Using decision tree in business collaborator

Chao Sun
University of Wollongong, chaos@uow.edu.au

Yu Zhang
University of Wollongong, azhang@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/commpapers

Part of the Business Commons, and the Social and Behavioral Sciences Commons

Recommended Citation
Using decision tree in business collaborator

Abstract
Purpose - Business collaboration is important for small and medium sized enterprises. The traditional method of choosing business collaborator is largely based on individual's experience and subjective criteria. However, the failure rate of business collaboration is still high for less experienced small firms. The purpose of this research is to find a different solution for managers in choosing business collaborators.

Methodology - Decision Tree is an advanced technology, which is used in different business and industry areas. It is adopted in this study to help the managers choosing business partners. This study using the data collected from 339 firms in Australia and China to examine the key determinants for successful collaboration.

Findings - The performance of business collaboration is influenced by different factors in different countries. Decision Tree gives good implications for small business decision making in collaborating strategy.

Value - This study adopted a new method in collaboration studies. It also distinguished the differences of key determinants for business collaboration in Australia and China.

Keywords
Using, decision, tree, business, collaborator

Disciplines
Business | Social and Behavioral Sciences

Publication Details
Abstract

**Purpose** – Business collaboration is important for small and medium sized enterprises. The traditional method of choosing business collaborator is largely based on individual’s experience and subjective criteria. However, the failure rate of business collaboration is still high for less experienced small firms. The purpose of this research is to find a different solution for managers in choosing business collaborators.

**Methodology** – Decision Tree is an advanced technology, which is used in different business and industry areas. It is adopted in this study to help the managers choosing business partners. This study using the data collected from 339 firms in Australia and China to examine the key determinants for successful collaboration.

**Findings** – The performance of business collaboration is influenced by different factors in different countries. Decision Tree gives good implications for small business decision making in collaborating strategy.

**Value** – This study adopted a new method in collaboration studies. It also distinguished the differences of key determinants for business collaboration in Australia and China.

**Keywords:** Collaboration, Decision Tree, Australia, China

1. **Introduction**

Business collaboration is important for small and medium sized enterprises. The traditional method of choosing business collaborator is largely based on individual’s experience and subjective criteria. However, the failure rate of business collaboration is still high for less experienced small firms. Zhang et al., (2009) proposed and adopted computer intelligence method in solving business collaboration problems.

The purpose of this research is to examine the key factors for business
collaboration using a new method, Decision Tree (DT). Most of the previous empirical studies focused on developed countries, especially U.S., Japan, and some Europe countries. However, there is a gap in the research to compare the differences of Australian and Chinese firms. Therefore, this paper focused on Australia and China.

2. Literature Review

Many researchers have studied inter-firm collaboration from different perspective. Collaboration among firms can be fruitfully examined from a wide range of theory. They include transaction cost economics, agency theory, network theory, the behavioral theories, property rights theory, economic empirical studies, strategic management, both in its positioning and resource based complementary perspectives, dynamic capabilities theory, real option theory, and institutional theories.

2.1 Transaction Cost Theory

One of the most important and basic economic theories of inter-firm relationships is Transaction Cost theory. Transaction cost theory regards the basic choice in organizing economic (Faulkner and Rond, 2000). Poppo and Zenger (2002) and Harrison (2004) regard transaction cost economics (TCE) as the common framework for understanding governance arrangements. Williamson (1975) highlighted the important influence of opportunism and bounded rationality on inter-firm collaboration. However, Williamson has been criticized for ignoring the role of power in markets and hierarchy (Francis, Turk, and Willman, 1983).

Transaction Cost theory is also criticized as it ignores many factors important to business collaboration (Doz and Prahalad, 1991; Gulati, 1998; Powell, 1990). Therefore, the important contributions of Resource Based View on exploring other types of collaboration, the dynamic of business transactions, and the key roles of trust become good supplementary to Transaction Cost Theory.

2.2 Resource Based View

Although generated from the discipline of economics, Resource Based View was also greatly contributed by the study of strategic management. Many researchers from economic studies and business and management studies did research on Resource Based View and contributed many profound results to this theory.

Resource based theories have examined the formation of collaboration (Pfeffer and Nowak, 1976) and shade a light on the dynamic of collaboration (Rumelt, 1991). Tallman (2000) linked the resource-based view with transaction cost theory and argued collaboration provides firms with complementary capabilities.
However, resource based view also received many criticizes. Gulati (1995) argued resource based view does not adequately account for alliance formation. Dyer and Singh (1998) also argued that according to resource based view, an individual firm should attempt to protect, rather than share knowledge. On the other hand some phrases are used loosely and interchangeably in Resource Based studies (Kale, 1999).

Among many theories that studied business collaboration, Transaction Cost Theory and Resource-Based Theory are two of the most important theories, which are closely tied with all the other theories. Transaction Cost Theory is the original and basic theory dealing with firms and enterprises. Resource-Based Theory, however, is widely used in recent researches and linked closely with many management and business studies.

2.3 Theoretical Framework

To study the key determinants of a successful inter-firm collaboration, the primary task is to determine “success” collaboration. Koh and Venkatraman (1991), Balakrishnan and Koza (1993), and Anand and Khanna (1997) used the event-study analyses on the stock market effects of alliance announcements. However, the majority of small and medium sized private firms are leaved out of the model. Baum and Oliver (1991, 1992), and Mitchell and Singh (1996) examined the relationship between firms in alliances and the likelihood of their survival. However, this may greatly depend on social environments and is hardly implemented in another nation or different period of time. The criterion may be very different for each industry and even for each firm (Gulati, 1998).

Managerial researchers took performance in terms of their overall satisfaction as another method used to study alliance results. Empirical results showed that both subjective and objective assessments are significant in measuring alliances’ performance and result (Garvis, 2000). Therefore, success will be measured by both objective and subjective method as supplementary to each other in this study.

From both economic and management literatures and empirical studies, the most important determinants to successful collaborations are categorized as in Figure 1 below.

2.4 Decision Tree Technique

The decision tree techniques are popular in the machine learning domain, as it's a natural and intuitive way to analyse problems with non-metric data. A decision tree grows from a root node, and this node is connected by successive (directional) links to other nodes. The nodes are connected similarly until leaf nodes are reached with no further links. Each node corresponds to a test on certain attribute value, and the corresponding links lead to the possible outcomes of the test. The
The 8th SMEs in a Global Economy Conference 2011: "Rising to the Global Challenge: Entrepreneurship and SMEs development in Asia"

links terminate at leaf nodes which contain class values as the decision from all the previous tests in the tree branch.

Many decision tree algorithms, such as ID3 [Quinlan, 1986], C4.5 [Quinlan, 1993] or CART [Breiman et al., 1984], recursively selecting the best attribute-value pair to construct current node according to the information entropy change after segmenting the whole training data by current node. A typical decision trees analysis includes two stages: Firstly, generate tree model based on a set of training data, which is essentially gathering the knowledge; then use testing data to evaluate the model or apply the tree model for predicting. The model can also be analysed for understanding the problem.

The decision trees techniques are reliable classifiers, particularly at modelling non-metric data. The attribute conditions used by a tree model generally reflect the natural characteristics of the problem. Since tree models are simple to understand and easy to interpret, it has been widely used in conjunction with other methodologies in various research domains for both classification and prediction purposes. The decision trees were actively used in areas with case-by-case data, such as medical research, financial prediction and marketing strategy. Xie and Zhao (2010) used ID3 algorithm to analyse the customers’ satisfaction degree for technology-supported company, in order to find out the main factors that are heavily related to the customers' satisfaction. Small Business Credit Scoring models were built with CART decision tree and a few other artificial intelligent methods on a Croatian bank dataset (Zekic-Susac et al. 2004). Guo et al. (2006) employed ID3 decision trees to analyse and build models for the customer churning in securities business. Due to the information digitalisation and data boost, the decision trees are used in more and more research areas, however not many examples in the economics domain.

![Figure 1: Framework of key determinants for successful business collaboration](image_url)
Table 1: Quantitative surveys in Australia and China

<table>
<thead>
<tr>
<th>Basic Descriptive Statistics</th>
<th>Total sample: 339</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Local</th>
<th>Foreign (Multination)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>96 (96%)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>China</td>
<td>209 (87.5%)</td>
<td>30 (12.5%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Size (* defined by country)</th>
<th>Micro</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia Size</td>
<td>91 (62.3%)</td>
<td>2 (1.4%)</td>
<td>2 (1.4%)</td>
<td>51 (34.9%)</td>
</tr>
<tr>
<td>China</td>
<td>79 (33.1%)</td>
<td>60 (25.1%)</td>
<td>25 (10.5%)</td>
<td>75 (31.4%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trust</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust level</td>
<td>3.75</td>
<td>0.95</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Risk level</td>
<td>2.32</td>
<td>0.88</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Similar goal</td>
<td>3.28</td>
<td>1.05</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Similar structure</td>
<td>2.53</td>
<td>1.17</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Similar process</td>
<td>2.63</td>
<td>1.17</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Reliable contact person</td>
<td>3.73</td>
<td>0.93</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Participation in business networks</td>
<td>3.22</td>
<td>1.23</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Openness on information</td>
<td>0.17</td>
<td>0.81</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Efficiency of communication</td>
<td>2.58</td>
<td>0.72</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Understanding of communication</td>
<td>2.70</td>
<td>0.79</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Frequency of communication</td>
<td>2.54</td>
<td>0.88</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Objective culture similarity</td>
<td>0.47</td>
<td>1.03</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Language similarity</td>
<td>3.96</td>
<td>1.30</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Religion similarity</td>
<td>3.56</td>
<td>1.51</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final Success</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective success</td>
<td>1.62</td>
<td>0.90</td>
<td>0</td>
<td>4.75</td>
</tr>
<tr>
<td>Subjective success</td>
<td>2.50</td>
<td>1.07</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Fulfill expectation</td>
<td>2.40</td>
<td>1.05</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
3. Methodology

3.1 Data collection

A quantitative survey was conducted in both Australian and China from 15th May to 6th Jul, 2010. The selected participants are taken from three sources:

1. Australian Telecommunications Industry Ombudsman (TIO) lists,
2. The researcher’s former business networks, and,
3. Extension of the researcher’s business networks.

An online survey system was developed by the researcher to save the costs and make it more convenience for the interviewees.

As a result, 342 online surveys were collected from both Australia and China, including 3 invalid (uncompleted) surveys. Therefore, the final valid surveys are 339, including 239 firms from China and 100 firms from Australia. The statistics of data are presented in Table 1.

Face-to-face interviews are also conducted in both Australia and China with 31 firms. The interviewees include CEOs, key managers, and senior executives, who have a good knowledge on collaboration and development strategy of the firm. This qualitative result is expected to provide complimentary evidence for quantitative study.

3.2 Data Structure

The survey data are structured as a matrix, in which each row represents a company and elements in the row are the answers for survey questions. Because many questions are designed to be answered in quantified scales, e.g. strongly disagree to strongly agree, or the time being established, they are converted into continuous numbers. But some other questions may produce non-quantitative answers such as the locations or business type, and they are also converted into numbers but associated to scattered type specified in a .names file.

Artificial variables are excluded from the dataset, for example the target variable Final Success is a computational result of the objective and subjective success, thus both objective and subjective successes cannot be used as inputs of the decision tree. The Final Success is quantified into five classes (1 to 5) for decision tree classification.

3.3 Decision Tree

A commercial decision trees application See-5 is used in this work to generate models from the data. The See-5 is produced by RuleQuest.com and it utilises improved algorithm of the widely used decision tree algorithm C4.5.

A simple See5 models can be viewed as a tree structure shown in Figure 2. In this example, the most important variable is the Trust, as the root node.
use *Trust* to separate the whole dataset into two sections. And similarly variable *pLocal* and *Size* are the next key factors for further purifying the cases. This indicates that the importance of variables can be sorted by analysing their usages in a See5 model.

The performance of a model can be presented as a confusion matrix, which summarises how the cases in data are classified by the model against their real classes. In a confusion matrix, rows stand for the real classes and columns is the classified result. For example, the 5-by-5 confusion matrix in Table 2.1 represents a dataset with 272 cases in total, and there are 5 classes from 1 to 5. If a case is classified as class-N and actually belongs to class-M, then element [M, N] (row, col) adds 1. And if a case is correctly classified, it will be located in the main diagonal of the confusion matrix [M, M].

This is to say that the main diagonal of the confusion matrix contains the number of correctly classified cases and the rest stands for the numbers of misclassifications.

![Figure 2: An example of See5 Tree Model](image)

**Table 2: Confusion Matrix on a See5 model**

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>&lt;classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>3</td>
<td></td>
<td></td>
<td>(a): Class 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>&lt;classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(a): Class 1</td>
</tr>
</tbody>
</table>
A model performs well on the training data doesn't guarantee it's also good for predicting. In order to evaluate the model for future cases prediction, the data normally are separated into training and testing data randomly at the ratio of 80:20. Models built on the training data are then evaluated with the testing data. As shown in Table 2, the accuracy is generally lower than evaluating with training data, as the testing data may contain unknown conditions. In the above example, the error rate rise from 20% to about 40% for evaluation the same See5 model on testing data.

### 3.4 Using Soft Threshold on Error Calculation

The above decision tree model doesn't provide a satisfying accuracy on testing data, and the reason may lie in the way how the classes are defined. It's understood that objectively there are only two results for collaborations: Success or Fail. However in this study the final success is a combination of both objective and subjective result, thus it has 5 classes.

The reason is that the differences between these classes aren't distinct. For instance, a case belongs to class-2 may have a real value of 1.85, which is closer to a class-3 case (real value 2.1) than other cases in class-2. In addition, subjective variances among participants make the borders between classes even more blurred, and one person's Totally Agree may only be Slight Agree to another person.

### Table 3: Soft-thresholds evaluation on testing data, Adjusted Error rate 19.8%

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>&lt;classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
<td>(a): class 1</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>3</td>
<td></td>
<td></td>
<td>(b): class 2</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>3</td>
<td></td>
<td></td>
<td>(c): class 3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>(d): class 4</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(e): class 5</td>
</tr>
</tbody>
</table>

Soft thresholds are employed to calculate the error rates in the rest of this paper to adapt this situation. With the soft thresholds, if a classification is only one class away from the real (green cells in Table 3), it's treated as only half wrong and contribute less to the error rate. Calculating this way, the error rate on previous testing data evaluation is reduced by half.

### 3.5 Models Stability and Variable Importance

With the same parameter settings, the decision tree models and their performance may still vary when different training and testing data are
selected. An exhaust modelling method is used to build 500 See5 models on randomly segmented datasets with 80:20 train/test ratio. The error rates from all testing data evaluation are recorded and plotted as a histogram in Figure 3.

![Figure 3: Histogram of 500 random See5 models](image)

In Figure 3, the error rates from all 500 models show a Gaussian distribution with mean value at around 21%. Since the exhaust test covers a large number of train/test situations, this indicates that the See5 modelling is generally effective and useful for modelling and predicting the business collaboration problem in this work.

The usage of variables can be summarised from the See-5 models, and here are a few statistical indices from the models generated in exhaust test:

1. **Variable Exposure Rate**: the number of models used certain variable divide by the total model number;
2. **Variable Usage**: how a variable contribute to a decision tree model and its Average and Standard Deviation over all models.

Because an important variable should heavily and stably contribute to the decision making procedure, we define the importance of a variable as its exposure rate divide by its coefficient of variation (std/mean):

\[
\text{Importance} = \frac{\text{ExpoRate} \times \text{mean(Usage)}}{\text{Std(Usage)}}
\]

### 4. Results and Discussions

#### 4.1 Quantitative results

The quantitative analysis on Chinese-Australian dataset shows business trust, communication, and firm size play important role in business collaboration. Table 4 lists the top 10 important variables sorted by their importance, based on 500 See-5 models in the exhaust test. *Trust, Size* and *CommFreq* all have significantly higher exposure rate and importance than other variables.

![Table 4: Top 10 Important Variables – CN-AU](table)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Exposure Rate (%)</th>
<th>Variable Usage</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td>98.2</td>
<td>89.4</td>
<td>20.7</td>
</tr>
<tr>
<td>Size</td>
<td>94.8</td>
<td>62.1</td>
<td>19.0</td>
</tr>
</tbody>
</table>

Khon Kaen University, Nong Khai Campus, T.Nongkomkor Muang Nongkhai 4300, Thailand
Tel: +66 42 415600 Ext. 46641 Fax: +66 42 415699  http://www.nkc.kku.ac.th/smesconference2011
In order to compare the firm collaborations between Australia and China, two exhaust modeling are conducted on Chinese and Australian only data respectively, and 100 models are built for each country. Figure 4 shows the histograms of error rates for both Chinese and Australian companies, and models for both datasets are showing consistent performance with reasonably low error rates. And the top important variables for Australian and Chinese business can be seen in Table 5 and 6 respectively.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Exposure Rate (%)</th>
<th>Average</th>
<th>Std Dev</th>
<th>Importance Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td>94</td>
<td>99.1</td>
<td>0.3</td>
<td>57.66</td>
</tr>
<tr>
<td>commFreq</td>
<td>90</td>
<td>89.6</td>
<td>11.9</td>
<td>1.359</td>
</tr>
<tr>
<td>commEfficient</td>
<td>25</td>
<td>26.6</td>
<td>4.1</td>
<td>0.324</td>
</tr>
<tr>
<td>infoShare</td>
<td>33</td>
<td>77.2</td>
<td>27.1</td>
<td>0.188</td>
</tr>
<tr>
<td>network</td>
<td>22</td>
<td>79.4</td>
<td>26.4</td>
<td>0.133</td>
</tr>
<tr>
<td>simTech</td>
<td>33</td>
<td>41</td>
<td>24.8</td>
<td>0.109</td>
</tr>
<tr>
<td>commUnderstand</td>
<td>14</td>
<td>55.9</td>
<td>22.2</td>
<td>0.071</td>
</tr>
<tr>
<td>risk</td>
<td>18</td>
<td>53.7</td>
<td>27.7</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 6: Top 10 Important Variables – CN
Comparing Table 5 and 6, it’s obvious that in both Australia and China, the most important variable is always Trust, and its importance are much higher than all other variables in both countries. Except this, the variance on See5 models is quite significant between both countries.

Table 6 suggests that in China, the success of business collaboration relates to multiple factors besides Trust, such as the size of company (Size), whether partner is local company (pLocal), previous experience (pExp), similarity in language (simLang), culture (simCulture) and company goals (simGoal). Communication frequency (commFreq) is also important because it’s used by 64 out of 100 models, however its usage varies a lot and not always stable.

However, decision making for firm collaboration in Australia seems to be much simpler than in China according to the See5 models. Trust and Communication Frequency are almost the only two variables used by the See5 models. 94 and 90 out of 100 models use Trust and commFreq to make decision, whilst all other variables are used here and there in less than one third models.

Table 7 gives a comparison of all the individual variable importance between China, Australia and overall models on all three datasets. In this table, not only important variables are shown, some variables are shown to be not important in every dataset, such...
4.2 Decision tree Analysis

One thing need to be noted is that the above statistical results only distinguish variables that are important and less important, but not how they actually affect the final success, e.g. positive or negative effects. This can be done by analysing the actual See-5 models. For example, List 1 is a See-5 model for Australian only data,

List 1: An example of Australian See-5 Model

<table>
<thead>
<tr>
<th></th>
<th>2:1 (5.1/0.1)</th>
<th>3 (53/11)</th>
<th>2 (13.9/2)</th>
<th>3 (8/2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>trust</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trust in [3,4,5]:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...commFreq in [3,4]:</td>
<td>3 (53/11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...commFreq in [1,2]:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...simTech in [1,4,5]:</td>
<td>2 (13.9/2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>simTech in [2,3]:</td>
<td>3 (8/2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the above List 1, *Trust* is the root node of the tree. If *Trust* is 2, then the collaboration will very likely to be fail (Class-1). If *Trust* is in mid-high levels (3,4,5), then the *Final Success* can be either class 2 or 3, i.e. *Trust* is having positive effect to *Final Success*.

*CommFreq*={3,4} leads to better success (3) and *commFreq*={1,2} leads to either class 2 or 3, this indicates *commFreq* also has positive effect. However things are different for variable *simTech*: when *Trust* is in mid-high levels (3,4,5), and *commFreq* is in low level (1,2), two companies having extremely similar technology (4,5) or extremely different technology (1) will both reduce the chance of success collaboration. Only when there are tech-wise difference between two companies (2,3), the collaborations are more likely to be success.

As mentioned before, Australian models are much simpler than Chinese models. One possible reason is that the Australian dataset has smaller size than Chinese dataset. Although the models are simple, the information gathered from the analysis is still valid and interesting. Because of the page limitation in this paper, analysis on Chinese models will not be discussed, however an example tree model is provided as List 2 in Appendix for reader’s interest.

4.3 Future Research

This paper focused on business collaboration in Australia and China. As the result supported that determinants for successful business collaboration are different in different countries, further researches should be conducted in other countries. However, methods and questions should be changed due to special environment and official definition of firm size in those countries. Further research should also take account in different cultures, industries, and technologies for the studied firms.

5. Conclusions

Different factors may provide very different contribution to collaborating result in different countries. Trust
plays a significant role in business collaboration in both Australia and China. Firm size plays more important role in business collaboration in China. The frequency of communication plays more important role in business collaboration in Australia. The results underscore the fact that collaboration research should be conducted separately for different countries. One factor that proved to be vital to business collaboration in one country does not necessarily important in another country.

It should be argued that the model for business collaboration should be adjusted to suit the different environments in different countries or regions. Decision Tree is a good supplement method for managers and decision makers when planning business collaborating strategies.

References


Appendix

List 2: An example of Chinese See-5 Model

\[
\begin{align*}
\text{trust in } \{1,2,3\} & : \\
\text{...SimCulture in } \{2,3\} & : \\
\text{...commUnderstand = 1} & : 2 (3.1/0.1) \\
\text{...commUnderstand in } \{2,3,4\} & : 3 (24.7/7) \\
\text{...SimCulture in } \{1,4,5\} & : \\
\text{...commUnderstand = 4} & : 2 (0) \\
\text{...commUnderstand in } \{1,2\} & : \\
\text{...Pmulti in } \{0,1,3,4\} & : 2 (28.6/10.6) \\
\text{...Pmulti = 2} & : 1 (2) \\
\text{...commUnderstand = 3} & : \\
\text{...SimCulture in } \{1,5\} & : 2 (4) \\
\text{...SimCulture = 4} & : \\
\text{...simGoal in } \{2,3\} & : 2 (5.4/1.4) \\
\text{...simGoal in } \{1,4,5\} & : 3 (5) \\
\text{trust in } \{4,5\} & : \\
\text{...Size > 2} & : \\
\text{...SimTech in } \{1,2\} & : 3 (17.5/1) \\
\text{...SimTech = 5} & : 4 (8.9/2.9) \\
\text{...SimTech = 3} & : \\
\text{...Psize <= 3} & : 4 (4.4/1.1) \\
\text{...Psize > 3} & : 3 (10.1/3.7) \\
\text{...SimTech = 4} & : \\
\text{...SimGoal in } \{1,5\} & : 4 (3) \\
\text{...SimGoal in } \{2,3,4\} & : 3 (22.3/3) \\
\text{Size <= 2} & : \\
\text{...infoShare in } \{1,2,4,5,7,8,9\} & : 3 (7/1) \\
\text{...infoShare = 3} & : 2 (2) \\
\text{...infoShare = 0} & : \\
\text{...SimLanguage in } \{1,2,3\} & : 3 (7/1) \\
\text{...SimLanguage = 4} & : \\
\text{...Pexp <= 3} & : 2 (5) \\
\text{...Pexp > 3} & : 3 (3) \\
\text{...SimLanguage = 5} & : \\
\text{...commFreq = 4} & : 4 (3/1) \\
\text{...commFreq in } \{1,2,3\} & : \\
\text{...simGoal in } \{2,4\} & : 3 (13/3) \\
\text{...simGoal in } \{1,3,5\} & : 2 (12/4)
\end{align*}
\]