Quality in peer production systems – impact of assortativity of communication networks on group efficacy

Agnieszka Rychwalska
University of Warsaw
a.rychwalska@uw.edu.pl

Szymon Talaga
University of Warsaw
stalaga@protonmail.com

Karolina Ziembowicz
Maria Grzegorzewska
Pedagogical University
klisiecka@gmail.com

Abstract

Many online peer production systems (e.g. Wikipedia or Open Source Software communities) strive to deliver high quality intellectual goods that could compare with commercial products. While quality is key to the communities' success – widespread adoption of their products – it is not clear what makes some succeed, while others provide subpar outcomes or fail entirely. Quality of Wikipedia articles has been previously related to the number of editors writing them or to the diversity of editors' competences. Here we tested the hypothesis that cohesiveness of private communication networks within collaborating groups increases the quality of their products. We analyzed communication within a sample of Wikiprojects on the English Wikipedia – groups of editors that coordinate their activities to improve articles related to a specific topic. We found that most Wikiprojects communicate in a highly hierarchical, disassortative way, but the successful ones break this trend and their communication networks are structured in a more egalitarian way.

1. Introduction

Widespread adoption of Information and Communication Technologies (ICT) in interpersonal and group interaction has spurred the growth of online peer production communities – collectives of individuals that produce (mostly intellectual) goods for common use [6]. As an example, Wikipedia is the 5th most visited website [35], Apache Webserver has been the dominant server software for online services for 20 years [36] and such knowledge bases as StackOverflow are in daily use by professionals and amateurs alike. While these communities have gained public acclaim and their products are in widespread, everyday use by the public at large, they constitute only a small percentage of the ever growing number of such production collectives. Most of them remain a niche for enthusiasts and never produce anything of sufficient quality to warrant public attention (e.g. many Open Source Software projects never take off or are never widely adopted). Moreover, the success cases mentioned here are also not immune to quality mishap. Many Wikipedia articles are underdeveloped, leaving the reader with questionable knowledge and some are downright erroneous. Similarly, bugs in the code of Open Source Software can lead to tremendous losses for the companies and institutions that rely on them – this was the case, for example, with the Heartbleed bug in OpenSSL encryption libraries which are the backbone of over half of online shopping services [39].

Given the growth of peer production communities and the increase in demand for their (free) products, it becomes more and more important to ensure that even in such grassroots production regimes that do not use monetary incentives to promote diligence and conscientiousness, quality of the end product remains high. Maintaining high quality of the outcomes – i.e. quality comparable to counterpart products developed in standard organizations with financial incentives – is critical for the success of many peer production communities and is often established as collaboration goal. Yet, it is still unclear what makes some of them efficacious in achieving this aim and allows them to deliver high quality products, while others provide subpar outcomes or fail entirely.

One of the hypothesized factors that can influence quality in online production communities is the size of the collaborating group. Success in peer production depends on self-selection to all the specific tasks – or parts thereof – that need to be completed [7]. For this to happen, tasks need to be modular and preferably of varied granularity, so different amounts of motivation on the part of volunteers can be matched with the different required efforts [6]. Thus, the bigger the initial pool of interested individuals, the bigger the chances that all parts of the project will be brought to a successful completion. Similarly, quality control can be peer produced, that is redistributed to volunteers, as

HĮCSS

"given enough eyeballs, all bugs are shallow" [25]. However, this optimism about the "strength is in numbers" or "wisdom of the crowds" may be overstated. A simple increase of editors in a Wikipedia article does not by itself lead to improvement of quality [16]. Another factor, diversity of competences brought to the table, might be important.

Diversity of knowledge, skills and mental models has been suggested as one of the most important factors leveraging the wisdom of the crowds phenomenon [24]. Differences in viewpoints and knowledge bases cause a "creative abrasion" among Wikipedia editors that can lead to improvement in article quality [1]. Moreover, the diversity in the subject pool of editors' competences predicts well whether an article would gain the Featured Article (FA) status – that is, would be considered by the community as one of the best quality articles, worthy of a notice on the front page of the Wiki [4, 30]. Yet, an even more important matter than diversity might be the way that those many little activities and tasks performed, coming from volunteers with different knowledge bases, get coordinated and combined into the final product.

Coordination of work in any organization that relies on multiple actors performing their tasks is a key to success, and coordination costs limit the size of standard organizations [8]. Seemingly, this limit is removed from online peer production communities because in part the coordination is enacted by the technology itself [28]. For example, human workers (coordinators) indexing search results (acquired from distributed providers) have been replaced by search engine algorithms. Similarly, algorithms enable filtering or summarizing content in any social platform, e.g. through selective digests, newsfeeds, hashtag filters, and so on.

However, for more complex tasks it seems that human coordination is still needed. When the number of editors of a Wikipedia article grows, coordination is crucial for quality [16]. In other communities also running as a wiki, the situation is similar – as the community grows, so grows coordination: through direct communication as well as through role creation (core / periphery), and indirectly through formulation of procedures and policies [17].

Interestingly, communication not directly related to articles (i.e. private communication) grows faster than task related discussions [17]. This may be important for more general coordination activities, such as maintaining high motivation of otherwise not incentivized contributors. Indeed, Wikipedians are more driven by the social motivations, such as altruism rather than self-development [23], and draw satisfaction from identifying with the community [27]. This is in line with studies on offline groups where cohesion has been identified as one of the factors impacting group efficacy [5, 11] and other online communities where group members' attractiveness and task identification both lead to higher engagement [18].

To foster such social constructs as identity and engagement, coordination through communication might be crucial. Such communication can be a vehicle to provide others with feedback on group processes - in virtual teams lack of such leads to a loss in efficacy [14]. This would suggests that communication directly related to task has different impact than private communication. Indeed, when analyzing the types of messages that editors of the English Wikipedia leave on the talk pages of articles (that is, the discussion spaces that serve as the main coordination point for each article) "Critiquers" had been found to have more impact on quality than "Encouragers" [12]. The situation is quite different on personal communication pages (user pages). There, receiving a message in itself increases motivation to contribute and when the message is positive (e.g. a recognition of contributions, or usage of smileys) this effect grows; at the same time, negative messages decrease the propensity to contribute [33].

Private communication – often only loosely related to particular tasks – is thus an important coordination activity that leads to higher engagement and motivation and hypothetically to group identity. However, it is still not clear how private messages impact quality of the product developed by a peer production group. Moreover, the studies mentioned so far focused on the content of the messages, i.e. what they transferred, and did not touch upon the structure of the group, i.e. who communicates with whom. Thus in the present study we set out to answer the research question of how the communication network properties, especially those related to group cohesion, impact the efficacy of a peer production group's work – product quality. To shed some light on this issue we turned to the English Wikipedia; however, we focused not on a selection of articles but rather on a selection of editor groups and their efficacy.

As our unit of observation we chose the so called Wikiprojects on the English Wikipedia. Wikiprojects are groups of editors that declare interest in a specific subject area and coordinate their activities to improve related articles. Each project has a dedicated space of a few pages on the Wikipedia (outside of the encyclopedic articles), where the group coordinates their activity – describes its long term goals and current tasks to be completed, stores the know-how and best practices in editing Wikipedia and monitors the changes in the

articles within their care. The articles to curate are marked as within interest of a project by its members or any other Wikipedian by placing an alert on the article's talk page. The topics that Wikiprojects curate vary greatly in scope, with the biggest – Wikiproject Biography – taking care of over 1.7 million articles, as well as in the size of the collaborating group – from few members to almost 900 for Wikiproject Military History. Yet, the one common trait they have is that all Wikiprojects as their goal set out to improve the quality of the articles they curate.

We gathered data on the private messages sent within a sample of Wikiprojects to extract communication networks. We analyzed network metrics typically associated with cohesiveness (density, clustering, assortativity, among others) and related them to the number of Featured and Good Articles that the projects have contributed to. These quality tags are assigned through a peer review process to articles that meet criteria established by the whole Wikipedia community and thus provide a fairly objective measure of a Wikiproject's efficacy in achieving its main goal. Based on the results of studies described above, pointing to the importance of private communication in increasing engagement, we hypothesized that successful projects would exhibit a denser communication structure with more clustering, with more editors linked into a single communication group, and with more egalitarian ties (i.e. neither assortative nor disassortative).

2. Method

2.1. Data

We extracted private, not directly task related communication between Wikipedia editors (registered users) from user talk pages for users who were active in the 90 days before data collection commenced (June – August 2018)[1]. The time limit was selected based on the Mediawiki database structure, where the "recent changes" table records edits in that precise time span. We set an additional requirement of at least 100 total edits (in any part of the Wikipedia) for the users to ensure that they were active and well socialized members of the community and that they would have higher chances of communicating with one another. Our initial set of editors numbered $N = 37487$.

To extract incoming communication we have parsed all of the selected editors' user talk pages and identified all usernames that posted messages on each of the editors' talk pages. To extract outgoing communication, we have analyzed all edits that the selected editors committed in the user talk namespace and identified the usernames on whose pages the edits were made. We have excluded messages that contained newsletters and other mass-messaging posts. For each editor we have thus gathered a set of incoming and outgoing links.

Next, we chose the 200 Wikiprojects whose pages were edited by the highest number of the selected users. For this we have analyzed all revisions committed by the chosen editors and identified which Wikiprojects' pages each of them edited. Then we have ranked the projects according to how many unique editors from our sample edited any page curated by them. The ranking excluded maintenance projects such as copyediting, disambiguation, referencing units and similar.

For the selected projects we have scraped member lists from their project pages. Each project has a dedicated space in the so called project namespace on the Wikipedia. This namespace – in contrast to the main namespace that is dedicated to the encyclopedic articles – holds all coordination and maintenance activities: policies and guidelines, help departments, noticeboards, dispute resolution discussions, idea labs for new community-wide solutions and also the Wikiprojects' pages. In this space each of the projects maintains a member list, which any Wikipedian can edit to add their user name and thus join the project.

Member lists are the most precise way to identify project members – since the editors listed had to manually edit the list to add their names, we could be sure that they had at least visited the project webpages and thus were more likely to have been active in the project's activities and also more likely to have been under the project's coordination. An alternative method of identifying project participants based simply on which pages of the encyclopedia they edited would have included editors not active in the projects, who might not be aware that projects exists and thus would have produced a much noisier dataset for assessing a project's communication network properties.

The format of the member lists as well as the URL format of their location vary considerably across projects and thus the lists had to be manually copied from the webpages and semi-automatically parsed into a uniform format. This necessitated limiting the sample

---

[1] The data and script developed to perform analyses are deposited at:

https://osf.io/cz4p8/?view_only=b16899235d0c4df6a523b03 8d6bcadee

to 200 projects instead of analyzing the whole set of 2097 Wikiprojects (as of data collection period).

Finally, we matched the scraped member lists with the list of selected editors and arrived at a set of $N = 4514$ of Wikiproject members with their within project communication networks.

To gather quality scores of Wikiprojects' work we have retrieved all encyclopedic articles from the main namespace using the Mediawiki API along with the list of the Wikiprojects that have them in their scope and the article ratings (i.e. quality tags).

2.2. Measures

2.2.1. Communication network measures. All network analyses were performed using *igraph* package in R [10]. Based on directed communication between editors we have constructed directed communication networks for all 200 projects. However, we then treated each connection – no matter if reciprocated – as an undirected link to reconstruct the networks into undirected ones. There is a custom on the English Wikipedia to reply to a message posted by another user on one's own talk page. That is, if one user messages another, the recipient replies on his or her own user page rather than on the talk page of the sender. The database used in this study did not contain messages posted by users on their own talk pages and thus we could not assess whether an incoming message was answered or not. Reciprocal posting of messages on each other user talk pages usually reflects two or more different discussion topics and thus treating only such cases as reciprocal connections would exclude many potentially important links.

For further analysis we have included only those projects whose communication networks had a giant component of at least 8 nodes in size to be able to estimate network measures ($N_{project} = 148$, $N_{editor} = 4259$). The cutoff value was chosen to balance the stability of network measures of cohesiveness (most of which are dependent on network size) and the size of the sample. Networks with giant components of less than 3 nodes in size (2 in the sample) were instantly excluded because they do not allow computation of the clustering coefficient. Giant components of size 3 to 5 produce measures which are highly dependent on random changes of even a single edge. For example, in a three element giant component (which has either 2 or 3 edges) clustering coefficient can only assume two values – either 1 or 0 – and density changes by one third with an addition or removal of a random connection. To avoid choosing an arbitrary cut off value we have chosen the

closest integer below the 1st quartile of the distribution of the size of giant components in the sampled projects (8 nodes). This cut off choice provided sufficient estimation stability while excluding less than 25% of data points.

For each network we have computed: node count (the number of editors from the project's member list that were active in the previous 90 days and had a total edit count of at least 100), fraction of nodes in the giant component (the largest fraction of editors who were linked together into a cohesive network), density (the fraction of existing connections in the network relative to the possible number of connections given the network size), global clustering coefficient (the relative number of closed triangles between any three nodes in the network) and degree assortativity (the level to which similarly connected nodes link to each other; assortativity is positive when highly connected nodes link to other highly connected nodes and poorly connected to other poorly connected, and negative when highly connected nodes link to poorly connected nodes).

Since editors can join multiple projects and we cannot establish whether their private communication (not directly related to any particular article) is related to a project, there can be cases where certain pairs of connected editors appear in many projects. To assess how much such overlap may affect the results, we have computed Jaccard distance between the sets of nodes in each project pair, i.e. the size of the intersection of the sets divided by the size of the sum of the sets; the measure ranges from 0 (identical sets) to 1 (completely dissimilar sets). On average the projects' distance was .99, and the minimal distance was .97. Thus we concluded that the networks were sufficiently dissimilar and their properties to a large degree affected only one project's efficacy.

2.2.2. Quality measurement. Efficacy of a Wikiproject can be estimated by looking at the quality scores of each of the articles under its curation. On the Wikipedia, articles' quality can range from "Stub" (a barebones placeholder for a topic) to Featured Article (recognized as top content on the encyclopedia). However, most of this scale is assessed by each Wikiproject individually, according to its specific criteria which might differ among projects and thus are difficult to compare. The two grades that are universal among all projects and all pages are Good Article (GA) and Featured Article (FA). These quality ratings are awarded in a peer review process by editors independent of the given Wikiproject, on criteria commonly agreed by the whole Wikipedia community [34, 37]. Therefore, we treated these as indicators of high quality.

For each project we computed the total number of FAs and GAs. However, it rarely is the case that a GA or FA is curated by a single project as the thematic scopes often overlap (only 14.1% of the pages within the scope of our selected projects were under a single project's care). To control for this overlap in the projects' products we have weighted each FA and GA by the number of projects that had it in its scope. Thus, an FA that was curated by 5 projects counted as one fifth of a FA to each, while one that was the sole responsibility of a single project, counted as one to its total score of FAs. The correlation between the number of FAs and GAs was .93 suggesting that these two scores are generated in a similar process of quality control. We have thus summed these weighted scores for each project.

Another thing to take into account when estimating a project's quality is the thematic scope of the project. A project that deals with a single city in a country (e.g. Wikiproject New Orleans) will naturally have fewer articles in its scope than project Biography which is tasked with taking care of all biographies, no matter the nationality or era of the person described. It can be assumed that the number of high quality articles is related to the total number of articles of a project. The total is an obvious limit on the number of FAs and GAs but also the more there are articles to choose from, the higher the chance that self-selecting volunteers will find some topic of true interest and invest effort to improve it. To counter for the differences in the scope of Wikiprojects we have normalized their quality scores by the total number of articles in their charge.

3. Results

3.1. Testing the model

Our hypothesis stated that cohesive private communication among volunteers in peer production groups would increase their efficacy, i.e. the products resulting from such collaboration would be of higher quality. To verify this hypothesis we tested a linear model with the computed quality score as the response variable and network density, fraction of nodes in the giant component, clustering coefficient and degree assortativity as predictors. We have also controlled for the number of editors on the projects' member lists to exclude the simple effect of project size on its capacities to produce high quality outcomes. The response variable had an approximately log-normal distribution and thus a logarithmic transform was used in the models. The

descriptive statistics of the predictors and the response variable are given in Table 1.

The model explained 12% of the variance in projects' quality scores; however, only one predictor, assortativity, was significantly related to quality ($F_{(5, 142)} = 4.02$, $p = .002$, Table 2., Model 1). We thus ran a model with only assortativity on the predictor side and this simplified model explained 11% of the variance of quality scores ($F_{(1, 146)} = 18.64$, $p < .001$, Table 2., Model 2). The difference in variance explained between these two models was insignificant ($F_{(4, 142)} = .45$, $p = .77$) suggesting that it is not justified to use the full model to explain differences in efficacy among Wikiprojects.

Table 1. Descriptive statistics

Variable	Mean (SD)	Median
Quality	.005 (.005)	.003
Density	.066 (.04)	.062
Fraction in giant component	.644 (.18)	.647
Clustering coefficient	.214 (.11)	.213
Number of project members	186.4 (161)	135
Degree assortativity	-.338 (.19)	-.327

Since our analysis of the communication network properties' impact on quality was exploratory, we wanted to test the robustness of the model. We repeated the single predictor model analysis on 20 randomly chosen sets of projects, each of which spanned 80% of the full Wikiproject list. The computed R^2 in the analyses ranged from .086 to .152, with the mean at .119 and the median at .122. R^2 from the full dataset was between the 1st quartile and the mean of the distribution suggesting that it reflects a true effect present in the data.

3.2. Assortativity and quality

Given that quality of peer produced work is influenced by a whole range of factors besides coordination activities and private communication, assortativity of communication networks explains a surprisingly large amount of variance in Wikiprojects' quality of outcomes. To see what mechanisms might lie underneath this relation we took a closer look at the distribution of assortativity in the sampled communication networks.

Wikiprojects' private messaging networks are strongly disassortative – assortativity ranged from -.81

to .23 with the mean at -.34 and median at -.33 (Table 1.) and only six projects had positive assortativity (4% of the sample). This means that Wikiproject members who have many connections to others (i.e. send and/or receive many messages) are predominantly connected to members who have few connections. Such a communication structure might be considered hierarchical with a few nodes – "leaders" or hubs – being the sociometric stars and dominating the communication. It is also different from structures commonly found in social networks which exhibit either positive assortativity, e.g. collaboration networks [21, 22], or have neutral or mildly disassortative structures: assortativity was close to zero for online dating networks [15], and between -.1 and 0 for peer to peer sharing networks [32] and social networking sites [13].

Table 2. Quality prediction models

Predictor	Model 1	Model 2
Density	.221 (3.298)	
Fraction in giant component	.507 (.591)	
Clustering coefficient	-.544 (.855)	
Number of project members	-.0004 (.0006)	
Degree assortativity	1.907*** (.453)	1.625*** (.147)
Constant	-5.305 (.331)	-5.256 (.147)
R^2	.124	.113

Note: Linear regression coefficients; standard errors in parentheses; *** $p < .001$

To test whether the strongly negative assortativity (i.e. disassortativity) stems simply from the particular degree distributions in Wikiprojects' networks, we compared the empirical assortativity scores to simulated networks. For each Wikiproject we have simulated 100 networks of exactly the same size and approximately the same degree distribution (a null, configuration model [20]) and computed the standardized distance (z score) of the empirical value from the mean of the distribution of the simulated networks. The histogram of the z scores (Fig 1. a) shows that Wikiprojects' disassortative communication structure is not solely the effect of the degree distributions – most networks have lower assortativity than the null model generated networks.

It is worth noting however that in some popular network generation models assortativity goes to zero with network size – e.g. in the Barabàsi-Albert preferential attachment (PA) model [21] which depicts a network growing by new nodes linking preferentially to those nodes that have more connections [3]. In our sample of relatively small networks, assortativity is weakly, positively related to the member list size (Pearson $r = .33$, $p < .001$) so it is possible that if the projects' size grew towards infinity assortativity would grow from the highly negative values towards zero. To test if this model of network growth accounts for the properties of the empirical networks, for each Wikiproject we have simulated 100 preferential attachment networks with the same size and similar density (at each step of the PA growth the number of new connections introduced with the new node was based on the empirical degree sequence) and computed the distance of the empirical assortativity from the mean of the simulations' distribution (Fig 1. b).
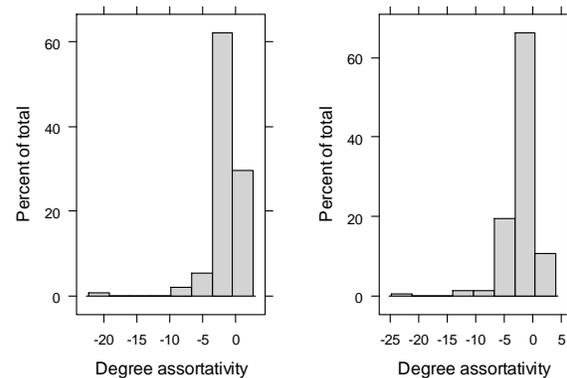

Figure 1. Comparison of Wikiprojects' communication networks assortativity to assortativity in simulated samples of *N*=100 networks generated from the null model (a) and a preferential attachment (b) model, maintaining network size and degree distribution. The distributions show the frequencies of differences of the empirical assortativity of each network as compared to the mean of the sample of simulated networks (*z* score). If the theoretical models were to fully explain the empirically observed assortativity, the z scores would be close to 0, yielding a single peak in the distributions around zero.

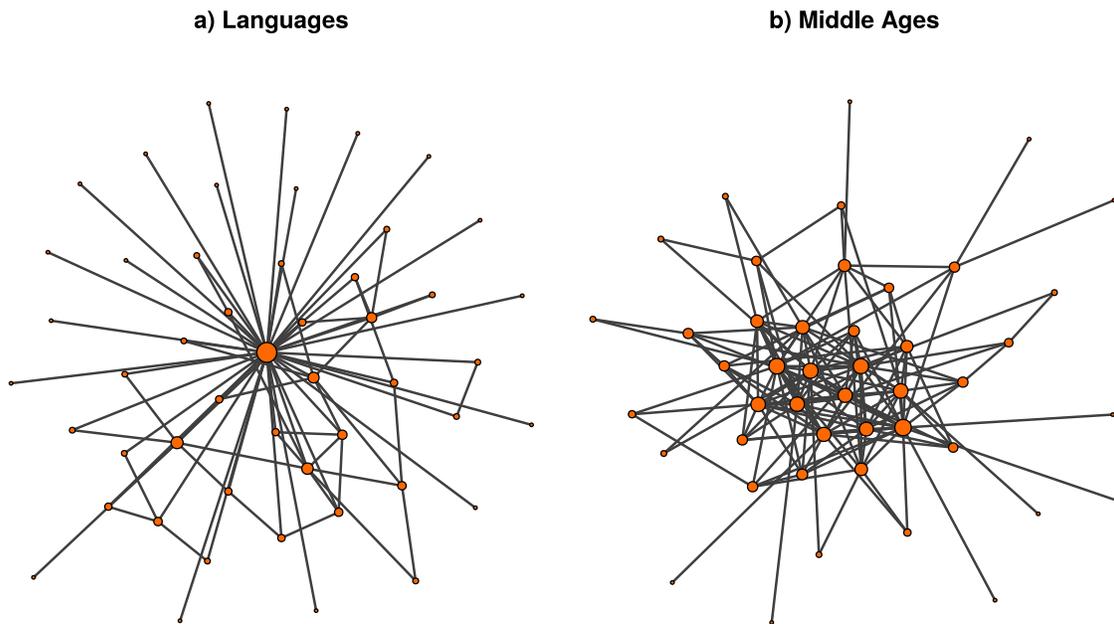**a) Languages** **b) Middle Ages**

Figure 2. Sample Wikiproject communication networks from a low scoring (a) and a high scoring (b) project in terms of product quality. Size of nodes is proportional to their degree.

Again, the observed assortativity scores fall well below what would be warranted by the network size and degree distribution in a preferential attachment model. We can conclude that the communication networks in Wikiprojects are generated by some other, nontrivial mechanism.

While in general the networks are disassortative the model showed that there is also a positive relation between assortativity and quality; that is, as assortativity grows from highly negative values towards zero, the quality of articles grows as well. In effect, Wikiprojects scoring highest on the quality score exhibit predominantly only mildly disassortative or entirely neutral connections: an editor's degree does not determine with whom he or she connects. This suggests that while there is a trend for the projects to maintain mostly links between highly connected nodes and poorly connected nodes, it is the projects which break this trend that become the most efficient in producing quality content. Figure 2 presents communication network visualizations for two sample projects of similar size of the giant component (45 nodes each) and similar density. The first, Wikiproject Languages (Fig 2. a) is ranked in the lowest ten projects in terms of quality: it curates over 10 thousand articles out of which 29 have reached either FA or GA status. The second project, Wikiproject Middle Ages (Fig 2. b), is in the top

ten most efficient projects, with almost 16 thousand articles under its care and over 800 articles with FA or GA status. The structures of these two projects' communication networks are different – the first consists of a clear hub connecting to many nodes of few connections, rendering an assortativity score of -.47, while the second has a larger fraction of nodes with many links going to both other highly connected nodes as well as to poorly connected ones and in effect its assortativity is -.15.

4. Discussion

The general hypothesis that the efficacy of a Wikiproject's work is influenced by the structure of its private, not directly task related communication network has been corroborated by the data. However, most of the group cohesion measures extracted from network structure proved insignificant. On the one hand, it might be surprising that e.g. clustering of the networks, usually reflecting the redundancy of ties in closely knit social circles, did not at all weigh in on the groups' efficacy in improving the quality of articles under their care. On the other, it is important to remember that the Wikipedia community as a whole, and the Wikiprojects themselves are typical task-oriented groups (Wikipedia's policies explicitly

underscore that it is not a place for socializing [38]). While private communication is usually only generally linked to users' work on the Wiki, it is definitely not typical socializing, as found on social media. Therefore, the impact of private communication can be expected to be limited, especially in comparison to other factors such as coordination on the articles' talk pages or dispute resolution.

Moreover, Wikiprojects have no ownership of any articles; the fact that they curate some does not limit anyone else from editing the pages. Therefore, quality of the product is impacted not only by the Wikiprojects' activities, including private communication, but by activities of all involved editors, some of whom might not be affiliated with any project or with the Wikipedia community as a whole.

Taking all these factors into account, the overall predictive power of the model is satisfying – it explains a substantial amount of variance of Wikiprojects' efficacy with a single predictor. This predictor, degree assortativity, explains how the structure of communication between hubs and other nodes in the network is different for successful and unsuccessful projects.

In general, Wikiproject networks are disassortative, which means that nodes with many connections link predominantly to nodes with few connections. Not only are the values of assortativity negative, they are also way below of what has been identified in other disassortative social networks [13, 15, 32].

The extreme values of disassortativity might be an effect of the function that Wikiprojects serve. They are task oriented groups, whose main purpose is coordination of editor activities. Thus, a disassortative structure, with clear hubs in the communication network that relay communication, might be more natural and effective, especially at the beginning, when a project is formed and requires efficient decision making structure. Another factor that might generate this type of communication structure is that Wikiprojects evolve from an initial, small number of users – project creators. Just as any task to be performed on the Wiki, users self-select to create a Wikiproject. From there on they promote, advertise and invite other users to join and contribute. In effect, just as in many other peer production systems, there exist one or two benevolent dictators for each Wikiproject whose large number of connections in the communication network is formed into a star like, disassortative structure.

The most important results of our study is, however, that successful projects diverge from this scenario – instead of the hubs dominating communication, the peripheral members are also connected among themselves, promoting a more "egalitarian" network in that the node degree does not strongly determine with whom the node communicates. Thus, peripheral, weakly connected editors are as likely to communicate with other peripheral project members as they are to communicate with editors who are highly connected to others. While the assortativity still remains slightly negative for effective projects, it is more similar to other online social networks than unsuccessful projects.

There might be various mechanisms that generate this network structure – it is not explained by either the specific degree distribution in the empirical networks or by a simple preferential attachment model. Most probably projects start with a hierarchical communication structure, dominated by messages sent by the creators but as they grow they can either maintain this pattern or promote communication between different types of contributors. It seems that the projects which take the latter path are more successful in terms of bringing the articles within their scope to high quality. This suggests that in peer production systems – which differ from many other task oriented groups in that they are based on intrinsic motivations and contributor self-selection – a strictly hierarchical communication structure resembling a set of sociometric stars is not preferable. Possibly, editors who work on different tasks, small and large, need to coordinate in order for the product to reach high quality. Moreover, the specifics of the tasks performed by Wikiproject members – complex and diverse – in standard organizations would be considered as requiring reciprocal interdependence between the editors and thus high levels of communication, mutual adjustment and ad hoc communication [31]. Therefore the core needs to connect to the periphery, but the peripheral members also need to communicate with one another.

On the other hand, a strongly assortative communication network, with a "rich" and "poor club" easily distinguishable might also be ineffective for such groups, even though, mathematically, the linear model might suggests otherwise. We cannot draw any definite conclusions about the impact of high assortativity on quality, as this would mean extrapolating beyond the observed data, but if high assortativity were beneficial, we might have encountered a higher number of projects with positive values of this measure in our sample. This is not the case, suggesting that some degree of hierarchy or a neutral connection structure is optimal.

Finally, what might be surprising is that the control variable, number of project members as declared on the projects' member lists, did not impact the quality of products. If the idea of wisdom of the crowd would work, the more contributors, the better should be the

results [29]. Moreover, the more participants, the higher the chances of encountering unique competences and of increasing diversity, which should also promote higher quality [24]. However, even in online peer production the coordination costs grow with the increase in number of contributors [16] and thus quality in such a case might be dependent on the availability of resources or skills specifically related to management and coordination.

We also have to take into account that one of the coordinative tasks that Wikiprojects undertake is the creation of new articles either in the drive to improve the general Wikipedia coverage of their respective areas or in response to requests for specific articles – made by other editors from their own or different projects or even from unregistered users who notice a lack of certain topics. In such a response, project members create article Stubs or Start level articles as placeholders for more content. Creating a basic version is of course much easier than getting the article to GA or FA status – this is a low granularity task that requires little effort but is still a valuable contribution from participants who lack resources (e.g. time) or specific knowledge to participate more [6]. Given data on the core / periphery structure of many peer production systems [2, 9, 19, 26] we can assume that this group of contributors will always be more numerous than those contributing to more demanding tasks. Such easier tasks are also less specific: they require general skills rather than precise knowledge and therefore more users would be likely to self-select to perform them. Thus, increasing the number of contributors would chiefly accelerate the creation of new, Stub level articles relative to the improvement of article status to FA or GA level. The latter process might speed up as well, simply not at the same pace and in effect the relative number of high quality articles does not scale with the number of participants.

To summarize, the results of the present study show that an important factor contributing to product quality in peer production systems is an egalitarian structure of communication network, wherein a collaborator's relations to others are not determined by their or their interlocutors centrality in the network. Such a relation structure may promote higher engagement and more contributions even in typical task oriented groups. In practice, this might suggest that platform design and community policies should allow for sufficient private communication to ensure that users can motivate and encourage each other to contributing to the common goal.

5. Acknowledgments

6. References

[1] Arazy, O., O. Nov, R. Patterson, and L. Yeo, "Information Quality in Wikipedia: The Effects of Group Composition and Task Conflict", *Journal of Management Information Systems 27*(4), 2011, pp. 71–98.

[2] Arazy, O., F. Ortega, O. Nov, L. Yeo, and A. Balila, "Functional Roles and Career Paths in Wikipedia", *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM (2015), 1092–1105.

[3] Barabási, A.-L., and R. Albert, "Emergence of Scaling in Random Networks", *Science 286*(5439), 1999, pp. 509–512.

[4] Baraniak, K., M. Sydow, J. Szejda, and D. Czerniawska, "Studying the Role of Diversity in Open Collaboration Network: Experiments on Wikipedia", *Proceedings of the 12th International Conference and School on Advances in Network Science - Volume 9564*, Springer-Verlag New York, Inc. (2016), 97–110.

[5] Beal, D.J., R.R. Cohen, M.J. Burke, and C.L. McLendon, "Cohesion and performance in groups: a meta-analytic clarification of construct relations", *The Journal of Applied Psychology 88*(6), 2003, pp. 989–1004.

[6] Benkler, Y., "Coase's Penguin, or, Linux and 'The Nature of the Firm'", *The Yale Law Journal 112*(3), 2002, pp. 369–446.

[7] Benkler, Y., "Peer production, the commons, and the future of the firm", *Strategic Organization*, 2016, pp. 1476127016652606.

[8] Coase, R.H., "The Nature of the Firm", *Economica 4*(16), 1937, pp. 386–405.

[9] Crowston, K., K. Wei, Q. Li, and J. Howison, "Core and Periphery in Free/Libre and Open Source Software Team Communications", *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, (2006), 118a–118a.

[10] Csardi, G., and T. Nepusz, "The igraph software package for complex network research", *InterJournal, Complex Systems 1695*(5), 2006, pp. 1–9.

[11] Evans, C.R., and K.L. Dion, "Group Cohesion and Performance: A Meta-Analysis", *Small Group Research 22*(2), 1991, pp. 175–186.

[12] Ferschke, O., D. Yang, and C.P. Rosé, "A Lightly Supervised Approach to Role Identification in Wikipedia

Talk Page Discussions", *Ninth International AAAI Conference on Web and Social Media*, (2015).

[13] Fu, F., X. Chen, L. Liu, and L. Wang, "Social dilemmas in an online social network: The structure and evolution of cooperation", *Physics Letters A 371*(1), 2007, pp. 58–64.

[14] Geister, S., U. Konradt, and G. Hertel, "Effects of Process Feedback on Motivation, Satisfaction, and Performance in Virtual Teams", *Small Group Research 37*(5), 2006, pp. 459–489.

[15] Holme, P., C.R. Edling, and F. Liljeros, "Structure and time evolution of an Internet dating community", *Social Networks 26*(2), 2004, pp. 155–174.

[16] Kittur, A., and R.E. Kraut, "Harnessing the Wisdom of Crowds in Wikipedia: Quality Through Coordination", *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, ACM (2008), 37–46.

[17] Kittur, A., and R.E. Kraut, "Beyond Wikipedia: Coordination and Conflict in Online Production Groups", *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, ACM (2010), 215–224.

[18] Kraut, R.E., P. Resnick, S. Kiesler, et al., *Building Successful Online Communities: Evidence-Based Social Design*, MIT Press, 2012.

[19] Mockus, A., R.T. Fielding, and J.D. Herbsleb, "Two Case Studies of Open Source Software Development: Apache and Mozilla", *ACM Trans. Softw. Eng. Methodol. 11*(3), 2002, pp. 309–346.

[20] Newman, M., *Networks: An Introduction*, Oxford University Press, 2010.

[21] Newman, M.E.J., "Assortative Mixing in Networks", *Physical Review Letters 89*(20), 2002, pp. 208701.

[22] Newman, M.E.J., and J. Park, "Why social networks are different from other types of networks", *Physical Review E 68*(3), 2003, pp. 036122.

[23] Oreg, S., and O. Nov, "Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values", *Computers in Human Behavior 24*(5), 2008, pp. 2055–2073.

[24] Page, S.E., *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton University Press, 2008.

[25] Raymond, E., "The cathedral and the bazaar", *Knowledge, Technology & Policy 12*(3), 1999, pp. 23–49.

[26] Rodríguez-Bustos, C., and J. Aponte, "How Distributed Version Control Systems impact open source software projects", *2012 9th IEEE Working Conference on Mining Software Repositories (MSR)*, (2012), 36–39.

[27] Schroer, J., and G. Hertel, "Voluntary Engagement in an Open Web-Based Encyclopedia: Wikipedians and Why They Do It", *Media Psychology 12*(1), 2009, pp. 96–120.

[28] Shirky, C., *Here comes everybody: The power of organizing without organizations*, Penguin, 2008.

[29] Surowiecki, J., *The Wisdom of Crowds*, Knopf Doubleday Publishing Group, 2005.

[30] Sydow, M., K. Baraniak, and P. Teisseyre, "Diversity of editors and teams versus quality of cooperative work: experiments on wikipedia", *Journal of Intelligent Information Systems 3*(48), 2016, pp. 601–632.

[31] Thompson, J.D., *Organizations in action.*, McGraw-Hill, New York, 1967.

[32] Wang, F., Y. Moreno, and Y. Sun, "Structure of peer-to-peer social networks", *Physical Review E 73*(3), 2006, pp. 036123.

[33] Zhu, H., R. Kraut, and A. Kittur, "Effectiveness of shared leadership in online communities", ACM Press (2012), 407.

[34] "Wikipedia:Good article criteria", *Wikipedia*, 2019. https://en.wikipedia.org/w/index.php?title=Wikipedia:Good_article_criteria&oldid=894072407

[35] "List of most popular websites", *Wikipedia*, 2019. https://en.wikipedia.org/w/index.php?title=List_of_most_popular_websites&oldid=894517978

[36] "May 2019 Web Server Survey | Netcraft", 2019. https://news.netcraft.com/archives/2019/05/10/may-2019-web-server-survey.html

[37] "Wikipedia:Featured article criteria", *Wikipedia*, 2019. https://en.wikipedia.org/w/index.php?title=Wikipedia:Featured_article_criteria&oldid=898591797

[38] "Wikipedia:What Wikipedia is not", *Wikipedia*, 2019. https://en.wikipedia.org/w/index.php?title=Wikipedia:What_Wikipedia_is_not&oldid=901366758

[39] "Heartbleed Bug", http://heartbleed.com/