Maastricht University

Purposes and challenges of legal network analysis on case law

Citation for published version (APA):

Document status and date:
E-pub ahead of print: 26/08/2020

Document Version:
Accepted author manuscript (Peer reviewed / editorial board version)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Download date: 07 Jan. 2021

Purposes and challenges of legal citation network analysis on case law

Dafne van Kuppevelt, Gijs van Dijck and Marcel Schaper

1. INTRODUCTION

Legal scholarship can pursue various objectives. It may aim to analyze what the law is and how it should be understood, to evaluate existing rules, to understand the law as a social phenomenon, or to understand the conceptual bases of legal principles (e.g. Hutchinson, 2010, pp. 7–8). Case law functions as a pivotal source of information in a substantial proportion of legal scholarship.

When analyzing court decisions, a variety of questions can emerge. Researchers may observe that textbooks differ in the sub-topics that can be distinguished. It may not be clear which decisions can be studied to obtain a proper overview of the relevant case law. There may be a discussion among scholars about which decisions should be considered the most relevant precedents. Researchers might be interested how the relevance of decisions changes over time (i.e. whether decisions gain or lose relevance over time). Or they might want to explore whether legal scholars have overlooked possibly relevant precedents.

Traditionally, legal scholars rely on human analysis when determining which court decisions are relevant or authoritative. Legal researchers search, read, and interpret legal documents to synthesize court decisions and to determine their relevance. The method of case synthesis is commonly applied by legal researchers and law students to assess the importance of a case (Nievelstein et al., 2013). It essentially entails that case outcomes are compared with the facts of the cases, with the purpose of explaining the differences in outcomes by the differences in facts (Gionfriddo, 2007).

With hundreds, thousands or tens of thousands of court decisions, even in a specific field, computational analysis has the potential to assist researchers in answering research questions regarding the trends, patterns, structure, and organization of the legal system. Network analysis in particular fits this purpose, as it allows mapping, measuring, and possibly visualizing relationships between information (Fowler et al., 2007; Schaper, 2014; van Dijck, 2017), which relates to the information legal scholars are often interested in when analyzing court decisions.

In this contribution, we explore the potential and challenges of Legal Network Analysis (LNA). After discussing for which research questions LNA may be used (Section 2), it will be explained that LNA lacks a proper reference point for evaluating the results (Section 3). Consequently, a methodology needs to be developed in order to produce results that are valid. This contribution is structured around four specific aspects, which are explored more in-depth: (1) how to select sub-networks, (2) which community detection method to select, (3) estimating the probability that the network and its relationships as observed in the data did not occur by chance, and (4) which centrality measure to select to determine the extent to which a decision is a precedent (Section 4). Issues regarding these topics will be discussed, and avenues for solutions will be explored. We conclude that the way in which sub-networks are selected can have a substantial impact on the findings, that community detection algorithms perform differently in different contexts, that statistical significance testing may allow for estimating the probability that the network and its relationships as observed in the data are not likely to be the result of chance, and that centrality measures commonly used in LNA produce different results in different contexts. By exploring the purposes and challenges of LNA, we aim to develop a research agenda for conducting LNA.

This contribution focuses on network analysis applied to case law, citation analysis in particular, as this is the most common type of data to apply LNA to. However, network analysis can also be applied to other domains or areas, including legal social networks (Beaverstock, 2004; Heinz and Laumann, 1982; Lazega, 2001; Lazega and Krackhardt, 2000; Lazega and van Duijn, 1997), networks of statutes and regulatory codes (Bommarito II and Katz, 2010; Boulet et al., 2011; Katz and Bommarito, 2014; Koniaris et al., 2017), and patent citations (Whalen, 2016, pp. 550–51 (providing an overview)).

When discussing the challenges and possible solutions for them, we draw upon literature on network analysis in other disciplines. Because there is an abundance of literature, a selection needed to be made. The references included in this contribution may be used to obtain more knowledge and insight in the challenges and solutions.

2. WHY CONDUCT LNA?

LNA seems to be predominantly used to identify relevant precedents (Olsen and Küçüksu, 2017; Olsen and Sadl, 2017), that is, decisions, rules, or rules set in a prior legal case are used to decide subsequent cases (Derlén and Lindholm, 2015, pp. 1075–6; Lupu and Voeten, 2012, pp. 416–17). Determining to what extent a court decision is a precedent is an important aim of legal scholarship. Computational analysis offers the advantage of scalability (compared to human analysis): hundreds, thousands or tens of thousands of cases can be investigated. Treating court decisions as nodes and court citations as edges (links), a precedent network can be construed that allows testing how central precedents are (Fowler et al., 2007, p. 325). For example, LNA has been applied in order to identify the most important precedents in the corpus of US Supreme Court majority opinions in the 1754–2002 period (Fowler et al., 2007; Fowler and Jeon, 2008). For this, the authors analyzed the complete network of 30,288 and the cases that cite them. Similarly, others have analyzed the body of case law of the Court of Justice of the European Union (CJEU), formerly known as the European Court of Justice. They found that the importance of judgments such as *van Gend en Loos*, *Costa v. ENEL*, *Brasserie du Pêcheur*, and *United Brand* is overemphasized, and the importance of other judgments like *Bosman*, *PreussenElektra*, and *Schumacker* underestimated (Derlén and Lindholm, 2014, p. 686). In similar fashion, LNA has been used to measure the centrality of CJEU rulings in the field of VAT (Knops and Schaper, 2014) or direct taxes (Schaper, 2014), of the International Criminal Court decisions (Tarissan and Nollez-Goldbach, 2016), of Canadian case law (Neale, 2013), and of Italian constitutional court case law (Agnoloni and Pagallo, 2015).

One assumption implicit in much LNA is that citation frequency is an indicator for how authoritative the decisions are: case law that is cited more frequently is presumed to be more important than cases cited less frequently. Thus, *centrality measures* are calculated based on, for example, *in-degree* (number of times the decision is cited) and *out-degree* (number of citations cited in the decision). Centrality, however, does not necessarily

coincide with relevance, importance, or authority. This has to do with the quality of the citations, that is, the reasons for citing case law. Court decisions may be cited as a matter of courtesy, to add authority for the interpretation of a rule or for solving a case, to indicate the same rule is applied that was laid out in previous cases, to distinguish the case at hand from other cases, or to stress that the deciding court departs from a rule laid down in a previous case (Derlén and Lindholm, 2014, p. 674; van Dijck, 2017, p. 28; Whalen, 2016, p. 556) Consequently, there might be a gap between the centrality of precedents and their relevance, importance, or authority.

Several studies have explored the relationship between centrality and relevance by comparing network centrality scores with expert rankings. (Fowler and Jeon, 2008, p. 21), in their study on US Supreme Court cases, report that the "top 10" cases in their analysis are considered in one or more of the legal sources they compared their results to. Similarly, the research that analyzed the centrality of CJEU decisions reported high levels of agreement regarding which decisions belong to the "top 10" judgments, although textbook authors significantly disagree as to which judgments are the most important:

> "65% of the 117 judgments only occur among the frequently cited in a single textbook, 76% of the judgments are among the frequently cited in only one or two textbooks, only 10 judgments are frequently cited in five or more textbooks, and only two cases are among the top 30 cited in all seven textbooks. (Derlén and Lindholm, 2014, pp. 668–9)"

The finding led the authors to the conclusion that textbooks "overemphasise the importance of van Gend en Loos, Costa v. ENEL, Brasserie du Pêcheur and United Brands", and that regarding van Gend en Loos, "its status as a source of law is questionable, at best" (Derlén and Lindholm, 2014, p. 686).

LNA can also be used to identify characteristics or events related to the centrality of a court decision. In another study on US Supreme Court decision citations, it was found that "decisions written in years when the mean judicial age is low and judges are more stable in

their use of precedent, more conservative in terms of the age of precedent cited, and the yearly citation network is less complex are more likely to be cited in future years" (Whalen, 2013). For CJEU decisions, LNA has shown that CJEU case law is particularly central in preliminary reference rulings, rulings that concern fundamental freedoms, competition law in particular (Derlén and Lindholm, 2015). The same study reports that the number of Member States submitting an observation in a preliminary reference procedure predicts the frequency with which a case will be referred to in subsequent decisions. LNA has also been used to measure how long decisions remain central in a network. A Canadian study reveals that case citation typically has a lifespan of 3 to 15 years, depending on the jurisdiction (Neale, 2013). Only the Supreme Court of Canada persist longer, namely 50 years, while reporting that 19 percent of Canada Supreme Court cases are regularly cited over time, as opposed to less than 3 percent in other jurisdictions.

Furthermore, LNA has been used to determine characteristics of overruled decisions. In a study that used a Supreme Court decisions dataset, it was found that overrulings are likely to be reserved for decisions that are more central than for decisions that remain salient longer (Whalen, 2014). Finally, LNA has been used to consider the structure network of decisions on a meso-scale, and thus discern clusters (or communities) of decisions that are closely related in the network. In a study of employer liability cases of Dutch Supreme Court decisions, several communities are identified and interpreted that seemingly provide a meaningful clustering from a substantive perspective (van Dijck and van Kuppevelt, 2018).

3. VALIDATING LNA RESULTS

Although, intuitively, the added value of LNA is large, a thorough evaluation of the new methodology requires a criterion to measure the success and usefulness of LNA. As the studies discussed in the previous sections illustrate, expert opinions are often used as the baseline against which the results of a LNA are evaluated. At the same time, the way in which the expert opinions are measured varies. Fowler and Jeon (2008, p. 20) mention the *Congressional Quarterly's Guide to the United States Supreme Court*, the *Oxford Guide to Supreme Court Decisions*, and the *Legal Information Institute* as sources that use opinions of judicial specialists to compile rankings of the most important Supreme Court decisions based, with the rankings being based on the case's historical and/or social significance, its

importance to the development of some area of law, and its impact on the development of American government.

If expert ratings are not available, they may be constructed based on existing sources. In their study of CJEU decisions, Derlén and Lindholm (2014) did so by counting textbook citations and citations by the Advocates General. Scholars that analyzed Canadian case law with LNA used the frequency with which the analyzed cases were viewed on *Canadian Legal Information Institute* (CanLII's) website. They concluded that in-degree centrality and PageRank scores are associated with the viewing frequency on CanLII's website (Neale, 2013). Another measure to compare centrality scores to is to count the number of scholarly opinions, comments, and annotations devoted to the cases analyzed. Applying this strategy to Italian constitutional law cases resulted in the finding that several cases discussed by experts turned out not to be central in the network analysis that was conducted, whereas a number of cases central in the network were scarcely debated or ignored by legal scholars (Agnoloni and Pagallo, 2015). Based on this result, the authors argued that results of LNA and the web of scholarly opinions should be combined in order to determine the relevance of court decisions.

Considering that LNA can bring insights that are not noted by previous scholars, the evaluation of LNA is inherently problematic. LNA that studies case law sometimes confirms scholarly commentary, and sometimes it deviates. As far as the purpose of LNA is to confirm scholarly commentary, the latter can serve as a proper baseline to measure LNA's effectiveness. This way, LNA can be seen as a faster and more powerful alternative to human analysis, at least when it comes to determining the centrality of decisions, which may be seen as an indication of relevance, importance, or authority of precedents. Interesting in this respect is the observation in the Fowler and Jeon study, where it was concluded that it had taken judicial specialists 18 years to recognize the significance of *Speiser versus Randall* (1958) as an influential decision (Fowler and Jeon, 2008, pp. 21–2), whereas the partitioning of the network to cases before 1979 allowed to detect the *Speiser* case as a central case in the network (although the network analysis was conducted 50 years after *Speiser versus Randall*).

In contrast to replicating scholarly expertise, LNA also provides a quantitative approach to answer research questions different from traditional legal research. When network statistics deviate from an expert baseline, this does not necessarily imply a shortcoming of the method, it can expose new insights in the data. An informed choice of methods and metrics is required to determine whether legal scholarship has overlooked information or whether LNA produced results that lack meaning. As will be further explained below, one approach is the formulation of clear hypotheses that can be statistically tested on the network data. In addition, awareness of the sensitivity of methods to noise in the data is required. When the proper method is used and significant findings do not match the expected result, additional analysis can reveal relationships between the network structure and expected structure.

LNA applied to case law is therefore partly confirmatory and partly exploratory, in that it produces new hypotheses that need to be tested by conducting additional analysis. The fact that it cannot be predicted beforehand whether an LNA will confirm or deviate from the findings in scholarly commentary further stresses the importance of a proper methodology for LNA and, consequently, of identifying and addressing challenges.

4. CHALLENGES OF LNA

When discussing the challenges that an LNA scholar might encounter, and the robustness of the methods available to address those challenges, we turn to other disciplines, as valuable lessons can be learned from other fields that have extensive experience using network analysis as a tool for research, most notably the social sciences (Social Network Analysis, or SNA).

We illustrate the challenges with examples of a citation network of Dutch court decisions. The data, which originates from the LiDO tool,[1] contain cases related to employer liability, which concern situations where the employee suffers a loss for which the employer is liable. Employer liability is covered by the Articles 7:658 and 7:611 of the Dutch Civil Code. All cases with computer-identified references to these articles were selected. An API provided by LiDO allowed for the automated selection of these references, as well as the references between the selected cases.

---

[1] http://linkeddata.overheid.nl.

4.1 Selecting a Sub-Network

The complete network of all available court decision is very large, which makes the network difficult to visualize and interpret. In addition, often in LNA, only a small portion of the decisions are relevant to the research questions. For example, a researcher may be interested in the topic of employer liability in case of traffic accidents instead of in all decisions on employer liability. The scope of the network can be decreased by selecting a sub-graph of (mostly) relevant cases. The smaller network is easier to visualize and interpret than a larger network, and results do not get confounded by irrelevant cases for the research question at hand.

Tarissan et al. (2016) show that selecting a sub-network of a certain body of case law reveals more landmark cases than the complete network. However, selecting a part of the network has consequences for the network properties of the sub-graph. Overall statistics, such as degree distribution and clustering properties, are likely to be different for the sub-graph than for the complete graph. Such deviations may be of interest to the researcher, as they reveal specific properties of the selection. An example of this is the work of Henrik Palmer Olsen, one of the other contributions in this book, where centrality scores within sub-networks is compared to centrality in the overall network, to highlight judgments that contain general legal principles. However, there are caveats to be aware of when using sub-networks.

Firstly, the method of selection needs to be carefully considered. One way of selecting a sub-graph is based on the network structure, for example the sub-graph of all nodes within a fixed degree from one node of interest. In the employer liability example, this would mean the researcher chooses one or more cases of interest and subsequently selects all decisions citing that decision, and possibly also the cases that refer to those decisions (and so on). It is also possible to sample based on network independent criteria, such as keyword search (e.g. searching for "employer" and "liability"). Finally, one may combine the methodologies, for example by selecting all nodes that satisfy a keyword plus nodes linked to those nodes. Any method used is likely to be imperfect, resulting in a selection that both excludes relevant documents and includes irrelevant documents. Consequently, conclusions drawn from the sub-network might be incorrect due to incomplete or incorrect data.

Secondly, the way in which the sub-graph is embedded in the larger network can provide valuable information. Suppose, for example, that one of the selected cases has many citations from cases outside the selection, but few from within the selection. Although the case clearly has some relevance (centrality), it receives a low centrality score in the sub-network. Of course, whether the case is relevant (central) is ultimately determined by the research question.

In the network for employer liability, we first selected a sub-network based on a network independent criterion, namely those cases that refer to statute articles related to employer liability. We will refer to these cases as the "seed cases". We then created a second network by extending the selection to include cases that cite the seed cases. These citing cases are expected to have some relevance related to the topic.

FIGURE 12.1: Network plots, created with case-law-app (DOI: 10.5281/zenodo.823668)

Note: Left: sub-network for employer liability; right: the sub-network, extended with cases citing the original sub-network.

The networks are plotted in Figure 12.1 and network statistics are shown in Figure 12.2. Note that the extended network is much denser, and indeed the average degree is higher in this network, especially in the later years. This can be explained by the fact that we extended the dataset specifically with cases that were citing one of the seed cases, which

gave them at least one outgoing citation by default. As a result, the "extended" cases have different properties than the original set of cases.

We also find that centrality measures are different for the extended sub-network. If, for example, we look at the "authorities" score (from the Hyperlink-Induced Topic Search (HITS) algorithm, Kleinberg 1999), the highest ranking case in the extended network is ECLI:NL:HR:2008:BD1847, with an authority score of 0.51. However, in the original, non-extended network, this node has an authority score very close to zero, even though in both networks it is the node with the largest in-degree. Apparently, the nodes that we added in the extended network often cite cases that cite ECLI:NL:HR:2008:BD1847, giving the node a different, more influential, role in the extended network. The substantive explanation is that the decision is on a topic not related to employer liability (yet still about article 7:611 Dutch Civil Code). The importance of the node is thus very different in the context of employer liability, than in a broader context, which coincides with the way the network was constructed. In other instances, a researcher may conclude that such a decision is an influential precedent if the sub-network is analyzed, but deemed irrelevant in the extended network. In summary, the method of selection should be chosen with care and match the research question.

The effects of selection can be measured, by comparing generic network statistics of the smaller network to the larger network. If large differences are found, it is worth investigating whether this is to be expected.

FIGURE 12.2: Network statistics of a sub-network selected on keyword and the same network, extended with all cases that cite one of the cases in the original sub-network
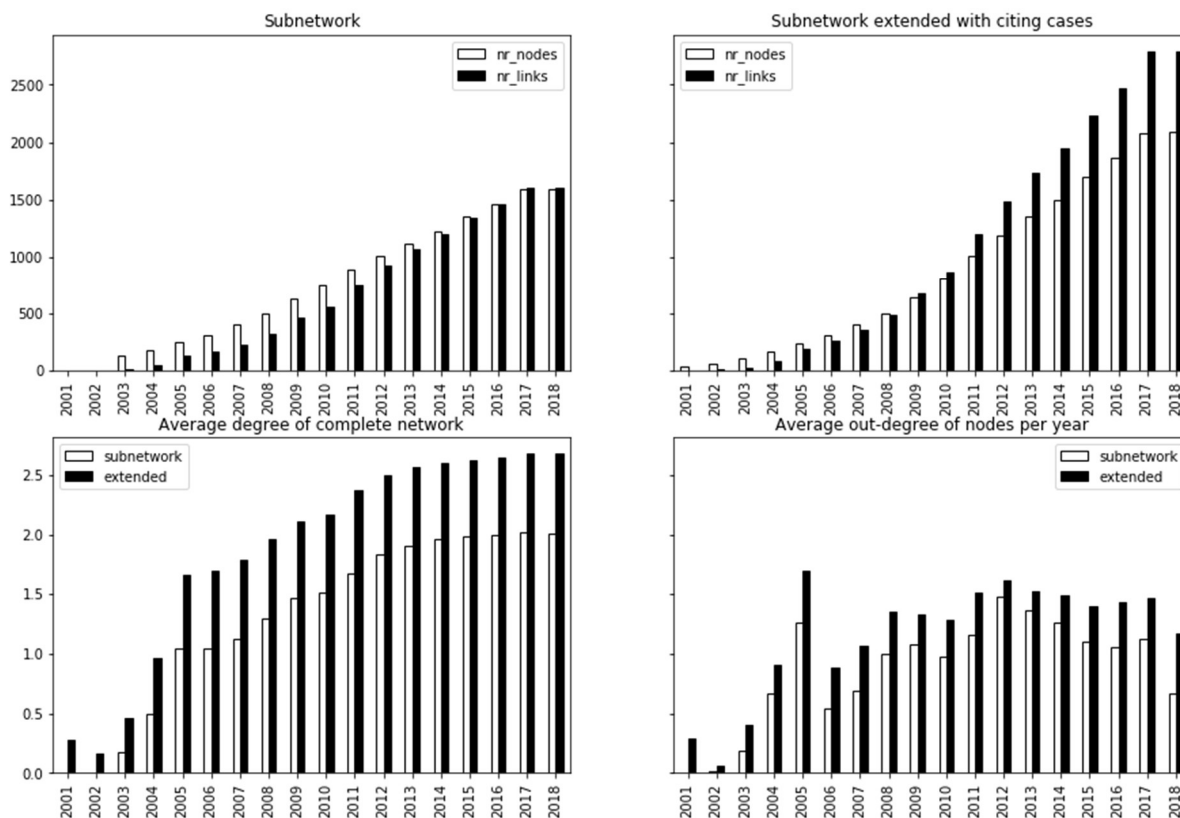
Decisions on network selection thus have substantial implications on the results of LNA, leading to differences in network structure and centrality scores. We therefore propose more extensive studies to the effects of sub-network selection in LNA, which to our knowledge have not been researched. Experiments need to be conducted in order to analyze the implications and to develop best practices as to how to select networks when one is interested in a sub-graph. Similar research has been done for social networks, where the effects of choice "network boundary" on centrality have been investigated (Laumann et al., 1983; Valente et al., 2013). A systematic approach in which network statistics for several sub-networks are calculated, can shed further light on sensible choices of selection.

4.2 Challenges of Community Detection Algorithms

Community detection methods provide a clustering of the nodes based on the network structure, giving an overview of the organization and meso-scale structure of the network. Each node in the network gets assigned to a group, so that within-group nodes have relatively many connections between them. In LNA, community detection can be used to

find structure in the legal domain, for example by defining legal sub-areas. It allows exploring what cases surround a certain precedent or landmark case, visualizing clusters of case law on similar topics, and testing whether certain decisions were cited more frequently in certain periods of time compared to other decisions.

There are many methods available for community detection, and there is an ongoing debate about the usefulness of the different methods and how they should be evaluated. In LNA, as well as in other domains, it is difficult to define a criterion that can serve as a reference point for how well a particular community detection performs. One strategy is to use metadata as reference point, for example metadata about the field the decision belongs to, but unfortunately it is very well possible that even high-quality metadata does not align well with the community structure. Moreover, the definition of what constitutes a "community" dictates how well a given algorithm performs, but "community structure" generally has no objective definition. This makes it difficult to distinguish whether the algorithm performed poorly, or whether the algorithm found a structure that is valid but fundamentally different from, for example, the traditional division made in terms of legal disciplines, topics, sub-topics, and so on. Nevertheless, comparing to metadata can be useful, particularly if the objective of community detection is to find a partition that is similar to the division into metadata categories.

Another issue with many community detection algorithms is the number of communities and sizes of the communities. Popular algorithms based on modularity optimization suffer from a resolution limit: the method prefers communities of sizes proportional to the square root of the number of nodes (Fortunato and Barthelemy, 2007). Consequently, the sizes of the communities depend on the number of nodes, with the expected community size being larger for larger networks.

As a result of resolution limits, the sizes of the communities that result from the algorithm do not always match the desired resolution, as small communities get merged into large ones when the network size increases. Several methods have been proposed to overcome this issue, including a resolution parameter in the equation that allows the user to control the sizes of communities (Reichardt and Bornholdt, 2006). A larger value for the resolution parameter results in larger, but fewer communities. Figure 12.3 shows network

communities for two different resolution parameter values. It can be useful to take into account multiple resolution communities and to "zoom in" until the desired resolution is obtained. However, with no reference point available, it is difficult to say what the desired resolution of communities is.

FIGURE 12.3: The largest component of a network of cases about employer liability
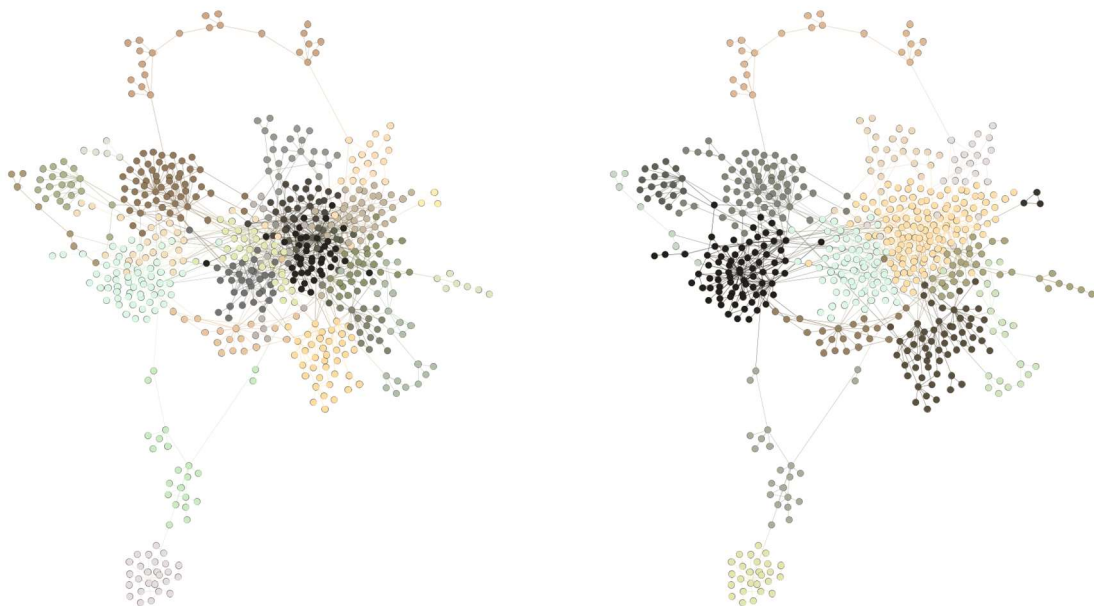
Note: The nodes are shaded based on the partition in communities with modularity optimization for different resolutions: left: resolution 0.5 (24 communities); right: resolution 1.0 (17 communities).

In addition to decisions about resolution size, decisions need to be made regarding the method of community detection. Each method has different designs and assumptions (Ghasemian et al., 2018; Schaub et al., 2017). It is therefore necessary to consider what assumptions are valid for legal citation networks. Directed links and ordering of nodes in time are not taken into account by many methods, whereas this is crucial in networks on court decisions, as citations by courts are only backward in time (i.e. to decisions that were already made, not to future decisions). This stresses the importance to consider what assumptions are valid for legal citation networks. Bommarito II et al. have signaled these

shortcomings and introduced an alternative clustering method aimed specifically at citation networks (Bommarito et al., 2010). Based on a qualitative evaluation, they claim it results in an informative clustering of US Supreme Court cases.

Additional methods have been specifically designed for citation networks, which consider the growth of the network over time. In Leicht et al. (2007), a method is proposed to cluster nodes based on in which years they either cite or are cited, to distinguish "eras" in the data. They apply the method on US Supreme Court cases.

We illustrate the difficulty of algorithm selection by comparing two methods of community detection, namely the widely applied Louvain method (Blondel et al., 2008) and the specialized method of Bommarito et al., which we refer to as the citation distance method. We apply both of these methods to the network of employer liability cases and we perform a qualitative assessment of the resulting clusters. We run the Louvain method with different resolution parameters. The citation distance method results in a hierarchical clustering, and we choose different cut-off values to obtain several resolutions of clusters, where a higher cut-off threshold corresponds to fewer communities.

First, we observe the number of communities that are produced by both methods for the employer liability network, which are plotted in Figure 12.4. The number of communities for the citation distance is shown, for different threshold values. The higher the chosen threshold value, the lower the number of communities, but the minimum number of communities for this network (threshold value=1) is 45. In addition, for threshold values higher than 0.8, more than half of the nodes are put in the largest community. In contrast, the number of communities for the different resolution values in the Louvain method is much lower.

FIGURE 12.4: The number of communities (left) and the fraction of nodes that is in the largest community (right) for different thresholds in the citation distance clustering method, and different resolutions in the Louvain method for modularity optimization
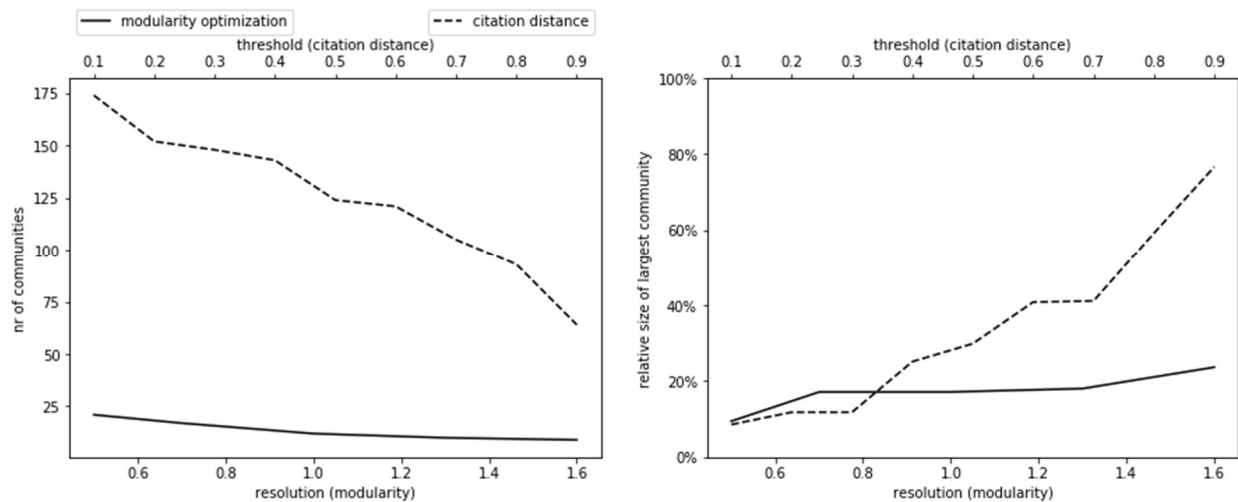
For the topic of employer liability, the low number of communities is closer to what is to be expected. In legal literature, between approximately 7 to 20 sub-topics can be distinguished on the topic of employer liability, depending on the textbook that is used as a reference point (van Dijck and van Kuppevelt, 2018). The substantially higher number of clusters in the citation distance method does not align with what would be expected from a substantive perspective.

We further inspect the individual communities to add meaning to the clustering. For this, we only select communities that contain Dutch Supreme Court decisions that have been cited at least ten times in order for the selection to include communities with at least one substantially relevant decision. The number of decisions that the selection resulted in was sufficiently small so that a human expert could inspect and categorize the decisions in the networks in order to determine precision levels. Accuracy and recall were beyond the feasible scope of this chapter, as they would require the expert to categorize a large number of decisions.

For each of the clusters in the complete network, we looked at the Supreme Court decisions in that cluster and manually classified the number of decisions that belonged to either the cluster of (a) duty to insure decisions and/or traffic accidents or (b) causality. The topics chosen resemble important topics in the field of employer liability in the Netherlands. The

results of the number of clusters and number of decisions per cluster for each topic is shown in Table 12.1.

Table 12.1: Results of community detection for employer liability network

Approach	Resolution	Topic	#clusters	True positives	True positives + false positives	Precision (%)
Citation distance	0.2	Causality	2	4	4	100
		Duty to insure/Traffic	2	9	14	64
	0.4	Causality	3	5	25	20
		Duty to insure/Traffic	1	11	21	52
	0.6	Causality	2	7	37	19
		Duty to insure/Traffic	2	11	38	29
	0.8	Causality	1	7	46	15
		Duty to insure/Traffic	1	10	46	22
Louvain	0.5	Causality	2	9	13	69
		Duty to insure/Traffic	2	7	10	70
	0.7	Causality	1	10	13	77
		Duty to insure/Traffic	2	8	12	67
	1.0	Causality	1	10	13	77

		Duty to insure/Traffic	2	8	13	62
	1.3	Causality	1	10	13	77
		Duty to insure/Traffic	2	10	19	53

The comparison reveals that, overall, the Louvain method of community detection results in higher percentages of positive observations (TP) compared to all observations in the selected clusters (TP+TN) (precision) than does the citation distance approach. For most of the resolutions, the citation distance approach produces a lower proportion of "on-topic" decisions in clusters than the Louvain method. Moreover, when the resolution size increases in the citation distance approach, it results in a limited number of large clusters that include a variety of decisions that are substantively different. Compared to the Louvain method, the citation distance approach often either overlooks relevant decisions in the networks that were analyzed, or it clusters together too many decisions that are substantively different.

Only the 0.2 resolution of the citation distance approach produced high precision levels, but this resolution also resulted in numerous small clusters that are not substantially meaningful. As to the Louvain method, the 0.7 and 0.5 resolution resulted in the highest precision levels, grouping together relevant decisions while distinguishing clusters that are substantively different. Nevertheless, higher resolutions (e.g. 1.3 or 1.6) with the Louvain method also result in clusters that include a variety of substantively different decisions, although not to the extent the citation distance approach does. This observation suggests the citation distance approach is more sensitive to resolution changes than is the Louvain method, at least in the network that was analyzed.

We suspected that the relatively small network size might explain why the Louvain method produced more substantively meaningful results than did the citation distance method – the latter produces a minimum of 40 communities, whereas the literature on employer liability indicates no more than approximately 20 communities – we subsequently turned to a larger and denser dataset. We constructed a citation network between 1233 rulings decided by

the CJEU on the subject matter "(FISC) Taxation" before 1 January 2018.[2] We again applied the Louvain and citation distance methods to the giant component of the network (1143 nodes, 5284 edges) at different resolution and threshold values. As an indication of the quality of the communities, we compared with the topic of each ruling, extracted from the metadata provided by EUR-Lex. EUR-Lex distinguishes between four topics in the subject matter Taxation (FISC), namely: Excise duties (ACCI), Indirect taxation (INDR), Internal taxation (INTE), Value added tax (TVA). We could not assign 193 rulings to any topic, whereas 922 rulings were assigned to one topic, 28 to two topics, and 1 to three. Most rulings were assigned to topic TVA, the fewest to ACCI. A meaningful community detection method would identify communities around these four topics, considering that they represent different taxes that relate to different EU legal acts, each with their own rules.

The citation distance method did not provide such meaningful results. Figure 12.5 gives the number of communities detected and the relative size of the largest community at different thresholds. At threshold 0.1 the results were very scattered: we found as many as 272 communities, but only 8 comprised ten or more rulings. A threshold value of 0.5 resulted in clearly fewer communities (82), but already placed 87 percent of all nodes in the same community. That largest community also comprised the largest proportion of cases on each of the four topics (TVA: 92 percent, ACCI: 92 percent, INDR: 71 percent, INTE: 71 percent). For higher thresholds (0.6–0.9), the proportions increased. Given the lack of proper results, we do not present the results of the citation distance method as we did for the Louvain method.

FIGURE 12.5: Communities in EU tax network: the number of communities (left) and the fraction of nodes that is in the largest community (right) for different thresholds in the citation distance clustering method, and different resolutions in the Louvain method for modularity optimization

---

[2] The data was gathered on 1 January 2018 from EUR-Lex: https://eur-lex.europa.eu.
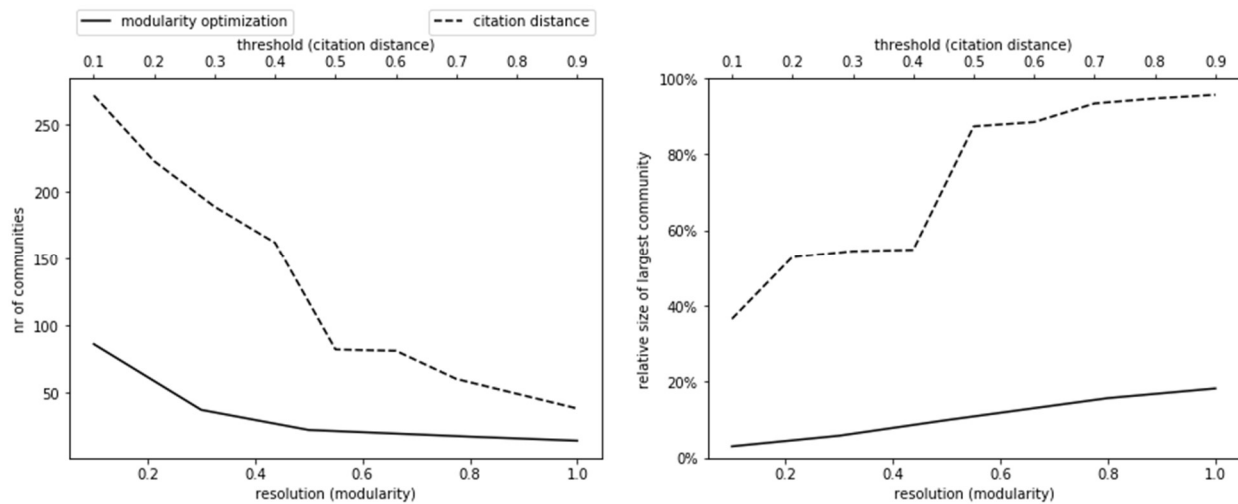
For the Louvain method, the results are different. It identified fewer communities than the citation distance method and the relative size of the largest community was smaller (see also Figure 12.5. At the lowest resolution that we tested (0.1), 56 of 86 communities had ten or more rulings (relative size of 1 percent or more). A resolution of 0.5 considerably reduced the number of communities (from 86 to 22) and increased the relative size of the largest community (from 3 percent to 10 percent) and the smallest community (relative size of 2 percent at 20 rulings). At resolution of 1.0, the smallest of the 14 communities comprised 24 rulings (relative size 2 percent) and the largest community had a relative size of 18 percent.

We analyzed how the Louvain method performed at these three resolutions. Table 12.2 illustrates in detail how the four topics were distributed over the 14 communities of Louvain for resolution 1.0. The distribution is very clear with regard to rulings on INTE (N=137) and INDR (N=68): 94 percent (129/137) of INTE rulings were in community 13 and 96 percent (65/68) INDR rulings were in community 2. The fact that TVA rulings can be found in several communities is consistent with the fact there are many sub-topics (e.g. exemptions, deductions, and so on) within Value added tax. Furthermore, in 11 of 14 communities, the proportion of rulings belonging to a specific topic comprised 70 percent or more (100 percent in 4 communities) of the rulings in that community. Although a lower resolution of 0.5 distributes the topics over more communities, 20 of 22 communities still have a high proportion (70 percent or more) of rulings belonging to a specific topic (100 percent in 10 communities). There are more and smaller communities at resolution 0.1, but we still found that in 74 of 86 communities the proportion of rulings belonging to a specific topic

comprised 70 percent or more of the rulings in that community (100 percent in 60 communities). We therefore conclude that the Louvain method provides meaningful community structure to this second citation network, at least from a substantive perspective and in this network.

Table 12.2: Distribution of topics over Louvain communities on tax law network

Community	TVA	ACCI	INTE	INDR	Total
0	73	2	0	0	75
1	3	1	1	2	7
2	5	0	4	65	74
3	144	3	0	0	147
4	83	2	2	0	87
5	188	0	0	0	188
6	38	1	0	0	39
7	0	13	0	1	14
8	17	15	0	0	32
9	25	0	0	0	25
10	3	4	1	0	8
11	83	0	0	0	83
12	27	0	0	0	27
13	37	9	129	0	175
Total	726	50	137	68	

It is not surprising that the two methods result in observably different clustering. Both methods are based on fundamentally distinctive principles. The citation distance method assumes that the topic of a decision is merely based on the decisions that preceded it, and uses the notion of "sink nodes" (decisions without outgoing citations) to determine what topics exist, and then clusters decisions based on these topics. The Louvain method does not have such assumptions, but optimizes the *modularity* value, which compares the

number of links within the communities to the expected number of links, which does not take into account the direction of the links.

Although the inventors of the citation distance method (Bommarito et al., 2010) show the value of the method on US Supreme Court cases, its success there does not appear to extrapolate to our datasets. Future research should investigate the relationship between the performance of the different methods and network properties, such as size, density, and number of triangles, to provide insight into which situations require which method.

Another difficulty with community detection is the apparent instability of the methods. As most community detection methods provide no significance score for the found partition, it is often unclear to what extent the methods are sensitive to noise in the data and randomization effects. For modularity optimization, Good et al. (2010) have shown that due to randomization in the algorithm, multiple iterations result in many different solutions which are all almost equally good. Depending on the application, conclusions drawn based on community detection are prone to be inaccurate, when not considering this uncertainty. Stability over multiple iterations can be achieved with consensus clustering (Lancichinetti and Fortunato, 2012). In this method, the results of multiple iterations of regular community detection methods are combined to find a stable community structure.

Note that there is no "perfect" community detection method, and the choice for one should always depend on the application. Some methods, such as modularity optimization, tend to overfit, meaning that they find patterns in a sample of the data that do not generalize well to the complete dataset, whereas other methods give a too generic description of the network (Ghasemian et al., 2018). We cannot give a complete overview of community detection methods here, and we refer to Fortunato (2010) for a survey on the subject.

4.3 Testing for Significance

Since a "truth" is not always available to validate the methodologies used in LNA, other techniques are required to provide trust in the result of the analyses. Not only do we need to choose an approach that matches our research question, the result of the methods should also be statistically significant if used beyond exploration. Statistical significance

testing allows for estimating the probability that the network and its relationships as observed in the data occurred by chance rather than by some systematic process.

Statistical testing in networks is not straightforward, because nodes are by definition not independent from each other. In addition, we generally have only one observed network instead of a large sample. Nevertheless, several methods have been developed for significance testing in networks, with some being more suitable for legal citation networks than others.

In SNA, two popular models for statistical testing are Stochastic Oriented Actor Models (SOAMs) and Exponential Random Graph Models (ERGMs). Both methods can be used to obtain the importance of local structures in the complete network, possibly related to external node-level attributes. These models are used in SNA to statistically test, for example, whether girls prefer to be friends with girls, or if one is more likely to become friends with a friend of a friend. A legal example is whether certain courts favor the citation of certain cases over others.

SOAMs (Snijders, 2001) model the formation of local network structures, with the assumption that nodes are actors that deliberately form or destroy ties. As an implication, SOAMs are specifically designed for dynamic ties in a fixed set of nodes (actors). This is a valid assumption in social networks, but less applicable to citation networks, where the number of nodes grows but links never disappear. Adaptation of the models to growing networks are possible, but applications of SOAMs do not seem to be a proper fit to analyze citation networks.

ERGMs (Robins et al., 2007) are a specific type of random graph model, which consider not only the observed network, but also the complete set of possible networks with the same network statistics. The model holds a free parameter for each of the chosen network statistics, or configurations, such as the degree sequence or number of triads. The values of these parameters are chosen in such a way that the probability of the observed network is maximal. This gives insight into the importance of the different configurations for the network: for example, a high value for the parameter for number of triads suggests that triads form a significant structure in the network.

Networks of court decisions have the special property that they have a particular ordering in time, based on their publishing date. This makes them Directed Acyclic Graphs (DAGs), meaning that the links have a direction, and no cycles are present. Cycles (e.g. A cites B and B cites A) do not generally occur in case law, because decisions can only cite decisions in the past. The DAGs that have been studied most extensively are scientific citation networks, so we can use methodologies from bibliometrics and scientific networks for LNA.

A model to generate random networks for DAGs was proposed by Karrer and Newman (2009), who showed that networks generated by this model appear similar to real-life networks, including the network of US Supreme Court cases. However, applications of random graph models in citation networks remain sparse, possibly due to the fact that software packages do not contain models for DAGs. The applicability of these models for LNA should be investigated further.

Furthermore, many existing community detection algorithms do not provide a significance value for their results, but there are some exceptions. Peel et al. (2017) propose a method to retrieve the relationship between community structure and metadata, by comparing the entropy of the network under the model against networks where the data is randomly permuted.

Significance of communities can also be estimated by rewiring the edges of a network, or by comparing to a sample of similar networks. In Mirshahvalad et al. (2012) this approach is applied to the European Court of Justice citation network. The random network model of Karrer and Newman (2009) was used by Speidel et al. (2015) to formulate a variant of modularity for DAGs.

We note that much of the research of statistical methods on DAGs is rather theoretical, and established research methods similar to the SOAMs and ERGMs for SNA are not present yet for citation networks. We propose further research to connect these domains and provide researchers in LNA with out-of-the-box statistical tools for legal citation networks.

4.4 Centrality Measure Selection

Algorithm selection is not only relevant with respect to the selection of communities, but also regarding the measures used to determine the centrality of decisions. The many available centrality measures are based on different assumptions and will thus capture

other properties of the network structure. Authors have evaluated several centrality measures on legal citation networks (Carmichael et al., 2017; Derlén and Lindholm, 2014; Fowler et al., 2007; Van Opijnen, 2012). Some have introduced specific centrality scores for citation networks, such as relative in-degree (Tarissan and Nollez-Goldbach, 2015) and a log-transformed degree measure (Van Opijnen, 2012).

Nevertheless, no guidelines exist in LNA as to which algorithms should be applied in which situations. Algorithm selection is important, as the application of different algorithms yields different results. This is partly caused by differences in underlying assumptions and calculation methods of the algorithms – they simply measure and capture different information. In addition, and importantly, the composition of the network also affects the results. To illustrate, we turn to the example of employer liability in the Netherlands.

When analyzing employer liability networks, we focus on commonly used centrality measures: degree centrality, in-degree centrality, out-degree centrality, relative in-degree, PageRank, HITS (authorities and hubs), closeness centrality, and betweenness centrality. We start with the non-extended sub-network that includes decisions at all levels (i.e. Dutch Supreme Court, courts of appeal, courts of first instance). In this network, which includes 338 nodes and 549 edges, several measures correlate in a statistically significant way, as is shown in Table 12.3. Out-degree centrality (6), in-degree centrality (5), PageRank (5), closeness centrality (5), and degree centrality (4) are related to various other measures (the number of statistically significant relationships reported between parentheses), with substantial variation in effect sizes across all measures.

<PLEASE INSERT TABLE 12.3 ABOUT HERE>

The correlations change when the network changes. Taking a smaller and less dense sub-sample of the network that counts 40 nodes and 40 edges between decisions from and to the Dutch Supreme Court, correlations between other measures emerge, as we see in Table 12.4. In contrast to the previous network, betweenness centrality (6), authorities (5), and hubs (4) are now also positively associated with various other measures.

<PLEASE INSERT TABLE 12.4 ABOUT HERE>

The results demonstrate that the outcomes that the centrality measures produce can be different for different networks. Consequently, in the practice of doing research, researchers may want to (1) inspect and even report the results of several network analysis measures and (2) compare these results to the ones found in the literature in order to provide a reference point to compare the results to. Assuming that centrality resembles relevance of precedents, at least to an extent, significant deviations from legal commentaries may either raise additional research questions (e.g. did legal scholarship overlook important decisions?) or indicate that a sub-optimal or even wrong centrality measure is chosen.

Table 12.3: Correlations between centrality measures (complete network)

Variables	authorities	betweenness_centrality	closeness_centrality	degree_centrality	hubs	in_degree_centrality	out_degree_centrality
authorities	**1.00**	−0.01	**−0.16**	−0.03	−0.03	0.02	**−0.15**
betweenness_centrality	−0.01	**1.00**	0.06	0.02	0.00	0.01	0.05
closeness_centrality	**−0.16**	0.06	**1.00**	0.10	0.05	**−0.20**	**0.96**
degree_centrality	−0.03	0.02	0.10	**1.00**	−0.07	**0.95**	**0.13**
hubs	−0.03	0.00	0.05	−0.07	**1.00**	−0.06	−0.01
in_degree_centrality	0.02	0.01	**−0.20**	**0.95**	−0.06	**1.00**	**−0.19**
out_degree_centrality	**−0.15**	0.05	**0.96**	**0.13**	−0.01	**−0.19**	**1.00**
pagerank	0.10	0.00	**−0.22**	**0.78**	−0.06	**0.84**	**−0.21**
rel_in_degree	0.01	0.01	**−0.15**	**0.90**	−0.06	**0.94**	**−0.15**

Note: Significant correlations (p < 0.05) in bold.

Table 12.4: Correlations between centrality measures (sampled network)

Variables	authorities	betweenness_centrality	closeness_centrality	degree_centrality	hubs	in_degree_centrality	out_degree_centrality
authorities	**1.00**	**0.60**	0.11	**0.67**	0.27	**0.70**	0.11
betweenness_centrality	**0.60**	**1.00**	0.25	**0.69**	**0.35**	**0.55**	0.29
closeness_centrality	0.11	0.25	**1.00**	**0.58**	**0.74**	−0.26	**0.98**
degree_centrality	**0.67**	**0.69**	**0.58**	1.00	**0.65**	**0.62**	**0.60**
hubs	0.27	**0.35**	**0.74**	**0.65**	**1.00**	0.07	**0.72**
in_degree_centrality	**0.70**	**0.55**	−0.26	**0.62**	0.07	**1.00**	−0.27
out_degree_centrality	0.11	0.29	**0.98**	**0.60**	**0.72**	−0.27	**1.00**
pagerank	**0.38**	**0.36**	**−0.37**	**0.38**	−0.12	**0.81**	**−0.37**
rel_in_degree	**0.45**	**0.39**	−0.20	**0.37**	0.14	**0.66**	−0.23

5. CONCLUSION

As we described in the previous sections, additional research is required to better understand the applicability of different network methods for the legal domain. It seems that different legal domains vary greatly in the type of networks, and consequently in the network methods and algorithms that are applicable.

An additional challenge in comparing methods and legal domains is the availability of data and software. The accessibility of citation data differs greatly for legal systems, making standardized comparisons difficult. In addition, it is not yet common practice among computational legal researchers to publish code and datasets used, along with written articles. We strongly advocate these "open science" practices, since it not only improves reproducibility, but it also allows for larger comparative studies over datasets and methods.

REFERENCES

Agnoloni, T., Pagallo, U., 2015. The case law of the Italian constitutional court, its power laws, and the web of scholarly opinions, in: *Proceedings of the 15th International Conference on Artificial Intelligence and Law: ICAIL '15*. ACM Press, New York, pp. 151–5. https://doi.org/10.1145/2746090.2746108.

Beaverstock, J.V., 2004. "Managing across borders": Knowledge management and expatriation in professional service legal firms. J. Econ. Geogr. 4, 157–79.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008, P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Bommarito II, M.J., Katz, D.M., 2010. A mathematical approach to the study of the United States Code. Phys. Stat. Mech. Its Appl. 389(19), 4195–200.

Bommarito II, M.J., Katz, D.M., Zelner, J.L., 2010. On the stability of community detection algorithms for longitudinal citation data. Procedia: Soc. Behav. Sci. 4, 26–37. https://doi.org/10.1016/j.sbspro.2010.07.480.

Bommarito, M.J., Katz, D.M., Zelner, J.L., Fowler, J.H., 2010. Distance measures for dynamic citation networks. Phys. Stat. Mech. Its Appl. 389, 4201–8. https://doi.org/10.1016/j.physa.2010.06.003.

Boulet, R., Mazzega, P., Bourcier, D., 2011. A network approach to the French system of legal codes-part I: Analysis of a dense network. Artif. Intell. Law 19, 333–55. https://doi.org/10.1007/s10506-011-9116-1.

Carmichael, I., Wudel, J., Kim, M., Jushchuk, J., 2017. Examining the evolution of legal precedent through citation network analysis. N. C. Rev. 96(1), 227-269.

Derlén, M., Lindholm, J., 2014. Goodbye van Gend en Loos, hello Bosman? Using network analysis to measure the importance of individual CJEU judgments. Eur. Law J. 20, 667–87. https://doi.org/10.1111/eulj.12077.

Derlén, M., Lindholm, J., 2015. Characteristics of precedent: The case law of the European Court of Justice in three dimensions. Ger. Law J. 16(5), 1073–98.

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486, 75–174. https://doi.org/10.1016/j.physrep.2009.11.002.

Fortunato, S., Barthelemy, M., 2007. Resolution limit in community detection. Proc. Natl. Acad. Sci. 104, 36–41. https://doi.org/10.1073/pnas.0605965104.

Fowler, J.H., Jeon, S., 2008. The authority of Supreme Court precedent. Soc. Netw. 30, 16–30. https://doi.org/10.1016/j.socnet.2007.05.001.

Fowler, J.H., Johnson, T.R., Spriggs, J.F., Jeon, S., Wahlbeck, P.J., 2007. Network analysis and the law: Measuring the legal importance of precedents at the U.S. Supreme Court. Polit. Anal. 15, 324–46. https://doi.org/10.1093/pan/mpm011.

Ghasemian, A., Hosseinmardi, H., Clauset, A., 2018. Evaluating Overfit and Underfit in Models of Network Community Structure, arXiv:1802.10582 [stat.ML].

Gionfriddo, J.K., 2007. Thinking like a lawyer: The heuristics of case synthesis. Tex. Tech Law Rev. 40, 1–36.

Good, B.H., de Montjoye, Y.-A., Clauset, A., 2010. Performance of modularity maximization in practical contexts. Phys. Rev. E 81, 046106. https://doi.org/10.1103/PhysRevE.81.046106.

Heinz, J.P., Laumann, E.O., 1982. *Chicago Lawyers: The Social Structure of the Bar*. Russell Sage Foundation.

Hutchinson, T., 2010. *Researching and Writing in Law*, 3rd ed. Thomson Reuters, London.

Karrer, B., Newman, M.E.J., 2009. Random acyclic networks. Phys. Rev. Lett. 102, arXiv:0902.4013v1. https://doi.org/10.1103/PhysRevLett.102.128701.

Katz, D.M., Bommarito, M.J., 2014. Measuring the complexity of the law: The United States Code. Artif. Intell. Law 22, 337–74. https://doi.org/10.1007/s10506-014-9160-8.

Kleinberg, J.M., 1999. Authoritative sources in a hyperlinked environment. J. ACM 46, 604–32. https://doi.org/10.1145/324133.324140.

Knops, M., Schaper, M., 2014. Holding complexity: Analysing the CJEU's VAT case law as a network. World J. VAT/GST Law 3(3), 141–65.

Koniaris, M., Anagnostopoulos, I., Vassiliou, Y., 2017. Network analysis in the legal domain: A complex model for European Union legal sources. J. Complex Netw. 1–29. ArXiv, abs/1501.05237. https://doi.org/10.1093/comnet/cnx029.

Lancichinetti, A., Fortunato, S., 2012. Consensus clustering in complex networks. Sci. Rep. 2, 336. https://doi.org/10.1038/srep00336.

Laumann, E.O., Marsden, P.V., Prensky, D., 1983. The boundary specification problem in network analysis. Appl. Netw. Sci 61, 18–34. https://doi.org/citeulike-article-id:430770.

Lazega, E., 2001. *The Collegial Phenomenon: The Social Mechanisms of Cooperation among Peers in a Corporate Law Partnership*. OUP, Oxford.

Lazega, E., Krackhardt, D., 2000. Spreading and shifting costs of lateral control among peers: A structural analysis at the individual level. Qual. Quant. 34, 153–75.

Lazega, E., van Duijn, M., 1997. Position in formal structure, personal characteristics and choices of advisors in a law firm: A logistic regression model for dyadic network data. Soc. Netw. 19(4), 375–97.

Leicht, E.A., Clarkson, G., Shedden, K., Newman, M.E.J., 2007. Large-scale structure of time evolving citation networks. Eur. Phys. J. B 59, 75–83. https://doi.org/10.1140/epjb/e2007-00271-7.

Lupu, Y., Voeten, E., 2012. Precedent in international courts: A network analysis of case citations by the European Court of Human Rights. Br. J. Polit. Sci. 42, 413–39. https://doi.org/10.1017/S0007123411000433.

Mirshahvalad, A., Lindholm, J., Derlén, M., Rosvall, M., 2012. Significant communities in large sparse networks. PloS One 7, e33721. https://doi.org/10.1371/journal.pone.0033721.

Neale, T., 2013. Citation analysis of Canadian case law. J. Open Access Law 1(1).

Nievelstein, F., Van Gog, T., Van Dijck, G., Boshuizen, H.P.A., 2013. The worked example and expertise reversal effect in less structured tasks: Learning to reason about legal cases. Contemp. Educ. Psychol. 38(2), 118–25.

Olsen, H.P., 2020. Needles in a haystack: Using network analysis to identify cases that are cited for general principles of law by the European Court of Human Rights, in:

Whalen, R. (ed.), Computational Legal Studies: The Promise and Challenge of Data-Driven Research. Edward Elgar. Cheltenham, UK / Northampton, MA, USA.

Olsen, H.P., Küçüksu, A., 2017. Finding hidden patterns in ECtHR's case law: On how citation network analysis can improve our knowledge of ECtHR's Article 14 practice. Int. J. Discrim. Law 17(1), 4-22.

Olsen, H.P., Sadl, U., 2017. Can quantitative methods complement doctrinal legal studies? Using citation network and corpus linguistic analysis to understand international courts. Leiden J. Int. Law 30(2), 327-349.

Peel, L., Larremore, D.B., Clauset, A., 2017. The ground truth about metadata and community detection in networks. Sci. Adv. 3, e1602548. https://doi.org/10.1126/sciadv.1602548.

Reichardt, J., Bornholdt, S., 2006. Statistical mechanics of community detection. Phys. Rev. E 74, 016110. https://doi.org/10.1103/PhysRevE.74.016110.

Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007. An introduction to exponential random graph (p *) models for social networks. Soc. Netw. 29, 173–91. https://doi.org/10.1016/j.socnet.2006.08.002.

Schaper, M.G.H., 2014. A computational legal analysis of acte clair rules of EU law in the field of direct taxes. World Tax J. 6(1), 77-108.

Schaub, M.T., Delvenne, J.-C., Rosvall, M., Lambiotte, R., 2017. The many facets of community detection in complex networks. Appl. Netw. Sci. 2, 4. https://doi.org/10.1007/s41109-017-0023-6.

Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. Sociol. Methodol. 31, 361–95. https://doi.org/10.1111/0081-1750.00099.

Speidel, L., Takaguchi, T., Masuda, N., 2015. Community detection in directed acyclic graphs. Eur. Phys. J. B 88, 203. https://doi.org/10.1140/epjb/e2015-60226-y.

Tarissan, F., Nollez-Goldbach, R., 2015. Temporal properties of legal decision networks: A case study from the International Criminal Court. Leg. Knowl. Inf. Syst. 111–20. https://doi.org/10.3233/978-1-61499-609-5-111.

Tarissan, F., Nollez-Goldbach, R., 2016. Analysing the first case of the International Criminal Court from a network-science perspective. J. Complex Netw 4(4), 616-634. https://doi.org/10.1093/comnet/cnw002.

Tarissan, F., Panagis, Y., Sadl, U., 2016. Selecting the cases that defined Europe: Complementary metrics for a network analysis. Proc. 2016 IEEEACM Int. Conf. Adv. Soc. Netw. Anal. Min. ASONAM 2016 661–8. https://doi.org/10.1109/ASONAM.2016.7752308.

Valente, T.W., Fujimoto, K., Unger, J.B., Soto, D.W., Meeker, D., 2013. Variations in network boundary and type: A study of adolescent peer influences. Soc. Netw. 35, 309–16. https://doi.org/10.1016/j.socnet.2013.02.008.

van Dijck, G., 2017. *MoneyLaw (and Beyond)*. Eleven International Publishing, The Hague.

van Dijck, G., van Kuppevelt, D., 2018. Deelonderwerpen en precedenten in het kader van werkgeversaansprakelijkheid ex art. 7:658/7:611 BW: een netwerkanalyse. Ned. Juristenblad.

Van Opijnen, M., 2012. Citation analysis and beyond: In search of indicators measuring case law importance, in: Frontiers in Artificial Intelligence and Applications. 250. pp. 95–104. https://doi.org/10.3233/978-1-61499-167-0-95.

Whalen, R., 2013. Modeling annual Supreme Court influence: The role of citation practices and judicial tenure in determining precedent network growth, in: Menezes, R., Evsukoff, A., González, M.C. (eds), *Complex Networks: Studies in Computational Intelligence*. Springer, Berlin, Heidelberg, pp. 169-176.

Whalen, R., 2014. Bad law before it goes bad: Citation networks and the life cycle of

overruled Supreme Court precedent, in: Winkels, R., Lettieri, N., Faro, S. (eds),

Network Analysis in Law. Edizioni scientifiche italiane, Napoli.

Whalen, R., 2016. Legal networks: The promises and challenges of legal network analysis.

Mich. State Law Rev., 539-565.