveal particular co-workers as communication partners; (3) post, edit, and sort text and files linked to themselves or others; and (4) view the messages, connections, text, and files posted, edited and sorted by anyone else in their organisation at any time of their choosing (Paul M Leonardi, Huysman, & Steinfield, 2013).

Another defining characteristics of ESN is their malleability and flexibility (Richter & Riemer, 2013). ESN have been associated with a variety of organisational practices such as communication, collaboration (Riemer, Richter, & Böhringer, 2010), knowledge management (Levy, 2009) crowdsourcing (Schlagwein & Bjorn-Andersen, 2014), open innovation (Dahlander & Gann, 2010), or open strategy (Tavakoli, Schlagwein, & Schoder, 2015). Similarly, research has shown that users in the same organisation often engage in a wide variety of ESN practices (Riemer, Altenhofen, & Richter, 2011). Given this flexibility, organisations face the challenge of keeping track and making sense of the emerging activity in their own ESN. This is the task of ESN Analytics.

ESN analytics is a sub field of social media analytics (Stieglitz, Dang-Xuan, Bruns, & Neuberger, 2014; Stieglitz, Mirbabaie, Ross, & Neuberger, 2018). We define ESN analytics as methods and practices for the identification and utilisation of metrics and models for measuring different aspects of user activity in enterprise social networks, including user activity levels and user profiles, network activity levels, structural network characteristics, and network health indicators, in support of support organisational goals and outcomes.

The number of studies contributing to establishing metrics or models to support ESN analytics is still limited (cf. Schwade & Schubert, 2017). So far two main areas of application for ESN analytics exist:

1. Metrics characterising the overall social network: Here, traditional social network analysis (SNA) techniques are being employed (Wasserman & Faust, 1994). For example, Riemer, Finke, and Hovorka (2015)) have shown how social capital metrics can be utilised to link certain network characteristics to employee performance. Behrendt, Richter, and Trier (2014) provide an overview of SNA metrics and studies for use in ESN contexts.
2. Development of dedicated ESN activity metrics: Recent research aims to develop metrics that characterise individual user behaviour and to generate models that classify user populations into distinct user types. Most notably is the research program by Hacker and colleagues (e.g. Hacker, Bodendorf, & Lorenz, 2017). Other works include a study by Cetto et al. (2018) who classified users by knowledge sharing and seeking behaviours, and Frank et al. (2017), who utilised log data from Exchange, Microsoft Office 365 and Sharepoint to identify user roles (Frank, Gimpel, Schmidt, & Schoch, 2017).

What is lacking so far in ESN Analytics are works that engage with ESN groups, the intermediate level of analysis between the network and that of individuals. Groups play an important role in ESN as they allow for the creation of dedicated spaces for conversations and information exchange between a sub set of users. Given their usefulness many companies find that the number of groups tends to proliferate over time, with some groups very active and many others abandoned. At the same time groups are used for all kinds of purposes, and they exist in different shapes and forms, from very small ones to large behemoths. We suggest that a better understanding of different group types, their structural features and activity patterns, will be useful for decision-makers in better harnessing their ESN for value.

We are aware of only one study engaging in detail with ESN user activity at the group level, classifying groups in the context of knowledge work (Riemer & Tavakoli, 2013). However this study is not useful in the context of ESN analytics, since the classification was based on a manual coding of user messages, which is impractical as the basis for analytics practices. Accordingly, we investigate which set of metrics discriminates best a population of ESN groups such that it results in a set of meaningful group types characterised by different activity patterns.

Study overview

We utilise ESN activity meta-data from three Yammer networks, obtained from Australian analytics company SWOOP Analytics Pty Ltd (in the following just: SWOOP). We set out to test a range of metrics to see which ones divide the sample of groups in our data set into distinct types. We briefly introduce our research setting and data set, before we outline our method and the metrics included in this study.

The SWOOP data set

SWOOP offers a cloud-based platform that provides analytics for organisations' Yammer, Facebook Workplace and Microsoft Teams networks. When given permission by an organisation to integrate with its network, SWOOP "provide[s] access to more than 30 measurement indicators giving organisations and individuals deep insights into collaboration across the enterprise." It uses these metrics to provide user profiles, in the form of a typology that classifies each user.

Generally, any action performed by an ESN user is stored in the backend database of the ESN system and available in the form of digital traces, "digitally stored, event-based, chronological records of activities of actors, which result in direct or indirect actor relations or content in different data formats" (Behrendt, Richter, & Riemer, 2014, 4). We distinguish usage data, or *meta-data*, collected *about* activities or interactions that indicates *how*, *when* and *where* an ESN activity was performed, the kind of interaction and *who* was involved from user-generated data, or *content,* which contains *what* was posted.

In order to ensure confidentiality SWOOP does not collect any content from organisations, only meta-data. Whereas the Yammer data model is organised around messages, SWOOP provides ESN activity data already organised as interactions between users. Moreover, SWOOP is able to collect from an organisation's ESN more information than is included in the Yammer database, such as information on 'Likes' or 'Mentions' of other users (tagging), each of which are represented in the SWOOP data model as particular interactions. SWOOP distinguishes the following interaction types: Post, Reply, Notification, Mention and Like. Table 1 shows the meta-data available for each interaction.

ID	Unique identifier for each interaction
Class	Type of interaction: Post, Reply, Notification, Mention or Like
From	User-ID of user initiating the interaction
To	User-ID of targeted user (not relevant if Class equals Post, as Post is undirected)
Thread ID	Unique ID for every thread, every interaction belongs to a thread, "Post" creates new thread
Date	Timestamp of the interaction
Group ID	Unique ID of the group in which an interaction takes place (if empty, not in group)

Table 1: Meta-data for each interaction in the SWOOP data model

For this study we utilised data from Yammer networks of three firms (two financial services and one professional services company). The data set contained meta data of all interactions in the various groups across these networks for a representative 10-week period. To protect user privacy SWOOP only shared anonymised meta-data, which was stripped of company, user and group names. Users, groups and all interactions remain traceable however through their unique IDs. In total, the data set contained 683,733 interactions by 40,304 users in 350 groups.

Method: Explorative clustering

Our aim was to identify those metrics that best discriminate the sample of ESN groups in a way that results in certain archetypes describing groups regarding their activity patterns. Much like individual user profiles and archetypes already provided by SWOOP, the question we explore in this study is thus, can we identify a set of metrics that provides a similar set of group archetypes?

Given the explorative nature of this question, our research approach needed to be 'creative' and iterate between identification and calculation of metrics and a clustering of groups based on varying sets of metrics. Hence, the steps in this process are: 1) identification of metric candidates, describing both the network structure of a group and user activity, 2) selection of metrics for inclusion in cluster analysis, 3) calculation of metrics for each group, 4) selection of clustering algorithm, 5) performing of cluster analysis, 6) interpretation of results. Steps 2 to 6 were repeated until a result emerged that a) discriminated well into distinct group clusters, and b) was interpretable in a way that corresponds with typical ESN use.

To identify clusters we used dendrograms, plotting of metrics and a three-dimensional plot of group locations according to their metric values. In turn, the requirement to judge and interpret the clustering result in each instance, meant that it was not feasible to include more than three metrics in each clustering attempt. Each clustering was thus done on the basis of triplets of metrics. This allowed surfacing first which individual metrics, and second which metric combinations discriminated the group sample most distinctively (given that some metrics correlate and didn't discriminate in distinct ways).

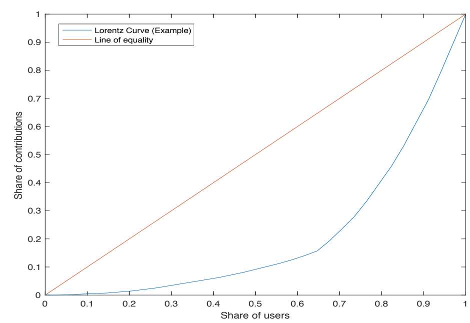
Background: Cluster Analysis
Cluster analysis is a method for semi-automated grouping of large numbers of objects based on their similarity described by a vector of quantified characteristics (Hartigan, 1975). Cluster analysis is 'semi-automated' because it is up to the researchers to determine whether or not a clustering was successful. According to Everitt (1993) success is given when the researcher, who is familiar with the data, can sensibly interpret the resulting clusters. A good set of clusters shows homogeneous and clearly separable clusters. Previous research already demonstrated that clustering techniques are useful for classifying complex networks of different kinds (Newman & Girvan, 2004; Strogatz, 2001). For this study we experimented with a number of clustering algorithms (Song, Di Matteo, & Aste, 2012). Ultimately agglomerative clustering, in particular the complete-linkage algorithm (Defays, 1977; Krznaric & Levcopoulos, 1998) with a standardised Euclidean distance measure (Pandit & al., 2011) produced the most useful results.

ESN social graph and metrics

For our study, SWOOP provided various types of interactions between users that can be utilised to construct network graphs for each group in our sample. At the same time, the inclusion of different interaction types in graph creation has implications for calculating and interpreting metrics; for example, does liking someone's post constitute a relationship with that person, or should a relationship only be considered based on a reply to a message, as this suggests that the respondent has actually read (and not merely seen) the message and found it stimulating enough to interact?

Background: The social graph
Any operationalisation of network metrics in ESN has to begin with the construction of the social network graph. Generally, a social network "consists of a finite set or sets of actors and the relation or relations defined on them" (Wasserman & Faust, 1994, 20). Whereas in public social networks, such as Twitter or Facebook, networks can be inferred from explicit friend or follower relationships, in ESNs relationships have to be constructed from user activity, as follower relationships either do not exist or are inconsequential to communication on the platform (Behrendt, Richter, & Riemer, 2014). At the most basic level a dyadic relationship between two individuals is said to exist when one user responds to another's message (Ahuja, Galletta, & Carley, 2003). This is in line with social network theory, which asserts that relationships emerge from interactions (Granovetter, 1973; Krackhardt, 1992). ESN meta-data can thus be utilised to infer the ensuing network (Behrendt, Richter, & Trier, 2014).

Drawing on existing research we identified a list of metrics candidates: 1) ESN group activity metrics describe different aspects of communication in each group, such as how many users post, how many interactions are carried out, how responsive users are in replying, how many replies each post elicits, how many users engage in each discussion. Our list (see table 2) was adapted from the metrics catalogue provided by Hacker, Bodendorf, and Lorenz (2016). 2) Social network metrics characterize structural properties of the social graph of a group, such as how densely users in a group are linked, how diverse the external links of users to other groups are, or to what extent the network is dominated by particular users, as measured by the Gini coefficient (Yakovenko & Rosser Jr, 2009).

Algorithm for calculating the Gini coefficient	
1. Count number of contributions for each active user of the group(Likes, Posts, Replies, Mentions), then sort them from low to high. 2. Calculate Lorentz Curve: Y-Axis: Proportion of total contributions that are made by the bottom x% of the users (see Figure) 3. Calculate size of area between red and blue line of Figure 4. Standardize by multiplying by 2 **5.** Get Value between 0 (if all users contributed equally) and 1 (if only one user contributed)	

Metric	Measurement	Interpretation
ESN group activity metrics		
# active users	Number of users who performed at least one interaction inside a group within a timeframe	Allows comparing groups according to different levels of user involvement
# interactions	Number of interactions inside a group within a timeframe	Allows comparing groups regarding different activity levels
Response rate (threads and posts)	The Share of Threads/Posts with at least one reply (Likes are not counted as reply)	Measure the level of engagement in a group
Response rate (includes likes)	Modified response rate that includes also Likes	Measures level of recognition, not just actual responses.
Replies per thread	Average number of replies per message thread	Measures extent to which group engages in longer discussions.
Passivity	Number of Likes divided by number of Replies	Measure the level of mere recognition relative to actual engagement
Users per Thread	Average number of different users that contribute to one thread	An alternative measure of engagement.
Group social network metrics		
Density of directed & undirected graph	Number of actual edges divided by the number of possible edges between nodes	Measures how evenly group members interact with each other.
User diversity (external links)	Average number of groups in which the users of a particular group are active	Measures how diverse the user population of a group is in terms of membership in other groups
Gini coefficient	General measure of equality applied to number of interactions per user (0=all users contribute equally, 1=all contributions by one user).	Measures how equal the contributions in a group are distributed among its users.
Key Player Index (SWOOP measure)	Relative number of users that perform 50% of all actions within a group.	Measures how dependent the group network is on certain individuals.

Table 2: Overview of key metrics candidates considered during the exploratory analysis

Findings: four ESN group types

Our explorative analysis 'tested' varying triplet combinations of the above metrics by running the cluster algorithm on the sample of 350 groups each time. The analysis converged on a set of three metrics that not only discriminate well within the group sample, but also differentiate the groups into four distinct clusters that are well interpretable and that correspond with known uses of ESN groups in organisations.

Metrics that best discriminate the groups sample

The metrics that best divided the sample of groups into clusters are as follows:

1 Density of directed Graph: for each group a directed graph is created by adding a node for each active user and a directed edge between all node pairs whose user-IDs appear as "From" and "To" in one or more transactions inside the group; the edge points to the node whose user-ID appears as "To". The density of this graph is defined as the number of existing edges divided by the number of possible edges. Density is a measure of the degree to which members of the group are connected, resulting from people talking directly to each other.

2 Gini Coefficient: this metrics stems from economics and was originally intended to measure wealth inequality, that is the unevenness of wealth distribution. In the ESN context, it measures how evenly activity in a group is distributed. The higher the Gini coefficient, the more uneven is the activity distributed in a group. A Gini of 1 means that only one person is responsible for all activity, a Gini of 0 means everyone contributes exactly the same amount of activity.

3 Thread reciprocity: thread reciprocity measures the share of all posts with at least one reply. It is thus akin to a response rate measure. Groups with a high thread reciprocity are more conversational. Note that a Like is not regarded as a Reply; a genuine response post is required.

Group types resulting from the cluster analysis

From these metrics the clustering algorithm derived a total of initially five clusters (chosen after visual inspection of the resulting dendrogram). After a further detailed analysis of the five clusters we decided to merge the two smallest of the clusters (shown as clusters 3 and 5 in figure 1, and in red and green in figure 2) as they turned out to be quite similar in terms of metrics. Figure 1 demonstrates for each of the three metrics separately how they discriminate between the clusters; figure 2 provides a three-dimensional plot which visually locates all 350 groups; and table 3 names and summarizes the metrics for each of the four clusters. In the following we interpret each of the clusters.

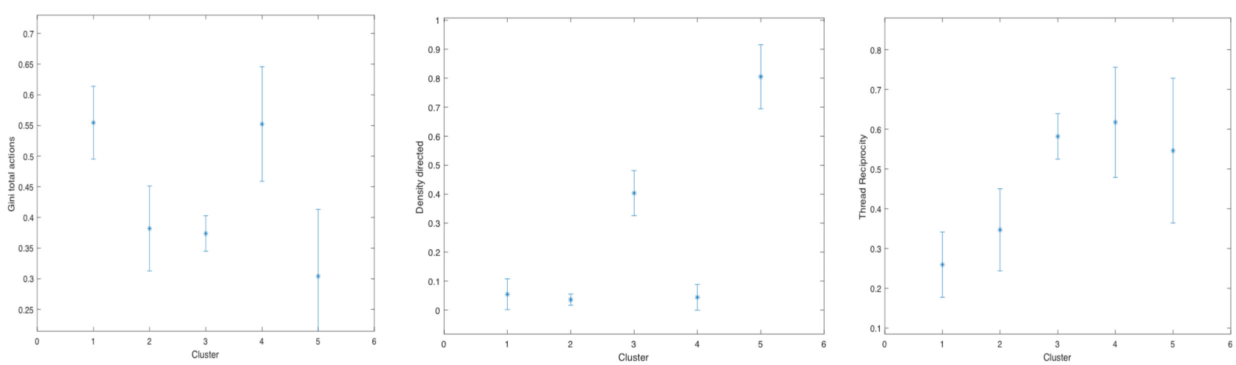


Figure 1: Metrics values for each of the resulting clusters

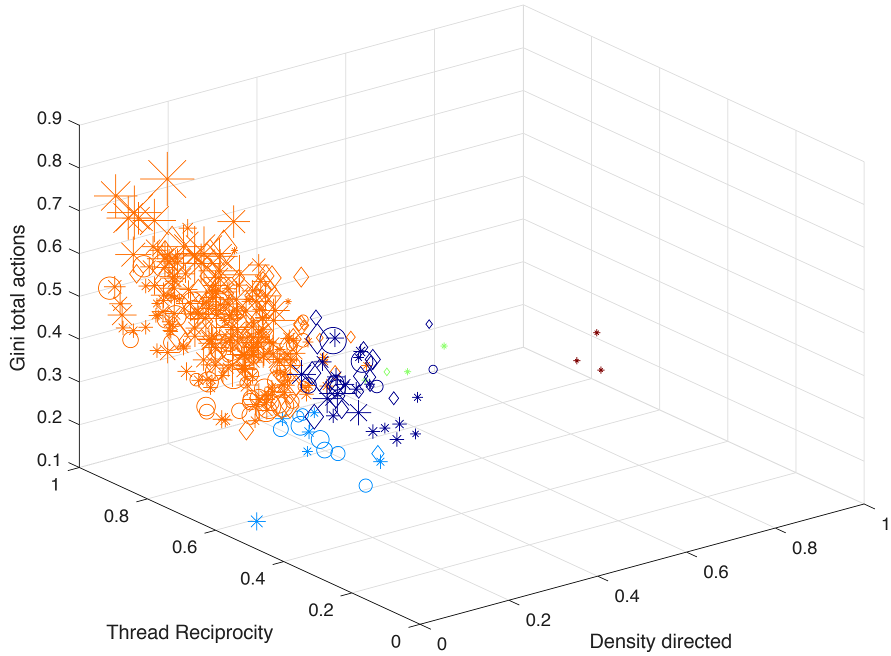


Figure 2: Cluster locations [shape of markers = company; marker size = group size]

Cluster			**Metrics**			**# of active users**		
#	**Colour**	**Name**	**Participation (Gini)**	**Density**	**Reciprocity**	**Avg**	**Min**	**Max**
1	Blue	Broadcast streams	uneven	low	low	80.5	13	352
2	Light blue	Information forums	even	low	low	59.0	36	107
4	Orange	Community of practice	uneven	low	high	125.9	9	1018
3/5	Green/red	Project teams	even	med/high	high	9.7	7	13

Table 3: Overview of group classification according to the three metrics

The following describes each of the four group types in detail (see also table 3):

1. Broadcast streams: These groups are quite large in terms of active users (those who interacted at least once in the 10-week period), yet they show only low levels of interaction and participation across the user population. Rather, they feature many single messages written by a small number of participants, and a large number of people who mostly read and only occasionally post. In addition, people are not well-connected with each other. Such characteristics are typical of groups used for announcements and the broadcasting of information. Typical uses are corporate communications or HR departments and business divisions pushing information to users in ways that resemble one-to-many ‘Intranet’ use. Such communication does not require responses from (reciprocity), or interaction among users (density). The relatively large number of active users is explained by ‘Likes’ acknowledging posts.

2. Information forums: Significant about this group type is that, while it shows rather even participation among users posting into the group, these posts do not solicit many replies from other users, or lead to interactions among users to build relationships. Such properties are typical of information forums, in which people post information, questions or requests for other users, but which are not home to many conversations or actual work interactions.

3. Community of practice: These groups show uneven participation but high reciprocity. This means that, while many posts receive replies from other users, these initial messages are written by a core group of members. In addition the overall network density is low in that people are not well connected among each other. The latter is partly explained by the fact that these groups are the largest on average in our sample. We term these groups 'Communities of Practice' (CoP). CoPs are groups of loosely connected members which often congregate around a particular topic and a core group of leaders or experts in the context of organisational learning and knowledge exchange, while a rather large number of group members follow the conversation as an audience and only occasionally participate.
4. Project teams: These groups are by the far the smallest in our sample and show significantly higher levels of connection between the group members than groups in the other three clusters. They are also highly interactive and conversational with even participation. Such properties are typical of project teams in which all group members are actively involved in performing joint work and all group members interact and converse with each other on a daily basis.

Discussion

We set out to investigate which set of metrics discriminates best in a sample of ESN groups such that it results in a set of meaningful group types characterised by different activity patterns. Our explorative analysis converged on three metrics that measure 1) reciprocity in terms of the proportion of messages eliciting replies, evenness of user participation, and density in terms of user connectedness in the group. Those metrics in turn distinguish four distinct group types, which we named broadcast streams, information forums, communities of practice and project teams.

Comparison with ESN group classification schemes

We note that our group types correspond to, yet extend in meaningful ways, the group categories used by ESNs such as Yammer or Facebook Workplace. For example, Yammer used to provide as a template for their users a classification of three group categories (see figure 3). Two of our types, project teams ('Project') and broadcast streams ('My Organisation') have direct equivalents with Yammer's categories, while Yammer subsumes all other use cases under a broad category 'community' intended for users to "share best practices, learn new skills and connect around shared interests." Yammer's recent decision to suspend the group classification feature, after feedback from users, indicates that the typology was not granular enough and thus unhelpful.

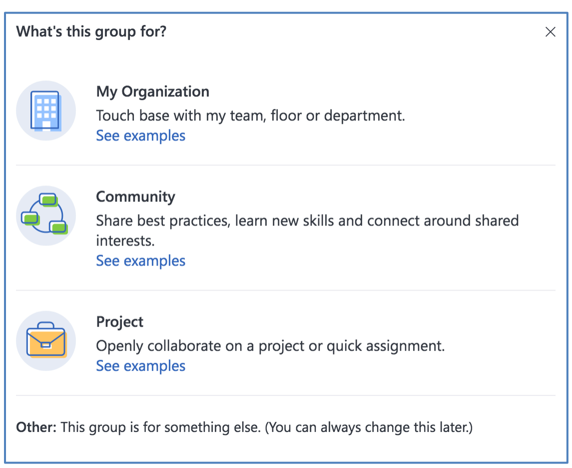


Figure 3: Yammer group template [note: Yammer recently suspended this classification feature]

Similarly, Workplace operates with three main categories, 'Teams & Projects', 'Open Discussions' and 'Announcements'. In addition, Workplace uses a group category 'Social & More' to separate out non-work-related communication and adds two more specialised groups ('Multi-Company' and 'Buy & Sell') that are out of scope in our context. We note that separating 'social' conversations from work-related ones, while appealing to certain executive managers, might send a questionable signal to employees that non-work-related conversations, while tolerated, are somehow 'second rate'. Our previous work has shown that healthy ESNs show about 40% communication that are not necessarily work-related but form the basis for any network community to exist in the first place (Riemer et al., 2011).

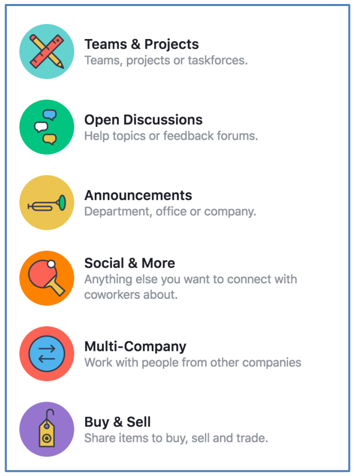


Figure 4: Workplace group template

Upon comparing our group typology with the two classifications used by Yammer and Workplace, we suggest that a distinction should be made between communities of practices and information forums to differentiate those groups that are intended to focus on sharing best practices and facilitating learning from those that revolve around sharing of interests and information. Neither Yammer nor Workplace make this important distinction, but lump these conversations together in their 'Community' and 'Open Discussions' categories.

Yet, communities of practice require more interaction and conversations between users (as measured by reciprocity) than information forums, but at the same time will show a certain un-evenness in participation (as measured by Gini), given that sharing of best practices and learning come with a differentiation in roles between experts/teachers and a broader audience of learners. This distinction is further supported by earlier, content-based studies that classify ESN use cases, where a strong distinction is made between communication genres that generate 'discussion and conversation' and those that are mainly one-way for 'providing input' for others (Riemer et al., 2011).

Utilising the group typology in ESN management practice

Initial feedback from SWOOP and its client base suggests that our typology will be helpful for ESN group leaders and community managers in managing groups within their ESN networks. Specifically, we suggest that measurement of group characteristics will allow group leaders to compare their aspiration for what the group intends to become with actual patterns. For example, a group that intends to support a project team might, upon application of our metrics, be classified as a community of practice, indicating a lack of density, resulting in unhelpful network fragmentation in the project team. Similarly, an intended CoP might be measured as a broadcast stream, indicating a lack of engagement (reciprocity) among its members. Finally, an intended information forum that lacks even participation becomes lopsided with a lack of diversity in contributions and perspectives (see figure 5). We suggest that knowledge of such discrepancies will allow group leaders to manage and counteract accordingly.

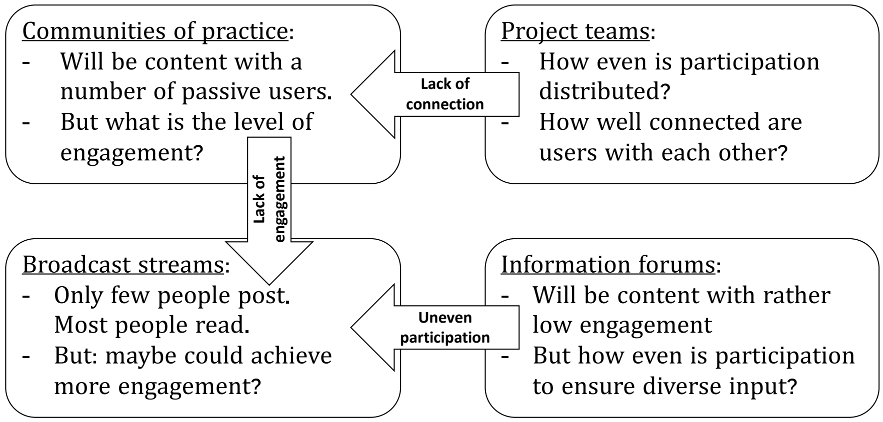


Figure 5: Examples of possible discrepancies between group aspiration and measured types

Conclusion

Our study contributes to ESN research in general, and the emerging field of ESN analytics more specifically, by extending ESN analytics practice to the group level. Specifically, we contribute a set of initial metrics and a first typology of ESN groups according to activity patterns, as the basis for broader research into understanding the role of groups in ESN networks.

Furthermore, our study contributes to ESN practice a method for ESN group leaders and network managers to measure group activity in a meaningful way, to visualise discrepancies between group aspiration and actual user activity, as measured by our metrics, and thus to improve group communication to achieve intended communication patterns. We envision that our metrics and classification could suitably be implemented in platforms such as that provided by SWOOP.

Future research is needed to corroborate the findings presented here, since our is merely a first, necessarily limited step in a broader research endeavour to extend analytics to the group level. We envision that future analyses will apply similar explorative analysis to different ESN networks to replicate our results, unearth additional useful metrics for discriminating group activity and extending our typology. Additionally, it will be worthwhile investigating the link between group-level and individual-level metrics and types, such as those identified by (Hacker, Bodendorf, et al., 2017). For example, will group of certain types benefit from the presence of certain individual user types among its members?

References

Ahuja, M. K., Galletta, D. F., & Carley, K. M. (2003). Individual centrality and performance in virtual R&D groups: An empirical study. *Management Science, 49*(1), 21-38.

Behrendt, S., Richter, A., & Riemer, K. (2014). *Conceptualisation of Digital Traces for the Identification of Informal Networks in Enterprise Social Networks*.

Behrendt, S., Richter, A., & Trier, M. (2014). Mixed methods analysis of enterprise social networks. *Computer Networks, 75*, 560-577.

Bughin, J. (2015). Taking the Measure of the Networked Enterprise. *McKinsey Quarterly, 51*(10), 1-4.

Dahlander, L., & Gann, D. M. (2010). How Open is Innovation? *Research Policy, 39*(6), 699-709. doi:10.1016/j.respol.2010.01.013

Defays, D. (1977). An efficient algorithm for a complete link method. *The Computer Journal, 20*(4), 364-366.

Dodd, L. (2011). Study reveals huge ROI when using Yammer. *Strategic Communication Management, 15*(6), 7.

Everitt, B. S. (1993). *Cluster Analysis (Third Edition)*. London Melbourne Auckland: Edward Arnold.

Frank, L., Gimpel, H., Schmidt, M., & Schoch, M. (2017). *Emergent User Roles of a Digital Workplace: A Network Analysis Based on Trace Data.* Paper presented at the 38th International Conference on Information Systems, Seoul, South Korea.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology, 78*(6), 1360-1380.

Hacker, J. V., Bernsmann, R., & Riemer, K. (2017). Dimensions of User Behavior in Enterprise Social Networks. In R. Helms, J. Cranefiels, & J. van Reijsen (Eds.), *Social Knowledge Management in Action: Applications and Challenges* (pp. 125-146). Cham, Switzerland: Springer International Publishing.

Hacker, J. V., Bodendorf, F., & Lorenz, P. (2016). A Framework to Analyze Enterprise Social Network Data. In M. Atzmueller, S. Oussena, & T. Roth-Berghofer (Eds.), *Enterprise Big Data Engineering, Analytics, and Management* (pp. 84-107). Hershey: IGI Global.

Hacker, J. V., Bodendorf, F., & Lorenz, P. (2017). A framework to identify knowledge actor roles in enterprise social networks. *Journal of Knowledge Management, 21*(4), 817-838.

Hartigan, J. A. (1975). *Clustering Algorithms*. New York: John Wiley & Sons.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. *Networks and organizations: Structure, form, and action, 216*, 239.

Krznaric, D., & Levcopoulos, C. (1998). Fast algorithms for complete linkage clustering. *Discrete & Computational Geometry, 19*(1), 131-145.

Leonardi, P. M. (2015). Ambient awareness and knowledge acquisition: using social media to learn "who knows what" and "who knows who" *MIS Quarterly, 39*(4), 747-762.

Leonardi, P. M., Huysman, M., & Steinfield, C. (2013). Enterprise social media: Definition, history, and prospects for the study of social technologies in organizations. *Journal of Computer-Mediated Communication, 19*(1), 1-19.

Levy, M. (2009). WEB 2.0 Implications on Knowledge Management. *Journal of Knowledge Management, 13*(1), 120-134.

McKinsey. (2012). The Social Economy: Unlocking Value and Productivity through Social Technologies. Retrieved from http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_social_economy

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review, 69*(2), 26-113.

Pandit, S., & al., e. (2011). A comparative study on distance measuring approaches for clustering. *International Journal of Research in Computer Science, 2*(1), 29-31.

Richter, A., & Riemer, K. (2013). Malleable End-User Software. *Business & Information Systems Engineering, 5*(3), 195-197.

Riemer, K., Altenhofen, A., & Richter, A. (2011, 2011). *What Are You Doing? - Enterprise Microblogging as Context Building.* Paper presented at the European Conference on Information Systems (ECIS), Helsinki, Finnland.

Riemer, K., Finke, J., & Hovorka, D. S. (2015, 16th December 2015). *Bridging or Bonding: Do Individuals gain Social Capital from Participation in Enterprise Social Networks?* Paper presented at the International Conference on Information Systems ICIS 2015, Fort Worth, United States.

Riemer, K., Richter, A., & Böhringer, M. (2010). Enterprise Microblogging. *Business & Information Systems Engineering, 2*(6), 391-394. doi:10.1007/s12599-010-0129-1

Riemer, K., & Tavakoli, A. (2013). *The role of groups as local context in large Enterprise Social Networks: A Case Study of Yammer at Deloitte Australia*. Retrieved from Sydney: http://ses.library.usyd.edu.au/handle/2123/9279

Schlagwein, D., & Bjorn-Andersen, N. (2014). Organizational Learning with Crowdsourcing: The Revelatory Case of LEGO. *Journal of the Association for Information Systems, 15*(11).

Schwade, F., & Schubert, P. (2017). *Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities.* Paper presented at the 50th Hawaii International Conference on System Sciences, Hawaii.

Song, W.-M., Di Matteo, T., & Aste, T. (2012). Hierarchical information clustering by means of topologically embedded graphs. *PLoS One, 7*(3), e31929.

Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). SocialMedia Analytics: An Interdisciplinary Approach and Its Implications for Information Systems. *Business & Information Systems Engineering*(2), 89-96.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management, 39*, 156-168.

Strogatz, S. H. (2001). Exploring complex networks. *nature, 410*(6825), 268.

Tavakoli, A., Schlagwein, D., & Schoder, D. (2015). *Open Strategy: Consolidated Definition and Processual Conceptualization*. Paper presented at the International Conference on Information Systems (ICIS), Fort Worth, USA. http://aisel.aisnet.org/icis2015/proceedings/ISstrategy/2

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge: Cambridge University Press.

Yakovenko, V. M., & Rosser Jr, J. B. (2009). Colloquium: Statistical mechanics of money, wealth, and income. *Reviews of Modern Physics, 81*(4), 1703.