1C5-4

Recommending Business Partners using Supply-Chains and Web Information

Junichiro Mori*1

*1Organization for Interdisciplinary Research Project, The University of Tokyo

Business development is vital for any firms. However, globalization and the rapid development of technologies have made it difficult to find appropriate business partners such as suppliers and customers, and build reciprocal relationships among them, while it simultaneously offers many opportunities. In this contribution, we propose machine learning approach to find plausible candidates of business partners using firm profiles and transactional relationships among them. We employ machine learning techniques to build a prediction model of customer-supplier relationships. We applied our approach to the large amount of actual business data and Web data. The proposed method contribute to developing one's own business in the complicated, specialized

1. Introduction

Business development is a perdurable issue for any firms pursuing sustainable development. In particular, organizational relationships among firms such as customer-supplier relationships and strategic alliance have an significant effect to develop business because they can work as a source of innovation. Firms in a cooperation network can utilize their network in a variety of ways. They not only share the costs and risks of their activities but also obtain access to new markets and technologies, make use of complementary skills, and share knowledge capabilities, and finally infiuence firm performance.

Firms must seek new business partners to acquire new opportunities and also to activate existing relationships, while long-term and dedicated relationships with trust is the vital source of innovation. But such an activity is laborious, time-consuming, and subjective. Traditional stakeholder theories and frameworks can enhance our understandings in the roles of business partners, but cannot serve as a guidance to find new partners and to keep reciprocity among them. And current experts employ a manual approach to selection of business partners and do not scale up to accommodate the rapid pace of change in business environment and market.

The aim of this paper is to find new business partners such as suppliers and customers, and build reciprocal relationships among them with the help of machine learning. To this aim, we propose a machine learning approach to model and predict customer-supplier relationships based on explicitly available profiles of firms and their transactional relationships and thier Web information. With the available firm data, we design several features that characterize customer-supplier relationships. We employ those features, which are represented with a very large dimensional vector, with the support vector machine (SVM) to learn a model of customer-supplier relationships. Consequently, our method can automatically predict potential business partners given firm profiles and its existing transactional relationships. We applied our approach to the vast amount of actual business data.

Our contributions in this paper are two folds. First, we design and examine meaningful features of explicit firm data for modeling business relationships using a machine learning approach. Second, we integrate our method into the Web-based system and demonstrate the applicability to practical services. These would lead to help find new and reciprocal business relationships in both theoretical and practical ways.

2. Related Works

Several studies have addressed to evaluate selected business partners by financial strength. In the previous literatures, attributional features such as financial data, product and service quality, and technological capabilities/compatibilities are utilized in the analysis to evaluate suppliers by introducing the implications of previous literatures in supply chain management.

However, it is not easy to obtain such data in advance before actual transactions is contracted and launched. On the other hands, we can easily obtain explicit data of firms such as the number of employee and date of foundation from the Web or commercial databases. Therein, one of important questions with a computational approach to find business partners is how to deal with the vast amount of firm data which is available from several data sources and therefore how to identify meaningful features among the data.

To exploit the opportunities acquired by partnerships, it is also crucial to establish reciprocal relationships in addition to finding new business partners. Business reciprocity is the term to describe business dealings between independent firms whereby they make mutual concessions designed to promote the business interests of each. Although researchers have studied the origin of business reciprocity from a point of buying power and profitability of customer, the current business situation is more complex and the understanding on the reciprocity is not enough. Therefore, for supporting to find business partners and build reciprocal relationships computationally, it is also important to

連絡先: Junichiro Mori, Organization for Interdisciplinary Research Project, The University of Tokyo, Hongo 7-3-1 Bunkyoku Tokyo, 03-5841-5726, jmori@platinum.u-tokyo.ac.jp

understand characteristics of firms building reciprocal relationships based on the actual data.

3. Method

Several studies have addressed a problem of customer/supplier selection using a machine learning approach. In the previous literatures, attributional features such as financial data, product and service quality, and technological capabilities/compatibilities are employed in the analysis to evaluate suppliers. However, it is often difficult obtain such information in advance. In contrast, we use explicitly available data of firms to design features. Additionally, previous works typically use a small corpus whose size is around 100. We use the large size whose size is larger than the previous works by the order of two. And we deal with a large dimensional vector of various features on top of the vast amount of data.

Moreover, in this paper we utilize customer-supplier relationships to predict the focal relationship. There is less studies utilizing such relational features in business data mining. Therefore, the issue tackled in this paper is closely related to the link-prediction problem. In the link-prediction problem, utilization of relational features intrinsic to the network itself can offer meaningful inferences from observed network data.

We regard selection of business partners such as supplier and customer as one of machine learning problems. As the machine learning problem, we try to learn a model to predict whether there exists a customer-supplier relation between a firm X_c and a firm X_s given a firm pair (X_c, X_s).

More formally, let denote by

$$(X_c, X_s) = x_{c1}, \cdots, x_{cl}, x_{s1}, \cdots, x_{sm}, x_{cs1}, \cdots, x_{csn}$$

the l attributes where each x_c represents the attribute about a customer X_c, the m attributes where each x_s represents the attribute about a supplier X_s, and the n attributes where each x_{cs} represents the relational attribute between the X_c and the X_s. Then, let denote by $Y := y$ the customer-supplier relation, where y takes a value either of +1 (which indicates "relation") or -1 (which indicates "non relation").

Our goal is to build a model f which relates the firm pair instances of (X_c, X_s) to their customer-supplier relations, i.e.,$Y = f(X_c, X_s)$. Several methods to build such models have been developed in the field of machine learning. In our study, we employ SVM which is one of the state-of-the-art predictive models.

4. Data

In our experiment, we focus on manufacturing firms in the Tokyo area which is the largest economic block in Japan. We select 30,660 firms on manufacturing industrial categories and obtain the data for each firm from commercial business databases (TEIKOKU DATABANK) The firm data includes attributes such as *capital*, *number of employee*, *date of foundation*, *number of competitors*, *ranking*, *address*, *CEO's native prefecture*, *CEO's old school*, *industrial categories*, *bankers*, *suppliers*, *customers*. The number of competitors shows how many other firms (nation-wide or prefecture-wide) exist in one's industrial category. Ranking shows the rank order of one's sales (nation-wide or prefecture-wide). Although profit data is available, we exclude profit from the attributes because we use profit data to evaluate reciprocal customer-supplier relationships as described in our experimental settings. In addition to those attributes, we extract keywords from the Web page of each firm and add to the attributes. The keyword is attached with a *TF-IDF* weight. We also add relational attributes that are based on commonly shared attributes between a customer and a supplier. As a result, one learning instance which is a supplier and customer pair is represented with 42 kinds of features.

5. Experiment

We conducted a first experiment with existing 34,441 supplier and customer pairs from our data set. The goal of this experiment is to learn a model to predict customer-supplier relationships in our data set. We used the pairs as positive instances. As negative instances, we randomly selected the same number of a firm pair that does not have any customer-supplier relationships. We created 5 set of the training data with 5 different set of negative instances.

With the data set, we compared the performance of different regularizations by using fivefold cross validation. As the performance measures, we employed Precision, Recall, and F-value. The optimal parameter for the regularization was chosen with a greedy search. Our preliminary experiment showed that the $L2$ regularization SVM performed well, which we employed in our evaluation.

We next conducted an experiment to learn a model to predict more reciprocal customer-supplier relationships. Therein, we regard a reciprocal relationship as such that both supplier's and customer's profit rates rank in the top N% of respective industrial categories.

We obtained 10,245 supplier and customer pairs as $N = 50\%$ from our data set, 3799 pairs ($N = 30\%$), and 358 pairs ($N = 10\%$). As negative instances, we randomly selected the same number of a customer-supplier pair for each data such that the supplier and the customer's profit rates do not rank in the top N% of respective industrial categories. We created 5 set of the training data with 5 different set of negative instances. With the data set, we compared the performance of a learner by using fivefold cross validation. We will be reporting our results in the conference.

参考文献

[Mori 12] Mori, J, Kajikawa, Y, Kashima, H, and Sakata, I: Machine Learning Approach for Finding Business Partners and Building Reciprocal Relationships, Expert System with Applications, Vol.39, No.12, 10402–10407 (2012).