

Metrics selection for group type identification in Enterprise Social Network (ESN) analytics

Kai Riemer

Discipline of Business Information Systems
The University of Sydney Business School
Sydney, Australia
Email: kai.riemer@sydney.edu.au

Laurence Lock Lee

SWOOP Analytics Pty Ltd
Sydney, Australia
Email: llocklee@swoopanalytics.com

Cai Kjaer

SWOOP Analytics Pty Ltd
Sydney, Australia
Email: cai.kjaer@swoopanalytics.com

Annika Haeffner

Ulm University
Ulm, Germany
Email: annika.haeffner@alumni.uni-ulm.de

Abstract

With the proliferation of Enterprise Social Networks (ESN) measurement of ESN activity becomes increasingly relevant. ESN analytics aims to develop metrics and models to measure and classify ESN user activity to support organisational goals and outcomes. In this paper we focus on a neglected area of ESN analytics, metrics for understanding activity in ESN groups. We engage in explorative research, utilising actual ESN meta data describing activity in 350 groups across three organisations, to identify a set of metrics that divides the group sample into distinct user types. By employing clustering techniques we derive a set of four group types: broadcast streams, information forums, communities of practice and project teams. For ESN analytics research we contribute a set of metrics and group types. For practice we envision a method that enables group managers to compare aspirations for their groups to embody a certain group type, with actual activity patterns.

Keywords: Enterprise Social Network, ESN groups, ESN Analytics, Social Media Analytics.

1 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 (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). More generally, 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 and Schubert 2017), which is part of social media analytics (Behrendt et al. 2014b; Stieglitz et al. 2018), aims to develop metrics and models to examine ESN log file and content data to gain a better understanding of actual ESN usage patterns. 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 et al. 2017a). While existing work has presented metrics at the network and individual level, the intermediate, or group level of analysis, has not been covered so far. Yet, groups play an important role in ESN communication and collaboration practices. In this paper we are concerned with identifying metrics to distinguish between ESN group types. We ask the following research questions: *Which metrics are best suited to discriminate ESN groups into distinct types?*

We utilise an ESN activity meta-data set, sourced from the Yammer networks of three organisations, provided by Australian analytics company SWOOP Analytics. We engage in explorative research. We use cluster analysis and explore 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.

We find that a combination of three metrics best divided our group sample into clusters that are not only well-interpretable, but relate to, and extend the classification used by Yammer itself: 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, which we named 1) broadcast streams, 2) information forums, 3) communities of practice and 4) project teams.

Our study contributes to ESN research and ESN analytics an initial set of metrics, and a typology of groups based on activity patterns, as the basis for broader research into understanding the role of groups in ESN. For practitioners, such as network managers and group leaders, we outline the scaffold of a method for visualising the discrepancies between aspiration and actual activity in ESN groups. This 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. Given its pioneering nature, the study points to various promising avenues for future research.

2 Background

Enterprise Social Networks (ESN) are part of the Enterprise 2.0 phenomenon (McAfee 2009) which describes the application of social technologies such as social networking, blogs, wikis, or microblogging services within organisations (Razmerita et al. 2014). As such, ESN present a variant of public social networking services (Richter et al. 2011). Typical services are Microsoft Yammer, Facebook Workplace, IBM Connections, Chatter by Salesforce, and smaller services such as Mumbacloud, Jive or Tibbr.

2.1 ESN definition and characteristics

Common to all social media is that they facilitate user participation, interaction, and the generation of content by users (Boyd and Ellison 2007). Specifically, ESNs are services that are 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 (Leonardi et al. 2013).

Another defining characteristic of ESN is their malleability (Richter and Riemer 2013b). Unlike more traditional information systems that are employed to solve a concrete problem and are thus associated with a concrete task or purpose, ESN are best understood as infrastructures that are intended to create potentials for new ways of communicating and working (Riemer et al. 2009). Hence, the proliferation of ESN often follows a bottom-up approach of implementation, a process that is more inclusive and

egalitarian than that of traditional IS implementations (Schneckenberg 2009). ESN have been associated with a variety of organisational practices such as communication, collaboration (Riemer et al. 2010), knowledge management (Levy 2009) crowdsourcing (Schlagwein and Bjorn-Andersen 2014), open innovation (Dahlander and Gann 2010), or open strategy (Tavakoli et al. 2015). This renders ESN both an interesting and important context for IS research, and a challenge for organisations as they have to keep track of the emerging activity within their ESN. The research presented below aims to help organisations understand important aspects of their own enterprise social networks.

2.2 Prior research on ESN

Prior research on ESN typically falls within one of four existing streams. The first stream captures conceptual work outlining typical ESN characteristics (Leonardi et al. 2013; Treem and Leonardi 2012), comparing ESN with traditional ways of relationship building in organisations (Kane et al. 2014), envisioning high-level benefits, such as for knowledge-sharing (Fulk and Yuan 2013; Majchrzak et al. 2013) or interacting in the workplace more generally (DiMicco et al. 2008; Zhang et al. 2010).

The second stream comprises concrete, explorative, usually qualitative case studies investigating usage patterns of ESN. Typical findings reveal benefits of ESN for information sharing and discovery (Zhao and Rosson 2009), for creating awareness within the organisation (Zhang et al. 2010), or for knowledge creation and sharing in professional service contexts (Riemer and Scifleet 2012).

The third stream represents studies that measure the benefits of ESN from different angles and in different contexts; these studies are mostly quantitative in nature, employing either survey-based approaches, e.g. to measure individual ESN benefits for knowledge workers (Mantymaki and Riemer 2016), or they use data obtained from the ESN services directly, such as for studying the connection between social capital and employee performance (Riemer et al. 2015).

Finally, the fourth stream, to which this study contributes, contains generative works that aim to identify metrics and develop analytics frameworks for measurement of ESN activity in organisations. Studies in this context focus on the application of social network analysis (SNA) to reveal emerging informal networks between ESN users (Behrendt et al. 2014a), identify particular value-adding users (Berger et al. 2014), characterise individual user activity profiles (Holtzblatt et al. 2013), or develop a comprehensive framework for identifying user types in knowledge work contexts (Hacker et al. 2017b).

2.3 ESN analytics and metrics development

As an emerging stream of research, ESN analytics is a sub field of social media analytics (Stieglitz et al. 2014; Stieglitz et al. 2018). Sometimes called social collaboration (or social software) analytics, it is “a specialized form of examination of log files and content data, to gain a better understanding of the actual usage of ESS.” (Schwade and Schubert 2017, 401). Here, 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.

ESN analytics becomes relevant for organisations for two main reasons. Firstly, due to their malleability, concrete implementations of ESN in organisations typically emerge in quite different ways, supporting a range of different use cases and activities (Richter and Riemer 2013a). It is thus important for organisations to be able to track and understand user activity on their own particular ESN. Secondly, ESNs are becoming more integral to organisational communication and collaboration practices, which renders gaining an understanding for the activity on the platform and ways in which to support individuals and social groups within the ESN more important (Schwade and Schubert 2017).

At the same time, there are opportunities in ESN analytics for both academia and industry. First, the particular nature of ESN affords detailed analysis of user activity, because all (or most) user interactions are logged and, in principle, available for analysis. Second, most ESN platforms currently do not provide sophisticated analytics capabilities, leaving room for new offerings by third-party providers and for academia to explore new ways of measuring and accounting for various aspects of ESN activity.

Given the newness of this field the number of studies contributing to establishing metrics or models to support ESN analytics is still limited (cf. Schwade and Schubert 2017). So far two main areas of application for analytics exist. The first area focuses on metrics characterising the social network as such. This is where traditional social network analysis (SNA) techniques are brought to bear (Wasserman and Faust 1994). For example, Riemer et al. (2015) have shown how social capital metrics can be utilised to link certain network characteristics to employee performance. Behrendt et al. (2014b)

provide an overview of SNA metrics and studies for use in ESN contexts. The second area aims to develop new dedicated ESN metrics to characterise individual user behaviours 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 et al. 2017b). 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 et al. 2017).

2.4 ESN group analytics

What is lacking so far 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.

However, 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 and Tavakoli 2013). Yet this study has a very different purpose, and was not useful in the context of ESN analytics, since the classification was based on manual coding of messages, which is impractical for ESN analytics. We re-formulate our research question in more detail as: *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?*

3 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.

3.1 Research setting and ESN data set

SWOOP offers a cloud-based platform that provides organisations with analytics for their 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 et al. 2014a, 4). We distinguish usage data, or *meta-data*, from user-generated data, or *content*. Meta-data indicates *how*, *when* and *where* an ESN activity was performed, what kind of interaction it was and *who* was involved; content data 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 collects from an organisation’s ESN information that is not 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 or Like. Table 1 shows what meta data is 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 had access to data from Yammer networks of three firms (two financial services and one professional services company). The data provided by SWOOP (with the firms' permissions) 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 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.

3.2 Methodology

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 a cluster algorithm and proximity measure, 5) performing of cluster analysis, 6) interpretation of results. A tool was implemented using the software package Matlab to facilitate iterating on steps 2 to 6 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.

A 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). Previous research already demonstrated that clustering techniques are useful for classifying complex networks of different kinds (Newman and Girvan 2004; Strogatz 2001). For this study we experimented with a number of clustering algorithms (Song et al. 2012). Ultimately agglomerative clustering, in particular the complete-linkage algorithm (Defays 1977; Krznaric and Levkopoulos 1998) with a standardised Euclidean distance measure (Pandit and al. 2011) produced the most useful results.

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.

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 turned out not to be 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). Judgements regarding which metrics combinations discriminate best were jointly made by the paper's two academic authors.

3.3 ESN social graph and metrics

Any operationalisation of network metrics in the context of ESN has to begin with the construction of the social network graph (Riemer et al. 2015). Generally, a social network "consists of a finite set or sets of actors and the relation or relations defined on them" (Wasserman and 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 platform activity (Behrendt et al. 2014a).

At the most basic level a dyadic relationship between two individuals is said to exist when one user responds to another's message (Ahuja et al. 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 et al. 2014b). 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?

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 et al. (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 and Rosser Jr 2009) (see Appendix 1 for Gini calculation).

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.

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

4 Findings: four ESN group archetypes

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.

4.1 Metrics that best discriminate the groups sample

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

1. **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.
2. **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.
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.

4.2 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.

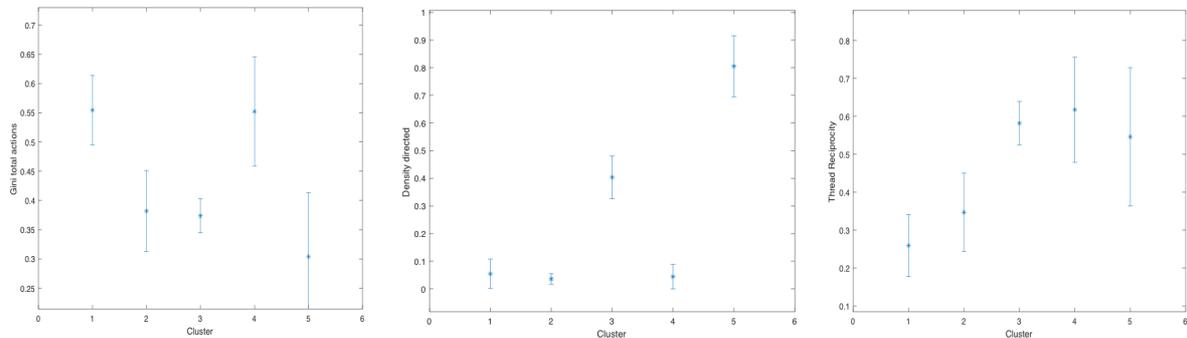


Figure 1: Metrics values for each of the resulting clusters

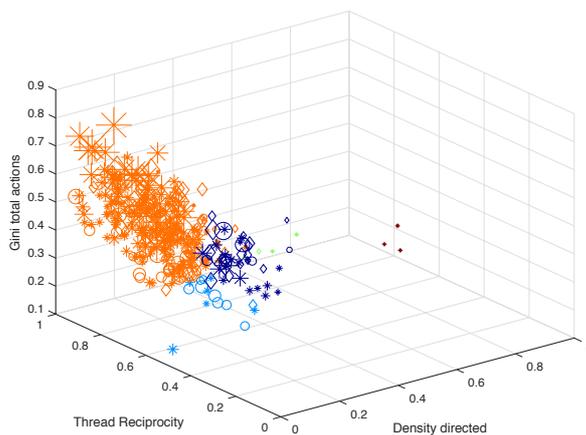


Figure 2: Cluster locations [shape of markers = company; marker size = group size]

Cluster			Metrics			# of active users		
#	Colour	Name	Particip. (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

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.

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 that they deem interesting for other users, or questions and requests, but which do not elicit many interactions or longer conversations.

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.

Project teams: These groups are by 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, and show 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 regular basis.

5 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, utilising activity meta-data from 350 groups from three organisations provided by SWOOP, converged on three metrics that measure 1) evenness of user participation, 2) density in terms of user connectedness in the group, and 3) reciprocity in terms of the proportion of messages eliciting replies. Those metrics in turn distinguish four distinct group types, which we named broadcast streams, information forums, communities of practice and project teams.

We note that these groups correspond to, yet extend in meaningful ways, the group types that Yammer used to provide (for a brief period of time) as a template for their users when creating new groups (see figure 3). Two of our categories, project teams (‘Project’) and broadcast streams (‘My Organisation’) have direct equivalents, 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.

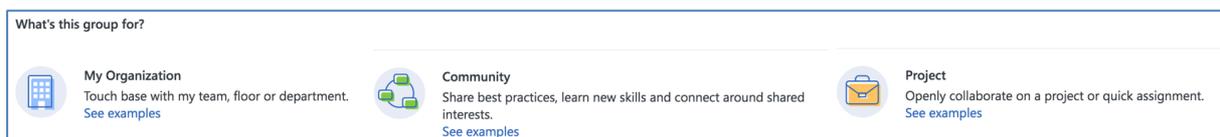


Figure 3: Yammer group template [note: Yammer recently suspended this classification feature]

Our finding suggests 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. The former, communities of practice, require more interaction and conversations between users (as measured by reciprocity), 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).

Initial feedback from SWOOP and its client base suggests that our typology will be helpful for ESN group leaders and community managers as a diagnostic tool for 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 be classified as a community of practice according to our metrics, which would indicate 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 4). We suggest that knowledge of such discrepancies will allow group leaders to manage and counteract accordingly, such as by motivating certain users to become involved in ESN activity.

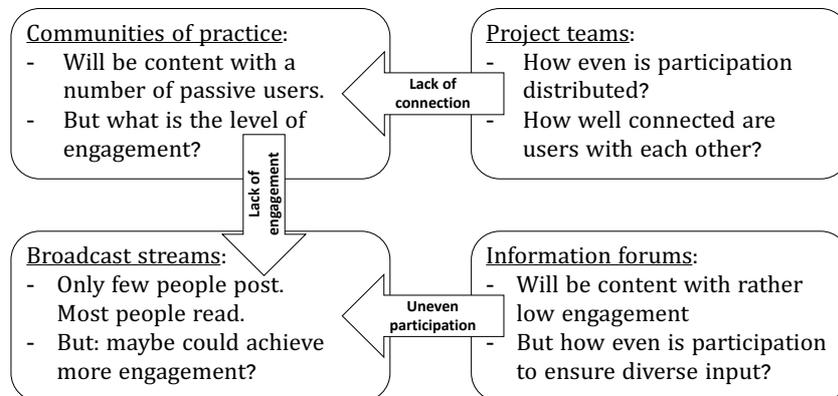


Figure 4: Examples of possible discrepancies between group aspiration and measured types

6 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 metrics 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. Such metrics could form the basis of a diagnostic tool with the aim to improve group communication so that activity patterns associated with the intended group type can be achieved. We envision that our metrics and classification could suitably be implemented in platforms such as that provided by SWOOP.

Without doubt, future research is needed to corroborate the findings presented here, since our study provides merely a first, necessarily limited step in a broader research endeavour to extend analytics to the group level. We envision that future research will apply similar explorative analyses 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 group types, such as those identified by (Hacker et al. 2017b). For example, will groups of certain types benefit from the presence of certain individual user types?

Moreover, qualitative research utilising interviews with network managers and group leaders might investigate the usefulness of our proposed typology, whether it captures all intended use cases for groups, and how it might be used to support decision-making. It would be particularly interesting to see how implementation of group metrics in platforms such as SWOOP will shape user behaviour.

References

- Ahuja, M. K., Galletta, D. F., and Carley, K. M. 2003. "Individual Centrality and Performance in Virtual R&D Groups: An Empirical Study," *Management Science* (49:1), pp. 21-38.
- Behrendt, S., Richter, A., and Riemer, K. 2014a. "Conceptualisation of Digital Traces for the Identification of Informal Networks in Enterprise Social Networks," *ACIS*.
- Behrendt, S., Richter, A., and Trier, M. 2014b. "Mixed Methods Analysis of Enterprise Social Networks," *Computer Networks* (75), pp. 560-577.
- Berger, K., Klier, J., Klier, M., and Richter, A. 2014. "Who Is Key?-Value Adding Users in Enterprise Social Networks," *Proceedings of the 2014 European Conference on Information Systems (ECIS)*.
- Boyd, D. M., and Ellison, N. B. 2007. "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Mediated Communication* (13:1), pp. 210-230.

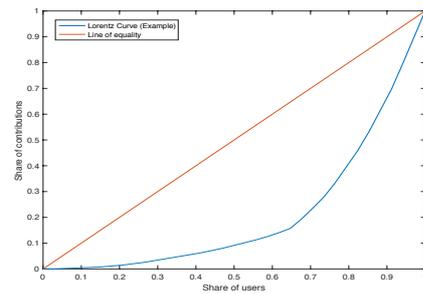
- Bughin, J. 2015. "Taking the Measure of the Networked Enterprise," *McKinsey Quarterly* (51:10), pp. 1-4.
- Dahlander, L., and Gann, D. M. 2010. "How Open Is Innovation?," *Research Policy* (39:6), pp. 699-709.
- Defays, D. 1977. "An Efficient Algorithm for a Complete Link Method," *The Computer Journal* (20:4), pp. 364-366.
- DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B., and Muller, M. 2008. "Motivations for Social Networking at Work," *Proceedings of the 2008 ACM conference on Computer supported cooperative work*: ACM, pp. 711-720.
- Dodd, L. 2011. "Study Reveals Huge ROI When Using Yammer," *Strategic Communication Management* (15:6), p. 7.
- Everitt, B. S. 1993. *Cluster Analysis (Third Edition)*. London Melbourne Auckland: Edward Arnold.
- Frank, L., Gimpel, H., Schmidt, M., and Schoch, M. 2017. "Emergent User Roles of a Digital Workplace: A Network Analysis Based on Trace Data," *38th International Conference on Information Systems*, Seoul, South Korea.
- Fulk, J., and Yuan, Y. C. 2013. "Location, Motivation, and Social Capitalization Via Enterprise Social Networking," *Journal of Computer-Mediated Communication* (19:1), pp. 20-37.
- Granovetter, M. 1973. "The Strength of Weak Ties," *American Journal of Sociology* (78:6), pp. 1360-1380.
- Hacker, J. V., Bernsmann, R., and Riemer, K. 2017a. "Dimensions of User Behavior in Enterprise Social Networks," in *Social Knowledge Management in Action: Applications and Challenges*, R. Helms, J. Cranefiels and J. van Reijssen (eds.). Cham, Switzerland: Springer International Publishing, pp. 125-146.
- Hacker, J. V., Bodendorf, F., and Lorenz, P. 2016. "A Framework to Analyze Enterprise Social Network Data," in *Enterprise Big Data Engineering, Analytics, and Management*, M. Atzmueller, S. Oussena and T. Roth-Berghofer (eds.). Hershey: IGI Global, pp. 84-107.
- Hacker, J. V., Bodendorf, F., and Lorenz, P. 2017b. "A Framework to Identify Knowledge Actor Roles in Enterprise Social Networks," *Journal of Knowledge Management* (21:4), pp. 817-838.
- Hartigan, J. A. 1975. *Clustering Algorithms*. New York: John Wiley & Sons.
- Holtzblatt, L., Drury, J., and Weiss, D. 2013. "Evaluating the Uses and Benefits of an Enterprise Social Media Platform," *Journal of Social Media for Organizations* (1:1), pp. 1-21.
- Kane, G. C., Alavi, M., and Labianca, G. J. 2014. "What's Different About Social Media Networks? A Framework and Research Agenda," *Management Information Systems Quarterly* (38:1), pp. 274-304.
- Krackhardt, D. 1992. "The Strength of Strong Ties: The Importance of Philos in Organizations," *Networks and organizations: Structure, form, and action* (216), p. 239.
- Krzmaric, D., and Levkopoulos, C. 1998. "Fast Algorithms for Complete Linkage Clustering," *Discrete & Computational Geometry* (19:1), pp. 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), pp. 747-762.
- Leonardi, P. M., Huysman, M., and 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), pp. 1-19.
- Levy, M. 2009. "Web 2.0 Implications on Knowledge Management," *Journal of Knowledge Management* (13:1), pp. 120-134.
- Majchrzak, A., Faraj, S., Kane, G. C., and Azad, B. 2013. "The Contradictory Influence of Social Media Affordances on Online Communal Knowledge Sharing," *Journal of Computer-Mediated Communication* (19:1), pp. 38-55.
- Mantymaki, M., and Riemer, K. 2016. "Enterprise Social Networking: A Knowledge Management Perspective," *International Journal of Information Management* (36:6), pp. 1042-1052.
- McAfee, A. 2009. *Enterprise 2.0: New Collaborative Tools for Your Organization's Toughest Challenges*. Boston: McGraw-Hill Professional.
- McKinsey. 2012. "The Social Economy: Unlocking Value and Productivity through Social Technologies." from http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_social_economy
- Newman, M. E. J., and Girvan, M. 2004. "Finding and Evaluating Community Structure in Networks," *Physical Review* (69:2), pp. 26-113.
- Pandit, S., and al., e. 2011. "A Comparative Study on Distance Measuring Approaches for Clustering," *International Journal of Research in Computer Science* (2:1), pp. 29-31.
- Razmerita, L., Kirchner, K., and Nabeth, T. 2014. "Social Media in Organizations: Leveraging Personal and Collective Knowledge Processes," *Journal of Organizational Computing and Electronic Commerce* (24:1), pp. 74-93.

- Richter, A., and Riemer, K. 2013a. "The Contextual Nature of Enterprise Social Networking: A Multi Case Study Comparison," *21st European Conference on Information Systems ECIS 2013*, Utrecht University, Netherlands.
- Richter, A., and Riemer, K. 2013b. "Malleable End-User Software," *Business & Information Systems Engineering* (5:3), pp. 195-197.
- Richter, D., Riemer, K., and vom Brocke, J. 2011. "Internet Social Networking," *Business & Information Systems Engineering* (53:2), pp. 89-103.
- Riemer, K., Altenhofen, A., and Richter, A. 2011. "What Are You Doing? - Enterprise Microblogging as Context Building," *European Conference on Information Systems (ECIS)*, Helsinki, Finland, pp. 1-13.
- Riemer, K., Finke, J., and Hovorka, D. S. 2015. "Bridging or Bonding: Do Individuals Gain Social Capital from Participation in Enterprise Social Networks?," *International Conference on Information Systems ICIS 2015*, Fort Worth, United States.
- Riemer, K., Richter, A., and Böhringer, M. 2010. "Enterprise Microblogging," *Business & Information Systems Engineering* (2:6), pp. 391-394.
- Riemer, K., and Scifleet, P. 2012. "Enterprise Social Networking in Knowledge-Intensive Work Practices: A Case Study in a Professional Service Firm," *Proceedings of the 23rd Australasian Conference on Information Systems 2012: ACIS*, pp. 1-12.
- Riemer, K., Steinfield, C., and Vogel, D. 2009. "Ecollaboration: On the Nature and Emergence of Communication and Collaboration Technologies," *Electronic Markets* (19:4), pp. 181-188.
- Riemer, K., and Tavakoli, A. 2013. "The Role of Groups as Local Context in Large Enterprise Social Networks: A Case Study of Yammer at Deloitte Australia," The University of Sydney, Sydney.
- Schlagwein, D., and Bjorn-Andersen, N. 2014. "Organizational Learning with Crowdsourcing: The Revelatory Case of Lego," *Journal of the Association for Information Systems* (15:11).
- Schneckenberg, D. 2009. "Web 2.0 and the Empowerment of the Knowledge Worker," *Journal of Knowledge Management* (13:6), pp. 509-520.
- Schwade, F., and Schubert, P. 2017. "Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities," *50th Hawaii International Conference on System Sciences*, Hawaii, pp. 401-410.
- Song, W.-M., Di Matteo, T., and Aste, T. 2012. "Hierarchical Information Clustering by Means of Topologically Embedded Graphs," *PLoS One* (7:3), p. e31929.
- Stieglitz, S., Dang-Xuan, L., Bruns, A., and Neuberger, C. 2014. "Socialmedia Analytics: An Interdisciplinary Approach and Its Implications for Information Systems," *Business & Information Systems Engineering*:2), pp. 89-96.
- Stieglitz, S., Mirbabaie, M., Ross, B., and Neuberger, C. 2018. "Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation," *International Journal of Information Management* (39), pp. 156-168.
- Strogatz, S. H. 2001. "Exploring Complex Networks," *nature* (410:6825), p. 268.
- Tavakoli, A., Schlagwein, D., and Schoder, D. 2015. "Open Strategy: Consolidated Definition and Processual Conceptualization," in: *International Conference on Information Systems (ICIS)*. Fort Worth, USA.
- Treem, J. W., and Leonardi, P. M. 2012. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association," *Communication yearbook* (36), pp. 143-189.
- Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- Yakovenko, V. M., and Rosser Jr, J. B. 2009. "Colloquium: Statistical Mechanics of Money, Wealth, and Income," *Reviews of Modern Physics* (81:4), p. 1703.
- Zhang, J., Qu, Y., Cody, J., and Wu, Y. 2010. "A Case Study of Microblogging in the Enterprise: Use, Value, and Related Issues," *28th annual SIGCHI conference on Human factors in computing systems*, Atlanta.
- Zhao, D., and Rosson, M. B. 2009. "How and Why People Twitter: The Role That Micro-Blogging Plays in Informal Communication at Work," *Proceedings of the ACM 2009 international conference on Supporting group work*: ACM, pp. 243-252.

Appendix 1

Algorithm for 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)



Copyright

Copyright: © 2018 Kai Riemer, Laurence Lock Lee, Cai Kjaer, Annika Haeffner. This is an open-access article distributed under the terms of the [Creative Commons Attribution-NonCommercial 3.0 Australia License](https://creativecommons.org/licenses/by-nc/3.0/australia/), which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and ACIS are credited.