Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10339

Uncovering collaboration dynamics in design projects using network analysis and log data

Lovro Sever, ✉, **Tomislav Martinec**, **Matthew Mueller and Stanko Škec**

University of Zagreb, Croatia

✉ lovro.sever@fsb.unizg.hr

ABSTRACT: This study explores the integration of network analysis and CAD/PDM log data to analyze collaboration and activity patterns in a multi-year engineering project. Using logs from a collaborative CAD platform with PDM features, the research examines team interactions and network evolution over time. Key findings reveal that early project stages featured smaller, denser networks, while later stages saw larger, less interconnected structures. Subteam formations were dynamic, with variations in size and number. Individual-level analysis showed that user influence, measured through eigenvector centrality, did not always align with activity volume. This work highlights the potential of CAD/PDM data for understanding collaboration dynamics and lays the groundwork for further studies on team interactions in design processes.

KEYWORDS: computer aided design (CAD), network analysis, collaborative design, teamwork, log data

1. Introduction

In today's rapidly evolving engineering landscape, effectively utilizing diverse data types is essential for gaining a comprehensive overview of the design process. Leveraging various forms of data enhances project management by enabling better-informed decisions and more efficient planning, ultimately leading to improved outcomes in engineering design and product development projects (Li et al., 2019). The rise of big data analytics has empowered researchers and organizations to collect and analyze vast amounts of information generated during the design process, such as log data, communication records, and project schedules to better understand workflows, team dynamics, and collaboration networks (see, e.g., Snider et al., 2017; Lan et al., 2018). Integrating these diverse sources allows project managers to better understand ongoing processes and fosters more effective collaboration and innovation.

Building upon this data-centric approach, network analysis has emerged as a powerful method for examining the complex interactions and dependencies existing across the data sources. Network analysis offers a framework for modeling and visualizing elements of the design process (actors, actions and artefacts (Vrolijk & Olechowski, 2024)) as nodes and their interactions as edges, thus facilitating the identification of critical moments, tasks, team members, bottlenecks and iterations, as well as opportunities for process optimization (see, e.g., Lan et al., 2018; Gu et al., 2014; Piccolo et al., 2019). This, in turn, facilitates a deeper understanding of team dynamics, timely identification of project issues, swift interventions, and improved decision-making.

A common problem, however, is that the contents of the actual networks need to be discovered and recorded. Many of the existing approaches rely on communication archives to identify these networks, which are not always available, whereas mapping people across multiple communication systems involves considerable manual effort (Jermakovics et al., 2013). One significant and non-intrusive source of data is the log files generated by collaborative design tools, especially those integrating Computer-Aided Design (CAD) and Product Data Management (PDM) systems. These environments inherently capture detailed records of user actions and interactions, providing continuous data without disrupting the design workflow. An example of such a platform is Onshape Enterise, which allows for a

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

non-intrusive collection of data on the traceable events occurring on specific documents and users, in a specified timeframe. This data is then summarized in the form of audit trail, whereas the specific event items are recorded in the form of audit trails, which include information such as the time of the event, the document (project) and document tab where it occurred (e.g., part studio, assembly, drawing), the user involved, and a general description of the action performed (e.g., user started editing a sketch, users opened a drawing, user dragged a part in the assembly).

Some researchers have already leveraged such CAD and PDM log data to investigate design team behaviors (Leonardo & Olechowski, (2021) and workflow patterns (Gopsill et al., 2016), as well as comparing the approaches of high- and low-performing teams (Celjak et al., 2023; Šklebar et al., 52024). However, the potential of utilizing CAD log data remains relatively underexplored, presenting significant opportunities to further exploit these rich data sources using network analysis techniques. Essentially, network analysis of CAD data could provide insights, not only about workload distribution and type of users characterized by recorded actions, but also about interactions in design team as well as combining these two types of information.

This paper hypothesizes that the combination of network analysis and log data from CAD and PDM systems can potentially uncover valuable insights into the design process, such as identifying key collaborators, critical tasks, document dependencies, periods of high and low activity, and the formation of sub-teams. The main premise is that much like the use of Version Control Systems in software development (Jermakovics et al., 2013), frequent access to and modification of the same data reflects collaborative activity in design projects. To test this hypothesis, the paper analyzes log data from Onshape, a cloud-based CAD platform with PDM features. The study explores how network metrics can reveal collaboration patterns and track the evolution of team interactions over time. By examining both global network structures and individual user interactions the paper demonstrates the potential of combining network analysis with CAD/PDM log data to uncover critical insights into design processes.

The structure of the paper is as follows: Section 2 reviews relevant background on CAD data and network analysis in design research. Section 3 outlines the methodology used for data preparation and network analysis. Section 4 presents and discusses the results of conducted analyses. Finally, Section 5 concludes the paper, highlighting the implications, limitations, and future research directions.

2. Background

Understanding the context and prior work in CAD data analysis and network analysis is essential for positioning this study. This section reviews existing research on CAD-related collaboration in design processes and explores the application of network analysis techniques to uncover patterns and dynamics in organizational and design data.

2.1. CAD data in design research

Existing research on CAD-related collaboration in design and product development has focused on utilizing user action data for analyzing the CAD activity process and performance, team member workload, as well as CAD roles and specific personas in teams. These types of analysis are enabled by non-intrusive data collection methods, enabling analysis of real-world design actions. For example, a study by Deng et al. (2022) proposes a framework for analyzing CAD actions, which categorizes actions into six categories: (1) editing, (2) creating, (3) deleting, (4) revising, (5) viewing and (6) other actions. This framework allows researchers to conduct analyses related to the distribution of each CAD activity category on both individual and team level. Šklebar et al. (2024) employed this framework to analyze four student teams during two days of design sprint. In this study, the team actions are analyzed per day and per percentage contribution of each action category with focus on identifying CAD pattern usage. The study by Leonardo and Olechowski (2021) explored an extensive user action dataset consisting of about 1.400.000 CAD actions. They used hierarchical clustering of collected action descriptions and identified a total of 14 clusters of CAD-related actions, which can be used for high-level analyses of CAD activity by reducing the otherwise vast number of specific CAD actions. For example, they used the 14 categories to investigate user behavior and identified three communities of users, whereas the user/team activity was analyzed in terms of number of actions (recorded events) per month and per user, as well as their contribution for each of the identified activity categories. The main purpose of the framework proposed by Leonardo and Olechowski (2021) is self-assessment and awareness for individual designers

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

by identifying types of individual designers and their contribution to different types of CAD actions. Celjak et al. (2023)'s study focuses on identifying patterns of CAD actions for high performing and low performing teams. This approach is different from the above-mentioned works because it uses Markov chains for finding the most frequent sequences of the categorized CAD actions.

2.2. Network analysis in design research

Many aspects of organizations and their actions can be interpreted as networks, and network analysis can be applied to organizational data. From examining collaborations in academia to exploring design processes, network-based approaches enable researchers to uncover hidden patterns. Various tools exist for organizational network analysis, with one of the specialized being the ORA toolkit (Carley, 2014), which includes analysis and visualization algorithms, community detection, and additional options. Some researchers (e.g., Schaefer & Rosen, 2013) use several metrics to describe network size and structure, such as the number of nodes (e.g. users are represented with nodes) and edges (links between two nodes) and communities (e.g. sub-teams). Metrics commonly used in network analysis are comprehensively described by Arif (2015). These key metrics include clustering coefficient (ability of node neighbors to create complete network), average degree (average value of degree centrality of all nodes), density (ratio of numbers of edges and possible number of edges), eigenvector centrality (measure of influence within a network), betweenness centrality (measure of strategic position of node), and closeness centrality (measure of closeness of node to all other nodes). For community detection (generating groups of nodes with defined rules, e.g. modularity maximization), several algorithms are widely used, such as the Louvain method, the Girvan-Newman method, and the Clauset-Newman-Moore algorithm (Schaefer & Rosen, 2013; Carley, 2014).

According to Chen et al. (2018), network analysis of design-related data generally falls into three subcategories: (1) networks and architecture, (2) networks and design decisions, and (3) networks and design ecosystems. Gyory et al. (2019) utilized data collected from 60 students tasked with generating concepts for a related problem. Participants provided descriptions and drawings for each concept, and semantic similarity was computed for each pair of concepts. For this network analysis, centrality and density metrics were used in order to characterize feature of design space. Piccolo et al. (2018) analyzed data from organizations involved in the design of power plants. To measure the robustness (resilience to change and tolerance to random failures) of the design process, the researchers used bipartite networks. Two types of network measures were employed: (1) global measures such as density, average shortest path, average degree, diameter, and clustering coefficient, and (2) local measures such as closeness centrality, betweenness centrality, and degree centrality. Similarly, Schaefer and Rosen (2013) collected data through participant interviews with the purpose of finding metrics for describing socio-technical networks (degree centrality, clustering coefficients, and closeness centrality). The Clauset-Newman-Moore algorithm was applied for community detection.

3. Methodology

As mentioned in the introduction, this study focuses on network analysis of data collected by log files of the Onshape platform, a cloud-based CAD and PDM solution. The data was sourced from a student team engaged in a multi-year project focused on developing a complex engineering product. The raw dataset (activity log) contains records of more than 32,000,000 actions conducted by a total of 619 users who worked on a total of 63803 documents. After sourcing the log data, it was cleaned and preprocessed (e.g. deleting of duplicate records) to ensure consistency and accuracy, then subjected to a general analysis of user activity, followed by various network analyses using established metrics from the literature.

3.1. Data preparation

The complete (multi-year) dataset was segmented into time periods to facilitate the tracking of metrics over time. Each period was defined as a 30-day interval, with a one-week overlap between consecutive periods to analyze time-distributed trends with reduced noise (moving-window analysis). The period duration and overlap can be adjusted based on specific analytical requirements. For smaller datasets, shorter time periods (e.g., less than a day) and finer overlaps (e.g., hourly or minute-level) can be used.

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

Given that the dataset contains user activity timestamps spanning three years, a 30-day period with a one-week overlap was selected as a practical compromise between computational efficiency and adequately capturing the dynamics of user actions.
Since the number of different types of user actions is very large (over 120 types of actions have been recorded across the timestamps), they were grouped and classified into activity categories based on two frameworks. The first framework was derived from the study by Leonardo and Olechowski (2021), who hierarchically clustered Onshape actions into 14 categories. Some additional actions, which exist in the sourced dataset but were not included in the clustered categories, were manually sorted. To further analyze the distribution of actions, the six-category classification framework proposed by Deng et al. was used and adjusted. It must, however, be noted that while this framework provides a clearer representation of the type of work performed on the documents (creating, editing, deleting, etc.), it accounts only for a smaller proportion of possible actions (different type of input data), and does not include actions such as 'Delete Team Event' or 'Update Task Metadata Event'. These two frameworks provide insight into dominant types of CAD actions conducted at different stages of analyzed different stages of the project.
Additional data preparation was required concerning the identification of relationships needed to perform network analysis. Since the log data captures interaction between users and documents, the relationships between the users themselves were reconstructed based on their work on the same documents. More precisely, if two or more users worked on the same document within a selected period, this was interpreted as an interaction between these users. User interactions for each period were captured in the form of a user-to-user matrix. It must be noted that this is only one of the many approaches for network creation that can be used based on the input dataset.

3.2. Network metrics and analysis

Network analysis was performed using the NetworkX library for Python, which has many built-in functions and algorithms for the creation, visualization and analysis of networks. Networks corresponding to the segmented project periods were analyzed using structural metrics such as the number of nodes, edges and communities, as well as using derived ratios, such as the number of nodes per community, the number of edges per node, and the number of edges per community. For global network metrics, user density and global efficiency were calculated. Additionally, for analyzing the activity of a specific user, the degree centrality and clustering coefficient were employed. Each of these metrics are described by another researchers (e.g., Arif, 2015; Schaefer & Rosen, 2013) and their definitions are used in this study. Following are the descriptions of the used network metrics:

- **Number of nodes, edges and communities:** Three fundamental metrics were utilized: the number of active users (nodes) during the current time period, the number of interactions (edges) identified between the active users, and the number of sub-teams (communities) of active users. The Louvain community detection algorithm was used to identify subteams of users. These results were additionally validated using the Clauset-Newman-Moore algorithm. Both algorithms are based on modularity principles: the Louvain algorithm optimizes modularity, while the Clauset-Newman-Moore algorithm focuses on modularity maximization. However, since the differences in the number of identified communities over time were insignificant between the two algorithms, Louvain algorithm was used for the subsequent analyses.
- **Number of nodes and edges per community and edges per node:** The number of edges and communities directly correlate with the number of active users. Since the number of active users in the sourced dataset varies significantly throughout different timestamps (project timeline), the number of edges was additionally divided by the number of nodes (active users) and communities (subteams of active users), whereas the number of communities was divided by the number of edges. This approach provides a more convenient insight into average activity, such as the number of active users per subteam, number of interactions per user, and number of interactions within subteams.
- **Density:** This metric shows how interconnected a network is by dividing the number of actual edges to the number of possible edges between all nodes. It reflects the overall level of interactivity within the network of active users, indicating how tightly knit or sparse the relationships among the users are.
- **Global efficiency:** This metric determines how effectively information or resources are exchanged across the entire network. It captures the average efficiency of all pairs of nodes, that is

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

the ease with which any node in the network can reach any other node through the shortest possible paths. Essentially, it's a measure of the network's overall ability to facilitate quick and efficient interaction between all pairs of active users.

- **Clustering coefficient:** This metric indicates the degree to which the nodes in the network tend to cluster together. It is a measure associated to individual nodes, but it can also be calculated for the overall network (as an average value of clustering coefficient for all nodes). In the context of this study, the clustering coefficient assesses the tendency of active users to create tightly knit subteams characterized by high density of interaction.
- **Degree centrality:** This metric quantifies the importance or influence of an individual node (active user), based on the number of edges (interactions) it has to other nodes in the network. It is most often calculated as a normalized value, where the number of edges for each node is divided by the network's maximum degree.
- **Eigenvector centrality:** Unlike degree centrality, which considers only the number of nodes' direct edges, this metric takes into account also the importance of these edges (whether the connected nodes are themselves highly important or influential). This recursive concept means that interacting with well-connected active users boosts an individual user's own centrality score. As such, eigenvector centrality reflects the active users' capacity to absorb information circulating within the network.

4. Results and discussion

The dataset was analyzed in three parts. The first was the analysis of user activity was conducted to grasp both the overall activity as well as the change in the distribution of activity type with the progress of the project. The second part was the analysis of the global network at different project timestamps, using metrics applicable to the overall network. Finally, the third part was an example of the network analysis for a single user using the corresponding metrics.

The following presentation of these results is accompanied by a brief discussion of possible implications. However, it must be noted that the goal of the paper is primarily to showcase the opportunities and applicability of network analysis for the analysis of CAD/PDM log data rather than deeply investigating the specific project, and that the interpreted insights have not been directly validated (e.g., via interviews, content analysis, etc.).

4.1. User activity analysis

To provide a general overview of the dataset, first the analysis of user activity for the entire project duration was conducted. For this purpose, the dataset was analyzed based on the previously described categories of collaborative CAD events.

Figure 1 shows the distributions of CAD activity categories on a weekly basis (moving window) for two frameworks. The graph on the left, which shows the distribution of categories as defined by Leonardo and Olechowski (2021), reveals that the users most often performed actions related to 'Element' and 'View/Scan' categories. Moreover, the 'Export/Import' category exhibits a small proportion, but from 135 week to 183 week its value spikes to a maximum 9% of total activity. Similar can be observed for the 'Add Part' category. This suggests that, on the begin of project, with small number of active users, they mostly create models. With a larger number of active users and a larger number of created documents, actions related to organization of user work and documents are more influential. At the last quartal, at the end of the project, larger number of completed documents are integrated to the final model (assembly) or exported to another format.

The right graph in Figure 1 shows the proportion of actions that fall into six main categories as defined by Deng et al. It reveals that 'Viewing' is the most present activity category throughout the project, whereas the 'Creating', 'Editing' and 'Revising' have a decreasing trend in favor of other actions as the project progresses. At the start of project, 'Revising' has one spike and after that their value decreasing with small variability. Similar can be observed for 'Editing' and 'Creating' categories. These categories have a smaller difference between minimum and maximum values and slightly expressed spikes in the first third of the project. One potential explanation for this could be that the more users join the project, and the more documents are created, the less CAD modeling work is done compared to other actions, such as administrating the project (PDM aspect of the Onshape platform).

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

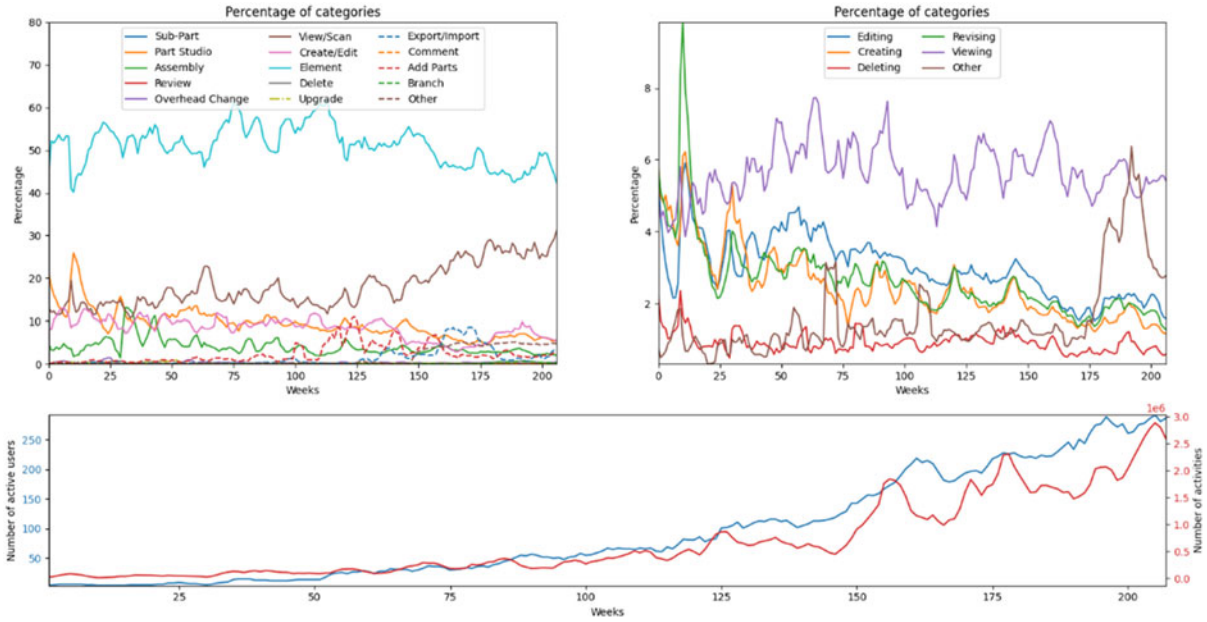


Figure 1. Moving window analysis of user activity distribution throughout the project

At the bottom of the figure, the number of active users and overall activity over time was plotted. It is visible that the number of actions is relatively proportional to the number of active users, which is expected for such a platform. During the first project quartal, both the number of users and their activity were relatively small (less than 20 users with several thousand actions). The activity changes drastically towards the end of the project, with more than 200 active users and more than 1 million actions on average in the last quartal. The maximum number of active users was 273, whereas the maximum number of actions in one period was 2,8 million. At the first 50 weeks, with small number of active users, categories like 'Editing', 'Creating' and 'Revising' (from the right chart) and 'Part Studio' (from the left chart) has larger impact then on rest of project. In period from 50 to 140 week, impact of early listed categories slightly fall. At the end of project, same categories have lowest values. Opposite situation is for categories 'Delete' (left chart) and 'Other' (right chart) which have maximum values. Increasing number of active users request better organization of works which is indicated by constantly falling of values of 'Editing' and 'Creating' categories and rising of category 'Other'.

The graphs plotted in Figure 1 showcase a simple, yet effective way to provide insight into the activity dynamics on a CAD/PDM platform. In addition to tracking the number of active users and their actions, the use of activity categories enables a more convenient overview of the type of work performed by the users. However, this type of analysis cannot capture the user interaction-related aspects of collaborative CAD work, such as forming of subteams and measuring importance of individual users. Therefore, these aspects are further explored by means of global and single user network analysis in the following subsections.

4.2. Global user network analysis

This subsection provides insight into the change of global user network throughout the project. For each period (moving window) a user-to-user matrix and the belonging network graph (visualization of nodes and edges) were created. In other words, an interaction snapshot was created for every one-week move of the one-month activity window, as a way of demonstrating the evolution of the global network of active users. Examples of networks graphs generated for various project stages are shown in Figure 2. These network graphs complement the activity graph shown at the bottom of Figure 1, by displaying the structure of users and their interactions.

During the first 50 weeks, the number of users is relatively small, with high network density (interactions between most pairs of users) and full network connectedness (there is an interaction path between any pair of users). During the rest of the project, with a larger number of active users, the density of the graph

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

has lower values. In Figure 2, the third snapshot from the left is the network graph for the 70th week. In the period from the 50th week to the end, network density varies more (e.g. 70th week and 207th week). Another noticeable change is the constant switching between connected (e.g., weeks 7, 25, 207 in Figure 2) and disconnected networks (e.g., week 70 in Figure 2).

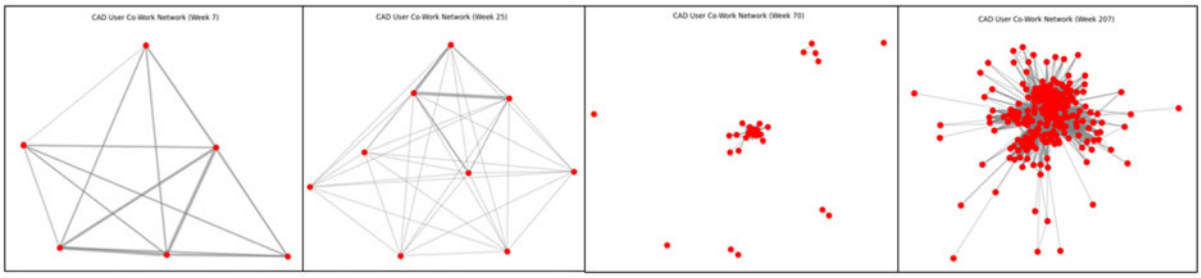


Figure 2. Snapshots of networks (visualization of user-to-user matrix) created at various stages of the project

Further insights were gathered by calculating the network metrics proposed in the Methodology section. The left side of Figure 3 shows the moving window plots of changes in the number of active users, interactions and subteams, whereas the right side shows the plots of changes in the number of users per subteam, edges per user, and edges per subteam.

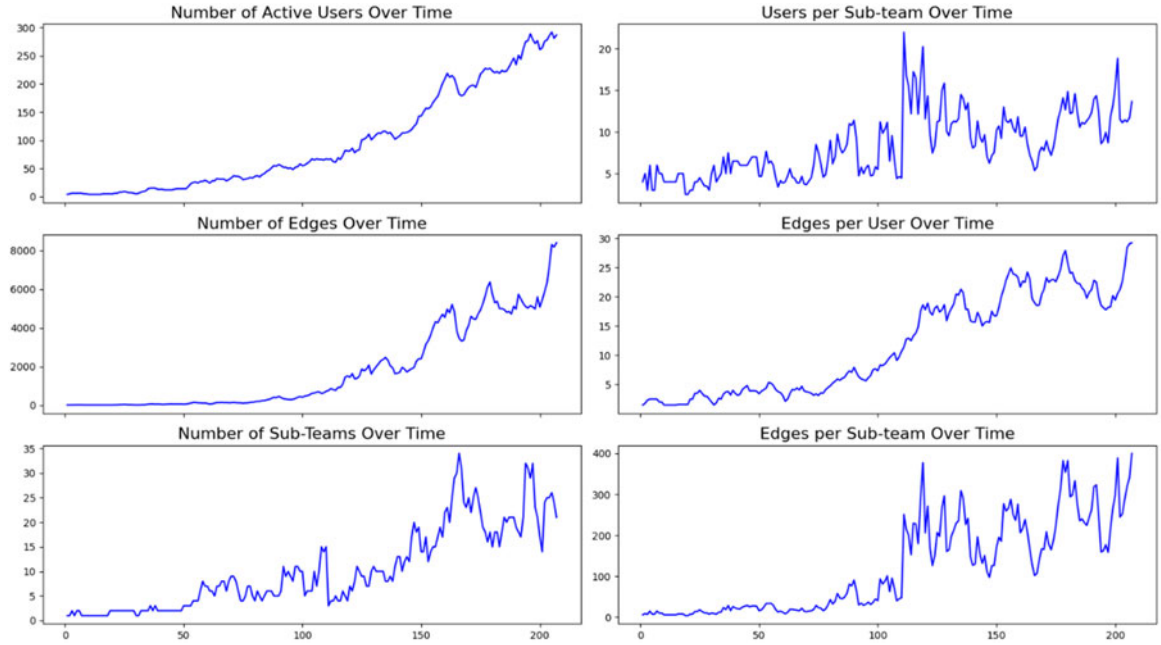


Figure 3. Moving windows analysis of number of active users, interactions and subteams (left) and number of users per subteam, edges per user, and edges per subteam (right)

The evolution of the network over time reveals notable patterns in the formation of subteams and the growth of edges (interactions). The number of identified subteams is at its lowest during the first quarter but fluctuates significantly afterwards. The number of edges, however, shows a weak dependence on the number of users. There were fewer than 100 edges during the first quarter, but this number grew to a maximum of 50,000 later in the timeline.

To address the differences in scale, these structural metrics were normalized by dividing them by the number of active users and subteams. Over the timeline, two distinct phases emerged: an initial period with relatively low values for these metrics, followed by a later phase where they are, on average, significantly higher. Nevertheless, high fluctuations are visible in the change of the number of users and edges per subteam, indicating that the increase in the number of active users does not necessarily imply

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

larger subteams. Rather, an interplay between the number of subteams and average subteam size is visible, thus providing interesting insights into the evolution of organizational structures at different project stages.
Additional moving windows analysis of metrics such as density, global efficiency and average clustering coefficient can be conducted to quantity aspects of collaboration related to interactivity/connectedness, the ability to facilitate quick and efficient interaction between all pairs of active users, and the tendency of active users to create closely connected subteams, respectively. Figure 4 demonstrates the change in these three metrics based on the moving windows analysis of project periods.

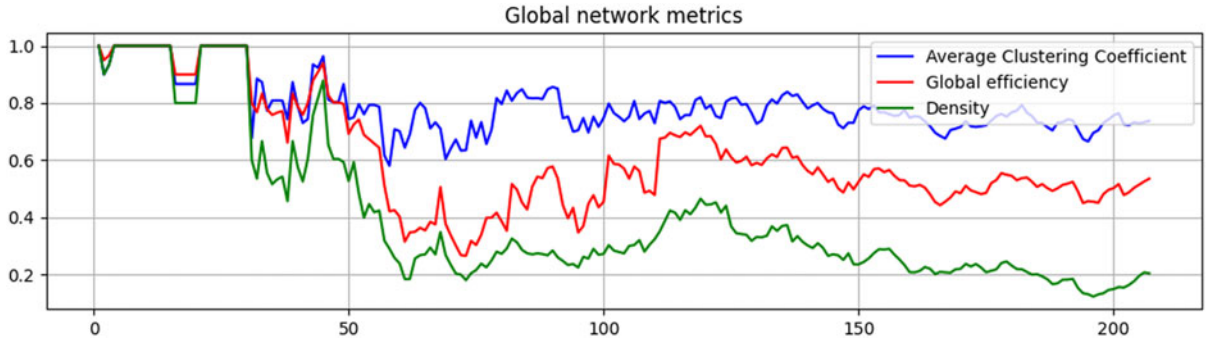


Figure 4. Moving windows analysis of global user network density, global efficiency and average clustering coefficient

The analysis of global efficiency and network density reveals distinct trends over time. During periods with a small number of active users, global network metrics, such as density and efficiency, tend to have higher values. This can be attributed to the fact that in smaller teams, users are more likely to collaborate directly with all other team members. In contrast, during periods with a larger number of users and documents, collaboration becomes more distributed, with users typically interacting with only a subset of the team. The size of subteams is directly indicated by number of users per subteam, where better connected networks (e.g., weeks 7 and 25 on Figure 2) are related with increasing subteam size. This shift results in lower values for both density and global efficiency, reflecting the less interconnected nature of the network in larger teams. The drop in values, although present, is less pronounced for the average clustering coefficient, which indicates that the tendency to form closely connected subteams remains relatively high and constant throughout the project timeline.
While these global metrics can provide insight into the dynamics of overall network structure and efficiency, analyses on the level of individual nodes are needed to determine the most important and influential users. The following subsection showcases the application of additional network metrics for a single user.

4.3. Single user network analysis

The graphs plotted in Figure 5 represent the moving window analysis of selected network metrics for a single user. More precisely, the graphs showcase the change in the values of eigenvector centrality, degree centrality, and clustering coefficient. Additionally, normalized activity distribution was plotted to describe the user's activity throughout the project, where the value represents their maximum activity during the project. While these metrics can be analyzed for each user, this study focuses only on a single example user—specifically, the one with the highest average degree centrality was selected, because it represents the most influential active user on average.
During the first quarter of the analyzed project, the degree centrality, eigenvector centrality, and clustering coefficient metrics for the selected user showed high values, even though this user's activity levels did not follow the same trend. This highlights the importance of network analysis, as the number of recorded actions does not necessarily correlate with the level of interaction between team members or the influence of specific users within the network.
Figure 5 shows three selected user metrics which indicate the selected user's influence. In the first quartal of the project, each of the metrics has higher values than the rest of time. While key users generally exhibit higher metric values, these values vary more compared to the global network metrics. Comparing

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

the metrics illustrated in the graphs (Figures 2 and 3), a slight dependence is observed. However, this conclusion relates only to the selected user, and it cannot be assumed that individual user graphs will align with global network trends.

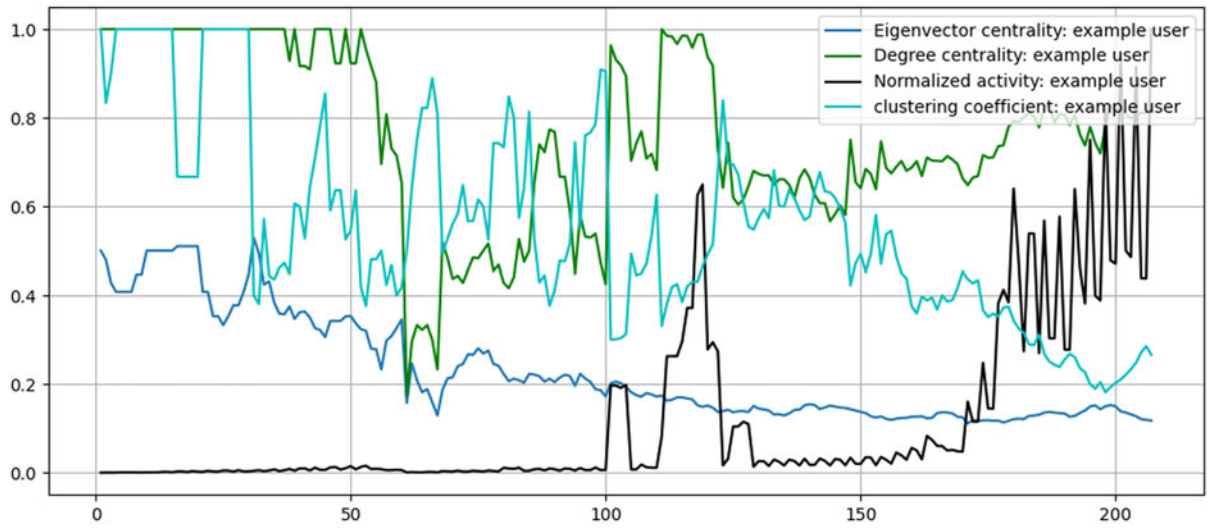


Figure 5. Moving average analysis of normalized activity, clustering coefficient, degree centrality and eigenvector centrality for a single user

5. Conclusions

This study aimed to showcase how the combination of network analysis and log data from CAD and PDM platforms can uncover valuable insights into the design process. The findings underscore the versatility of network analysis as a tool for exploring collaboration patterns, user interactions, and the dynamic evolution of design processes. For example, the analysis revealed that global network metrics, such as density, global efficiency, and clustering coefficient, provide insights into the overall structure and dynamics of collaboration over time. On an individual level, the potential of metrics such as degree centrality and eigenvector centrality for measuring the influence and importance of specific users was demonstrated, illustrating that the level of activity does not always correlate with a user's significance in the network.

The study highlights the research potential of using detailed log data to examine team dynamics and process structures in engineering design. The findings suggest new directions for investigating the alignment between inferred and actual user interactions, as well as exploring the relationship between network metrics and design project outcomes. In practice, the ability to analyze collaboration networks at both global and individual levels offers tangible benefits for engineering project management. Managers can use these insights to identify bottlenecks, improve team composition, and enhance communication strategies. By pinpointing key collaborators and understanding the structure of subteams, organizations can allocate resources more effectively, support high-performing teams, and address underperforming areas. Moreover, integrating network analysis into CAD/PDM platforms could enable real-time monitoring of collaboration dynamics, providing actionable feedback to improve project workflows.

The study also has several limitations that future work should address to enhance the validity and applicability of the findings. The interpretations of network metrics were not validated through methods such as user interviews and surveys or document reviews, which could provide critical context for understanding the observed trends. Additionally, the recorded edges represent inferred interactions through shared documents rather than direct user-to-user communication, and their accuracy as proxies for real-world interactions remains unconfirmed. Future research should validate these proxies and explore their alignment with actual collaboration dynamics. Lastly, the study analyzed a single dataset from a specific CAD/PDM platform, limiting generalizability. Expanding the approach to diverse datasets and platforms could strengthen the applicability of the proposed methods and insights.

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press

Acknowledgement

This research was funded by the project NPOO.C3.2.R3-I1.04.0121: Generative Design for Mass Personalization of Dental Implantoprosthetic Abutments (GENKON).

References

Arif T. (2015). The Mathematics of Social Network Analysis: Metrics for Academic Social Networks. *International Journal of Computer Applications Technology and Research*, 4(12), 889–893. https://doi.org/10.7753/ijcatr0412.1003

Carley K. M. (2014). *ORA: Toolkit A. for Dynamic Network Analysis and Visualization*.

Celjak R., Horvat N., & Škec,S. (2023). COMPARING COLLABORATIVE CAD MODELLING PATTERNS OF HIGH-PERFORMING AND LOW-PERFORMING TEAMS. *Proceedings of the Design Society*, 3, 1007–1016. https://doi.org/10.1017/pds.2023.101

Chen W., Heydari B., Maier A. M., & Panchal J. H. (2018). Network-based Modeling and Analysis in Design. *In Design Science* (Vol. 4). Cambridge University Press. https://doi.org/10.1017/dsj.2018.8

Deng Y., Mueller M., Rogers C., & Olechowski A. (2022). The multi-user computer-aided design collaborative learning framework. *Advanced Engineering Informatics*, 51. https://doi.org/10.1016/j.aei.2021.101446

Gopsill J., Snider C., Shi L., & Hicks B. (2016). COMPUTER AIDED DESIGN USER INTERACTION AS A SENSOR FOR MONITORING ENGINEERS AND THE ENGINEERING DESIGN PROCESS. *In No. DS 84). Faculty of Mechanical Engineering and Naval Architecture. INTERNATIONAL DESIGN CONFERENCE-DESIGN* (Vol. 84). http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/

Gu N., Watanabe S., Erhan H., & Haeusler M. H. (2014). APPLYING SOCIAL NETWORK ANALYSIS TO DESIGN PROCESS RESEARCH, A CASE STUDY. *Rethinking Comprehensive Design: Speculative Counterculture, Proceedings of the 19th International Conference of the Association of Computer-Aided Architectural Design Research in Asia CAADRIA* 2014

Gyory J. T., Goucher-Lambert K., Kotovsky K., & Cagan J. (2019). Exploring the application of network analytics in characterizing a conceptual design space. *Proceedings of the International Conference on Engineering Design, ICED, 2019-August*, 1953–1962. https://doi.org/10.1017/dsi.2019.201

Jermakovics A., Sillitti A., & Succi G. (2013). *Exploring Collaboration Networks in Open-Source Projects*. https://doi.org/10.1007/978-3-642-38928-3_7ï

Lan L., Liu Y., & Feng Lu W. (2018). Learning from the Past: Uncovering Design Process Models Using an Enriched Process Mining. *Journal of Mechanical Design*, 140(4). https://doi.org/10.1115/1.4039200

Leonardo K., & Olechowski A. (2021). Identifying computer-aided design action types from professional user analytics data. *Proceedings of the ASME Design Engineering Technical Conference*, 3B-2021. https://doi.org/10.1115/DETC2021-72102

Li Y., Roy U., & Saltz J. S. (2019). Towards an integrated process model for new product development with data-driven features (NPD 3). *Research in Engineering Design*, 30(2), 271–289. https://doi.org/10.1007/s00163-019-00308-6

Piccolo S. A., Lehmann S., & Maier A. (2018). Design process robustness: A bipartite network analysis reveals the central importance of people. *Design Science*, 4. https://doi.org/10.1017/dsj.2017.32

Piccolo S. A., Maier A. M., Lehmann S., & McMahon C. A. (2019). Iterations as the result of social and technical factors: empirical evidence from a large-scale design project. *Research in Engineering Design*, 30(2), 251–270. https://doi.org/10.1007/s00163-018-0301-z

Schaefer D., & Rosen D. W. (2013). Cloud-Based Design and Manufacturing Systems: Social A. Network Analysis. In *INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN*.

Šklebar,J., Martinec T., Škec,S., & Štorga,M. (2024). Analysis of collaborative CAD user actions in design sprint: insights from an educational setting. *Proceedings of the Design Society*, 4, 2963–2972. https://doi.org/10.1017/pds.2024.300

Snider C., Škec,S., Gopsill J. A., & Hicks B. J. (2017). The characterisation of engineering activity through email communication and content dynamics, for support of engineering project management. *Design Science*, 3. https://doi.org/10.1017/dsj.2017.16

Vrolijk,A.-P., & Olechowski A. (2024). Computer-aided design as a research instrument for engineering design: supporting multiple methods by observing artifacts, actors, and actions. *Journal of Engineering Design*, 1–32. https://doi.org/10.1080/09544828.2024.2432835

https://doi.org/10.1017/pds.2025.10339 Published online by Cambridge University Press