



UvA-DARE (Digital Academic Repository)

Determining authority of Dutch case law

Winkels, R.; de Ruyter, J.; Kroese, H.

Published in:
Frontiers in Artificial Intelligence and Applications

DOI:
[10.3233/978-1-60750-981-3-103](https://doi.org/10.3233/978-1-60750-981-3-103)

[Link to publication](#)

Citation for published version (APA):
Winkels, R., de Ruyter, J., & Kroese, H. (2011). Determining authority of Dutch case law. *Frontiers in Artificial Intelligence and Applications*, 235, 103-112. <https://doi.org/10.3233/978-1-60750-981-3-103>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Determining Authority of Dutch Case Law

Radboud WINKELS^{a,1}, Jelle de RUYTER^a and Henryk KROESE^b
^a*Leibniz Center for Law, University of Amsterdam*
^b*Science Faculty, University of Amsterdam*

Abstract. In this paper we present the results of two studies to see whether the analysis of the network of citations between cases can be used as an indication of the relevance and authority in the Dutch legal system. Fowler e.a. have shown such results for the US common law system, but given the different status of case law in continental tradition it is not clear whether this will hold in the Netherlands. Moreover, we introduce a way to validate the results using selections made by human experts for legal education. We discuss the results and conclude that network analysis of cases is a useful tool for legal research.

Keywords. Network analysis, reference parsing, relevance, case law, citations

1. Introduction

"Law is a seamless system with its own autonomy. It provides one correct answer to any case, difficult or not, by application of its rules, precedents, principles and spirit." [1]. In Dworkin's theory, a judge finds *one* solution in each case, however difficult it is. He starts looking for the solution with the best 'fit', searching for rules, cases and other sources of law that are most 'on point'.

With the growing number of published sources of law, this task is becoming increasingly difficult for legal practitioners. What are the relevant sources of law for a particular case? And which ones are the most important ones?

In this paper we discuss the result of two studies we conducted on Dutch case law to see whether we can aid in determining the importance of a case by analysing its place in the network of sources of law. Network analysis has already been used to determine importance of scientific publications [2], patent requests [5] and to get an idea of the complexity of French legal code [7]. Fowler e.a. [3, 4] showed that network analysis can be used to get a handle on the authority of US Supreme Court judgements. Given the different status of case law in our continental law tradition, we want to investigate whether this also holds for Dutch case law.

Not all cases are being published², courts and commercial publishers make selections and the importance of the ruling will play a role in these decisions, but not all published cases are of equal importance. Our hypothesis is – not surprisingly – that

¹ Corresponding author: Radboud Winkels, Leibniz Center for Law, University of Amsterdam, PO Box 1030, 1000 BA Amsterdam, Netherlands; Email: winkels@uva.nl

² In 2004 not even 1% of all cases were being published in the Netherlands on the Dutch official portal (rechtspraak.nl) [8].

people refer more to important cases than to other ones. In order to test this hypothesis we need some other way to decide upon the importance of cases. It seems plausible to ask legal experts what they consider to be important ones.

We conducted two studies on two different data sets: The first study uses only Dutch Supreme Court case data of a legal publisher that uses its own system for identifying and referring to case law – the so called NJ-number. The second study uses cases from all kinds of courts published on the official Dutch portal (rechtspraak.nl). In the next sections we will discuss and compare the two experiments and draw some conclusions.

2. The First Study

For the first study we limit ourselves to cases from the Dutch Supreme Court (“Hoge Raad”) as published in the Dutch periodical NJ (“Nederlandse Jurisprudentie”) between January 1st 1965 and December 31st 2008. In total this collection contains 15,053 Supreme Court cases. References to other cases in these decisions are unfortunately not marked or otherwise machine readable. A first step is to detect these references automatically and build up the network of citations.

2.1. Building the Citation Network

There are several places in a Supreme Court decision where references to other cases can be found. First, the history of the case in earlier instances is summarized, the attorneys or the public prosecutor plead their case, then the court explains its decision, the Attorney General sometimes gives his opinion, the decision may be annotated and finally the editorial board of the publisher may add links to other cases deemed relevant. Most of these parts of a decision can be detected quite easily by their header, but surprisingly the actual opinion of the court is sometimes harder to detect. Often it starts with the text: “The Supreme Court, etc.”, and ends where the opinion of the Attorney General starts, signalled by the text “[OPINION]”, or the annotations or the editorial additions start, or where the complete text ends.

Table 1: Validation of the approach for finding references

Number of Cases	100		
Correct	862	F1	98,7 %
False Positives	19	Precision	97,84 %
False Negatives	4	Recall	99,54 %

In earlier research we have built a parser to detect references in legislation using a context free grammar approach [6]. The present corpus is simpler since the publisher uses a more or less standard way to identify and refer to cases in their own database by a so-called NJ-number³: The letters “NJ” followed by one or more spaces, the year of

³ This excluded potential references to cases not published in NJ.

publication, followed most of the time by a comma, backslash or forward-slash and then a follow-up number. We used regular expressions to find these references.

We found 106,559 references this way. To validate the approach we examined a random sample of 100 cases by hand. **Table 1** presents the results; an F1 score of almost 99% is good, especially recall is very good. Since the database only contains Supreme Court decisions, only references to Supreme Court decisions can be resolved.

2.2. A First Analysis of the Network

A network of case law is different from most other ones like social or computer networks; it evolves rather simply. Prior cases do not disappear, though they may be cited less and less, links between cases remain and a new case can only cite older cases.⁴ **Figure 1** below shows the average number of in- and outgoing references used by the Supreme Court itself per year. The first observation is that this average is low, about 1 till the 1980s and increasing afterwards to about 2.⁵ Not surprisingly the average number of incoming references decreases at the end, because there are fewer years from which these citations could come.

The Supreme Court also does not refer much compared to other parties. **Table 2** shows these figures. The Advocate General uses the most citations (40%), followed by the editorial board of the publisher. This last category is added later on and may grow over time and *can* refer to future cases. This category will be ignored in the remainder of this paper, as will the Introduction and the ‘Essence’ or summary of the decision that is also written by the editorial board.

2.3. The Top-10 Cases in the Network

Based on the number of incoming references we can determine a top-10 of cases. **Table 3** gives an overview of these cases. It does not display the number of times the case was cited, but the relative ranking for a particular actor or party. The overall ranking is based on the average ranking over the four actors: Supreme Court, Advocate General, Annotator and Other parties. So the first place is taken by the “Zwolsman” case, a decision concerning police investigation methods in criminal procedures. These methods led to the so-called “IRT affair”, a parliamentary investigation and finally an addition to the Dutch penal code for special police investigation competences (title IV added in 1999). Since that code became effective, the number of citations of the Zwolsman case also drops as can be seen in **Figure 2**. The number of citations peaks in 1999 and drops after that, especially for the court itself and the annotators.

In the Zwolsman case and the Heesch/Van den Akker case (nr 3), there was no applicable law and the Supreme Court forms law itself. In the Haviltex case (nr 2) there is an applicable article of law (art. 1378 of the old Dutch Civil code), but the court interprets it differently from before.

⁴ Not completely true, the publisher may later on add links to newer cases.

⁵ If we compare this to the Supreme Court of the US [3] we notice that they refer about ten times as much.

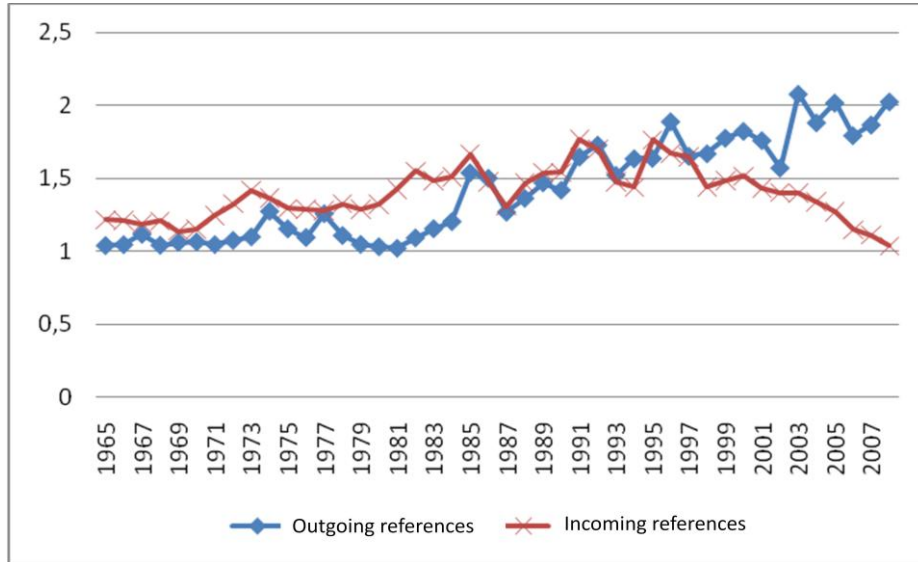


Figure 1: Average in- and outgoing references of the Supreme Court per year

Table 2: Number of references per section of decision

Position in the decision	References	%
Introduction	1.440	1,4%
Essence	2.528	2,4%
Prior instances	5.930	5,6%
Supreme Court	7.992	7,5%
Advocate General	42.858	40,2%
Annotations	10.246	9,6%
Editorial links	35.565	33,4%
Total	106.559	

Most of the cases in the top-10 are about procedural law, except for the cases Boon/Van Loon (nr 4) and case nr 10, which has no name. Both these cases deal with family law issues and in both cases the court breaches a line of thinking from the past. They are pointed out in **Figure 3** by arrows in a network that shows two clusters around them. The cases that connect the two, that cite either Boon/Van Loon or case nr 10, are all cases in which the court also changes a line of reasoning. Especially case NJ 1980/353, the so-called bull-calf case in which a calf injured a farmer that crossed its meadow, in which the court moves from a ‘suspicion of guilt’ to a ‘risk liability’ line of reasoning. The Supreme Court itself only refers once to the bull-calf case (in this dataset), in another case about risk liability (NJ 1984/2), but the attorney general cites the case 19 times and the annotators 3 times, mostly to point to the breach in reasoning and not to the content of the case.

Table 3: Top-10 cases based on incoming references

Nr	NJ-number	Case Name	Other parties	Court	Advocate General	Annotator	Area
1	1996/249	Zwolsman	1	3	2	1	Criminal procedure
2	1981/635	Haviltex	5	5	1	4	Civil procedure
3	1986/723	Heesch/Van den Akker	10	10	3	2	Civil procedure
4	1982/503	Boon/Van Loon	33	10	4	3	Family law
5	2000/721	Reasonable period	10	1	8	21	Civil procedure
6	1986/242	Enka/Dupont	5	15	5	51	Criminal procedure
7	1994/427	Limits oral testimony	3	8	10	33	Criminal procedure
8	1993/659	Vredo/Veenhuis	33	23	14	6	Civil procedure
9	1989/4	HBM/Wielenga	3	33	22	79	Civil procedure
10	1986/3	- <no name>	7	8	143	51	Family law

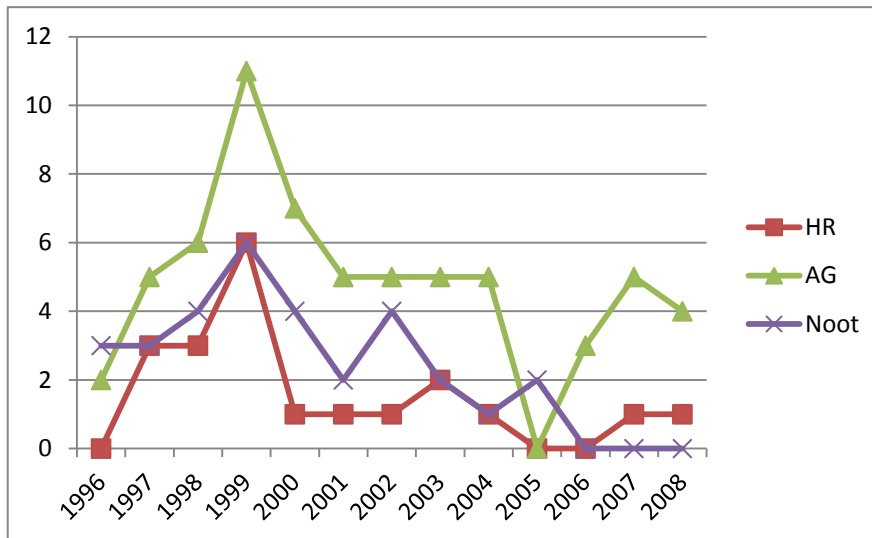


Figure 2: Citations of the Zwolsman case over time for the court (HR), attorney general (AG) and annotators (Noot).

2.4. The Top-10 Cases according to Legal Experts

To get an idea which cases of the Supreme Court are considered important by legal experts, we decided to analyse collections that are being used in teaching law at Dutch universities. We analysed five of these collections from two different publishers and only considered decisions by the Supreme Court between 1965 and 2008. In total we selected 376 cases and ranked them according to the number of citations to these 376 cases in the database of 15,053 cases we created as discussed above. Results are presented below in **Table 4** the same way as in **Table 3** above.

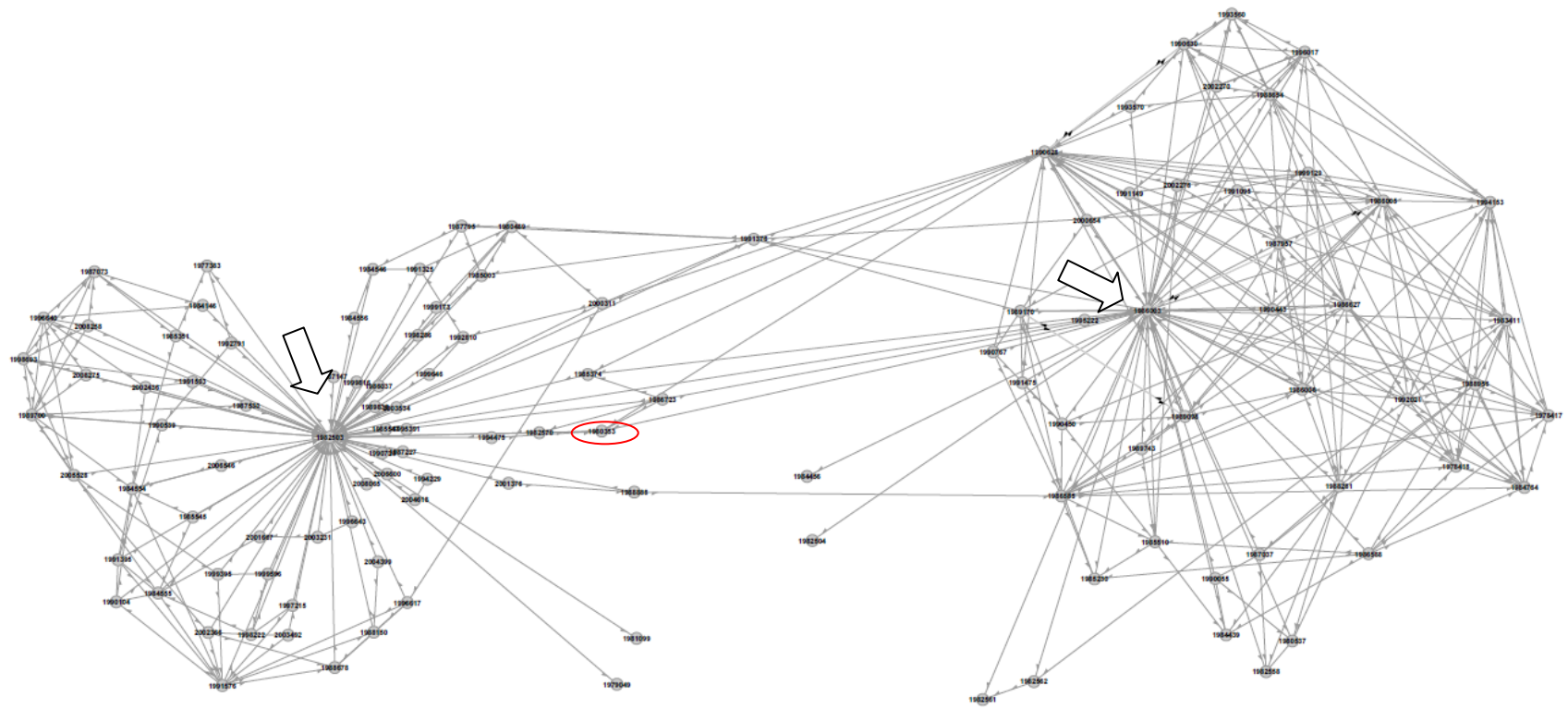


Figure 3: Network of cases that cite Boon/Van Loon (left arrow) and case nr 10 (right arrow). The bull-calf case is circled.

Table 4: *Top-10 cases based on case law collections*

Nr	NJ-number	Case Name	Other parties	Court	Advocate General	Annotator	Area
1	1996/249	Zwolsman	1	3	2	1	Criminal procedure
2	1981/635	Haviltex	5	5	1	4	Civil procedure
3	2000/721	Reasonable period	10	1	8	21	Civil procedure
4	1994/427	Limits oral testimony	3	8	10	33	Criminal procedure
5	2004/376	Waste pipe	33	2	16	198	Criminal procedure
6	1977/241	Bunde/Erckens	1231	764	7	115	Private law
7	1982/411	Roof tiler	19	369	14	581	Criminal procedure
8	1974/450	Meer en Vaart	217	51	16	581	Criminal procedure
9	2006/393	Pot nursery	1231	23	10	5370	Criminal procedure
10	1991/393	State/Windmill	217	15	328	8	Civil procedure

The first four cases are also in the top-10 of our network cases: Zwolsman, Haviltex, ‘reasonable period’ and ‘limits oral testimony’ (1, 2, 5 and 6 in **Table 3**), the other six are not. Cases 7, 8 and 9 are all about the ‘duty to respond’ for judges, i.e. duty to motivate a rejection of an appeal by a defendant. Case 5 is about rules for evidence in criminal cases. Cases 6 and 10 have in common that there is no article directly applicable and general principles of private law apply.

In general it is not the case that these collections of cases only contain cases that are cited a lot by the Supreme Court. 46 of the 376 selected cases were not cited at all in our database, but the overall citation distribution over the 376 cases resembles that of the complete database.

None of the top-10 network cases from **Table 3** appear in collections of both publishers, but 4 of them appear in collections of the same publisher as that of the database (Kluwer). The other publisher also uses different sources and citation methods, i.e. not the NJ-number. If all publishers would use the recent standard way for citing cases that has been developed for the Netherlands,⁶ it will be easier to compute and compare citation rates (cf. [9]).

The family law cases do not appear in any of the collections.

3. The Second Study

For the second study we use data from the official Dutch portal for case law: rechtspraak.nl. We took 89,179 cases from 1999 (when the portal started) till 2008 from all types of courts and all kinds of legal areas. 12% of these cases were from the Dutch Supreme Court, 34% from the lowest courts. Again, the cases do not contain machine readable references to other cases. The first step is to build the network.

3.1. Building the Citation Network

Although the portal uses the recent standard way of identifying Dutch case law (LJN-number, [9]), references to cases use all kinds of mechanisms for identification,

⁶ Juriconnect, <http://www.juriconnect.nl/>

including the NJ-numbers of the first study. This time we adapted a later version of the reference parser we developed for Dutch legislation [6], to find case citations. This version uses patterns in GATE⁷ (General Architecture for Text Engineering) to find and tag references. Again, we validated the approach by parsing 100 random cases and checking these by hand, resulting in an F-score of 95.3%. We analysed the false negatives and added one missing journal to the patterns and deleted one pattern that caused too many false positives with hardly any true positives.

To resolve the references we need to translate all identification methods to the standard LJN number. This service should be provided by the portal, but despite their help, we did not get it to work. We decided to only use the LJN and NJ references, which are the most widely used ones anyway. This left us with 64,042 references in total (14,931 LJN-numbers and 49,111 NJ-numbers).

3.2. The Top-10 Cases in the Network

The top 10 cases in this network based on the number of incoming references are presented in **Table 5**.

Table 5: Top-10 cases 2nd network based on incoming references

Nr	LJN/NJ	Court	Name	PageRank
1	AA 7309/ NJ 2000/721	Supreme Court	Reasonable period	1
2	AG 4158/ NJ 1981/635	Supreme Court	Haviltex	3
3	AU 9130/ NJ 2006/393	Supreme Court	Pot nursery	5
4	AM 2533/ NJ 2004/376	Supreme Court	Waste pipe	9
5	AQ 4473/ NJ 2005/149	CoJ EU	-	2
6	AD 5163/ NJ 2002/317	Supreme Court	-	10
7	AO 1427/ NJ 2005/493	Supreme Court	Pension fund DSM/Fox	6
8	ZC 2536/ NJ 1999/285	Supreme Court	Brok/Huberts	4
9	ZD 0328/ NJ 1996/249	Supreme Court	Zwolsman	-
10	AD 4922/ NJ 2002/82	Supreme Court	Schwarz/Gnjatovic	-

Instead of just looking at the number of incoming references, we can also take the ‘weight’ of the referring source into account like in the PageRank algorithm. The result of such an approach is presented in the last column of **Table 5**. The main difference between plain incoming references and PageRank, besides some shifts of position, is the disappearance of the last two cases from the top-10. These cases have a lot of incoming references from cases that have no incoming references themselves. The average weight of nodes referring to Zwolsman is 1.28 which is low compared to other cases. The case that came in the place of Zwolsman for example, NJ 2004/376 Waste pipe, has an average weight of 2.02 of incoming references. Besides, the Waste pipe case has an incoming reference from NJ 2004/376 Pot nursery, which is also a top ten case.

⁷ <http://gate.ac.uk>

4. Comparing the two studies

The first thing to notice is that although the second database contains cases from all kind of courts with only 12% of cases from the Supreme Court, the top-10 consists of nine Supreme Court cases and one case of the European Court of Justice. This supports the intuition that Supreme Court cases are more authoritative than lower court cases.

Secondly, the average number of references in the second database – 64,000 references for 89,000 cases – is even lower than in the first, even though we also use the LJN-type references besides the NJ-number. This is probably caused by a lot of low court decisions (34% of the cases are of first instance courts) that are not cited at all.

Thirdly, the top-10s of both studies have three cases in common: Zwolsman, Haviltex and “reasonable period”, but except for Haviltex in different positions. If we compare the top-10 of the second study to the top-10 from the case collections used in teaching, we see again an overlap of 4 cases: Zwolsman, Haviltex, “reasonable period” and “pot nursery”. Moreover, 8 of the top-10 cases of the second study appear in some collection used in teaching, while this is only true for 5 of the top-10 cases of the first study. This may be due to the fact that the case set of the first study contains older cases (from 1965 till 2008) and so does its top-10 (6 of the 10 are from before 1990), while the case set of the second study contains cases from 1999-2008 and only 1 case in its top-10 is from before 1990.⁸ Perhaps case collections used in teaching prefer more recent cases.

5. Conclusion

Does the number of citations tell us anything about the importance of case law and can it be used to help legal practitioners find relevant authoritative cases?

The answer to the first part of the question is yes, but it is not the complete story. The cases that are cited most seem to have certain things in common: They are mostly about procedural law, which is not surprising since most of these are from the Dutch Supreme Court, the highest and last instance court whose primary task is to judge judicial procedures and not so much material facts. Secondly, the cases ‘fill holes in legislation’, i.e. the court forms law where the legislator has not (yet) done so. This was best illustrated by the citation pattern over time of the Zwolsman case; once the legislator had ‘repaired’ the hole, the number of citations dropped. So the number of citations is not the only criterion that determines relevance, recency is another one. The type of court matters, but even if we do not take that explicitly into account, all most cited cases turn out to be from the Supreme Court and the European Court of Justice.

Network analysis of sources of law seems to have potential as a research tool for legal scholars. Fowler e.a. [3,4] already showed this for US Supreme Court cases, but given the different status of case law in our continental law tradition, we wanted to see if this would also be the case for Dutch cases. Our courts cite much less than in the US, but despite that the number of citations seems significant.

Fowler e.a. used prediction for future citation to validate the outcome of their network analysis. Given the fact that case citation is less important in the continental legal tradition and our courts cite much less and less long over time, we could not rely

⁸ Remember, cases from 1999-2008 may *cite* older cases and so older cases may turn up in the top-10.

on future citation prediction. Comparison to the selection by publishers and teachers for case collections used in educating law students seems a good alternative.

A ‘sudden’ drop in the number of citations of a case may be an indication of codification, as we have seen in the Zwolsman case. It also seems worthwhile to further investigate the network structures around cases where the court changes a line of reasoning. Perhaps the structure presented in **Figure 3** has some features in common with similar cases that can be used to ‘mine’ for these breaches in case law.

Using the PageRank algorithm instead of just incoming references does not seem to help in finding authoritative cases, i.e. the relative importance of the referring case does not seem to predict the relevance of the cases it refers to. On the Internet, where PageRank was developed by the Google founders to rank search results, *trust* is an issue, and a reference from an important source may enhance trust. In the legal field of professionals this is probably less important. Future research may try out other network analysis measures like ‘betweenness centrality’ of nodes.

Finally, the (*material*) *content* of a case is of course also an important criterion for its relevance in a particular case. Outcomes of network analysis need to be combined with more traditional or other ways to search and match cases on their content (cf. “reason for citing” as used in [10]).

References

1. Dworkin, D., 1978. *Taking rights seriously*, Harvard University Press, USA.
2. Egghe, L. & Rousseau, R. 1990. *Introduction to Informetrics*, Elsevier, Amsterdam.
3. Fowler, J.H. & Jeon, S. (2008). The authority of Supreme Court precedent, *Social Networks* 30: 16-30, Elsevier.
4. Fowler, J.H. Johnson, T.R. Spriggs II, J.F. Jeon, S. & Wahlbeck, P.J. (2006). “Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents” in *Political Analysis, Vol. 15, No. 3*, pp. 324-346, Oxford Journals, Cary, USA.
5. Jaffe, A. & Trajtenberg, M. (2002). *Patents, Citations and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge (MA), USA.
6. Maat, E. de, Winkels, R., Engers, T. Van (2006). Automated Detection of Reference Structures in Law. In: T. van Engers (Ed.), *Legal Knowledge and Information Systems - JURIX 2006*, pp. 41-50. IOS Press, Amsterdam.
7. Mazzega, P., Bourcier, D., Boulet, R. (2009). The network of French legal codes. *ICAIL 2009*, pp. 236-237, ACM.
8. Mey, J.M. de, 2006. Toegang tot rechterlijke uitspraken. Rapport van de VMC-studiecommissie Openbaarheid van rechtspraak, *Mediaforum 2006-4*. (in Dutch)
9. Oppijnen, M. van (2006). A Public Index of Case Law References - the End of Multiple and Complex Citations. *Proceeding of JURIX 2006: The Nineteenth Annual Conference*. IOS press, Amsterdam, pp. 71-80.
10. Zhang, P. & Koppaka, L. (2007). Semantics-based legal citation network. *ICAIL 2007*, pp. 123-130, ACM.