Michael J. Nelson
Faculty of Information and Media Studies
University of Western Ontario
London, Ontario, Canada

The Visualization of the Citation Patterns of Some Canadian Journals

Abstract:
In order to easily see the citation patterns of a journal or subject area it is very useful to use a graphical diagram to visualize all the connections between journals. Using data from derived from the *Journal Citation Reports*, this study investigates the citation patterns of three Canadian Journals in three subject areas: library and information science, psychology and mathematics.

Résumé : Afin de mieux voir les modèles de citations d'un journal ou d'un domaine, il peut être très utile d'utiliser un schéma graphique pour visualiser toutes les relations entre les journaux. Utilisant les données dérivées du *Journal Citation Reports*, cette étude examine les modèles de citations de trois journaux canadiens dans trois domaines : la bibliothéconomie et les sciences de l'information, la psychologie et les mathématiques.

1. Introduction

One of the applications of citation studies is to explain the patterns of journal usage by researchers. An early study (Bishop, 1987) used the impact factor from JCR to show the rank of Canadian Scientific Journals in various subject areas. Garfield (1993) has also studied Canadian scientific research by using citations to compare countries, universities in Canada and highly cited authors and papers. On a more general scale Leydesdorff has studied the structure of both the science journals (Leydesdorff, 2004a) and social science journals (Leydesdorff, 2004b) by using data provided in the Journal Citation Reports from ISI. Our objective here is to study some example Canadian journals to see how there citation patterns describe their connections to other journals.

In this study we will compare two visualization methods for describing these connections, one based on social network analysis (de Nooy, 2005) and the other a two dimensional display from multidimensional scaling. Three example journals are chosen from a wide range of subject areas to see if the methods are applicable and useful across such a variety of areas. They are the Canadian Journal for Information and Library Science, Canadian Pyschology and the Canadian Journal of Mathematics. We purposely did not choose very specialized journals but wanted to look at more general patterns of journals.

Both visualization methods are based on data from a method developed by Leydesdorff (1986). The Journal Citation Reports are used create an asymmetric matrix of the number of citations from one journal to another. This matrix is then used to calculate a journal to journal citation matrix of similarities. In the original

analysis the Pearson correlation coefficient was used. Leydesdorff has provided this data for 2003 on a journal by journal basis for both the journals citing an individual journal and those journals cited by the same journal. The new 2003 data (Leydesdorff, 2005) uses cosine similarities which is more appropriate for our visualizations (Ahlgren, Jarneving and Rousseau, 2003). The data is limited to by the following conditions:

> "The citation environments are limited to those journals which cite the seed journal more than one percent of its total citation rate or which are cited by this journal to more than one percent of its citation pattern. (These two criteria operate independently.)" (Leydesdorff, 2005)

In addition, only cosine similarity values greater than 0.2 are provided for the analyses. So for each journal we have two symmetric matrices of cosine correlations, one for journals that have cited the core journal and one for the citing pattern of the core journal. The citing data gives us information on the journals used by authors in their research and cited data is related to the influence of a journal and it's authors. We can use this individual journal data to visualize the patterns using the network analysis program Pajek (Network/Pajek, 2005) which is free for non-commercial use and SPSS for the multidimensional scaling.

2. Visualization from Social Network Analysis

For the first visualization we use the Kamada-Kawai method from the Pajek software. This is the method used by Leydesdorff (2004a, 2004b) to visualize subject areas in his research. The idea behind the algorithm is to minimize the energy in the system when the connections are modeled by springs.

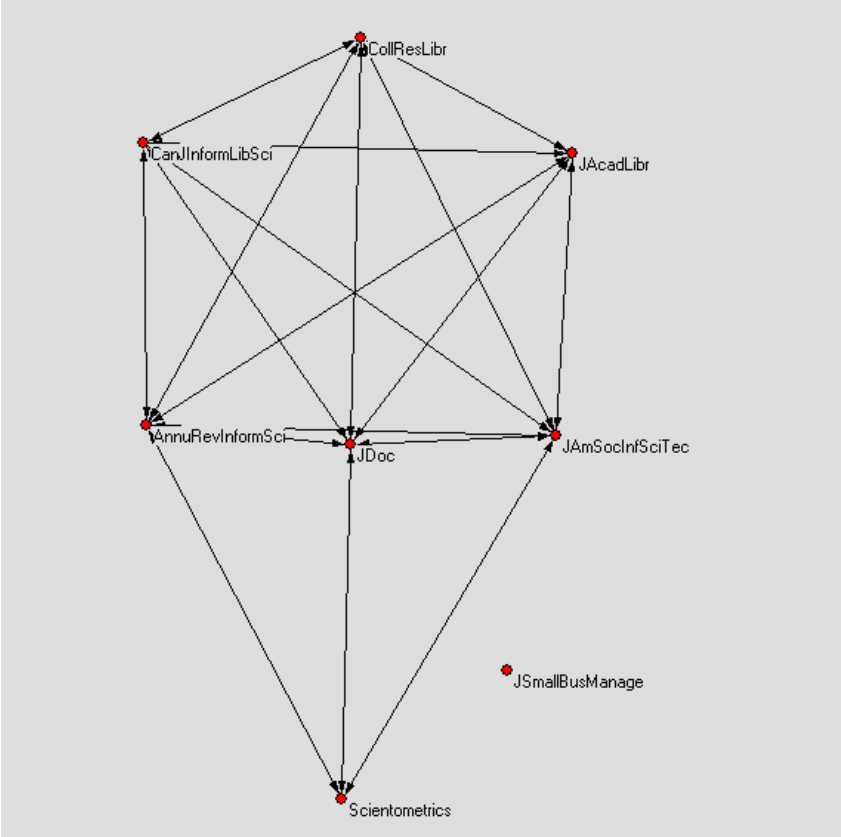

Figure 1: Canadian Journal for Information and Library Science. Cited.

This analysis only uses the threshold to calculate connections, so that journals with cosine similarity greater than 0.2 are connected and those with similarity less than 0.2 are not connected.

The visualizations tend to produce a symmetric diagram of connections but do not reflect the strength of the connection between journals. For the first example we will investigate the Canadian Journal for Information and Library Science (CJILS). The total number of journals included for the library and information science analysis is eight, since the data only includes journals which cite the seed journal more than one percent of its total citation rate or which are cited by this journal to more than one percent of its citation pattern and there are many journals that only have a small number citations so are not included. The network diagram for the cited is in Figure 1. The Journal of Small Business Management is not connected because of the threshold and because it is connected in the citing dimension in Figure 2. In the cited dimension Journal of Documentation is in the centre only because it has the most connections. In the citing dimension CJILS is in the centre because it has the most connections, which is natural as it is the seed which brings in the other journals.

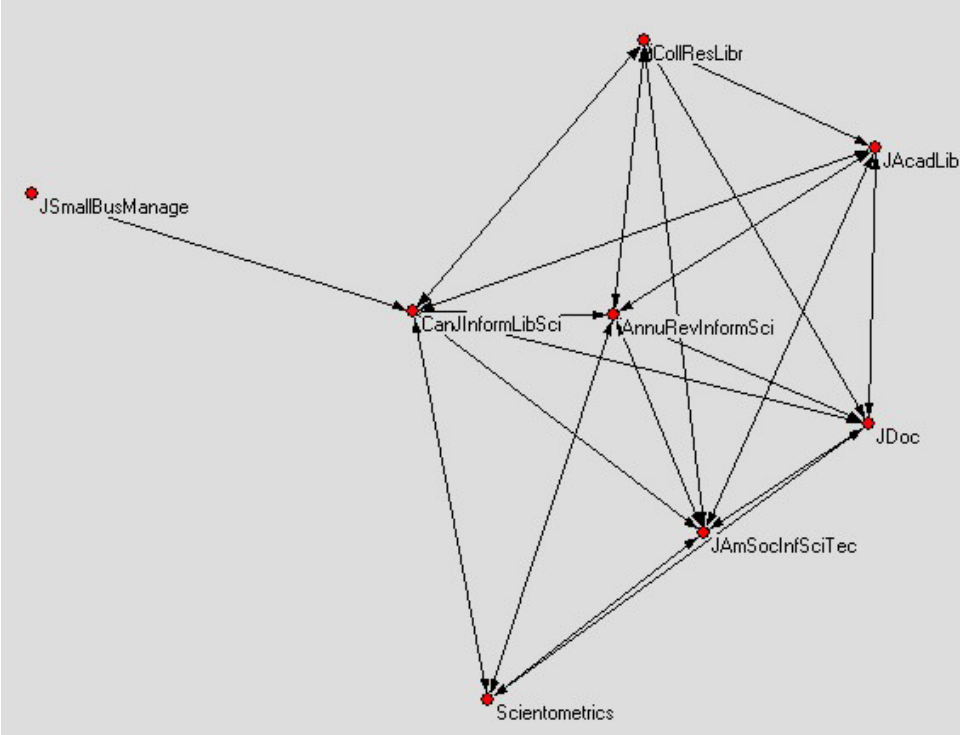

Figure 2. Canadian Journal for Information and Library Science. Citing.

Remember that a connection means that the *pattern* of being cited or citing is being used as the measure of similarity or connection.

The visualizations for the other journals are in figures 3 to 6.

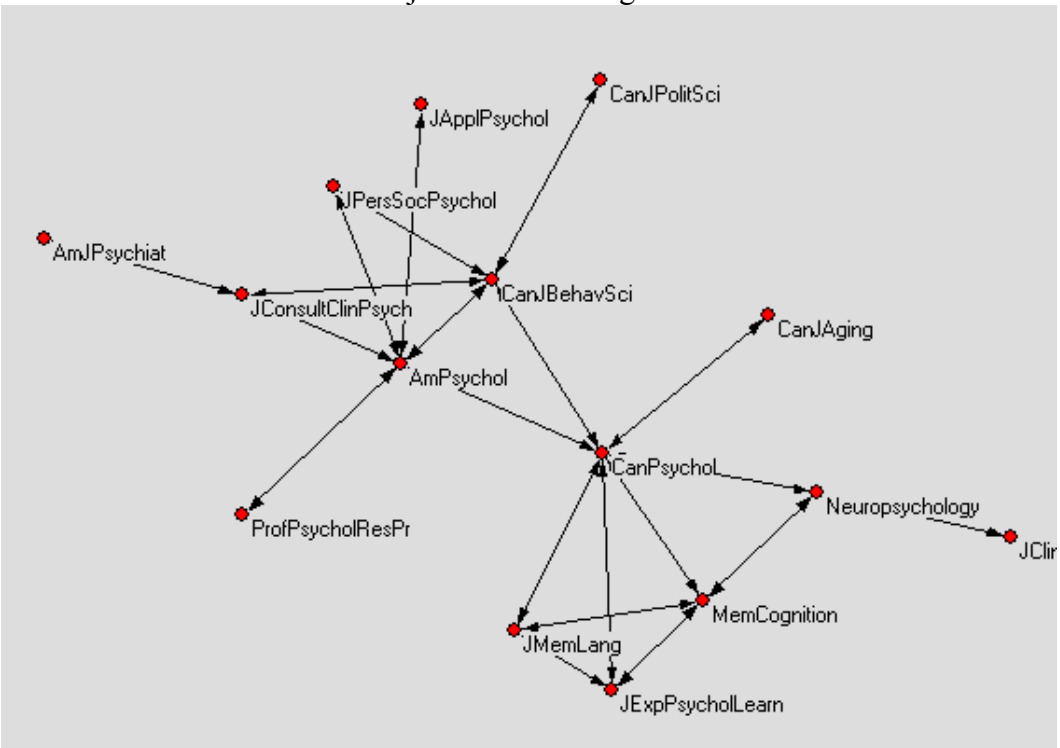

Figure 3. Canadian Psychology. Cited

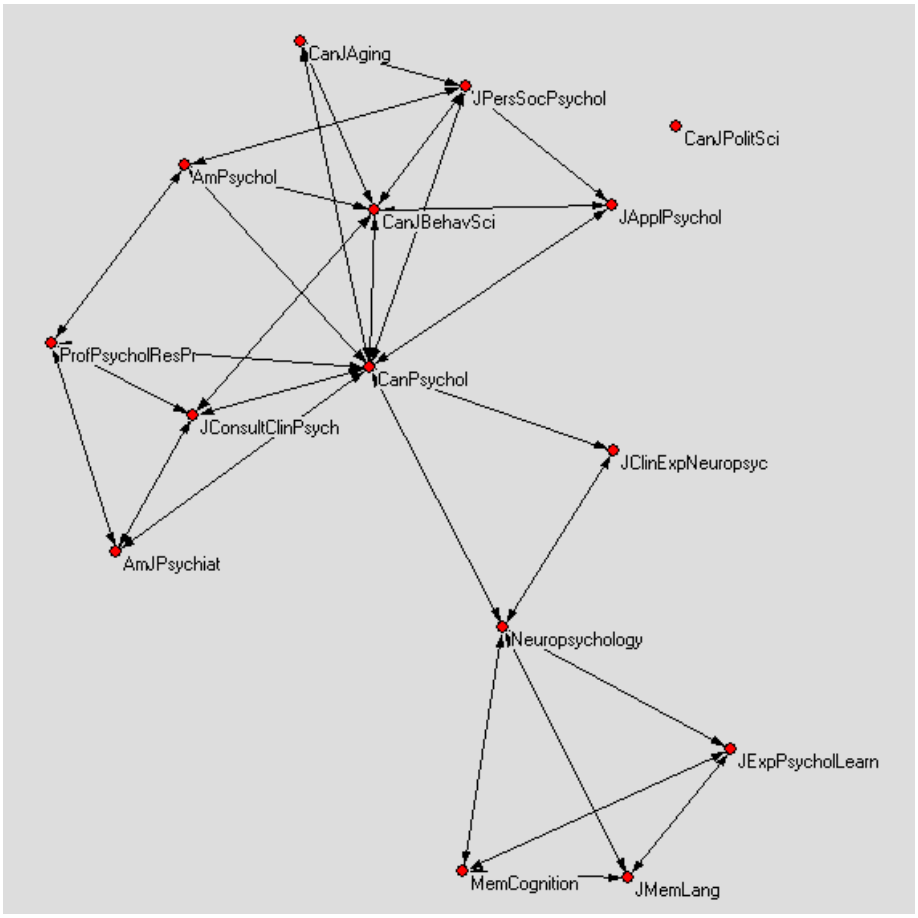

Figure 4. Canadian Psychology. Citing.

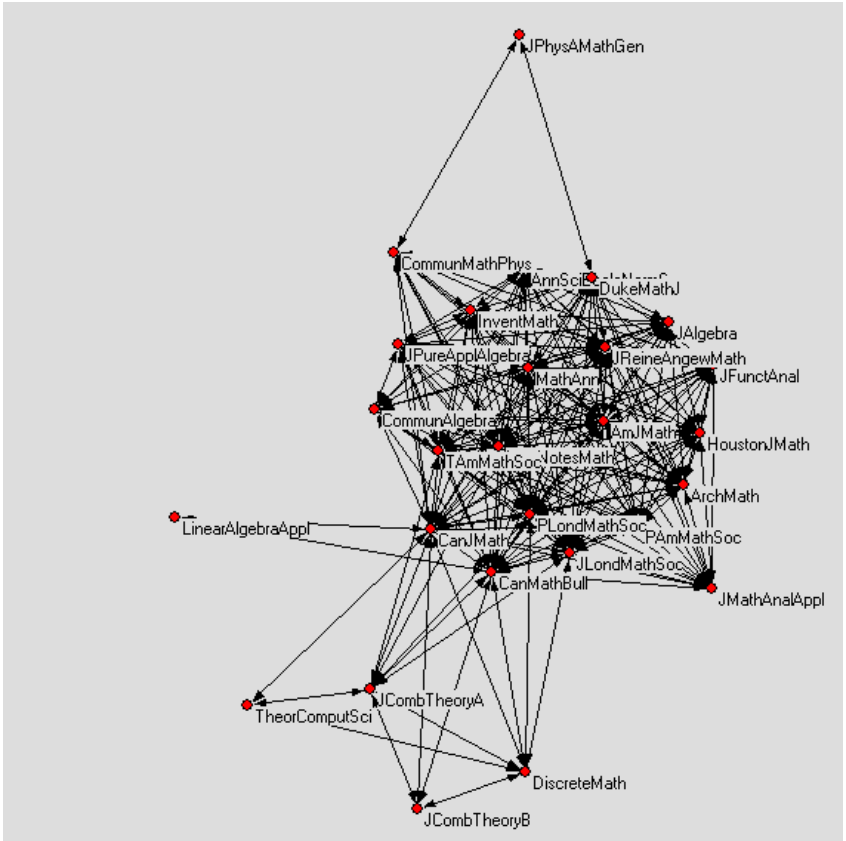

Figure 5. Canadian Journal of Mathematics. Cited.

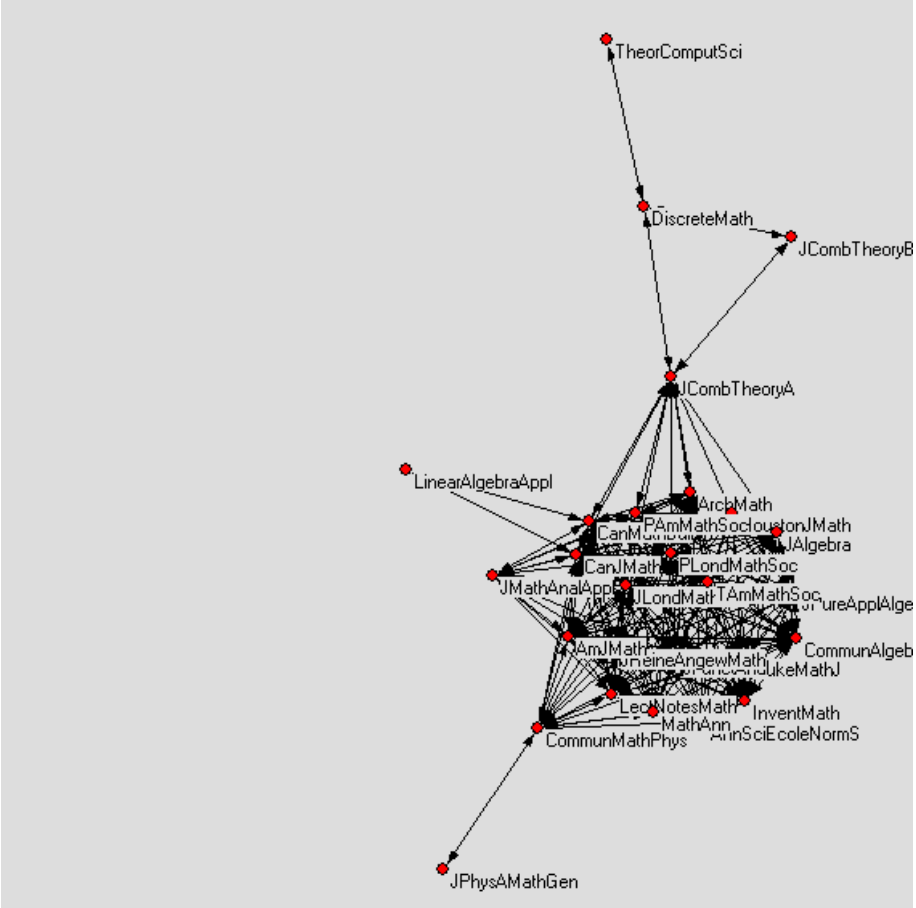

Figure 6. Canadian Journal of Mathematics. Citing.

The Canadian Psychology visualizations include 15 journals and do show that the journal is a pivot point between two sets of journals, the two sets are quite different but Canadian Psychology has connection to both. One set seems to be the area of neuropyschology and memory the other set is more general psychology. The Canadian Journal of Mathematics diagrams are not so revealing except there are a large number of mathematics journals in the core which have very similar citation patterns and a few journals not directly connected. The peripheral journals include combinatorics and discrete math on one end and mathematical physics on the other end. There are 27 journals in the mathematics diagrams, which makes the interpretation more difficult.

3. Visualization using multidimensional scaling.

One of the main weaknesses of the social network visualizations is that the strength of the relationship is not used in the analysis so is not reflected in the diagrams. To create a visualization which uses this data, multidimensional scaling (MDS) is used to produce a two dimensional map reflecting the Euclidean distances between journals as closely as possible when projecting onto two dimensions. (See for example Cox and Cox, 1994.

In order to use SPSS for the visualization, the correlations where transformed to distances with the formula: distance = 1.0 – cosine. Another advantage of this method is that clusters of journals are more easily seen on the diagram. If we

look at CJILS in Figure 7, we see that there are four clusters around the journal which could be described broadly as library science, information science, Scientometrics (by itself) and the Journal of Small Business Management (by itself). When the data for citing patterns is used we get almost the same visualization with exactly the same clusters.

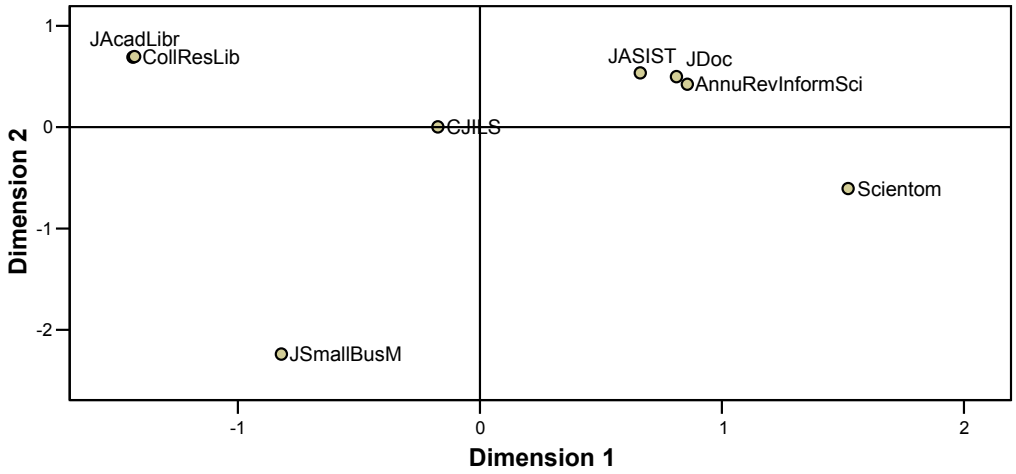

Figure 7. CJILS. Cited. Multidimensional Scaling

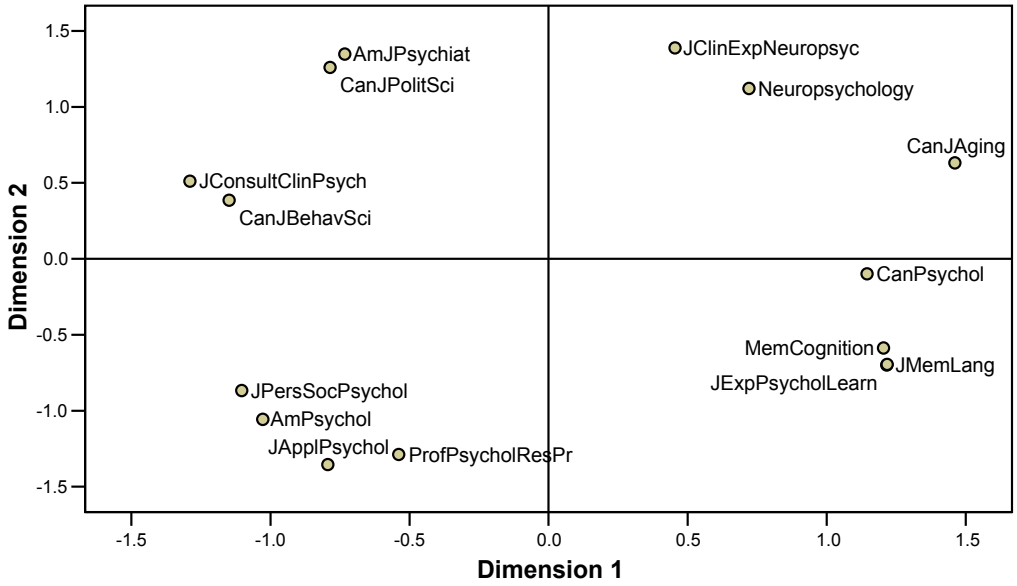

Figure 8. Canadian Psychology. Cited. Multidimensional Scaling

For Canadian Psychology we also see fix or six clusters of journals, but this time the seed journal is not in the centre of dimension 1. In multidimensional scaling these dimensions are not very meaningful so it is difficult to give any further explanation.

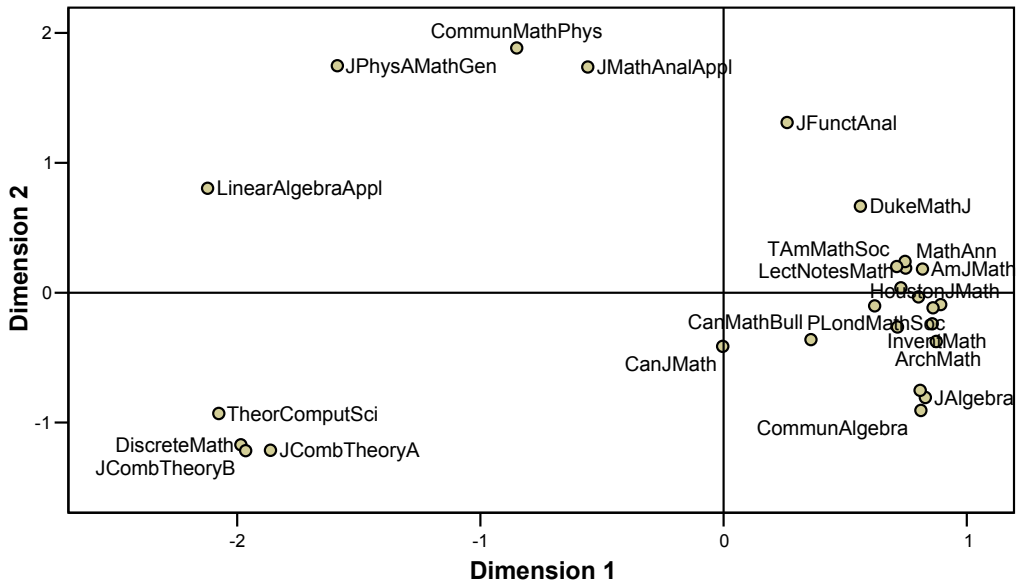

Figure 9. Canadian Mathematics Journal. Cited. Multidimensional Scaling

The Canadian Mathematics Journal shows much the same pattern as in the social network but we see that there are several journals which are not very close to the core cluster.

For MDS the diagnostic statistic the stress statistic. For CJILS the stress is a respectable 0.12 which means the map onto two dimensions fits the data reasonably well. For the other two journals however, the stress was 0.33 and 0.28 which indicates there were problems in trying to project the data onto two dimensions. This is partly because there were more journals to map, but it also means the data provided conflicting requirements to the algorithm.

4. Conclusions

Using the data provided by Leydesdorff provides some limitations as it only gives us similarities based on the whole citation pattern of the journals and only includes values of the cosine correlation greater than 0.2. It would be interesting to see if there are any differences in the MDS if the full set of correlations are used. Another possibility is to return to the basic asymmetric matrix of journal-journal citations to see if it could be used in other types of visualizations.
The social network analysis can sometimes give insights into the general patterns, for example the Canadian Psychology case, but I think that MDS gives a better visualization of the clusters of journals around a seed journal, despite some of the fit problems.

References:

Ahlgren, Per; Jarneving, Bo and Rousseau, Ronald. 2003. Requirements for a Cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*. 54(6): 550-560

Bishop, Claude T. 1987. An overview of Canadian scientific publications. *Canadian Journal of Information Science*. 12(1): 20-34.

Cox, Trevor F. and Michael A.A. Cox. 1994. *Multidimensional Scaling*. Chapman and Hall.

Garfield, Eugene. 1993. What citations tell us about Canadian research. *Canadian Journal of Information Science.* 18(4): 14-35.

Leydesdorff, Loet. 1986. The development of frames of reference. *Scientometrics*. 9 (3-4): 103-125.

Leydesdorff, Loet. 2004b. Clusters and maps of science journals based on bi-connected graphs in Journal Citation Reports. *Journal of Documentation*; 60 (4), 371-427.

Leydesdorff, Loet. 2004a. Top-down Decomposition of the Journal Citation Report of the Social Science Citation Index: Graph- and factor-analytical approaches. *Scientometrics* 60 (2): 159-180

Leydesdorff, Loet. 2005. Journal citation patterns. http://users.fmg.uva.nl/lleydesdorff/jcr03/cited/index.htm [Accessed 20050415]

Networks / Pajek , Program for Large Network Analysis. http://vlado.fmf.uni-lj.si/pub/networks/pajek/ [Accessed 20050113]

de Nooy, Wouter; Andrej Mrvar and Vladimir Batagelj. 2005. *Exploratory Social Network Analysis with Pajek.* Cambridge University Press.