2001

# Mining Social Network Analysis Data

Debbie Richards
*Division of Information and Communication Sciences*, richards@ics.mq.edu.au

Phillip Higgins
*Division of Information and Communication Sciences*, phiggins@ics.mq.edu.au

# Mining Social Network Analysis Data

Debbie Richards and Phillip Higgins

Department of Computing
Division of Information and Communication Sciences
richards@ics.mq.edu.au, phiggins@ics.mq.edu.au

**Abstract**

*This project was motivated by the need to meaningfully display large amounts of Social Network Analysis (SNA) data from an exploratory case study into the existence of technological convergence in Australia. We found that many tools used for the display of SNA data did not handle large datasets well due to the denseness of information, a typical problem in the display of large graphs. The approach we offer in this paper does not address our motivating problem. Instead of handling large graphs we sought an alternative approach that would allow us to use the tools we had by mining the dataset for interesting concepts and displaying that subset. This is work in progress and we intend to do more work on our graphics tool and exploring alternative algorithms. Our initial results show that a machine learning approach can provide useful information from SNA data that may not have been apparent from typical SNA techniques.*

**Keywords**

Rule-based deduction, Organisational value chain, Power in Organisations, Organisational design, Information Flows, IS models, Simulation and modelling, Data Mining, Graphics, Machine Learning, Social Network Analysis

## INTRODUCTION

This paper describes some work on the visualisation of results from an exploratory case study into the existence of technological convergence in Australia. The case study will be briefly discussed to provide a description of the data, its collection and processing. What became apparent from the case study was that the tools accessible to us were rather old, (in computing years), and lacked a feature set necessary for the display of the complex and extensive Social Network Analysis (SNA) data we had. While commercial software did exist, such as NetMap by NetMap Solution Pty Ltd (Sbarcea, 1999), that we believed would have handled our data and provided valuable insights, such software was not available for this project. What we essentially faced was the problem of how to handle large graphs. The approach we offer in this paper does not address our motivating problem. Instead of handling large graphs we sought an alternative approach that would allow us to use the tools we had by mining the dataset for interesting concepts and displaying that subset. In the next section we will introduce the case study and SNA. In section 3 we describe the tools we explored and our reasons for dissatisfaction with them. In Section 4 we describe the theoretical underpinnings, project, results and future work. The conclusion is given in the final section.

## THE CASE STUDY

The internet era has resulted in old-style markets and organizations being forced into a new world with new markets, evolving business processes and changing governmental policies. This single-case exploratory case study utilises visualisation and SNA methods to analyse data collected in a literary review covering the period from 1995 to 2000. The research question relates to the role a huge Australian telecommunications provider, which we will refer to as Gigante, has on convergence in Australia and which, if any, hypothesis can be applied to other significant companies.

Digital technology is seen to be the driving force behind the acceleration of change evident in traditional industry and regulatory structures. The established industries generally remained static throughout the eighties and early nineties mainly due to National and State regulators being able to control domestic service providers on an industry-based basis (Convergence Report, p. 17). Now, however:

> "information and communication technologies are allowing companies to move into entirely new markets – indeed the boundaries between entire industries are disappearing" (BT World Communications Report 1998/99)

The blurring of boundaries between industrial sectors due to some convergent agent is what convergence is all about. Convergence, by nature, forces industries to restructure in order to adapt to the new landscape. An illustration is the convergence of business and the internet to form a new sector known as e-commerce, a sector that is based in both computers and business but belongs wholly to neither. The case study was interested in first determining if convergence existed and then to consider the impact of convergence on industry and the public in general. We were interested in discovering convergence because of the restructuring that occurs.

The data for this study has been gathered over the last 2 years using Fairfax Publications news archive and search facility as it's major source. Information freely available in the press and from the companies concerned has been collated resulting in a comprehensive 'snapshot' of the industries over the 5 years from 1995 to the year 2000. The main dataset takes the form of tables indicating the ties or links between Gigante and 'partner companies' with which it entered into an alliance over the 5-year study period. The output of this work was a set of tables in UCINET[1] each containing some aspect of the companies or the alliances made. They were placed in separate tables initially to aid the analysis of singular aspects and as the data lent itself to this layout. The dataset was also categorised using: the type of alliance, the core business of the company, the purpose of the alliance, the scope of the alliance and the location. Analysis of the datasets was performed manually and electronically. The main impetus for the collection of research data is a parallel multi-faceted study into the strategy Gigante is implementing in terms of alliances (More and McGrath 2001). This exploratory case study was concerned with the development of independent hypotheses in search of original conclusions.

### Social Network Analysis (SNA)

Social Network Analysis originated in the 1930s with Harvard sociologists, Manchester anthropologists, and the Gestalt theory (principally associated with Wolfgang Köhler, (see Köhler, 1947)). It has since evolved through three main lines: sociometric analysis (graph theory); interpersonal relations and the formation of cliques; and finally the structure of community relations, particularly in tribal villages. From these beginnings Harrison White, (again from Harvard), and his students had united social network analysis into a complex but coherent framework by the 1960s (Scott, 1992).

Today SNA is used as a set of methods for analysis into social structures and their otherwise hidden relationships. It is still evolving and only with the computing age is this mathematically based social methodology about to realise its full potential. SNA concerns itself primarily with relational aspects of social structures. Based on matrix algebra it utilises various algorithms and theories that 'massage' the data into revealing hidden structures and henceforth suggesting possible solutions to observed events.

SNA was chosen for this study due to it's ability to analyse networks based on links or relationships; to analyse data in different ways and because SNA permits searching for characteristics (such as social factors) embedded in the data that are otherwise not reachable in any other way. The focus of this study was the identification of technological convergence within Australia. A key to this identification was to find out who is owned by, affiliated with, collaborating with, etc., whom. SNA considers the various types of contact one party may have with another and the importance of those relationships. The ideas such as centrality and the identification of cliques were seen to be highly relevant in determining if convergence was occurring and to what extent and which parties are involved and in what way. The final but less important reason for choosing SNA over other analytical methods proved to be the flexibility of the data format required by SNA. The format fitted our data well and also ensured a high degree of portability between other programs and utilities, which permitted different tools to run separate analyses and/or create comprehensible graphs from within other applications.

Definition of a few key concepts used in SNA is in order. We here describe the sociogram, degree, centrality and cliques.

A *Sociogram* **-** is "A way of representing the formal properties of social configurations" (Scott 1992, p7) which allows a researcher to visualize the flow of information across the network. A sociogram is essentially a set of nodes (or actors) interconnected according to observed social rules. Figure 1 shows a simple sociogram. The matrix that this sociogram is based on is given in Figure 2. A 1 (one) represents a link between the nodes and a 0 (zero) means no link. The direction of the link is determined by the row actor's value. ie. George talks to Ken, (hence the 1 on the first row under Ken), but Ken doesn't talk to George, (2nd row, first column). Conversely, an undirected sociogram results in a symmetric matrix as every actor talks back to anyone talking to them.

[1] UCINET is software written for SNA purposes. Running in Window/98 it can perform most of current SNA methods on spreadsheet-like data. It supports the import/export of a number of formats however it does not support visualisation of the data, the main concern of this paper.

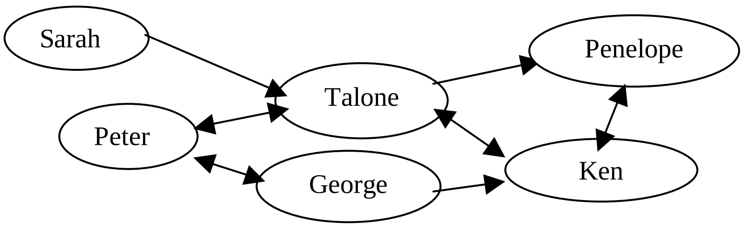


Figure 1: Example of a Simple Sociogram

| | George | Ken | Penelope | Peter | Sarah | Talone |
|---|---|---|---|---|---|---|
| George | 0 | 1 | 0 | 1 | 0 | 0 |
| Ken | 0 | 0 | 1 | 0 | 0 | 1 |
| Penelope | 0 | 1 | 0 | 0 | 0 | 0 |
| Peter | 1 | 0 | 0 | 0 | 0 | 1 |
| Sarah | 0 | 0 | 0 | 0 | 0 | 1 |
| Talone | 0 | 1 | 1 | 1 | 0 | 0 |

Figure 2:  Matrix for the simple sociogram shown in Figure 1

*Degree* **-** is an SNA term where the number of relationships or connections into a node are summed either as a whole or according to the direction of the link. Referring to Figure 1, Talone has degree of 4, an in-degree of 3 and an out-degree of 3.

*Centrality* **-** is a measure of a node's degree, which indicates an actor's centrality or power potential. In Figure 1, Talone's centrality is 4 (same as degree) and is thus more prominent than any of the other actors.

*Cliques* -can be thought of as a sub-grouping of a network. This sub-grouping is built up, or developed out of the combining of dyad and triads into larger, but still closely connected structures. Studying the role of cliques can be insightful in understanding how the network as a whole does or is likely to behave.

## SNA TOOLS

In previous sections we have introduced our goals, the data and SNA. When we began this case study we did not anticipate that although SNA contained the features we desired, the volume of data we needed to display was not manageable with the currently available tools. The first product we considered is known as Krackplot, an SNA visualisation tool on high standing in the SNA world. While it's display and organization of small datasets is admirable, even exemplary in certain circumstances as in Figure 3, once the data goes beyond a given size, the screen crowds up, clusters become messy shapes and information is lost to the user.

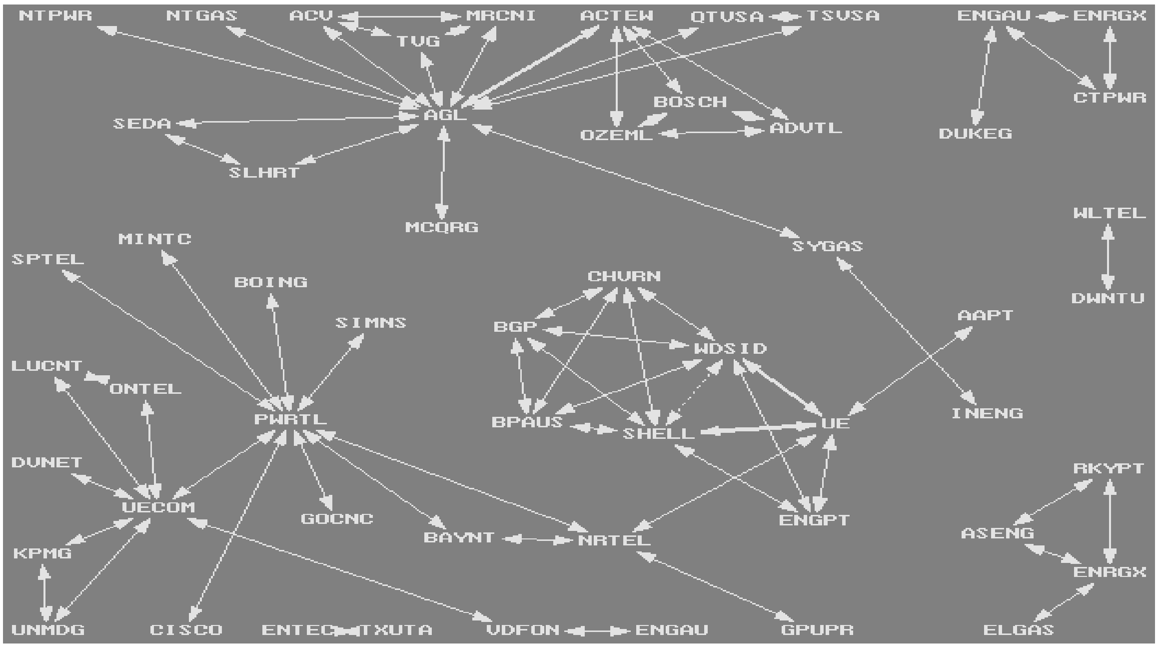


Figure 3: Good example using a small dataset with Krackplot

Figure 4 shows a typical example of the views we were given. On the positive side, the graph in Figure 4 does show convergence in the centre which reveals the organization/s with high centrality (importance). However, it is not conclusive which organization is being pointed to. Specific questions such as who is connected to IBMAU (we challenge you to find that node) are impossible to answer without at least some wrong interpretations.

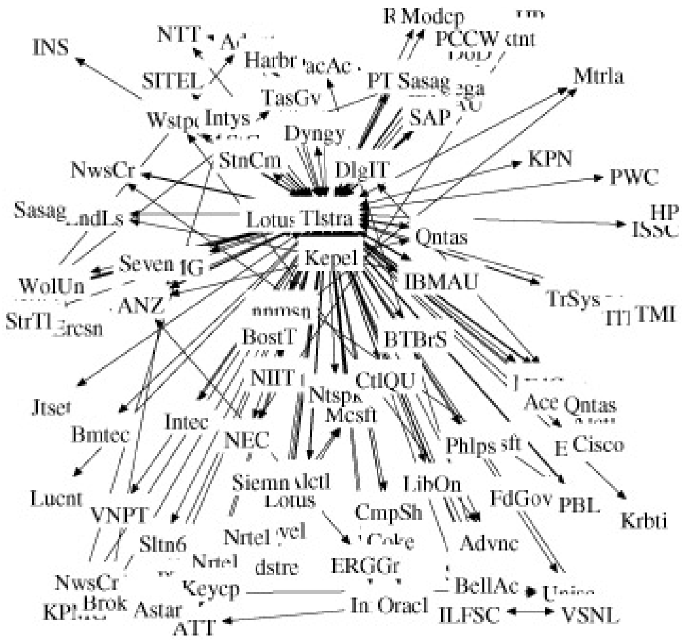


Figure 4: Example of a typical graph showing clutter and poor readability

Most of these flaws are simply inherent in the operating system it was built on (DOS) and the age of the product (no upgrades in the last 5 years). There is no scrolling and limited zooming and the clusters refused to be ordered intelligently. These features may sound like bells and whistles for the case study, but since SNA is also a visualisation technique, poor visualisation results in poor analysis. With a dataset the size of this study display proved almost useless on anything attempted. Only the most basic of inferences could be made amongst the confusion of lines and nodes.

It thus became apparent that a more powerful tool than Krackplot was needed. There are other tools, mostly ported from Unix/Linux or of lesser standing than Krackplot. Whilst searching for a solution to the drawing and navigation of large social networks, we found numerous papers (eg *www.isi.edu/~papers/gd95.html*) citing the usefulness of colour, zooming, intelligent clustering and time lapse. It appeared that others also desired the same capabilities. Our preference was to find some research or public domain software with a solution. Mage was been recommended as a good solution for visualising SNA data (Freeman 1998). Mage was actually designed as a chemical modelling tool, but as it allows flexible rules between nodes it is quite capable of displaying SNA sociograms. The interesting feature of this program was its ability to show the data in pseudo-3D. The sociogram could be rotated and explored via zoom and time lapse. This seemed to be the holy-grail. Alas, despite downloading a conversion program from Freeman (1998), we could not get the dataset to work in any functional way. It was back to the drawing board.

The problem we faced was not just related to SNA sociograms. The problem exists for other fields interested in displaying large graphs, eg in the field of software visualisation (Churcher, Keown and Irwin 1999). Even if we have a tool that displays the graph clearly and we can move around the graph we have problems such as losing our starting point or context. Zooming and scrolling are good features but users often find that they get lost can can't comprehend the data because they lose the context of their original query. Techniques such as the fish eye approach (Hollands et al. 1989) which displays the graph around the focus of attention clearly and becomes denser away from the focus does not solve all problems. If we use different levels of abstraction or just work on sections of the graph at a time we may miss important detail or connections. As humans we have difficulty comprehending a large graph that spans over many virtual pages. It is similar to the disorientation that occurs when we get lost in cyberspace where we no longer remember how we got to the spot we are at, but want to go back to a previous site. We decided that although we would still pursue better graph display and manipulation, we would try another technique that is similar to the ideas behind data mining. Data mining has been found to be a valuable way of extracting knowledge from databases. The goal is to discover patterns in the data that can be formed as association rules. These associations can be used for such purposes as to learn about the purchasing trends of its consumers (Deogun, Raghavan and Sever 1998), decision making or risk management (Fortin, Liu and Goebel 1996). In this study we are also looking for trends in the data. It is these trends that we wish to visualise rather than the complete data itself. We wanted to discover the centrality, degrees of importance and cliques using the SNA sociograms. However, cluttered screens made recognition of these features impossible. If we could classify the data into meaningful groups then we could explore individual cliques and who the central players were. Maybe we would find cliques within the larger cliques. By identifying those with high centrality we could draw individual graphs for each major player. To be able to do this we first had to identify what the key selection criteria might be. We can also identify uninteresting features in the data that can be ignored and not included in the display. For example, company code is obviously uninteresting as a factor affecting convergence. There may be other fields in the data that are also irrelevant and unnecessary to include in a graph. To this end we began to explore the application of a machine learning (ML) algorithm to extract the interesting concepts. The set of interesting concepts will be smaller than the total set and thus allow us to develop more comprehensible graphs such as that in Figure 3 rather than the poor graph in Figure 4. The usefulness of a particular algorithm or technique will often depend on the data (Bradzil, Gamma and Henery 1994). We would like to offer a number of algorithms that could be used depending on the features of the dataset, just as graph layout software tends to offer different layout algorithms such as springer, hierarchical, orthogonal, etc. However, at this stage we have only considered use of C4.5 (Quinlan 1993) which is still considered to be the gold standard in the ML community.

## Mining the SNA data using C4.5

C4.5 and its predecessor ID3 were developed by Ross Quinlan for inducing decision trees from a set of examples. Rules can be extrapolated from the decision trees. C4.5 develops classification rules. An association rule is one which associates the presence of one set of items with the presence of another set of the form $X \rightarrow Y$ where $X, Y \subseteq I$ and $X \cap Y = \emptyset$ and I is the set of items or literals. (Fortin, Liu and Goebel 1996). A classification rule is in the same form but instead implies that one set of items (precedents, conditions or clauses) implies the truth of the consequent, antecedent or conclusion. In our approach we alter the rules generated by changing the class we are trying to learn. The basic ideas behind the ID3 technique include:

- In a decision tree each node corresponds to a non-categorical attribute and each arc to a possible value of that attribute. A leaf of the tree specifies the expected value of the categorical attribute for the records described by the path from the root to that leaf. [This defines what is a Decision Tree.]
- Each node in the decision tree should be associated with the non-categorical attribute which is most informative among the attributes not yet considered in the path from the root. [This establishes what is a "Good" decision tree.]
- Information theory is used to assist in choosing a useful property on which to split the tree by determining which property provides the most information gain. The elements you need to distinguish are the different values in the target concept, and the probabilities that are obtained from the proportion of each value in the training set at each node (Poole, Mackworth and Goebel 1998). Entropy is the notion used to measure how informative is a node. [This defines what we mean by "Good".]

C4.5 is an extension of ID3 that takes into account unavailable values and continuous attribute value ranges, performs pruning of decision trees and rule generation. The following formulae are used in the calculation of which attribute to split on.

**Formula 1-** the expected information requirement after T has been partitioned in accordance with the n outcomes of a test X.

$$\text{infoX}(T) = \sum_{i=1}^{n} \frac{|T_i|}{|T|} \times \text{info}(T_i)$$

**Formula 2 -** the average amount of information needed to identify a class of a case in T.

$$\text{info}(T) = \sum_{j=1}^{k} \left(\frac{\text{freq}(C_j,S)}{|S|}\right) \times \log_2 \left(\frac{\text{freq}(C_j,S)}{|S|}\right) \quad \text{bits.}$$

**Formula 3 -** the information gained by partitioning T in accordance with the test X.

$$\text{gain}(X) = \text{info}(T) - \text{infoX}(T)$$

**Formula 4 -** potential information generated by dividing T into n subsets.

$$\text{split info}(X) = \sum^{n} \underline{|T_i|} \times \log_2 (\underline{|T_i|})$$

i=1 | T | |T |

**Formula 5** = proportion of information generated by the split that is useful.

gain ratio(X) = gain (X) / split info (X)

Where C = a class, outcome, k = an attribute, n = number of different classes, outcomes,
j = the value of the attribute, i = the value of the class, T = set of training cases, X = a test

We had the C4.5 (version 8) software available to us. The formatting of the data involved converting from a UCINET style spreadsheet into two comma delimited files. One file was for the names of the attributes, values and classes to be tested, while the other held all the data in coded tables. This was achieved by first designing the table to be tested, ie. the attributes necessary and selecting the class. Following the technique used typically in ML we separated our data into a training set and a test set. The training set is used to deduce the rules, then these rules are run on a set of test data containing unseen objects to evaluate their accuracy.

In our first experiment we included the company type as one of the attributes, but soon realised, that one, it wasn't going to lead to any useful classifications since that attribute alone was sufficiently distinct to classify examples, and two; that the format of the dataset precluded this approach as there could be more than one value for this attribute per alliance. To overcome this problem we decided to group the company types into their respective sectors and have them repeated three times to allow for multiple partners for each alliance from the same sector. This initially was tested against the year each alliance was started, as it would have been interesting to see if any information or rules could be formulated that showed any trends towards the types of alliances formed each year and the sectors involved. This however resulted in a huge number of different classes and a high error rate rather than the small set we were looking for. A second attempt used the location of the alliances as the class. This proved more enlightening, from which real conclusions could be drawn. The output is given in Appendix 1. This result is quite interesting even without visual display as it indicates that the location of each alliance is almost exclusively dependant on whether or not a telecommunications company is involved. If the company is a supplier or manufacturer (coded as TM) then the focus is local on 10 of the alliances, but this is outweighed, if only slightly, by the 12 alliances where the alliance is international in focus (coded as TF). The 'n' in the above example indicates alliances where a telecommunication company was not involved in any way.

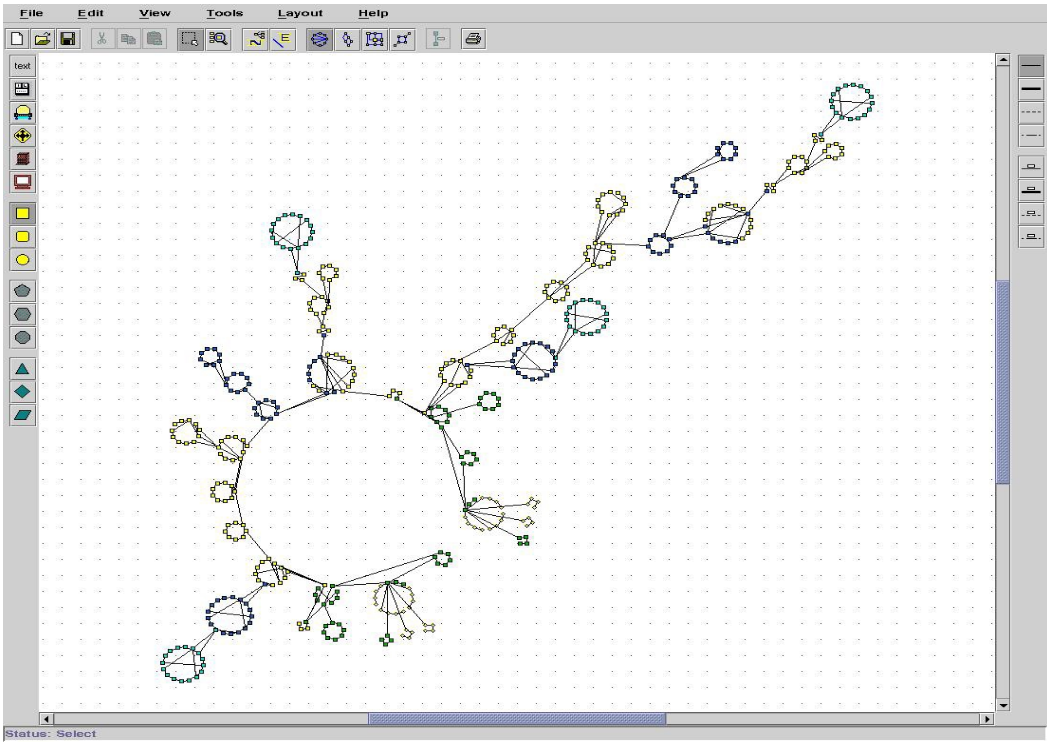


Figure 5: A graph from Tom Sawyer showing clear clustering

Similarly, we can rerun our experiments using different attributes as the class to be learnt. The tree that is drawn also shows the interaction between attributes. We can then choose to draw a diagram that only includes the interesting attributes and does not include the irrelevant ones. For example, based on the result in Appendix 1 we would look at a graph for companies where Telco=TM and another graph for companies where Telco=TF.

Within those less cluttered graphs we can discover who the central player/s are and if cliques exist. We may discover which features are shared and tend to encourage or identify convergence.

As part of this project we have also begun using the Tom Sawyer Graph Editor Toolkit which includes a Java Class Library of useful graph routines so that we may develop our own display tool to be used in conjunction with the ML algorithm. Using the software and our own Java code we plan to develop a tool that gives the user flexibility to build a graph using all the data, which might give some initial insights which can be used as the class to be learnt by C4.5. Below is an example of what Tom Sawyer can do and the effect we are trying to achieve in this project. You can see the clear clustering of the nodes and the layout is such that inferences can be drawn quite easily. This software also has scrolling and zooming capabilities, right down to a single node on screen. The nesting function should also prove handy for related sets of data, as it would have been possible to have a second-tier sociogram *within* a related node! This can be seen in the example below.

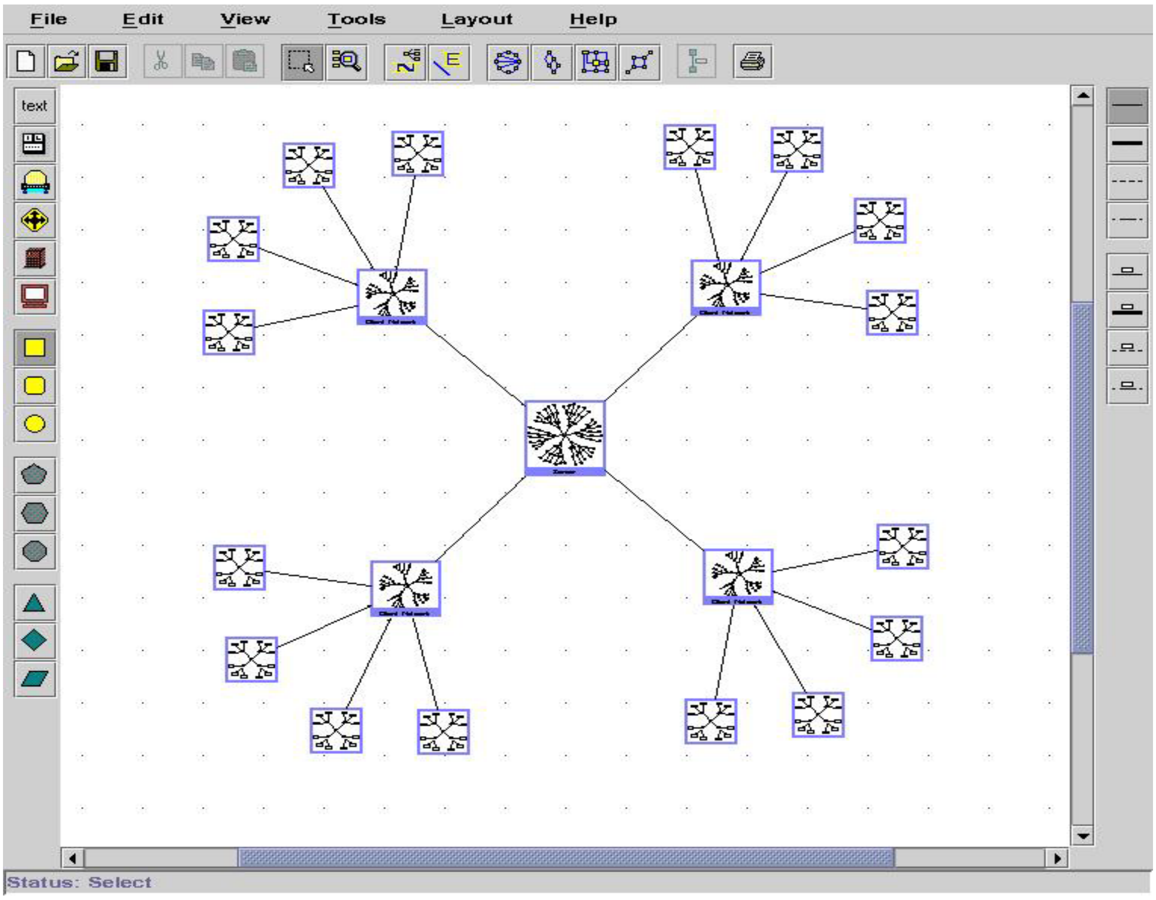


Figure 6: Another graph layout from Tom Sawyer

Here each node comprises a sociogram itself, each of which can be exploded and examined. Of course, these in turn could have nested nodes, thus layering information conceptually and reducing the amount of crowding on the screen.

We need to do further work with the Tom Sawyer Toolkit so that the rules output by C4.5 are displayed. We also need to design and implement an interface that will allow the user to select and manipulate the data in different ways. Figure 7 shows a pipe and filter system architecture which will support the use of different ML algorithms in the fourth subsystem.

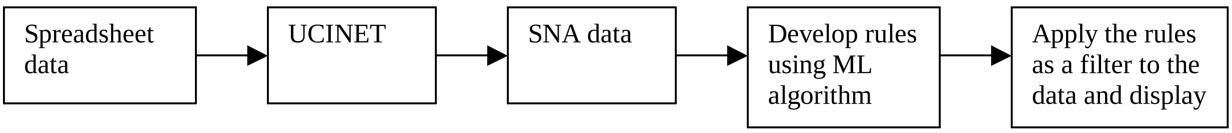


Figure 7: Proposed system architecture

## CONCLUSION

This project is concerned with the problem of handling large amounts of data and displaying it meaningfully. It is motivated by an exploratory case study into the existence of technological convergence in Australia using SNA data from a larger study. We found that display of the data was problematic and decided to add a layer of intelligent processing before display to reduce the amount of data to be shown. This is work in progress and we need to do more work on our graphics tool and exploring alternative algorithms. We would also like to look at the relationship between the SNA and ML algorithms and combine the techniques where possible. For example, we could use the centrality measure to identify which actor is the most important and then use the machine learner to identify what features in this actor make them more outstanding than the other actors. Our initial results show that a data mining approach can provide useful information that may not become apparent from the typical SNA techniques simply because it provides a way of focusing what is of interest to include in the analysis.

## REFERENCES

Bradzil, P. Gama, J. and Henery, B. (1994) Characterising the Applicability of Classification Algorithms Using Meta-level Learning *Technical Report LIACC University of Porto*, Rua Campo Alegre 823, 4100 Porto, Portugal.

BT World Communications Report 1998/99: http:www.bt.com/global_reports/bt_mci/section1a.htm.

Churcher, N, Keown, L and Irwin, W. (1999) Virtual Worlds for Software Visualisation *Proceedings of the Third Workshop on Software Visualisation*, Dec 3-4, 1999, University of Technology, Sydney, 9-16.

Convergence Report. May 2000 National Office for the Information Economy.

Deogun, J., Raghavan, V. and Sever, H. (1998) Association Mining and Formal Concept Analysis, In *Proceedings Sixth International Workshop on Rough Sets, Data Mining and Granular Computing*, Vol 1: 335-338.

Fortin, S., Liu, L. and Goebel, R. (1996) Multi-Level Association Rule Mining: An Object-Oriented Approach Based on Dynamic Hierarchies, *Technical Report TR 96-15*, Dept. of Computing Science, University of Alberta.

Freeman, L. (1998) Exploring Social Structure using Dynamic Three-Dimensional Colour Images In L. Freeman, C. Webster, D. Kirke (eds) *Social Networks*, 20:109-118.

Hollands, J.G., Carey, T.T., Matthews, M.L. and McCann, C.A. (1989) Presenting a Graphical Network: A Comparison of Performance using Fisheye and Scrolling Views *Proceedings of 3rd International Conference on Human Computer Interaction* September 1989, 313-320.

Köhler, W. (1947) *Gestalt Psychology*. New York: Liveright.

More, E. and McGrath, M. (2001) Strategic alliances as collaborative strategy or a method of implementing strategy: A case study in Australia's communications sector, Communications Research Forum, Canberra, 26-27 September, 2001, Department of Communications, Information, Technology and Arts.

Poole, D., Mackworth, A. and Goebel, R. (1998) *Computational Intelligence: A Logical Approach*. Oxford University Press.

Quinlan, J. Ross (1993) *C4-5 Programs for Machine Learning.* San Mateo: Morgan Kaufman Publishers Inc.

Sbarcea, K. (1999) Mapping the Pathways to Knowledge *Image and Data Manager* September/October 1999: 32-33.

Scott, J. (1992) *Social Network Analysis: A Handbook* SAGE Publications, California, USA.

## ACKNOWLEDGEMENTS

The authors wish to thank Mike McGrath and Elizabeth More for their supervision and assistance with the case study described in this paper and for providing the dataset used in the analysis.

## APPENDIX 1

**Results from C4.5 when trying to learn a decision tree/rules for the Location Class.**

```
C4.5 [release 8] decision tree generator
```

```
      Options:
        File stem <tel-directloc>
Read 82 cases (15 attributes) from tel-directloc.data

Decision Tree:
Telco = TE: Int (1.0)
Telco = TM: Loc (10.0/1.0)
Telco = TF: Int (12.0)
Telco = n:
      Other = OG: Loc (4.0/1.0)
      Other = OF: Loc (3.0)
      Other = OE: Loc (3.0)
      Other = OC: Int (2.0)
      Other = OT: Loc (3.0)
      Other = OU: Loc (0.0)
      Other = OM: Loc (2.0/1.0)
      Other = n:
            Media = MB: Loc (4.0/2.0)
            Media = MP: Loc (2.0)
            Media = MD: Int (1.0)
            Media = MM: Loc (0.0)
            Media = ME: Loc (0.0)
            Media = n:
                  Computing = CI: Loc (6.0/2.0)
                  Computing = CC: Loc (13.0/4.0)
                  Computing = CS: Loc (6.0/3.0)
                  Computing = CE: Loc (4.0/1.0)
                  Computing = CM: Loc (5.0/2.0)
                  Computing = n: Int (1.0)

Simplified Decision Tree:
Telco = TE: Int (1.0/0.8)
Telco = TM: Loc (10.0/2.4)
Telco = TF: Int (12.0/1.3)
Telco = n: Loc (59.0/23.1)
Tree Saved

Evaluation on training data (82 items)
      Before Pruning                After Pruning
      Size        Errors            Size        Errors      Estimate
      25          17(29.7%)         5           21(25.6%)   (33.7%)
```

## APPENDIX 2

### Classification of Organisations

| | |
|---|---|
| T | Telecommunications Sector |
| TC | Traditional carriers – long established licensed providers of telephone, cable, satellite, network communication services |
| TE | Emerging carriers – relatively new licensed providers of telecommunications services, may have started out as resellers or rebillers |
| TM | Telecommunications Equipment Manufacturers and Suppliers – businesses engaged in the manufacture and distribution of telco equipment; businesses engaged in cable laying and transmission line construction |
| TF | Foreign Carriers – licensed providers of telecommunications services based outside Australia |
| C | Information and Communications Technology Section |
| CI | Internet Services – businesses that include Internet SPs, Applications SPs, Content SPs |
| CC | Computer Hardware Wholesalers and Software Developers – includes e-commerce solutions |
| CS | Communications Services – includes call centre services, outsourcing |
| CE | E-commerce |
| CM | Multi-IT |

| | |
|---|---|
| M | Cultural Recreational Services |
| MB | Broadcasting Media – includes television operators (free-to-air, pay-TV, cable TV) and radio operators |
| MP | Print Media – includes newspapers, magazines |
| MD | Diversified Media |
| MM | Multimedia |
| ME | Entertainment – movies |
| U | Utilities and Construction Services |
| UE | Energy |
| UO | Oil |
| UG | Gas |
| UW | Water |
| UM | Multi-Utility |
| UP | Postal and Courier Services |
| UD | Developer/constsruction |
| O | Others |
| OG | Government Bodies |
| OF | Financial Institutions |
| OE | Electronics |
| OC | Management/Consultants/Administration |
| OT | Travel |
| OU | Universities |
| OM | Manufacturers/Retailers |

**Scope of Alliance**

| | |
|---|---|
| Loc | Local |
| Int | International |

## COPYRIGHT