Cluster Ambiguity in Networks as Substantive Knowledge

Mathieu Jacomy[1], Tommaso Elli[2], Andrea Benedetti[3], Guillaume Plique[4], Benjamin Ooghe-Tabanou[4], Paul Girard[5], and Alexis Jacomy[5]

[1] Tantlab & MASSHINE, Aalborg University, Copenhagen, Denmark
[2] Dipartimento di Design, Politecnico di Milano, Italy
[3] Università degli Studi di Milano, Italy
[4] médialab, Sciences Po, Paris, France
[5] OuestWare, Nantes, France

Abstract

Visual network analysis has become increasingly used by scholars in the social sciences and humanities (SSH). While ambiguity is an inherent characteristic of this methodology, most currently available tools lack strategies to make such ambiguity transparent, particularly in relation to non-deterministic algorithmic results. The Louvain modularity algorithm, commonly used to detect communities within networks, is a good example of the issue: repeated executions can result in different community assignments for the same nodes. This paper introduces a novel technique for the visual inspection of results produced by the Louvain modularity algorithm. The proposed method involves an edge-centric analysis that evaluates how consistently pairs of connected nodes are assigned to the same community. This consistency metric is then visualised through a dedicated technique that uses colour-coding to highlight both stable and ambiguous relationships between nodes and clusters. The paper demonstrates the effectiveness of this approach with a proof-of-concept applied to benchmark datasets frequently used in the evaluation of network analysis tools. Finally, the contribution reflects on how this visual technique can support and enhance the heuristic practices of SSH scholars.

Keywords: visual network analysis, ambiguity, community detection, hermeneutics

1 Introduction

Network analysis has become a valuable tool for humanities scholars seeking to understand complex relationships within cultural, historical, and literary phenomena. Like with any other applied computational method, scholars face the challenge of meaningfully engaging with algorithmic processes while maintaining the interpretive depth that characterizes humanistic inquiry.

Community detection algorithms, which identify clusters of densely connected nodes within networks, exemplify this challenge. While these methods can reveal meaningful patterns in humanities data, they are *non-deterministic*: the same algorithm applied to identical data may produce different community structures, reflecting the complex and often ambiguous nature of the underlying relationships.

This article demonstrates how to measure and visualize ambiguity in the specific case of cluster detection in networks. We argue that ambiguity is not a secondary qualification of the results but constitutes primary knowledge about data and algorithmic processes. We address the problem of its quantification and visualization in ways relevant to humanities scholars. Our approach recognizes

Mathieu Jacomy, Tommaso Elli, Andrea Benedetti, Guillaume Plique, Benjamin Ooghe-Tabanou, Paul Girard, and Alexis Jacomy. "Cluster Ambiguity in Networks as Substantive Knowledge." In: *Computational Humanities Research 2025*, ed. by Taylor Arnold, Margherita Fantoli, and Ruben Ros. Vol. 3. Anthology of Computers and the Humanities. 2025, 119–130. https://doi.org/10.63744/f3L9hsFcGqVc.

that variability in computational processes is a substantive aspect of knowledge production that warrants systematic investigation and representation.

This perspective requires shifting from a broad understanding of uncertainty, which often encompasses ambiguity in existing literature, to a more specific definition distinct from ambiguity. For us, uncertainty refers to the limitations of observations and measurements, including issues of accuracy and completeness, while ambiguity refers to properties of phenomena that resist definitive categorization. Crucially, ambiguity is measurable, as it is inherent to a phenomenon, while uncertainty qualifies the measurement process itself. To illustrate this distinction: a blurry picture of a sharp thing must not be confused with a sharp picture of a blurry thing. Studying fuzzy realities requires distinguishing between uncertain observations of well-defined phenomena and accurate observations of inherently ambiguous phenomena.

By making the ambiguity of community detection interpretable, we propose that it constitutes knowledge relevant to a hermeneutic approach to network analysis. This perspective shifts the focus from seeking singular, definitive answers to exploring the range of plausible interpretations that computational methods can generate.

Our computational approach involves running a non-deterministic community detection algorithm multiple times and building ambiguity metrics by analyzing result variations. We then compile these metrics into a visualization designed to be actionable by humanities scholars, enabling them to identify areas of stability and ambiguity within their network data. This methodology provides researchers with a more nuanced understanding of network structures.

This short paper presents a design exploration and proof-of-concept implementation. It allows us to establish conceptual and practical landmarks necessary for future empirical validation while contributing immediately to the ongoing dialogue about ambiguity representation in computational humanities.

2 Related work

Humanities scholars encounter uncertainty in both their data and methods, requiring them to engage with it to make sound methodological decisions and communicate their findings effectively. Uncertainty, understood broadly to include ambiguity, has been identified as worthy of visualization, provided that its evaluation is "performed by objective and reproducible methodologies" [13]. This requires systematic characterization and framing of uncertainty (ibid). However, uncertainty visualization remains poorly understood within humanities scholarship [17].

Some humanists have proposed and discussed practical solutions to visualize uncertainty [6; 22], but importantly, those scholars also often contest that the humanities should have the same goals as the natural sciences, notably when it comes to minimizing uncertainty. Less certainty might on the contrary bring us "closer to the practice of humanistic hermeneutic traditions", as "visual argument structures such as contradiction, ambiguity, parallax, … are fundamentally hermeneutic in character" [7].

Uncertainty in networks in particular has been discussed. Venturini et al. have argued that it could be an asset for network analysis [19]. Different solutions for visualizing uncertainty at the level of nodes and edges have been explored [9; 20] and more recently, the discussion has been extended to other kinds of uncertainty and more visual variables [5]. They identify two main strategies, uncertainty integrated into the graph, or visualized as a supplement to the network. Our work draws on the former.

While those works aim to discuss the visualization of uncertainty in a relatively broad manner, we target the specific case of variability in the outcome of a community detection algorithm, in this case the "Louvain" method [3].

Many humanities scholars are unaware that popular cluster detection algorithms produce non-deterministic results. Existing user interfaces and visualization techniques do not indicate which nodes fluctuate between different clusters across multiple runs. Yet this variability has been documented by network analysis researchers. We know that the landscape of modularity clustering solutions is "degenerate" [8], meaning that many equally valid[1] yet distinct solutions exist for the same network. These solutions typically cannot be unified into a single compromise because they are too heterogeneous [16], and any single solution may miss relevant structural information [4].

However, this research has not emphasized that degeneracy in modularity maximization is unevenly distributed across the network. Not all nodes and edges participate equally in the variation between solutions. Some subgraphs consistently appear in the same community across multiple runs, which reveals stable structural features of the network. Building on our argument that ambiguity represents knowledge relevant to hermeneutic inquiry rather than noise to be mitigated, we propose a method for visualizing community detection ambiguity as an integral component of network analysis.

3 Method

3.1 Metric computations

The steps of the metric computations are: (1) sample the community detection solution landscape; (2) compute the "edge co-membership score" for each edge; (3) compute the "ambiguity score" for each edge; (4) aggregate it into the "ambiguity score" for each node.

Sampling the solution landscape. Set sample size n (default 50) and run n instances of the Louvain method in its non-deterministic variant [12], which is the implementation featured in the popular network analysis tool *Gephi* [2]. This presents unexpected challenges (cf. section 5).

Compute edge co-membership score. This score c_{ij} in the [0, 1] range measures the probability that the two nodes of an edge get placed in the same cluster across the sampled partitions. It could be computed on any node pair, but we only compute it for connected pairs (cf. section 5).

$$c_{ij} = \frac{1}{n} \sum_{k<=n} \delta_{ij}(k) \tag{1}$$

Where i and j are connected nodes; k is a sample partition; $\delta_{ij}(k)$ is 1 if i and j are in the same cluster in partition k, and 0 else.

Compute edge-level ambiguity. We use a straightforward heterogeneity metric that can be seen as the Gini coefficient for two populations. The underlying idea is that ambiguity is null at the two ends of the spectrum: whether the two nodes are *always* in the same cluster or *always* in different clusters. Ambiguity is maximal in-between, when the chances are equal. This score a_{ij} is normalized to range in [0,1].

$$a_{ij} = c_{ij} * (1 - c_{ij}) * 4 \tag{2}$$

Where i and j are connected nodes.

Compute node-level ambiguity. We simply aggregate the edge-level score at the node level by averaging it. It therefore also ranges in [0,1].

$$a_i = \sum_{j \text{ connected to } i} \frac{a_{ij}}{d_i} \tag{3}$$

[1] "Valid" meaning, in this context, maximizing a network partition metric known as "modularity" [14].

Where d_i is the degree of node i.

3.2 Visualization

Our solution aims to communicate at the same time which clusters are consistent and where (i.e. for which nodes and edges) clusters are ambiguous.

Visual variable	Represented quantity or quality
Node size	Ambiguity a_i
Edge hue	If $(c_{ij} > 0.5)$: cluster-specific hue Else: black (no hue) (the edge is a bridge)
Edge lightness	Ambiguity a_{ij} (high score in white)
Edge thickness	Ambiguity a_{ij} (high score thicker)
Edge drawing order	Ambiguity a_{ij} (high score on top)

Table 1: visual variables

In addition, the background is set to gray and all nodes set to white.
Remarks:

- We need to commit to a given partition to determine the edge hue, and for that we use one of the samples chosen arbitrarily. Remark that this choice is most problematic for edges with a low co-membership or a high ambiguity, but the former are visualized in black and the latter in white, which circumvents the problem (black and white have no hue, and therefore do not visualize any cluster).

- The edge drawing order is not technically a visual variable, but it is key to the success of the demonstrated technique.

- Conversely, node placement (spatial proximity) does not appear in this table even though it is a prominent visual variable, because the layout is technically independent of our design. However, it interacts strongly with it (cf. section 5).

Perceptive rationale. For simplicity, let us assume a sharp distinction between bridging edges, in-cluster edges, and ambiguous edges (Figure 1). In reality, the distinction is continuous and determined by c_{ij} and a_{ij}. First, the bridging edges can be identified because they contrast in black over the gray background. Second, the in-cluster edges contrast in color over the gray background. The cluster-specific hue gives a sense of where each *stable* cluster is located. Third, the ambiguous edges contrast in white over the gray background, the black bridging edges, and the colored in-cluster edges. Fourth and finally, the ambiguous nodes contrast in white over everything else except ambiguous edges. Like ambiguous edges, they indicate where community detection produces inconsistent results. We do not display non-ambiguous nodes (size set to zero).

4 Results

Our method provides new insights to interpret cluster detection results. It offers to contextualize cluster attribution outputs by displaying which edges are in-cluster, bridging, or ambiguous. Our design maintains a basic notion of communities but prioritizes the identification of ambiguity over it. It is not meant as a replacement, but as a complement to the usual community visualization where ambiguity is typically absent.

The resulting visualizations follow the intuition of experts while communicating simple facts to beginners. When the network consists of well-defined clusters (Figure 2), community detection is typically stable and few nodes will be ambiguous (Figure 2 has few white nodes and edges). There is nevertheless value in identifying those nodes for further investigation.

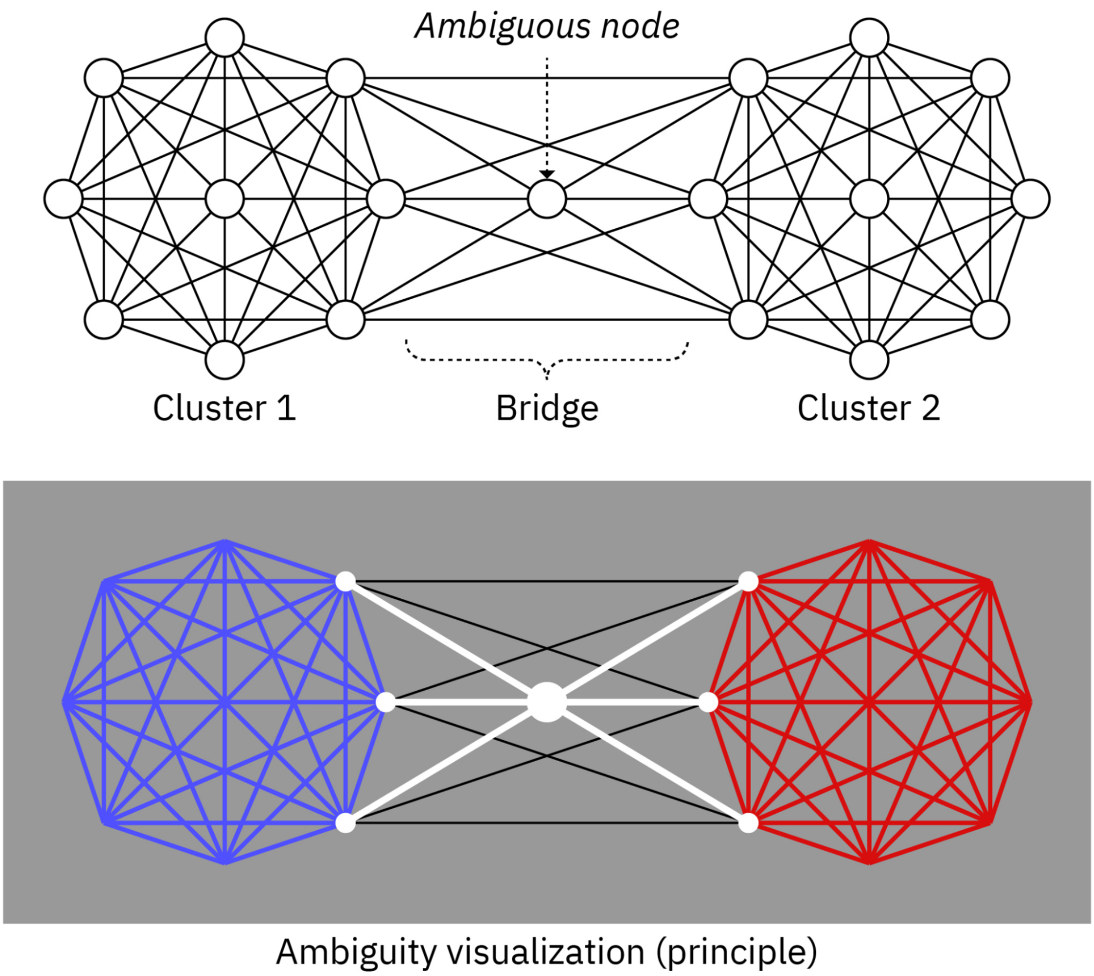

Figure 1: Principle of the visualization. Black edges are consistently bridging over different clusters. Colored nodes and edges are consistently within the same cluster. White nodes and edges are sometimes bridging, sometimes within the same cluster, depending on random factors in the community detection algorithm.

When the network does not have well-defined clusters (Figure 3), ambiguous nodes and edges are much more common (Figure 3 has more white than Figure 2). Nevertheless, the ambiguity is unevenly distributed and often concentrated in the middle of the network, while some areas, often on the sides, remain relatively stable (colored subclusters in Figure 3). This follows the intuition that nodes in the middle are disputed by different clusters.

The effect can be seen more clearly in Figure 4, representing a finite square lattice, where communities are consistently found in each corner (color in Figure 4) while the middle and sides are ambiguous (white in Figure 4).

Our visualization's most immediate impact is revealing how clustering results are far more ambiguous than users typically recognize. Rather than viewing clusters as fixed underlying structures that algorithms struggle to capture accurately, users we observed came to understand clusters as inherently fluid realities that resist complete categorization. The visualization also prompted them to examine "fluctuant nodes" more carefully, improving the accuracy of their analysis.

When users can interpret the visual encoding effectively, they can quickly identify which nodes shift most frequently between clusters and evaluate how consistent the detected communities actually are. This leads to more accurate network interpretation and creates valuable opportunities

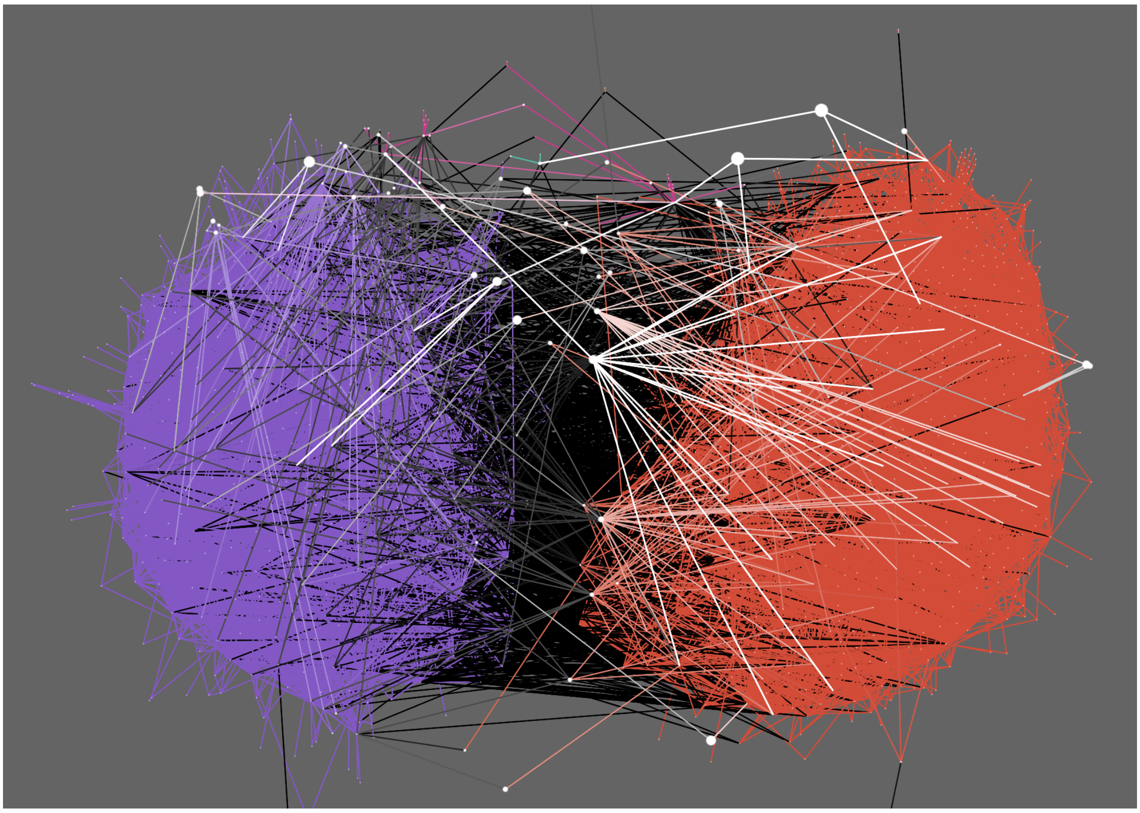

Figure 2: Visualization of the ambiguity of modularity clustering on the network from Divided they Blog [1], representing the US political blogosphere in 2024, structured in two clusters. Layout Force Atlas 2 [10].

for a more hermeneutic dialogue with the data and the computational processes mediating them.

We implemented our method in an open access interactive network visualization web application whose code source is licensed under the GPL-3.[2]

5 Discussion

5.1 Interactions between layout and visual variables

It is worth remarking that the visualization draws effectiveness from the fact that in-cluster edges and bridging edges are often spatially separated. Our graphic design is harder to read when in-cluster edges and bridging edges overlap. Those situations are typically rare because community detection is consistent with force-driven layout algorithms [15] which isolates communities in separate spatial regions. Consequently, other layout algorithms like Hive Plot [11] may not be compatible with our graphic design.

5.2 Omitting disconnected node pairs

Sampling all node pairs is quadratic and therefore computationally expensive, but we can focus instead on all edges, seeing them as node pairs (i.e., omit the disconnected pairs). We made this choice to keep the algorithm's computational budget in check, but also because we visualize the metric on edges, which prevents rendering it on disconnected pairs anyway.

[2] https://lite.gephi.org/

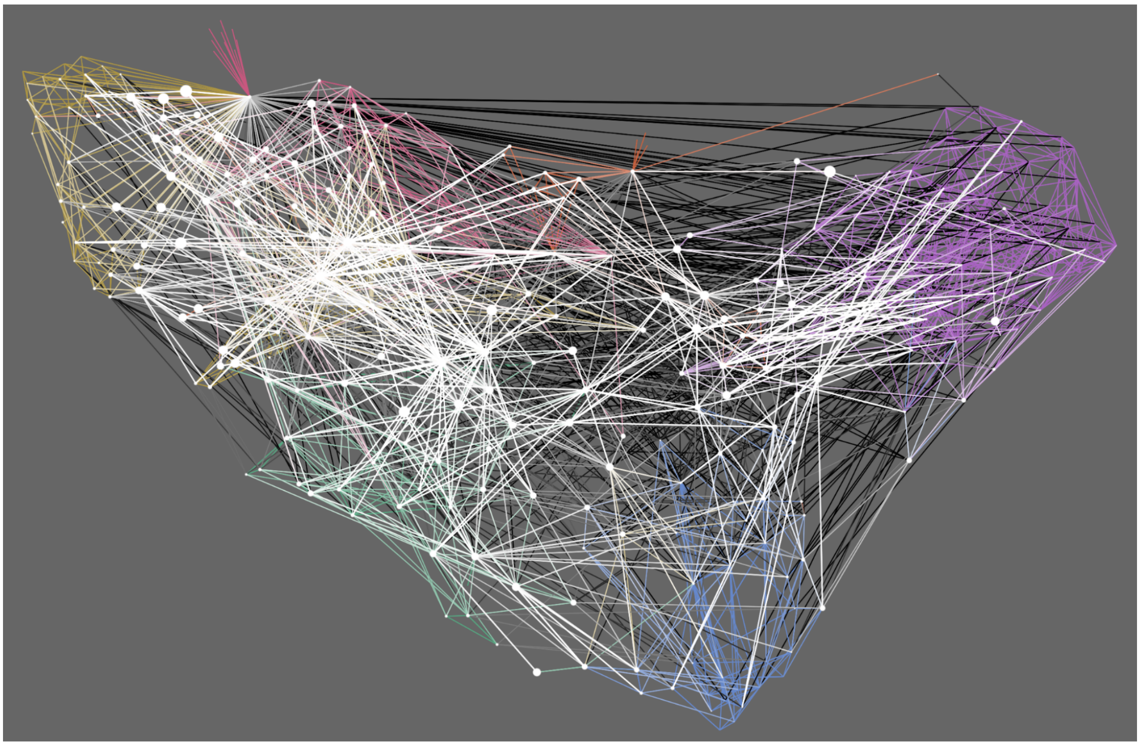

Figure 3: Visualization of the ambiguity of modularity clustering on the C. Elegans network, consisting of a nematode's neurons and their connections [21]. Layout Force Atlas 2.

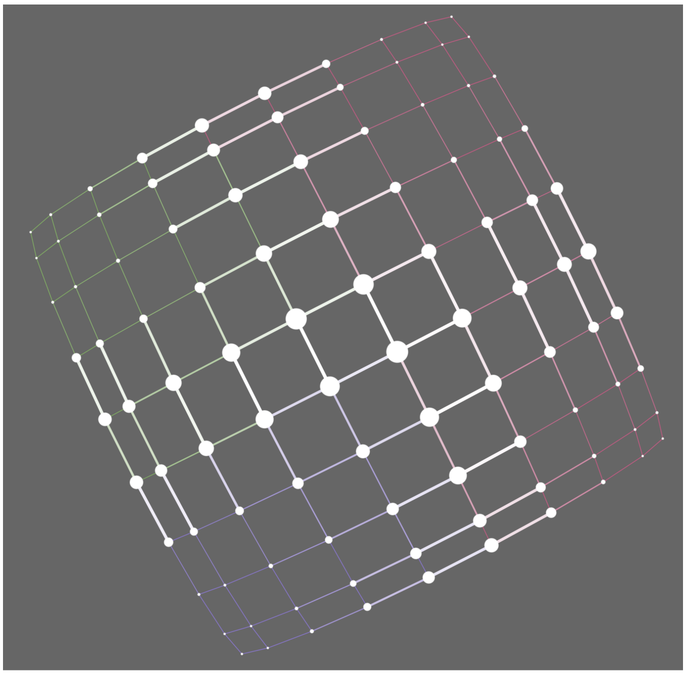

Figure 4: Visualization of the ambiguity of modularity clustering on a square lattice (a simple mathematical grid). Layout Force Atlas 2.

As a result, the ambiguity score at the node level reflects the ambiguity of the pairs it forms with its neighbors, not with every other node. This bias is aligned with the bias of modularity clustering, which aims to bring *connected* nodes in the same cluster, but it is not aligned with other forms of clustering like the stochastic block model. This is not a problem since our algorithm is specific to the Louvain method anyway.

It has at least another identified drawback. The ambiguity of any orphan node cannot be computed, because it has no neighbors. Similarly, the ambiguity of poorly connected nodes is measured less accurately than highly connected nodes.

5.3 Sampling Louvain results is unexpectedly challenging

Despite the high-level description offered by the Louvain method article [3], implementation is mostly an exercise left to the reader. We had to engage with it and change the implementation. We report here on why and how.

Let us start by remarking that anyone implementing this algorithm has to think carefully about which data structures to use and which computation strategies to leverage in order to avoid dramatically deteriorating its performance guarantees. We observed that implementers had relied on common techniques to optimize running time all while ensuring result accuracy, notably in the Gephi version.

Unfortunately for us, this had the effect of narrowing down the solution landscape, presumably unintentionally. This problem is neither concerning nor apparent if you run the algorithm only once, because modularity remains effectively and demonstrably maximized. However, when we endeavored to sample the landscape of results, we realized that the samples were unexpectedly contained within a region around local optimum, albeit a good one. Although this was not a problem to the user seeking a single result, it meant that our sampling was sensibly biased and far from representative.

Acknowledging that implementing the Louvain algorithm for sampling is not equivalent to implementing it for one-shot clustering, we had to make three interventions:

1. Optimizations that tie-break neighboring communities when they yield an equivalent modularity delta in a deterministic way had to be made random.

2. We had to make the graph traversal truly random, i.e. over a permutation of the node list. Multiple implementations (notably in Gephi) only select the starting node at random and then iterate over the node list in source data order, which is far from random.

3. As introduced in [18], some implementations maintain a queue of neighbors that are relevant to visit later, in order to avoid needless delta computations when performing multiple traversals of the graph at each iteration of the algorithm. This optimization has to be dropped to correctly sample the landscape of solutions.

The necessity of such intervention yields a valuable lesson. Visualizing ambiguity as foreground information conflicts with the common practice of arbitrarily reducing sources of randomness for optimization purposes. Indeed, in our situation, randomness is not an inconvenience to mitigate but a valuable source of knowledge about algorithmic behavior. We leverage randomness as a proxy for degrees of freedom within the algorithm that we aim to probe, and it is therefore important to preserve them in order to study the algorithm's behavior accurately.

5.4 Extending our design to other clustering methods

Our algorithm could be easily extended to algorithms that use similar strategies to the Louvain method, which includes the Leiden method [18] and Bayesian inference [23]. The two requirements for applying our approach are the non-deterministic nature of the algorithm, and its goal to capture *assortative* clusters.

5.5 Applications

Our system allows identifying whether a network is self-consistently partitioned by a given clustering algorithm, and if there is inconsistency, where it lies. First, as context for node classification, it highlights areas of the network that should not be considered accurately classified (e.g., which neurons should not be reduced to a role in a local cluster?). Second, and more importantly, it allows identifying areas of interest for scholars whose research questions involve the ambiguity, polyvalence, or contradiction of node categories (e.g., which parliamentarians vote unlike their political affiliation?). In both cases, it brings nuance to the interpretation of the network's community structure.

6 Conclusion

This work demonstrates that ambiguity in community detection algorithms can constitute substantive knowledge rather than incidental context. By systematically measuring and visualizing the variability in clustering outcomes, we provide humanities scholars with a tool to embrace interpretive multiplicity rather than seeking singular, definitive categories to describe the multifaceted phenomena they study.

Our visualization method reveals what is typically hidden: algorithmic degrees of freedom that produce variability in the presence of undefined cluster boundaries. This transparency enables more nuanced interpretations of network structures and encourages scholars to engage critically with computational methods. Rather than treating algorithmic ambiguity as a limitation to overcome, we propose it as a valuable dimension of analysis that can deepen the hermeneutic understanding of relational phenomena.

Future work should extend this approach to other clustering methods, conduct empirical studies to refine our design, and explore how ambiguity visualization can be integrated into the hermeneutic practices that define humanistic scholarship.

References

[1] Adamic, Lada A and Glance, Natalie. "The political blogosphere and the 2004 US election: divided they blog". In: *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005, pp. 36–43. ISBN: 1-59593-215-1. URL: `http://www.cyberjournalist.net/news/001461.php`.

[2] Bastian, Mathieu, Heymann, Sebastien, and Jacomy, Mathieu. "Gephi: An Open Source Software for Exploring and Manipulating Networks". In: *Third International ICWSM Conference*. Citation Key: Bastian2009 ISSN: 14753898. 2009, pp. 361–362. ISBN: 978-1-57735-421-5. URL: `http://www.aaai.org/ocs/index.php/ICWSM/09/paper/viewFile/154/1009/`.

[3] Blondel, Vincent D., Guillaume, Jean-Loup, Lambiotte, Renaud, and Lefebvre, Etienne. "Fast unfolding of communities in large networks". In: *Journal of Statistical Mechanics: Theory and Experiment 2008*, no. 10 (2008). arXiv: 0803.0476. ISSN: 17425468. DOI: `10.1088/1742-5468/2008/10/P10008`.

[4] Calatayud, Joaquín, Bernardo-Madrid, Rubén, Neuman, Magnus, Rojas, Alexis, and Rosvall, Martin. “Exploring the solution landscape enables more reliable network community detection”. In: *Physical Review E 100*, no. 5 (Nov. 2019). arXiv: 1905.11230 publisher: American Physical Society, p. 052308. ISSN: 24700053. DOI: 10.1103/PhysRevE.100.052308.

[5] Conroy, Melanie et al. “Uncertainty in humanities network visualization”. English. In: *Frontiers in Communication 8* (Jan. 2024). ISSN: 2297-900X. DOI: 10.3389/fcomm.2023.1305137. URL: https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2023.1305137/full.

[6] Drucker, Johanna. *Humanities Approaches to Graphical Display*. 2011.

[7] Drucker, Johanna. “Non-representational approaches to modeling interpretation in a graphical environment”. In: *Digital Scholarship in the Humanities 33*, no. 2 (June 2018), pp. 248–263. ISSN: 2055-7671. DOI: 10.1093/llc/fqx034.

[8] Good, Benjamin H., De Montjoye, Yves Alexandre, and Clauset, Aaron. “Performance of modularity maximization in practical contexts”. In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 81*, no. 4 (Apr. 2010). arXiv: 0910.0165, p. 046106. ISSN: 15393755. DOI: 10.1103/PhysRevE.81.046106.

[9] Guo, Hua, Huang, Jeff, and Laidlaw, David H. “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”. In: *IEEE Transactions on Visualization and Computer Graphics 21*, no. 10 (Oct. 2015), pp. 1173–1186. ISSN: 1941-0506. DOI: 10.1109/TVCG.2015.2424872.

[10] Jacomy, Mathieu, Venturini, Tommaso, Heymann, Sebastien, and Bastian, Mathieu. “ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software”. In: *PLoS ONE 9*, no. 6 (2014). Citation Key: Jacomy2014 ISBN: 10.1371/journal.pone.0098679, pp. 1–18. ISSN: 19326203. DOI: 10.1371/journal.pone.0098679.

[11] Krzywinski, Martin, Birol, Inanc, Jones, Steven JM, and Marra, Marco A. “Hive plots—rational approach to visualizing networks”. In: *Briefings in Bioinformatics 13*, no. 5 (Sept. 2012), pp. 627–644. ISSN: 1467-5463. DOI: 10.1093/bib/bbr069.

[12] Lambiotte, R., Delvenne, J.-C., and Barahona, M. “Laplacian Dynamics and Multiscale Modular Structure in Networks”. In: *IEEE Transactions on Network Science and Engineering 1*, no. 2 (July 2014). arXiv:0812.1770 [physics], pp. 76–90. ISSN: 2327-4697. DOI: 10.1109/TNSE.2015.2391998.

[13] Levontin, Polina, Walton, Jo, Aufegger, Lisa, Barons, Martine, Barons, Edward, French, Simon, Houssineau, Jeremie, Kleineberg, Jana, Mcbride, Marissa, and Smith, Jim. *Visualising Uncertainty: A short introduction*. Jan. 2020. ISBN: 978-1-912802-05-0.

[14] Newman, M. E. J. “Modularity and community structure in networks”. In: *Proceedings of the National Academy of … 103*, no. 23 (2006). arXiv: physics/0602124 Citation Key: Newman2006 ISBN: 0027-8424 (Print)r0027-8424 (Linking), pp. 8577–8582. ISSN: 0027-8424. DOI: 10.1073/pnas.0601602103.

[15] Noack, Andreas. “Modularity clustering is force-directed layout”. In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 79*, no. 2 (2009). arXiv: 0807.4052 ISBN: 1539-3755. ISSN: 15393755. DOI: 10.1103/PhysRevE.79.026102.

[16] Peixoto, Tiago P. “Revealing Consensus and Dissensus between Network Partitions”. In: *Physical Review X 11*, no. 2 (Apr. 2021), p. 021003. DOI: 10.1103/PhysRevX.11.021003.

[17] Therón, Roberto and Wandl-vogt, Eveline. "Uncertainty in Digital Humanities track summary". In: *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*. TEEM'18. New York, NY, USA: Association for Computing Machinery, Oct. 2018, pp. 815–818. ISBN: 978-1-4503-6518-5. DOI: 10.1145/3284179.3284317. URL: https://dl.acm.org/doi/10.1145/3284179.3284317.

[18] Traag, V. A., Waltman, L., and Van Eck, N. J. "From Louvain to Leiden: guaranteeing well-connected communities". en. In: *Scientific Reports 9* (Mar. 2019), e5233. DOI: 10.1038/s41598-019-41695-z.

[19] Venturini, Tommaso, Jacomy, Mathieu, and Jensen, Pablo. "What do we see when we look at networks: Visual network analysis, relational ambiguity, and force-directed layouts". en. In: *Big Data & Society 8*, no. 1 (Jan. 2021), p. 20539517211018488. ISSN: 2053-9517. DOI: 10.1177/20539517211018488.

[20] Wang, Yong, Shen, Qiaomu, Archambault, Daniel, Zhou, Zhiguang, Zhu, Min, Yang, Sixiao, and Qu, Huamin. "AmbiguityVis: Visualization of Ambiguity in Graph Layouts". In: *IEEE Transactions on Visualization and Computer Graphics 22*, no. 1 (2016), pp. 359–368. ISSN: 10772626. DOI: 10.1109/TVCG.2015.2467691.

[21] Watts, Duncan J. and Strogatz, Steven H. "Collective dynamics of 'small-world' networks". en. In: *Nature 393*, no. 6684 (June 1998), pp. 440–442. ISSN: 1476-4687. DOI: 10.1038/30918.

[22] Windhager, Florian, Salisu, Saminu, and Mayr, Eva. "Exhibiting Uncertainty: Visualizing Data Quality Indicators for Cultural Collections". en. In: *Informatics 6*, no. 33 (Sept. 2019), p. 29. ISSN: 2227-9709. DOI: 10.3390/informatics6030029.

[23] Zhang, Lizhi and Peixoto, Tiago P. "Statistical inference of assortative community structures". In: *Physical Review Research 2*, no. 4 (Nov. 2020), p. 043271. DOI: 10.1103/PhysRevResearch.2.043271.