

Visualization of the evolution of collaboration and communication networks in wikis

Youssef El Faqir

Facultad de Informática
Universidad Complutense de Madrid
Madrid, Spain
yelfaqir@ucm.es

Javier Arroyo

Facultad de Informática
Universidad Complutense de Madrid
Madrid, Spain
javier.arroyo@fdi.ucm.es

Abel Serrano

Facultad de Informática
Universidad Complutense de Madrid
Madrid, Spain
abeserra@ucm.es

ABSTRACT

Commons-based peer production communities can be analyzed with the help of social network analysis. However, since they are fluid organizations that change over time, the time dimension needs to be taken into account.

In this work we present a web application, *WikiChron networks*, to facilitate the study of the evolution of wiki communities over time. The tool displays three different community networks depending on the pages considered for the interactions: articles, talk pages of articles or talk pages of users. The consideration of these three networks offer complementary views of the same community, while the time dimension makes possible to observe how the network structures changes over time and the changes in the network role experimented by some editors. We illustrate the usefulness of our tool analyzing the evolution of a wiki community in different moments and showing network structures that can be seen in other wiki communities.

WikiChron networks is open source and is publicly available. We hope that it will stimulate research on the evolution of collaboration and communication in wiki communities.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing design and evaluation methods; Wikis*; • **Information systems** → *Wikis*.

KEYWORDS

Wiki, Open Collaboration, Network Analysis, Visualization, Web application

ACM Reference Format:

Youssef El Faqir, Javier Arroyo, and Abel Serrano. 2019. Visualization of the evolution of collaboration and communication networks in wikis. In *The 15th International Symposium on Open Collaboration (OpenSym '19)*, August 20–22, 2019, Skövde, Sweden. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3306446.3340834>

1 INTRODUCTION

Wikis are extraordinary webtools to support collaborative content creation in the Internet. They were invented by Ward Cunningham in 1995 to reduce the barriers to write web pages and to enable collaboration among contributors. The idea succeeded and today thousands of wikis populate the web.

Wiki communities have a common need: content has to be created and maintained by a sufficient number of contributors. However, wikis allow great differences in how production is organized [9]. This depends on factors such as the community goals, the norms governing behavior, the policies agreed upon, the nature of the work, or the characteristics of individual contributors.

In addition, the community of a wiki is highly dynamic: the number of contributors evolves, casual contributors come and go, long-term contributors may leave the community, etc. Furthermore, the way wikis are organized usually changes over time, as change the roles adopted by wiki contributors [1]. For example, wiki communities typically start as small group with informal coordination, but can change towards centralized structures [17], or form subgroups [10]. Regarding governance, most wiki communities tend to adopt oligarchic structures, but some others may be more democratic [17]. Regarding participation, most of the work is usually done by a small group of editors, but the distribution of work can change over time [8, 14], even towards more participation of the rest of the community [15].

From a general perspective, Faraj et al. propose to consider online communities as fluid objects where boundaries, norms, participants, artifacts, interactions, and foci continually change over time [5]. They argue that the tensions and the responses that fluidity creates in the community make knowledge collaboration in online community fundamentally different from that in traditional organization structures.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

OpenSym '19, August 20–22, 2019, Skövde, Sweden

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6319-8/19/08...\$15.00

<https://doi.org/10.1145/3306446.3340834>

As a result, online communities are social structures that need to be analyzed considering the temporal dimension.

Social network analysis helps discover the contours of structure in communities that have no formal organization and the detection of different behaviors based on that structure. Social networks have been extensively used to analyze the evolution of the structure of online communities of friends, scientists, professionals, or FLOSS developers, e.g. [2, 11, 12], and also to analyze wiki communities [6, 7, 19].

As it happens in other social media, in wiki communities the network position of an individual can be related to the its role in the community. As an example, Welser et al. identify structural signatures both in the edit distribution across wiki namespaces and network features in the user talk page network, and relate them to different role behaviors [18].

Regarding the changes in the collaboration network of a wiki, Zhang and Wang study the case of Chinese Wikipedia before and after it was banned by the Chinese government and determines a causal relationship between the editor network position and her contribution behavior [20].

However, we believe that the evolution of collaboration networks in wikis have not been properly studied. To stimulate this kind of analysis, we present a web application to visualize how the structure of a wiki community evolves and to track the network role of editors over time. Since wikis provide several context for editors to interact, the tool provides three different networks: co-edition of articles, co-edition of article discussions and editor-to-editor communication through user talk pages.

The tool presented herein extends WikiChron [16] to incorporate social networks analysis. WikiChron¹ aims to visualize the evolution of wiki activity. Formerly, this was done by means of time series of metrics, but now it is also possible to visualize the network of wiki community and its evolution over time.

This new feature serves to analyze how community dynamics evolve over the lifetime of the wiki project. Similarly, we can analyze how a community is affected by different events, for example: external events that cause activity peaks, the appearance or disappearance of highly-active editors, etc. Furthermore, it is also possible to observe how the nature of the wiki project is related with the collaboration structures. For example, a wiki of a dictionary or a wiki of questions and answers tend to show different collaboration patterns than those where the content is more verbose.

The rest of the article proceeds as follows. Section 2 introduces the tool developed and Section 3 describes how community interaction is represented by means of three different networks and how to interpret network metrics. In Section 4 a wiki community is analyzed over time with the

help of our tool and other wiki communities are used to show interesting community interactions. Finally, Section 5 concludes the article.

2 WIKICHRON NETWORKS: A TOOL TO VISUALIZE THE EVOLUTION OF WIKI COMMUNITIES

The present work is an extension of our web application for wiki visualization: WikiChron [16]. WikiChron has been developed to help researchers and wiki administrators to analyze the evolution of wikis

The first version of the tool represented wiki evolution as monthly time series of a rich set of metrics and enabled the user to compare the evolution of a set of wikis. We used that version for exploring how the distribution of participation of a wiki may significantly change over time [15].

Recently, we released a new major version— WikiChron 2.0— that makes the tool extensible with new modes. For example, the time-series metrics former mode, renamed now as *WikiChron compare* mode, and the new mode presented here, *WikiChron networks*. Each mode offers a complementary view of the evolution of a wiki, but they all use time as the main dimension of analysis.

WikiChron networks enables the study of the evolution of a wiki community with the help of network analysis, as it will be described in Section 3. While the network of a wiki community is computed for a given time window, the user can move the time window to observe how the network evolves. The time dimension turns out to be highly relevant for appreciating the changes in the community dynamics, as we will see later in Section 4.

Being part of WikiChron, *WikiChron networks* is publicly available and comes with a set of preloaded wikis ready to be analyzed. However, any user can upload a new one.²

WikiChron networks, as part of WikiChron, incorporates features to facilitate wiki analysis. For example, it incorporates a feature for downloading the data of the visualization shown. More precisely, in the case of *WikiChron networks* you can download a GML file of the network you have selected, and then easily import that data with another network analysis tool, such as Gephi, Cytoscape or any other tool that can read such format. Similarly, an analysis displayed in WikiChron can also be shared with others or bookmarked by means of the URL address.

3 WIKI COMMUNITIES AS NETWORKS

Wikis can be analyzed with the help of network analysis in many different ways depending on the object of the analysis

¹Available online at <http://wikichron.science/>

²The process of uploading a new wiki is documented in the wiki of the project: <https://github.com/Grasia/WikiChron/wiki/How-to-add-a-new-wiki>

[19]. For example, a network can represent the hyperlinks between articles [21]. However, in the present work we are interested in wikis as communities and will consider editors as the focus of our analysis.

In *WikiChron networks*, editors are nodes and we will infer an edge between editor *A* and editor *B*, if they interact during the time window *W* considered. The time window *W* determines the period when interaction among editors occurs. We will consider a minimum value of one month and a maximum value that can potentially be the lifespan of the wiki. The longer the time window, the longer the potential gap between the considered interactions. A short time window ensures that the interaction took place close in time. *WikiChron networks* allows the user to adjust the size of the time window depending on the object of the analysis and the editing frequency of the period considered.

In addition, the time window can be moved along the temporal dimension to observe changes in the interaction network shown.

Our networks are intended to represent the community interactions. Thus, we discarded bot editions since they do not correspond with human activity. Anonymous editors are neither taken into account since they are identified by an IP address, which is an ambiguous identifier. Neither anonymous nor bot editors can be regarded as community members of a wiki and they are consequently excluded from network representations.

Wiki content is categorized in different *namespaces*. Each of them can be considered as a different space for editor interaction and used to construct a network. Similar approaches have been already proposed for studying collaboration [19] and communication networks [13].

In *WikiChron networks*, three different networks can be displayed considering the following interaction spaces:

- Collaboration network: A collaboration network among editors can be inferred from interaction in content pages, commonly known as articles, which are located in the main namespace of the wiki.
- Article talk network: A communication network among editors can be constructed considering the interaction in the discussion page of each article, that is, the *page:talk* page.
- User talk pages network: A communication network for direct communication through personal talk pages of editors, that is, the *user:talk* pages that are intended for receiving direct messages from other editors and answering them.

The first two networks are undirected graphs, where an edge between editor *A* and editor *B* is inferred if they edit in the same page during the time window *W*. The edge weight correspond to the number of pages co-edited by *A* and *B*

during *W*. The article talk pages has been used to propose a directed network of one-to-one communications in other works [6, 19]. However, we consider that discussions in talk pages are better represented as a product of group conversation instead of direct communication.

On the other hand, the User talk page networks that we propose are directed graphs. We will infer a directed edge from editor *A* to editor *B*, if *A* edits in *B*'s talk page during time window *W*. The weight of that edge corresponds to the number of editions made by *A* in *B*'s page during time window *W*. In this case the network reflects the flow of direct communication among users.

The three networks shown in *WikiChron networks* illustrate key aspects of a wiki community, as success for a wiki community is not only achieved by crowds writing content, but also by coordination between peers [9].

Network configuration

WikiChron makes possible to color nodes using a color gradient to emphasize different activity indicators (edits in article talk, edited articles, etc) or network metrics (node degree, betweenness, etc). In this sense, network coloring is intended to provide complementary views to the analyst of the editor activity and their role in the network.

Similarly the size of the node of each network can be changed to be proportional to activity indicators or network metrics. However, the size of the node has a default configuration for each network, namely:

- In the collaboration network, the default size of a node *A* is proportional to the sum of all edits made by editor *A* in article pages during time window *W*.
- In the article talk pages network, the default size of a node *A* is proportional to the sum of all edits made by editor *A* in article talk pages during time window *W*.
- In the user talk pages network, the default size of a node *A* is proportional to the number of edits by editor *A* in her own talk page during time window *W*.

Generally speaking, the default node size is related with the edit intensity of the editor in the time window selected.

Metric interpretation

Network metrics can help to understand the wiki community represented [20]. We provide both node metrics and global network metrics of the network shown.

Node metrics characterize the role of the editor in the community. We considered the following:

- Degree: It measures the number of collaborators (links) of an editor in the period considered.
- Closeness centrality: It averages of the shortest distance from an editor to all other editors considering the weight of the ties. We provide a normalized value

between 0 and 1, where higher values represent editors that collaborate with more people and that have a more central role in the network.

- Betweenness centrality: It is related with the number of times an editor lies on the shortest path between other editors. A higher value represents more prominent the *bridge* role the editor plays in the networks, meaning that she connects people that works in different parts of the wiki.

The value of node metric is better contextualized if compared with those from the rest of nodes. As a result, the tool provides a ranking of the nodes according to these network metrics. Furthermore, node metrics are summarized by global network metrics which serve to characterize the behavior of the wiki community. More precisely, *WikiChron networks* offer:

- Density. It measures network cohesion as the number of observed edges in the network divided by the number of potential edges, hence it ranges from 0 to 1. Higher density levels imply a higher degree of collaboration among editors.
- Gini of degree. Gini coefficients are inequality measures that rank from 0 to 1. A value close to 0 means that all editors have roughly the same number of co-editors. It could happen if all editors collaborate in the same articles, or if groups of the same number of editors collaborate in non-overlapping articles. On the other hand, a value close to 1 means that some editors collaborate with a lot of counterparts (they would be hubs, according to the network analysis jargon) while the majority collaborate with few or none editors.
- Gini of betweenness. It is used to measure the inequality of the *bridge* role among the community. Hence, a value close to 0 represents that all editors play a similar role, while a value close to 1 means that some editors have a prominent role.
- Assortativity degree [3]. It ranges from -1 to 1 and represents how editors associate among themselves. A value close to 1 —corresponds to an *assortative network*— means that hubs (i.e. editors with many co-editors) tend to be connected among them, and non-collaborative editors are also linked among them. On the other hand, a value close to -1 means that hubs avoid to each other and they link to small degree editors. Finally, a value close to 0 —corresponds to a *neutral network*— means that there is no clear pattern of connection among editors according to their degree.

Besides network metrics, the tool also provides activity metrics for the time window W , such as number of edited articles, or number of edits by the whole community and for each editor. This information complements the view of the

network metrics and helps to better understand what happened in the community during the period being analyzed.

4 ANALYSIS OF COMMUNITY NETWORKS WITH WIKICHRON

This section shows how to use WikiChron to analyze a wiki community. The first subsection analyzes the English community of Hitchwiki, a hitchhiking wiki, in different periods of its almost 15 years of life. This will serve to illustrate how community behavior vary considerably over time and how the role of individual editors also changes over time. The evolution can be easily appreciated with the help of WikiChron.

The second subsection shows two very different collaboration structures in the collaboration network that were found in different communities at moments of intense participation. These cases exemplify that collaboration is a rich phenomenon and that the network representation and the tool proposed serve well to analyze it.

To reproduce the analysis shown here in *WikiChron networks* is really easy. It is only required to select the wiki, the network and the time window. This is done in the home screen in three steps. Once a network is generated for a given time window, the time window can be moved a given number of months and the network shown can be changed without going back to the home screen.

All the plots shown and metrics included in this section has been calculated with WikiChron. Furthermore, as *WikiChron networks* enables sharing an analysis using an URL that registers the selected wiki, network and time window, we use this feature across this section to include hyperlinks to each of the networks analyzed for the perusal of the interested readers.

The evolution of Hitchwiki

Hitchwiki is a wiki for hitchhiking in English. The project has over 4600 articles, a community of over 5,100 registered editors and it is a useful resource for hitchhikers all over the world. There are smaller Hitchwiki projects in other languages including German, French and Spanish.

The project has gone through different periods in its almost 15 years of life. Below we will analyze five moments with the help of *WikiChron networks*. Each moment consists of a time window of one month.

The moments considered serve as an example of the differences that can be observed in the community dynamics. Table 1 shows a summary of the edit activity and the network metrics of the collaboration network in the five moments analyzed. The first moment — October 2007— corresponds to a first activity burst with intense edit activity, few editors, but intense collaboration. The second moment — August 2008— shows more editors involved but similar productivity

	Oct-07	Aug-08	Aug-10	Aug-12	Aug-16
Article edits	554	508	326	442	253
Edited articles	281	173	156	250	97
Editors	25	34	42	78	66
Links	57	116	40	57	17
Density	0.19	0.21	0.05	0.02	0.01
Assortativity	-0.67	-0.27	-0.26	-0.28	-0.03
Gini between.	0.83	0.80	0.87	0.95	0.96
Gini degree	0.54	0.54	0.66	0.63	0.79

Table 1: Activity and network metrics of the collaboration network in Hitchwiki for the different moments analyzed

in terms of article edits and in terms of intensity of collaboration, according to the density measure. The three last moments—August 2010, 2012 and 2018—still show a healthy activity level by a higher number of editors, but exhibit less collaboration, according to the density level. The increase of both Gini coefficients reveal that the community is less egalitarian in terms of role played by the editors in the network. Regarding the assortativity coefficient, the negative value in October 2007 means that in that month editors with a high number of collaborators worked mainly with editors with few collaborators, in other words, core editors associated with casual ones. However, the neutral assortativity value in August 2016 reveals that these associations do not take place anymore and that there is no clear association pattern in the collaboration network.

These conclusions can be drawn by the global metrics provided by *WikiChron networks*. However, a more detailed analysis of each considered moment has been carried out and is shown below.

October 2007: First activity burst. The wiki starts in December 2004, but until November 2006 it only shows activity of two registered editors that worked in more than 300 articles. The first moment that we will analyze is an activity burst in October 2007 that involves 25 editors.

Figure 1 shows [the collaboration network](#) as displayed by our tool. The size of the nodes shown is the default value for the collaboration network, i.e. the number of edits in articles, while the color intensity of a node was associated with the tenure in the wiki of the corresponding editor. For the sake of clarity, the metric associated to color remains constant in all the figures shown in this article, and the one associated to node size is the default one, as detailed in Section 3.

In the collaboration network in Figure 1 we can see that most active editors are strongly linked among themselves, which means that co-edited together many pages. According to their color, they are mostly veteran editors with the exception of Patschi and Worldhitch.

Regarding the community metrics in Table 1, the density of the network in that month was 0.19, which is a considerable value for this kind of networks given the number of editors. The density of the network reveals that collaboration (co-edition) was intense. In addition, the Gini coefficient of the node degree is 0.54, which means that the distribution of degrees is unequal, but not in a strong way.

The assortativity value is -0.67 which means that editors with high-degree tend to link to editors with low-degree (isolated editors). While the network does not reveal the order of the edits, it could be possible that high-degree editors, who are also those that edited more and that are more tenured, tend to work in pages that were edited first by occasional editors.

On the other hand, the Gini of the Betweenness is high (0.83) revealing that there are a few editors with an extremely central role in the network. In a collaboration network it means that a few editors have the role of bridging the rest of editors. In particular, *Platschi* is the editor that ranked higher in terms of degree, betweenness and in productivity statistics such as number of edited articles and edits, followed by *Speckmade* and *MrTweek*.

Regarding the communication networks, the activity in the [article talk pages](#)—not shown here—only involved 7 editors and 9 talk pages, revealing that the coordination mechanism is not extensively used. Interestingly, the editors with higher degrees and more edits are again *MrTweek* and *Platschi* in that order, but *Speckmade* did not participate.

[The user talk page network](#) in Figure 2 mainly exhibit several star shapes from tenured editors (darker colored) to new ones (lightest color) that after an inspection on the pages reveal a welcoming mechanism to new editors. The ones in charge of writing the welcome message were again *MrTweek* and *Platschi*, and also *Guaka*, one of the longest standing editors. In this case the size of the node is proportional to the number of edits in their user page.

August 2008: Intense participation of the newbies. Now we will analyze the community activity in August 2008 that exhibits intense activity and collaboration of several new editors.

Figure 3 shows [the collaboration network](#) that involved 34 editors. Again the size of the nodes is related with the number of edits and the color with tenure in the community. Surprisingly, several new editors (*Terrehappy*, *Sitarane*, *Uptojoe*) had intense edit activity, similar to that from long-standing editors. Some of them have even strong links (i.e. have co-edited several pages) with more veteran editors. In this sense, their behavior does not reflect the casual collaboration typical of new members, but a strong involvement in the community.

According to the betweenness value, *MrTweek* has the most central position co-editing with both “satellite” new

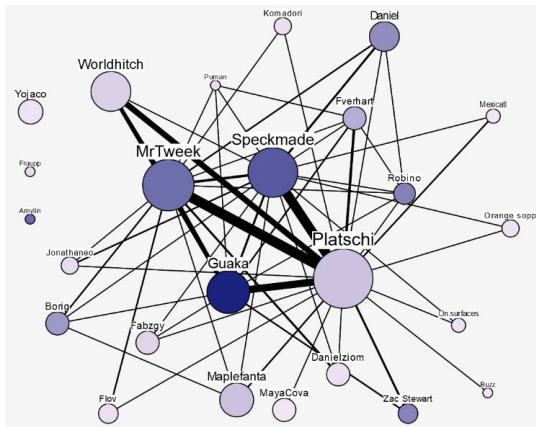


Figure 1: Hitchwiki collaboration network in October 2007 shows stronger collaboration between the most active editors.

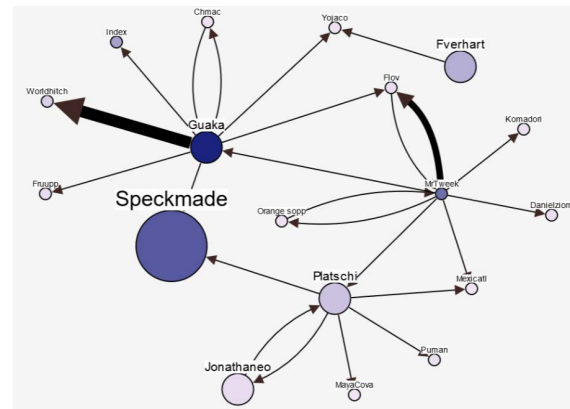


Figure 2: Hitchwiki user talk pages network in October 2007 reveals star-shaped communication from relevant editors to new ones.

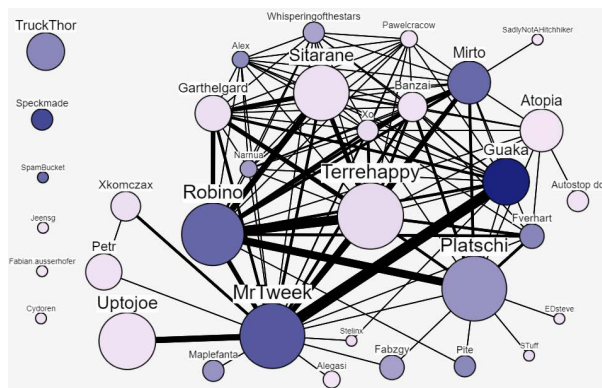


Figure 3: Hitchwiki collaboration network in August 2008 exhibits dense collaboration and a prominent role of some newbie editors.

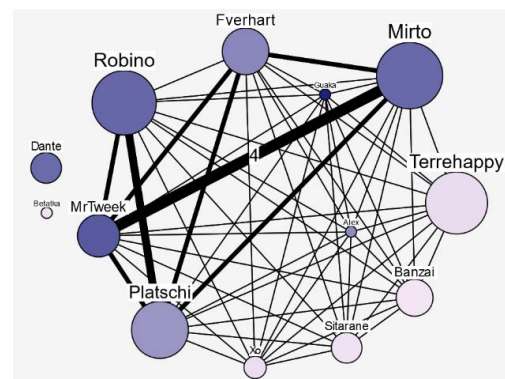


Figure 4: Hitchwiki article talk network in August 2008 exhibits dense coordination of a group of editors, including some newbies.

editors and with the rest of the network. He also has a very strong link with the most tenured node (*Guaka*), which reinforces the idea of his importance in the community.

The network now is slightly more dense (0.21) despite having 6 non-connected editors. There is a densely connected area in the upper part of the network with most edges having a weight of 1. This means that this intense collaboration takes place only in just one page that involves a lot of editors.

The network is less assortative, which means that the linking pattern between low-degree editors and high-degree editors is not as clear as in the previous collaboration network analyzed.

Figure 4 shows the collaboration network in the article talk pages. In this network, the size of the nodes is related with the number of edits in article talk pages, while the color

is related with editor tenure. According to the color of the nodes, some new editors participated in article discussion with more veteran editors. However, veteran editors have stronger links among themselves, which means that they were involved in discussions in more articles of the wiki. In fact, while the total number of talk pages edited are 16 most links are of weight 1, and the dense connections reveal that a single page concentrated most of the talk involving veteran and new editors, as happened in the collaboration network.

As it can be expected, veteran editors have stronger links among themselves. *MrTweek* has strong connections with all of them, which reinforces the impression that he played an important role in this period due to his presence.

August 2010, 2012 and 2016: The way towards stigmergy. In August 2010 there are 42 active editors and 326 edits. According

to the indicators, the activity is greater than in the previous moment analyzed. However, [the collaboration network](#) (Fig. 5) is substantially less dense: 40% of the editors did not collaborate with anyone, and density is 0.05. Interestingly, editors that did not collaborate include not only newbies or occasional collaborators, but also well seasoned active editors, as we can see by the color intensity and the node size.

The main connected component in the collaboration network is not dense. There are no large groups of nodes connected among themselves, which means that most articles were edited by around 2 or 3 editors. However, we can find some strong links between editors which means that some editors worked together in different articles. Again, we can see strong collaboration between editors of different tenure. In this occasion, the editor with higher betweenness is *Zenit*. *Zenit* also has strong links with other very active editors, such as *Guaka* and *CRCulver*, and consequently we can infer that he had a central role during that period.

In [the article talk network](#) in Figure 6 we see the central role of *Zenit* again, who is the most active editor as well with 12 edits in article talk pages. The network has few editors (12) and very few links, which means that the intense edit activity of the wiki did not required much coordination.

[The user talk page network](#) (not shown for the sake of brevity) is sparse and involve only 14 nodes and 10 edges, including several welcome messages to new users.

The high edit activity together with the few interactions in the talk networks and the low density of the network point towards the fact that the wiki mostly is working without explicit collaboration. There are some explicit collaboration such as some aspects of “netiquette” are explained in some articles of the wiki, and the discussion that takes place in some articles. In the case of veteran editors, they probably can work with implicit coordination, because each of them adjust their behavior observing the behavior of the others and few explicit coordination. However, given the number of editors involved it seems that another important part of the collaboration is stigmergic. It would mean that the wiki articles and the wiki itself is the communication mechanism in most cases, without the need to recourse to other coordination mechanisms. Editors try to adjust their contributions to produce a result which matches that of other articles.

In Figure 7 we show [the collaboration network during August 2012](#). The activity of the wiki is remarkable, with a collaboration network of 78 editors and 452 edits. *WikiChron compare* reveals that the community welcomes this month 49 new registered editors — the highest to date— and that anonymous editors were not allowed in the wiki since May 2011. Nevertheless, stigmergy seems stronger.

The network is very sparse with a density of 0.02 and 27 editors not connected (not shown in the figure). Veteran

editors (dark colored nodes) have an important role in terms of edit activity (node size), but the most relevant editor is relatively new — *GeorgDerReisende*— who joined in [June 2011](#). In August 2012 he carries out 120 edits in 110 articles. Despite the notable activity, he is just linked with 16 editors and the star shape around him reveals that most of his collaborations involved a few editors. Star shapes can also be seen around veteran editors (*Guaka*, *CRCulver* and *Prino*). Furthermore, we also can see several pairs of editors not connected with the main connected component. Despite the activity level, collaboration is sparse.

In [the article talk network](#) we can see only 9 nodes involved and just one link. [The user talk page network](#) is slightly more dense — with 15 nodes and just 15 links — and it mostly involves veteran editors including some of the most tenured, being *Prino* the most active with an out-degree of 9, but an in-degree of 1.

Given the little communication in the talk pages and the sparsity of the collaboration network, we believe that the wiki is probably mature and makes possible stigmergic coordination. As a result, the community seems to work efficiently without much explicit coordination or collaboration.

According to [the time series in WikiChron compare](#), from 2013 onwards, the wiki remained active and with a constant input of new editors. In that period, activity was not as high as in the 2010-2012 period that had a mean around 400 article edits per month. We can see a decrease on the level of the time series to 250 edits per month in 2013-2015 and to 150 edits since then. Besides some occasional peaks, talk activity is not specially remarkable because it mostly happens in the form of welcome messages from veteran editors to new ones. Interestingly, the participation of editors with at least 100 edits in the wiki is much less intense. Furthermore, the 10:90 ratio decreases slowly, but steadily, which means that the contributions of the 90% of the community are gaining in weight with respect to those from the 10% of most active editors. For the sake of comparison with the previous collaboration networks, Figure 8 shows the main connected components of [the collaboration network](#) in August 2016, when there were 45 unconnected editors as well.

WikiChron networks together with *WikiChron compare* helped us to understand the evolution of Hitchwiki. In the beginning it was a project managed tightly by a small group of people. As it successfully grew, it was more loosely managed by the core editors, and in the late period, it is being managed by the community base with very little intervention of the core editors without jeopardizing its quality.

Collaboration structures with high participation

In the previous subsection we analyzed a community with a significant participation over time. In wiki communities, it is difficult to find dense collaboration networks as the activity

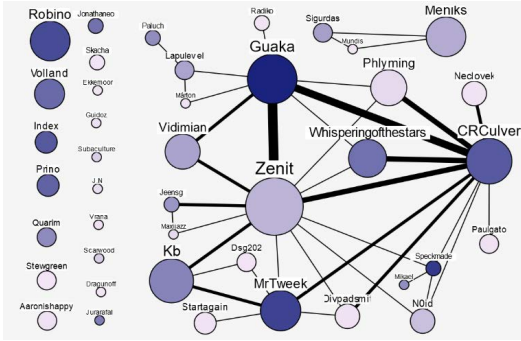


Figure 5: Hitchwiki collaboration network in August 2010 is less dense and Zenit plays a central role in it.

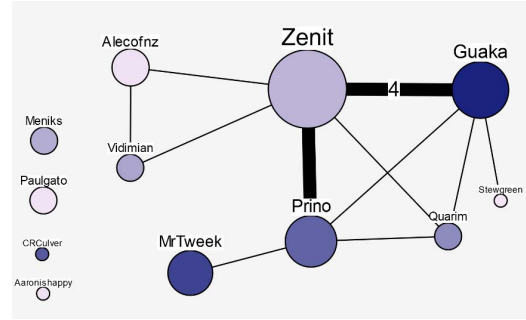


Figure 6: Hitchwiki article talk network in August 2010 shows that coordination was scarce and the important role of Zenit.

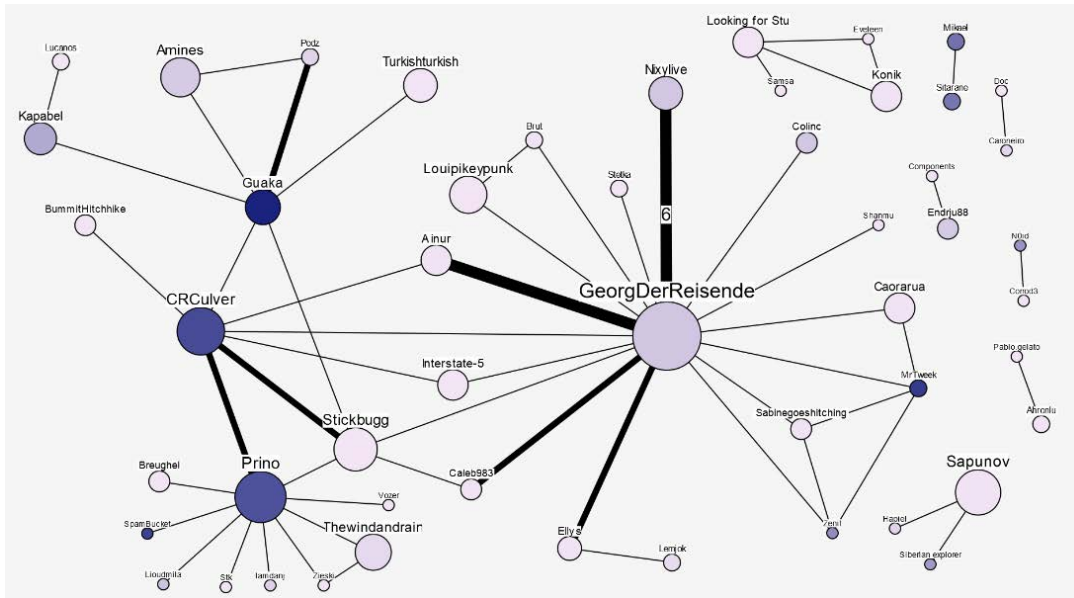


Figure 7: Hitchwiki collaboration network in August 2012 is sparse and reveals strong stigmergy.

of the wiki grows. Sparsity seems a natural consequence of growth. However, the collaboration structures that can be observed still can exhibit interesting variations that can be studied with the help of *WikiChron networks*.

In some cases, a few core editors tend to collaborate with numerous occasional editors that edit each of them one or few articles with little or no overlap with other editors. In these cases, collaboration networks are star-shaped and have a negative assortative degree that reflects that core editors tend to associate with occasional ones (the coefficient is not extremely high because some collaboration among core editors and among occasional editors usually takes place). Typically, we can observe non-connected editors as well. An example of this kind of network can be seen in Figure 9

that shows [the Zeldapedia collaboration network in June 2017](#) with 82 active editors that worked in 136 articles. This wiki has Gini coefficient of betweenness of 0.94 as one user has a very high degree and a remarkable central position in the network. The assortativity degree is -0.53 and density is 0.08. In this case, the color intensity of the nodes is proportional to the edits in article talk pages and we can see that communication is also managed by the active editors.

A very different collaboration structure can be seen when an article is the result of shared participation of a group of editors, as that collaboration is represented as each editor involved in the article is connected with the rest of the group. If this collaboration structure occur in several articles the result is a dense network with many links and where most

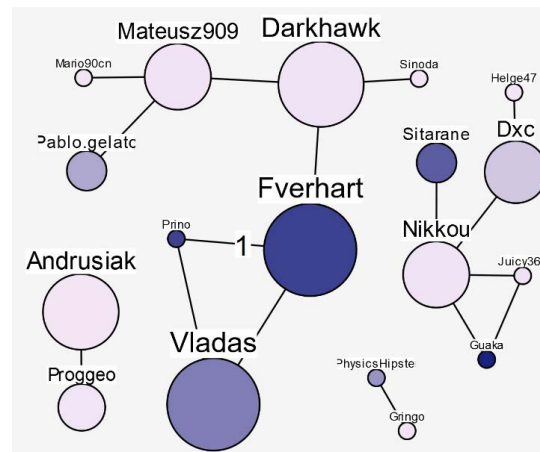


Figure 8: Hitchwiki collaboration network in August 2016 is highly sparse and stigmergic.

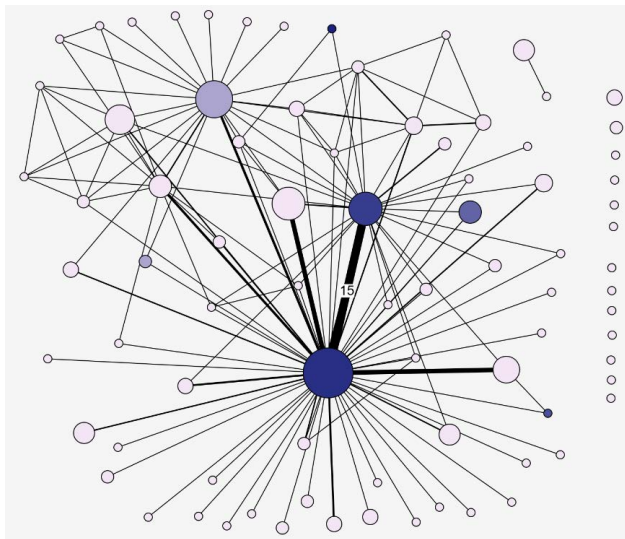


Figure 9: Zelda collaboration network in June 2017 shows star-shaped collaboration structures.

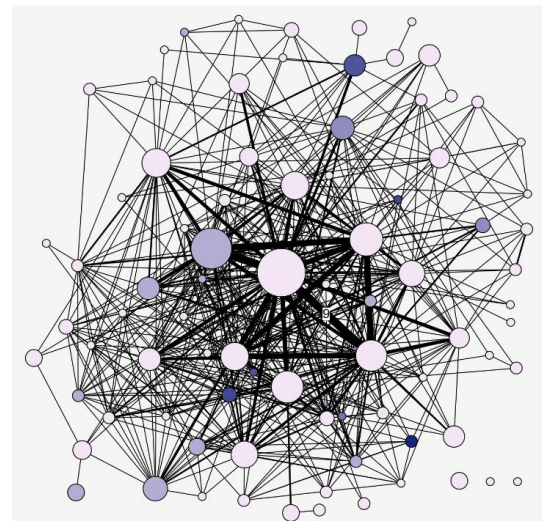


Figure 10: Good Luck Charlie collaboration network in July 2013 exhibits intense collaboration by different groups of editors.

active editors are densely connected among themselves with strong links. This is the case of the network structures of the Good Luck Charlie wiki community, which tend to exhibit this kind of highly collaborative behavior. In Figure 10 we see [the Good Luck Charlie collaboration network in July 2013](#) where there were 83 active editors that worked in 114 articles. These numbers are similar to those from the Zeldapedia example, but in this case network density is much higher (0.15), the Gini coefficient of betweenness is considerably lower (0.78) and the assortative degree is more neutral (-0.24). Again the color intensity of the nodes is proportional to the edits in article talk pages, but in this case we see

that communication is not monopolized by core editors. All these indicators suggest that the community is much more horizontal than the Zeldapedia one.

The Harry Potter wiki constantly showcases remarkably dense collaboration structures given the high number of active editors and its productivity in terms of edited articles. For example in August 2016 there was a peak of over 7,200 edits in 2520 articles made by 420 editors. While there was a small group of extremely active editors, there was also a large group of recurrent editors that intensely collaborate among themselves. [The network metrics of the Harry Potter wiki in August 2016](#) reflect that intense collaboration and shared

participation given the large number of active editors. More precisely, network density is 0.1, the Gini of betweenness is 0.86 and the assortativity degree is -0.32. These numbers are between those of the Zeldapedia and the Good Luck Charlie cases, but given the scale of the community and the wiki, they indicate a very collaborative community.

5 CONCLUSIONS

Commons-based peer production communities have enormously developed in the contemporary networked environments [4]. Network analysis seem to be an adequate tool to study this forms of production that are more complex than traditional hierarchical organization structures. Furthermore, these structured are fluid and hence they should be analyzed taking into account the time dimension. Alas, network analysis has not been extensively applied on wikis and those works that use it often overlook or “flatten” the time dimension.

WikiChron networks addresses this gap and provides a means for visualizing the evolution of the collaboration and communication structures in wiki communities. Our tool can help community members or wiki administrators trying to gain insight on how their communities work. It may also help to formulate research hypotheses on collaboration or communication patterns to be validated with statistical tools in a representative sample of wikis.

WikiChron networks is part of the *WikiChron* system, which aims to facilitate research on wiki communities. *WikiChron* is open-source, well-documented and publicly available for the benefit of the wiki research community.³

ACKNOWLEDGMENTS

This work was partially supported by the Spanish Ministry of Science, under the grant RTI2018-096820-A-I00, and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 759207.

REFERENCES

- [1] O. Arazy, J. Daxenberger, H. Lifshitz-Assaf, O. Nov, and I. Gurevych. 2016. Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Coproduction. *Information Systems Research* 27 (2016), 792–812.
- [2] A.L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. 2002. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications* 311, 3 (2002), 590 – 614.
- [3] A.L. Barabási and M. Pósfai. 2016. *Network Science*. Cambridge University Press.
- [4] Y. Benkler. 2006. *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.
- [5] S. Faraj, S. L. Jarvenpaa, and A. Majchrzak. 2011. Knowledge Collaboration in Online Communities. *Organization Science* 22 (2011), 1224–1239.
- [6] J. Foote and N. Contractor. 2018. The Behavior and Network Position of Peer Production Founders. In *Transforming Digital Worlds*, G. Chowdhury, J. McLeod, V. Gillet, and P. Willett (Eds.). Springer International Publishing, Cham, 99–106.
- [7] B. Keegan, D. Gergle, and N. Contractor. 2012. Do Editors or Articles Drive Collaboration?: Multilevel Statistical Network Analysis of Wikipedia Coauthorship. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 427–436.
- [8] A. Kittur, E. Chi, B.A. Pendleton, B. Suh, and T. Mytkowicz. 2007. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World Wide Web* 1, 2 (2007), 19.
- [9] A. Kittur and R. E. Kraut. 2010. Beyond Wikipedia: Coordination and Conflict in Online Production Groups. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10)*. ACM, New York, NY, USA, 215–224.
- [10] A. Kittur, B. Pendleton, and R. E. Kraut. 2009. Herding the Cats: The Influence of Groups in Coordinating Peer Production. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration (WikiSym '09)*. ACM, New York, NY, USA, Article 7, 9 pages.
- [11] R. Kumar, J. Novak, and Andrew Tomkins. 2010. *Structure and Evolution of Online Social Networks*. Springer New York, New York, NY, 337–357.
- [12] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins. 2008. Microscopic Evolution of Social Networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*. ACM, New York, NY, USA, 462–470.
- [13] P. Massa. 2011. Social Networks of Wikipedia. In *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia (HT '11)*. ACM, New York, NY, USA, 221–230.
- [14] F. Ortega, J. M. Gonzalez-Barahona, and G. Robles. 2008. On the Inequality of Contributions to Wikipedia. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. 304–304.
- [15] A. Serrano, J. Arroyo, and S. Hassan. 2018. Participation Inequality in Wikis: A Temporal Analysis Using WikiChron. In *Proceedings of the 14th International Symposium on Open Collaboration (OpenSym '18)*. ACM, New York, NY, USA, Article 12, 7 pages.
- [16] A. Serrano, J. Arroyo, and S. Hassan. 2018. Webtool for the analysis and visualization of the evolution of wiki online communities. In *Proceedings of the 26th European Conference on Information Systems (ECIS)*.
- [17] A. Shaw and B. M. Hill. 2014. Laboratories of oligarchy? How the iron law extends to peer production. *Journal of Communication* 64, 2 (2014), 215–238.
- [18] H. T. Welser, D. Cosley, G. Kossinets, A. Lin, F. Dokshin, G. Gay, and M. Smith. 2011. Finding Social Roles in Wikipedia. In *Proceedings of the 2011 iConference (iConference '11)*. ACM, New York, NY, USA, 122–129.
- [19] H. T. Welser, P. Underwood, D. Cosley, D. Hansen, and W. Black L. 2011. Wiki Networks: Connections of Creativity and Collaboration. In *Analyzing Social Media Networks with NodeXL*, Derek L. Hansen, Ben Shneiderman, and Marc A. Smith (Eds.). Morgan Kaufmann, Boston, 247 – 271.
- [20] X. Zhang and C. Wang. 2012. Network Positions and Contributions to Online Public Goods: The Case of Chinese Wikipedia. *Journal of Management Information Systems* 29, 2 (2012), 11–40.
- [21] V. Zlatić, M. Božičević, H. Štefančić, and M. Domazet. 2006. Wikipedias: Collaborative web-based encyclopedias as complex networks. *Physical Review E* 74 (Jul 2006), 016115. Issue 1.

³The source code is available at <https://github.com/Grasia/WikiChron>