Statistical Disclosure Control: To Trust or Not to Trust

Helen Giggins Ljiljana Brankovic
School of Electrical Engineering and Computer Science
The University of Newcastle
University Drive, Callaghan NSW 2308, Australia
{Helen.Giggins, Ljiljana.Brankovic}@newcastle.edu.au

Abstract

The statistical analysis of data stored in data warehouses is an important phase in the organisation's strategic planning process. For the maximum benefit to be gained from such data warehouses, a relationship of trust needs to exist between all parties involved. In this paper we investigate its importance with respect to the statistical security problem. Understanding trust relationships in this context is particularly crucial since an individual's privacy cannot be guaranteed using traditional security mechanisms.

1. Introduction

The past few decades have witnessed ever increasing amounts of data being collected and stored by both industry and governments alike. The potential benefits that arise from the analysis of these data range from improved market analysis, via strategic planning, to scientific research and development. Due to the typically multidimensional and hierarchical nature of data warehouses, OLAP operations allow for the analysis of data at varying levels of aggregation [6]. In this paper we focus on the statistical analysis of the data in a warehouse at its lowest level of aggregation. This view of the data is equivalent to microdata in Statistical Database literature [17]. In this context users can only retrieve aggregate statistics such as SUM, COUNT, MEAN and AVERAGE. The aim of Statistical Disclosure Control (SDC) is to provide the highest quality statistics while also preventing the disclosure of values from individual records.

One application of a statistical data warehouse in the area of health administration is Australia's controversial HealthConnect initiative [13]. The aim of the scheme is to implement electronic health records that can be linked and shared across various organisations. However, as such data warehouses will contain personal or otherwise sensitive information about patients, there is a potential for the invasion of the individual's privacy. In general, whenever private and sensitive information is collected about an individual, there is a potential to breach that person's privacy.

Recent work in the area of computer security has seen a shift from so-called hard security methods, such as authentication and access control, to soft security or social control mechanisms [8]. In this current environment there is a movement towards the analysis of trust and risk in the field of Trust Management [7]. The main objective of this paper is to apply the principles of trust management to the area of SDC so as to better understand the important role trust plays in such systems.

The structure of the remainder of the paper is as follows. In the next section we define trust and discuss various types of trust and the difference between trust and distrust. We then introduce the statistical security problem and stress the importance of trust to the collection and management of data. We also discuss the trust relationships that exist in the context of a secure statistical data warehouse system. In Section 4 we present a model of a secure statistical database system with trust as an essential component. We finish with a discussion of future research directions and concluding remarks.

2. Trust

Trust is an intrinsic part of the human experience. Yet, for a concept so fundamental to our very existence, it is a mysterious beast, both challenging to pin down and awkward to model. In order to gain a better understanding of trust we first establish the context and the stakeholders within it. For the purpose of this paper we define a trust relationship as an asymmetric relationship occurring between two parties, the trustor and trustee. The trustor is the trusting party, while the trustee is the trusted party. In effect, the trustor is placing their trust in the trustee. In our secure statistical data warehouse system we have three stakeholders, namely the *Data Source*, *Data Manager*, and *Data User*.

The Data Source is the person, or system, providing their information to the Data Manager. The Data Manager is

978-0-7695-3428-2/08 $25.00 © 2008 IEEE
DOI 10.1109/CSA.2008.12

responsible for the collection of data and the creation and management of the data warehouse. The Data User relies on the Data Manager to provide high quality data on which they then perform various statistical queries. We discuss the complex trust relationships between these various stakeholders in more detail in Section 3 and Section 4.

2.1. Trust Considered

To aid our discussion of the various properties and causal factors related to trust, we give the definition of trust proposed in [7].

> **Definition 1 (Trust)** *Trust is the extent to which a given party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible.*

The notion of *dependence* is at the heart of trust, and is reflected in the "willing to depend" component of the above trust definition. The need to delegate a task to another person is a necessary condition for trust [7, 5]. By trusting another, we are delegating an important task which we would otherwise be unable to complete ourselves, or at least not easily. Closely related to the notion of dependence is *reciprocity* defined by Mui et al. as a mutual exchange of deeds, both positive and negative in outcome. The authors argue that reciprocative actions help a person to acquire a *reputation* [14].

Reputation and credentials are also closely linked to trust, particularly in relation to open distributed systems such as the Internet [1, 7]. These concepts embody the "feeling of relative security" component of the above trust definition. A reputation allows us to infer something about future behaviour based on an informed observation of past actions [1]. Reputation systems allow for a reputation to be accrued and assessed on a global scale without the usual face to face personal interactions [7].

Perhaps the most discussed component of any trust definition is the notion of *risk*, embodied in the above trust definition as the possibility of "negative consequences". While risk is not directly part of any trust definition, it is strongly related to the likelihood of cooperation occurring between the trusting parties.

We note that *context* is very important to any definition of trust, as shown by the excerpt "in a given situation" in Definition 1. In his formalism of trust for artificial agent systems, Marsh [11] discusses the importance of context to trust in a given situation, and particularly the utility and importance of the outcomes of the trusting relationship. Another quality of trust relationships is that they are fluid or dynamic, not only depending on context, but also changing over time [11].

Before looking at different types of trust, we briefly outline work done on defining trustee characteristics. McKnight and Chervany [12] have compared cross-disciplinary definitions of trust and extracted four high-level trustee or trust referent characteristics of the trustee; *benevolence*, *integrity*, *competence*, and *predictability*. A *benevolent* party acts in a caring way in order to achieve the greater good, rather than acting opportunistically. *Integrity* refers to acting truthfully, and acting in good faith. *Competence* is the ability or power to achieve what is needed. *Predictability* refers to a consistency of trustee actions in a given situation, be they of good or bad consequence.

Having discussed various important aspects of trust and its importance in forming collaborative relationships, we now briefly analyse the fundamental differences between trust and distrust.

2.2. Trust and Distrust

Just as trust is fundamental to the formation of cooperative endeavours, distrust can just as quickly break down relationships and stall collaboration. Distrust has often been described as the opposite end to trust along a continuum, where trust values fall in the range [-1,1) with total distrust being -1, no trust being 0 and close to 1 being high trust [10]. This definition precludes the idea that we can have both a high level of trust and distrust co-existing. However, it is easy to find a counter-example, for instance we can imagine allies in war simultaneously trusting and distrusting each other [12].

Marsh and Dibben [10] perform a detailed analysis of various aspects of trust and distrust and come up with four separate terms: *Trust*, *Untrust*, *Distrust* and *Mistrust*. They describe trust in a similar way to Definition 1, and distrust is defined as a negative measure of how much the trustee is actively working against the trustor to prevent the completion of the delegated task [10]. Trust is placed in the trustee only when the level of trust is higher than some cooperation threshold [11]. This means that there is a gap between trust and distrust, termed untrust, which describes how little the trustee is trusted [10]. A distinction between misplaced trust and distrust is also conjectured. What Marsh and Dibben term mistrust results from a default of trust in a given situation [10].

The more accepted view is that trust and distrust are opposites of one another, yet also separate constructs [9, 12]. Lewicki et al. have outlined a theory of trust whereby high trust and distrust can coexist in complex relationships [9]. However, if dealing with a specific context, it no longer makes sense to talk about concurrently high trust and high distrust. The same applies for the opposite end of the spectrum, when we have coexisting low trust and low distrust [12].

3. Statistical Disclosure Control

In this section we introduce the Statistical Disclosure Control (SDC) problem as it relates to statistical data warehouse systems, and discuss the complex trust relationships that exist between the system stakeholders.

The way in which data in the data warehouse is collected can impact greatly on the accuracy and completeness of the information stored; without quality data it is impossible to extract quality statistics. We argue that when a person has a low level of trust in an organisation, they are likely to provide false information. On the other hand, when they stand to gain nothing from participating and face no possible punishment, they are likely not to provide their data at all.

So what are the potential consequences for Data Managers when their Data Source has a low level of trust or high level of risk involved in providing their data for future statistical analysis? Clearly, having a reduced number of people willing to participate and provide information is a problem as it can impact on future data collection [16]. However, a potentially more damaging outcome for Data Managers is when individuals provide false or misleading information, which is likely to occur when the individuals do not have the option to withhold their information.

3.1. SDC Problem

The type of system we consider is a data warehouse that only allows for statistical queries to be performed on the data. There are two key conflicting goals in such a system. Firstly, the Data Manager wants to ensure that sensitive information relating to individual records in the data warehouse is not disclosed by answering queries. Secondly, the Data Manager wants to ensure the highest accuracy of released statistics are provided to the Data User. These goals are by their very nature in conflict, a higher level of security (privacy) implies a lower quality/amount of released statistics and vice versa. The real problem faced by the Data Manager is how to find the best balance between these conflicting goals. This is known in literature as the Statistical Security problem or the Statistical Disclosure Control (SDC) problem. Comprehensive overviews of this topic can be found in [2, 4, 3, 16, 17].

To ensure the integrity of the data warehouse system, it should be impossible for users to infer confidential values from any sequence of aggregate values. A situation where a user is able to determine some or all individual values is termed a database compromise or statistical disclosure. An obvious first step in protecting the statistical data warehouse from such an individual would be to remove all direct identifiers from the database. However, this alone is not enough to anonymise the data [17], hence the statistical security problem is typically dealt with in one of the following two ways: restricting queries that users can pose to the system or adding noise to the data. Query restriction methods prevent compromise from occurring, while when noise addition techniques are applied compromise is still possible, although the intruder has a degree of uncertainty about the exact values. In either case it is important to find the right balance between security and usability of the database, where the latter is measured by the number and quality of released statistics.

3.2. Trust Relationships in the SDC Context

Before presenting our trust model, we first examine the trust relationships that exist between the three system stakeholders of our statistical data warehouse system: Data Source, Data Manager and Data User.

Data Source and Data Manager. The Data Source trusts that the Data Manager will not misuse their data, that is, they will only use it for the purposes previously agreed upon and seek permission before using it for anything else. They also trust that the Data Manager will not on-sell their information to a third party without their specific consent. In addition to trusting that the Data Manager will only use data for previously agreed upon purposes, the Data Source also trusts that the Data Manager will properly manage their data and keep sensitive information private.

In the reverse trust relationship between the Data Source and Data Manager, the Data Manager trusts that the Data Source will provide them with complete and accurate data.

Data Manager and Data User. The Data Manager trusts that the Data User will not misuse the data provided to them. One obvious way in which this type of trust could be eroded is where the Data User agrees with the Data Manager to only use the provided data for specific purposes and then uses it for another purpose. The Data User is also trusted by the Data Manager to keep private information confidential.

In the reverse relationship, the Data User trusts that the Data Manager will provide them with quality data.

Data Source and Data User. The Data User trusts that the Data Source will provide accurate data, albeit not directly. Conversely, the Data Source trusts that the Data User will not misuse their data, nor compromise the privacy of any individual's record in the data warehouse. The nature of the delegation task itself impacts on the willingness of the Data Source to participate and place their trust in the Data User.

4. Model of Trust for SDC

To gain insight into the interactions between the stakeholders we first model the trust relationships to illustrate

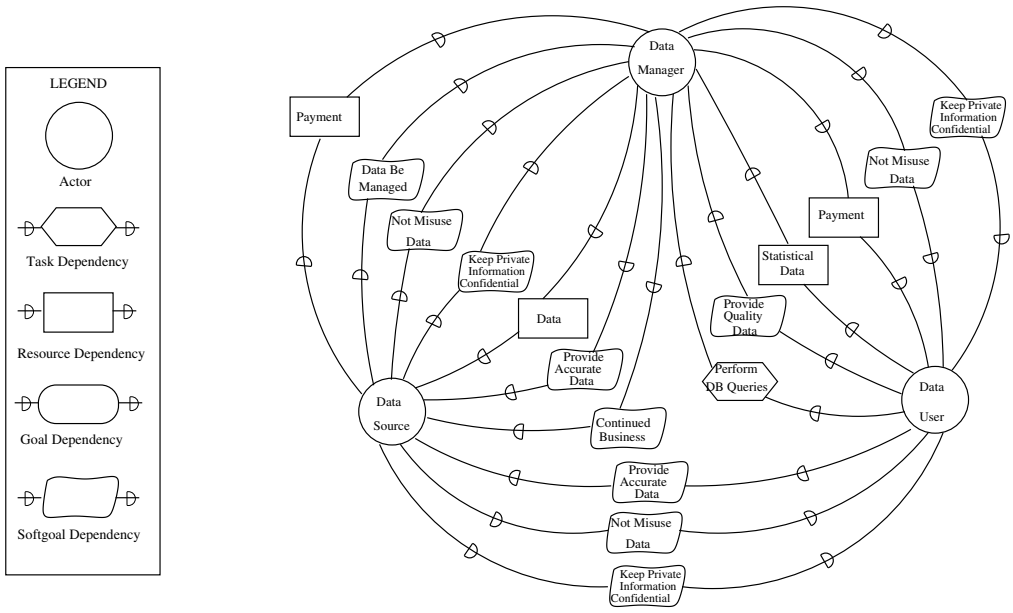

Figure 1. Strategic Dependency (SD) Model of a Statistical Data Warehouse System

the extent and importance of trust in a well managed statistical data warehouse. When a breakdown of trust occurs between one or more of the system stakeholders, there is clearly a need to ensure that cooperation still occurs. So how does a Data Manager decide when the system is operating effectively? We provide a trust model designed to assist Data Managers in evaluating the relative levels of trust in the system, so as to better recognise potential problems. We advocate that the Data Manager then use a privacy protection framework to ensure that the needs of all stakeholders are adequately managed.

4.1. Modelling Trust Relationships

We now model the trust relationships among various system stakeholders using components of the so-called ***i**** framework, which was developed as a requirements engineering tool for early stage system development [19]. The framework allows for qualitative reasoning about opportunities and vulnerabilities of system stakeholders. This framework has been previously applied to the modelling of the role trust plays in system design highlighting areas where an erosion of trust may develop [18].

In this paper we use a subset of the ***i**** framework as presented in [18]. This subset is sufficient to model the intentional dependencies within a network of system stakeholders (actors), via a Strategic Dependency (SD) model. Figure 1 shows a Strategic Dependency model for a statistical data warehouse system, with the three system stakeholders modelled as actors. As discussed in Section 2.1, dependence and delegation are at the heart of trust, and with the SD model we are able to capture four separate types of dependencies: task dependency, resource dependency, goal dependency and softgoal dependency. These dependencies should be self-explanatory, except possibly for the softgoal dependency, which we define as follows.

A softgoal dependency exists when there is no specific a priori principle for what constitutes meeting the goal but rather the dependor and dependee must decide on an individual basis if the goal has been sufficiently accomplished [18]. For instance, in Figure 1 the Data Source trusts the Data Manager to keep their sensitive information private, while there is no clear-cut standard for how this confidentiality will be achieved. Modelling trust as an ***i**** soft goal was first presented in [18].

Figure 1 incorporates all of the trust relationships that were discussed in Section 3.2. We can use this figure to reason about how low levels of trust greatly impact the management of the secure statistical data warehouse system. For instance, if a Data Source were to lose trust in the Data Manager, they may then decide to falsify their data in any future dealings with them. The Data Manager may not become immediately aware of the change to data quality; however, the Data User may notice a drop in quality of the new data they receive. This would lead to a reduction in the level of trust between the Data User and Data Manager. It is not difficult to imagine the cyclic (feedback) affect of such drops in trust levels between the system stakeholders. The ultimate outcome could easily be that the Data User no longer relies on the Data Manager to provide them with data, causing loss of business to the Data Manager.

It is clear from this discussion that the various trust relationships that exist between the three system stakeholders are vital to the optimal operation of any statistical data warehouse system. When there has been a breakdown of trust between two parties it is essential that some mechanism be employed to ensure that goals are still achieved and the system runs smoothly. We now present a trust model for a statistical data warehouse system that incorporates all of the issues we have discussed thus far.

4.2. Trust Model

Our trust model in Figure 2 incorporates the spirit of the trust definition (Definition 1) from Section 2, and is based on several existing models of trust in the literature, namely that of Mayer et al. [15], Marsh [11] and McKnight and Chervany [12].

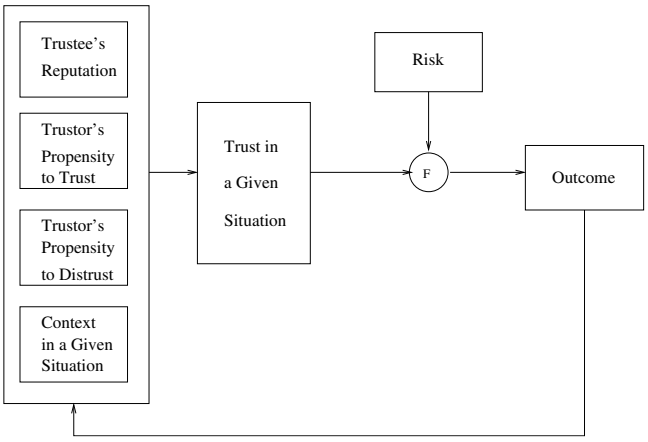

Figure 2. Trust model for Statistical Data Warehouse system.

In Figure 2 we can see that the *trust in a given situation* is a function of the *trustee's reputation*, the *trustor's propensity to trust*, the *trustor's propensity to distrust* and the *context in a given situation*. This resultant trust is measured against the *perceived risk* via a *cooperation function* (F), such as that proposed in [11], to decide if the trust delegation will occur. The *outcome* of this trusting relationship, either negative or positive, is then used to update the constructs on the left-hand side of the model. We now examine each individual component of the model in more detail.

The trustee's reputation is based on the four key trustee characteristics presented in [12], which are competence, integrity, benevolence and predictability. The perceived levels of each of these contributing factors will be constantly updating via the feedback loop shown in Figure 2.

The trustor's propensity to trust is based on McKnight and Chervany's 'Disposition to Trust' [12] or Marsh's 'basic' trust [11]. This refers to a person's general trusting disposition, or how they trust in general, regardless of the situation or the person being trusted. As with any type of trust, the level of trusting disposition can change over time, depending on the outcomes of trusting relationships. Similarly, the disposition to distrust relates to a general tendency to not be willing to depend on others in general [12]. At first glance it may appear that there is no difference between these constructs, but as discussed in Section 2.2 we prescribe to the view that trust and distrust are opposite yet separate constructs [9, 12].

The fourth factor that contributes to trust in a given situation is the context of the situation itself. The situation a person finds themselves in will alter the level of trust that they feel towards another. Two important factors to consider here are the utility and the importance of the outcome of the trusting relationship [11].

When the level of trust in the given situation has been established it needs to be compared to the level of perceived risk and a decision made as to whether to cooperate or not. The function used to decide if cooperation will occur is analogous to the so-called 'cooperation threshold' proposed by Marsh [11]. When the level of trust is higher than the cooperation function, then the trustor will delegate to, or place their trust in, the trustee. Conversely, when the level of trust is below the cooperation threshold, then they will choose not to cooperate. In the case of the Data Source being our trustee, their choosing not to cooperate would result in them withholding or altering their information.

Marsh defines a cooperation threshold as the level above which situational trust must be for cooperation to occur [11]. He incorporates the level of perceived risk, the trustee's competence, the trustor's general trust towards the trustee, and the importance of the situation in his calculation of the cooperation threshold. Our cooperation function is somewhat simpler than this because we have incorporated most of these elements into the calculation of trust in a given situation. When the level of trust is higher than the perceived risk, the trustor will delegate to the trustee, and when it is lower, they will choose not to delegate.

The outcomes of the cooperation between the trusting parties is used in a feedback loop to incorporate the dynamic nature of trust over time. If a trustee is able to achieve their delegated task, then this could conceivably lead to an increase in their reputation and hence an increased level of trust in future interactions. A perhaps less obvious effect of a negative outcome would be when the trustee's reputation increases despite them being unable to complete their delegation task.

5. Future Work: Evaluating Reputation, Trustworthiness and Risk

In order for the type of model presented in the previous section to be of any use to Data Managers, we need to be able to quantify all of the trust constructs presented. The following is a discussion on some issues and evaluating difficulties arising when the levels of perceived reputation, trustworthiness and risk in a Statistical Data Warehouse system.

We firstly look at how to evaluate the reputation and perceived trustworthiness of the Data Source, that is, the person providing information, some of which is of a sensitive nature, to the Data Manager. One of the difficulties in assessing the reputation of the Data Source is the lack of feedback that can be obtained. This means that the application of any traditional reputation system, such as those presented in [1], would be challenging and leads us to require a different approach to assessing reputation.

Evaluating the trustworthiness of the Data Manager from the point of view of the Data User and Data Source can also be difficult. One reason for this is that often there is no direct contact between the parties, particularly when talking about the Data Source. When it comes to evaluating the trustworthiness of the Data User, their reputation will in part be dependant on who the Data User is. That is, some occupations are naturally more trustworthy than others. It is also important to know for what purpose the data will be used. For example, a recognised research project attached to a well respected University is likely to invoke more trust than a market research survey.

Evaluating the level of risk in a situation is just as important as evaluating the trust. One of the risks that a Data Manager has to consider when dealing with a Data Source, is whether they will withhold their information, or potentially provide false information. When the Data Manager is dealing with the Data User, they need to consider the risk that the Data User may misuse the information. For the Data User the main risks involve the perceived quality of the information they receive from the Data Manager. The risk includes not only the Data Source withholding, or providing false information, but also the Data Manager's incorrect collection and management of the data before passing it on to the Data User.

In conclusion, we note that evaluating reputation and risk in a SDC context are challenging tasks, primarily due to the lack of direct contact and/or feedback between the parties involved.

References

[1] A. Abdul-Rahman and S. Hailes. Relying on trust to find reliable information. In *Proceedings of the International Symposium on Database, Web and Cooperative Systems (DWACOS '99)*, Baden-Baden, Germany, 1999.

[2] N. R. Adam and J. C. Wortmann. Security-control methods for statistical databases: a comparative study. *ACM Comput. Surv.*, 21(4):515–556, 1989.

[3] L. Brankovic and H. Giggins. *Security, Privacy and Trust in Modern Data Management*, chapter Statistical Database Security. Springer, 2007.

[4] L. Brankovic and M. Miller. Introduction to statistical database security. *Communications of the CCISA*, 9(4):1–30, 2003. In: Special Issue, Selected Topics of Cryptography and Information Security.

[5] R. S. B. Denise M. Rousseau, Sim B. Sitkin and C. Camerer. Not so different after all: a cross-discipline view of trust. *The Academy of Management Review*, 23(3):393–404, 1998.

[6] J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufman Publishers, 2001.

[7] A. Jøsang, C. Keser, and T. Dimitrakos. Can we manage trust? *Lecture Notes in Computer Science*, 3477:93107, 2005.

[8] S. J. Lars Rasmusson. Simulated social control for secure internet commerce. In C. Meadows, editor, *Proceedings of the 1996 New Security Paradigms Workshop*. ACM, 1996.

[9] R. J. Lewicki, D. J. McAllister, and R. J. Bies. Trust and distrust: New relationships and realities. *The Academy of Management Review*, 23(3):438–458, 1998.

[10] S. Marsh and M. R. Dibben. Trust, untrust, distrust and mistrust - an exploration of the dark(er) side. In P. Herrmann, V. Issarny, and S. Shiu, editors, *Proceedings of Third International Conference on Trust Management (iTrust)*, volume 3477 of *Lecture Notes in Computer Science*, pages 17–33. Springer, May 23-26 2005.

[11] S. P. Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, Department of Computing Science and Mathematics, University of Stirling, April 1994.

[12] D. H. McKnight and N. L. Chervany. Trust and distrust definitions: One bite at a time. *Lecture Notes in Computer Science*, 2246:27–54, 2001.

[13] B. McSherry. Ethical issues in health*Connect*'s shared electronic health record system. *Journal of Law and Medicine*, 12(1):60–68, 2004.

[14] L. Mui, M. Mohtashemi, and A. Halberstadt. A computational model of trust and reputation. In *Proceedings of the 35th Annual Hawaii Conference on System Sciences*, pages 2431–2439, 2002.

[15] J. H. D. Roger C. Mayer and F. D. Schoorman. An integrative model of organizational trust. *The Academy of Management Review*, 20(3):709–734, 1995.

[16] L. Willenborg and T. de Waal. *Statistical Disclosure Control in Practice*. Lecture Notes in Statistics. Springer, New York, USA, 1996.

[17] L. Willenborg and T. de Waal. *Elements of Statistical Disclosure Control*. Lecture Notes in Statistics. Springer, New York, USA, 2001.

[18] E. Yu and L. Liu. Modelling trust for system design using the i* strategic actors framework. *Lecture Notes in Computer Science*, 2246:175–194, 2001.

[19] E. S.-K. Yu. *Modelling strategic relationships for process reengineering*. PhD thesis, University of Toronto, Canada, 1996.