New Zealand Legislation Network

Neda Sakhaee [a] Mark C.Wilson [a] and Golbon Zakeri [b]
[a] *Computer Science Department, University of Auckland*
[b] *Engineering Science Department, University of Auckland*

Abstract.
This paper concerns the recently introduced concept of Legislation Networks, with an application focus on the New Zealand legislation network. Legislation networks have some novel features which make them an excellent test case for new network science tools. We develop several such networks, compute relevant centrality measures, and apply community detection algorithms. We study the relationship between the legislation network measures and legal/political factors.

Keywords. Legislation Network, Directed Graph, Citation Network, Centrality, Community Detection

1. New Zealand Legislation Network (NZLN)

Citation networks of scientific articles have been relatively well-studied [1] but other application areas much less so. One interesting application with a legal flavour has been a study of cases citing majority opinions of the United States Supreme Court [2], which used importance scores to find the most legally relevant precedents. Another study proposes a semantics-based legal citation network viewer as a new tool for legal professionals [3]. The network of French legal codes has been described [4], and another study compares several network representations of the corpus of US Supreme Court decisions [5]. More recently, the corpus of European Union legislative documents has been described as a citation network [6].

We use the idea from [6] of describing the entire legislative system as a complex network, specifically for the jurisdiction of New Zealand. We present a quick review of the network construction process. Then we calculate general network science measures to show the structural difference between Legislation Network and other popular citation networks. We go beyond the basic descriptive studies of [6], with more in-depth network science analysis involving centrality (importance) of nodes and communities in the network.

The New Zealand legislation system includes several types of legislative documents: Bills, Acts, Regulations, and Case Laws which results in a complex multilayer network. In this paper for simplicity we focus on the network of Acts. The latest version of each current Act (with dates of enactment from 1267 to 2015) is available in XML format from the NZ Government Legislation website `www.legislation.govt.nz`. Specific XML tags are used to denote links between documents, as well as metadata such as title, type, date of enactment, etc. We extracted the relevant data using a custom-written C#

program. Errors in the XML tags (which were rather common, over 10%) were detected and corrected by regular expression searching and human evaluation.

We now describe the networks under study. The nodes are precisely the current Acts (the latest version of a law passed by Parliament that has not expired). Some Acts have a special form — they are Amendment Acts whose sole purpose is to change another Act. There are two types of links in these documents. *Reference links* occur when an Act refers to another Act in order to define a concept or to delineate the boundary of application of the original Act. *Amendment links* occur when an Act amends another Act in order to add new information, repeal a section of the original Act, or change the current law. Note that both types of links are directed.

The NZ Act network (denoted ACT) has all nodes and a directed edge from X to Y if and only if there is at least one link from Act X to Act Y. We also study several networks derived from ACT. By restricting only to amendment links, we obtain the network denoted AMEND. The other network, denoted CITE, is more complicated to construct. We restrict only to reference links, and also merge some nodes: the "X Amendment Act" is merged with the Act "X". Edges can be considered as either binary (as above) or weighted, where the weight of the edge from X to Y is the number of links from Act X to Act Y. Thus we also build the weighted version of the three networks above, and denote them with the prefix W.

Table 1. Network General Measures

	RN [1]	ACT	W-ACT	RN	CITE	W-CITE	RN	AMEND	W-AMEND
Nodes	3856	3856	3856	2142	2142	2142	3856	3856	3856
Edges	33884	33884	33884	20124	20124	20124	9030	9030	9030
Average Degree	9.711	8.878	13.233	10.112	9.394	17.257	3.206	2.342	3.648
CCcyc	0.003	0.223	0.25	0.004	0.491	0.551	0.001	0.000	0.000
CCmid	0.003	0.304	0.339	0.004	0.655	0.726	0.001	0.006	0.007
CCin	0.003	0.528	0.554	0.004	0.413	0.438	0.001	0.03	0.003
CCout	0.003	0.506	0.526	0.004	0.374	0.389	0.001	0.033	0.034
Average CC	0.003	0.39	0.416	0.004	0.483	0.525	0.001	0.04	0.043
Average Path length	6.123	3.569	3.569	7.253	3.346	3.346	1.816	4.43	4.43
Small world	No	Yes	Yes	No	Yes	Yes	No	No	No
Acyclic	-	No	No	-	No	No	-	Yes	Yes
Indegree 0, Outdegree >0	-	751	751		431	431	-	799	799
Indegree >0, Indegree $=0$	-	128	128		96	96	-	505	505
Isolated Nodes		21	21		131	131		1512	1512

The constructed networks exhibit two important features. First, they are directed. Second, they may contain cycles. For example, there can be a cycle of references: Section A of Act X cites Section B of Act Y, and Section B of Act Y cites Section C of Act X. We compute general network science measures of degree, average path length, clustering coefficient, and small world property [7,17]. Unlike previous studies which apply methodology appropriate only for undirected graphs, we consider a legislation network as a directed graph (with cycles) for all the calculations. This affects measures of clustering. In [8] the clustering coefficient for directed graphs is defined as the average of the four measures *CCcyc*, *CCmid*, *CCin*, and *CCout* which together cover all possi-

[1]In Table 1 RN is a random graph $G_{n,p}$ chosen according to the Erdős-Rényi model with a specific number of vertices n and connection probability of p chosen to match the network in question [18]. The indexes are calculated based on means from a sample of 100 graphs.

ble directed triangles. If we consider the network as an undirected graph, the clustering coefficient would be smaller.

Table 1 illustrates the general measures for all six networks. As can be seen about one third of the edges are amendment, and two thirds of them are reference links. As illustrated, ACT and CITE networks have the small world property owing to the high clustering coefficient and low average path length compared to random networks.

2. Centrality and Communities

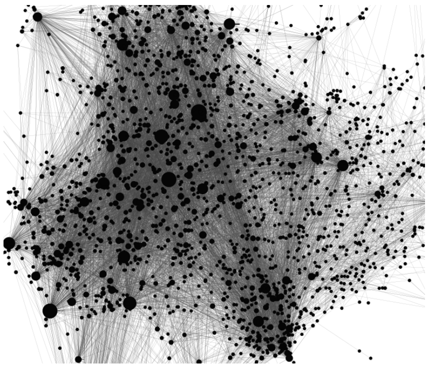
There are many centrality measures, all of which attempt to determine the importance of nodes [9]. Each is based ultimately on a model of flow along edges [10]. We select eight measures that we believe to be most appropriate for our networks. These are PageRank [11], Katz prestige [12], indegree, eigenvector centrality [11], Kleinberg authority [13], outdegree, Kleinberg hub [13], and total degree, The first five are defined in terms of inward links, the next two in terms of outward links, and the last in terms of both types of links. We compute the rank of each act according to each measure . Table 2 represents the most important nodes of CITE, and their rank based on different measures. In CITE, the top ranked nodes based on the Katz Prestige are highly ranked by other measures. These acts appear to be essential references for students of NZ law, which gives us confidence in our methodology. The lowest ranked nodes (not shown) are indeed very obscure.

Table 2. Top-ranked nodes in CITE by Katz prestige

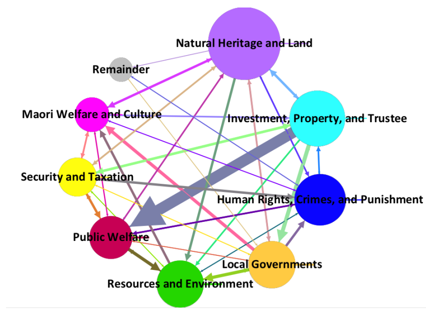
Act	Katz Prestige	PageRank	In-degree	Degree	Eigenvector	Kleinberg Authority	Out-degree	Kleinberg Hub
Public Finance Act 1989	1	1	1	2	13	9	6	23
Criminal Procedure Act 2011	2	4	4	1	33	17	1	6
Summary Proceedings Act 1957	3	7	3	3	4	3	14	8
State Sector Act 1988	4	9	9	16	6	6	69	39
District Courts Act 1947	5	3	5	6	3	4	12	5
Crimes Act 1961	6	11	12	13	9	8	21	19
Companies Act 1993	7	4	7	11	18	11	25	16
Local Government Act 1974	8	8	2	3	30	13	19	22
Judicature Act 1908	9	6	6	7	5	5	11	3
Privacy Act 1993	10	19	15	17	17	10	22	20

Considering that the legislation network is directed and cyclic, three algorithms are appropriate: The Girvan-Newman algorithm based on clustered edge betweenness measure [15], maximizing modularity using extended Louvain algorithm for directed weighted graphs [14], and maximizing map equation [16]. Figure 3(a) shows the structure of the network W-CITE in detail, and the node size represents the in-degree of each node. If we consider the graph as undirected there is one giant connected component. Figure 3(b) shows how Louvain algorithm detects meaningful communities. To label them, we used a keyword search across the communities. Then in each community we search for the list of keywords, that appear in at least 85% of its nodes and don't appear in more than 10% percent of the nodes in the remaining communities.

Space limitations preclude discussion of other topics (to appear elsewhere), including correlating network features with social and political factors and with the time evolution of the network (which we also intend to model). For the latter, new techniques will be needed, because the historical versions are not available in a structured machine-readable format.



(a) W-CITE



(b) Communities of W-CITE

Figure 1. Louvain Algorithm result for W-CITE

References

[1] Hummon, Norman P., and Patrick Dereian. “Connectivity in a citation network: The development of DNA theory.” *Social networks* 11.1 (1989): 39-63.

[2] Fowler, James H., et al. “Network analysis and the law: Measuring the legal importance of precedents at the US Supreme Court.” *Political Analysis* 15.3 (2007): 324-346.

[3] Zhang, Paul, and Lavanya Koppaka. “Semantics-based legal citation network.” *Proceedings of the 11th international conference on Artificial intelligence and law. ACM*, 2007.

[4] Mazzega, Pierre, Daniéle Bourcier, and Romain Boulet. “The network of French legal codes.” *Proceedings of the 12th international conference on artificial intelligence and law. ACM*, 2009.

[5] Bommarito II, Michael J., Daniel Katz, and Jon Zelner. “Law as a seamless web?: comparison of various network representations of the united states supreme court corpus (1791-2005).” *Proceedings of the 12th international conference on artificial intelligence and law. ACM*, 2009.

[6] Koniaris, Marios, Ioannis Anagnostopoulos, and Yannis Vassiliou. “Network Analysis in the Legal Domain: A complex model for European Union legal sources.” *arXiv preprint arXiv:1501.05237* (2015).

[7] Newman, Mark EJ. “The structure and function of complex networks.” *SIAM review* 45.2 (2003): 167-256.

[8] Tabak, Benjamin M., et al. “Directed clustering coefficient as a measure of systemic risk in complex banking networks.” *Physica A: Statistical Mechanics and its Applications* 394 (2014): 211-216.

[9] Rusinowska, Agnieszka, et al. “Social networks: Prestige, centrality, and influence.” *International Conference on Relational and Algebraic Methods in Computer Science. Springer Berlin Heidelberg*, 2011.

[10] Borgatti, Stephen P. “Centrality and network flow.” *Social networks* 27.1 (2005): 55-71.

[11] Bonacich, Phillip. “Power and centrality: A family of measures.” *American journal of sociology* (1987): 1170-1182.

[12] Katz, Leo. “A new status index derived from sociometric analysis.” *Psychometrika* 18.1 (1953): 39-43.

[13] Perra, Nicola, and Santo Fortunato. “Spectral centrality measures in complex networks.” *Physical Review E* 78.3 (2008): 036107.

[14] A.L. Barabási, “Network Science”, *Cambridge University Press* (2015).

[15] Girvan, Michelle, and Mark EJ Newman. “Community structure in social and biological networks.” *Proceedings of the national academy of sciences* 99.12 (2002): 7821-7826.

[16] Rosvall, Martin, Daniel Axelsson, and Carl T. Bergstrom. “The map equation.” *The European Physical Journal Special Topics* 178.1 (2009): 13-23.

[17] Watts, Duncan J., and Steven H. Strogatz. “Collective dynamics of small-world networks.” *nature* 393.6684 (1998): 440-442.

[18] Erdős, Paul, and Alfréd Rényi. “On the strength of connectedness of a random graph.” *Acta Mathematica Hungarica* 12.1-2 (1961): 261-267.