

Constructing Knowledge Maps of A Manager's Managerial Logic by A Text Mining Approach

Jen-Ying Shih¹ Yu-Jung Chang^{2,3}

¹ Department of Business Administration, Chang Gung University, Taiwan, ROC

² Department of Computer Science and Information Engineering, National Taiwan University

³ Institute of Information Science, Academia Sinica, Taipei, Taiwan

Abstract

The objective of this research is to represent the managerial logic of Mr. Yung-Ching Wang, the Chairman of Formosa Plastics Group (also known as the “God of Business” in Taiwan) through the construction of knowledge maps using a text-mining approach, including automatic key phrase extraction, term identification, document vector modeling, and a clustering method named growing hierarchical self-organizing map (GHSOM). A collection of speeches given by Mr. Wang in various occasions is used to construct the topic maps, representing his managerial logic. The maps can be browsed on the web site (http://synteny.iis.sinica.edu.tw/kmap/m_logic).

Keywords: Managerial logic, topic map, knowledge map, GHSOM.

1. Introduction

From the upper echelons perspective, an organization can be viewed as a reflection of its top managers. Organization outcomes—both strategies and effectiveness – are viewed as reflections of the values and cognitive bases of powerful players in the organization [5]. Top management and firm leaders play critical roles in conceptualizing business models, shaping culture, defining performance measures and rewards, defining hiring criteria, and developing the overall strategy of the firm [1]. Therefore, it is a useful way to analyze why an enterprise can achieve success by studying its leaders' managerial logics (dominant managerial logic), a common way of observing how best to do business as a manager in the firm.

In the past, to learn the dominant managerial logics of these leaders, we either rely on storytelling by face-to-face interview or reading their autobiographies. That usually is time-consuming and costly and feasible only if you get chances to have in-depth interviews or their autobiographies do exist. Recently, some text mining approaches have been proposed to help extracting knowledge from various

document sources with some degree of success [2,8]. Therefore, in this paper, some text mining tools will be used to observe and understand the dominant managerial logic of Formosa Plastics Group (FPG), the largest private business conglomerate in Taiwan. The businesses of FPG cover petrochemical, textile, high technology, autos, hospitals, health care, transportation, and education. Its revenue counts 12% of GDP in Taiwan.

Yung-Ching Wang is the founder and leader of FPG. The corporate culture of FPG is deeply affected by his managerial logic. Moreover, he is one of the major contributors in Taiwan's early economic miracle, and is thus renowned as the “God of Business” in Taiwan. Hence, we select his managerial logic as our test bed and represent his management philosophy through constructing his knowledge maps.

2. Dominant Managerial Logic and Tacit Knowledge

Each manager brings business situation a set of biases, beliefs and assumptions about the market that his/her business serves, whom to hire, what technology to use to compete in the market, who the firm's competitors are, and what types of business models are successful in the industry. This set of biases, beliefs and assumptions is a manager's managerial logic. It defines a mental model that a manager is likely to use as a basis in making decision. A firm's strategies, systems, technology, organizational structure, culture, and success will determine its dominant managerial logics. It is a common way of observing a firm's best practice [1]. However, managerial logics are usually tacit and not easy to get. They are usually delivered and shared through business activities. Therefore, it is a challenge to get the whole picture of a firm's dominant managerial logics.

In knowledge management, knowledge can be categorized as explicit knowledge and tacit knowledge in terms of the visibility of knowledge. Tacit knowledge is understood and applied through subconscious, and it is thus hard to communicate and is often developed by experience and action. It is

communicated, described and shared by high-degree interaction [10]. Explicit knowledge can be coded, stored and diffused [6]. Hence, the efficient and effective collection, storage, creation and diffusion of tacit knowledge have long been important research issues in knowledge management.

Dominant managerial logic is a kind of tacit knowledge in terms of its property. It is a key factor affecting business models of enterprises and can be used to understand the success of businesses [1]. It's now an important research issue in discovering the dominant managerial logics of successful enterprises. To explore these logics, most experts use questionnaires, interviews, or some related readings and materials to induce them. Few research apply artificial intelligence methods to explore these logics. In this paper, an automatic term extraction and growing hierarchical self-organizing maps are used to represent Mr. Wang's logic.

3. Constructing Mr. Wang's Knowledge Maps

In this section, we describe each step of the system flow of constructing Mr. Wang's knowledge maps.

3.1. Data Collection

The data set we used in this paper is composed of 292 documents that are Mr. Wang's speech contents about business and external environment of business; these speeches have been given to employees, students and media during past 20 years. The data will be used in computing knowledge to represent the knowledge maps of Mr. Wang's dominant managerial logic. Based on the knowledge maps, his management philosophy will be explored and analyzed.

3.2. Automatic Key Phrase Extraction

We use a key phrase extraction mechanism which has the ability to update the PAT-tree after removal and the use of stop words [7] for generating a lexicon of Mr. Wang's speech contents. We have extracted a total of 6,795 terms from Mr. Wang's speeches that we studied. Some special new terms or phrases that represent Mr. Wang's managerial logic are identified by this extraction approach, such as "thin goose spirit", "hen's behavior in breeding chickens", and "healthy-life-style culture village". These are his special managerial logic that is often described as his management philosophy and practiced in his businesses.

3.3. Term Identification

It is not easy to identify terms in Chinese articles, because there are no spaces between Chinese characters [10]. Some methodologies have been developed to resolve the problem, such as dictionary, linguistic, and statistical approaches [7]. We identify terms by implementing a program using a maximum matching rule [3], and by basing on a lexicon that is generated by the step described under section 3.2.

3.4. Document Vector Models

We calculate the frequencies of a term in each document and the inverse document frequencies of such term. Then, a $tf \times idf$ weight vector (1) is calculated from term frequency multiplying inverse document frequency [9],

$$w_i(d) = tf_i(d) * \log\left(\frac{N}{df_i}\right) \quad (1)$$

where $w_i(d)$ represents the weight of a term (i) in a document (d),

$tf_i(d)$ represents the occurrence frequencies of such term (i) appearing in the document (d),

N represents the total number of all documents, and df_i represents the number of the documents which contain term (i).

The $tf \times idf$ weight vector is well-known in information retrieval and is considered the state of the art for content representation. This weighting scheme assigns high values to terms that are considered important for purposes of describing the contents of a document and discriminating between various documents. We select features of the document vector from the lexicon described in section 3.2. The input data of the clustering model is represented by document vectors that can be described as the coding schema or the professional interpretation of a particular document.

3.5. Growing Hierarchical SOM

In this paper, we use an improved self-organizing map (SOM) algorithm, the growing hierarchical SOM (GHSOM) [4], to construct the knowledge maps for representing the managerial logics. The GHSOM algorithm can be summarized as follows:

1) Initialize all parameters of GHSOM, including the learning rate, the neighborhood range, the initial map size for the training process, the growing-stopping criterion, the hierarchical stopping criterion, the maximum number of labels, and the label threshold.

2) Start with a virtual layer 0 consisting of only one unit whose weight vector is initialized as the average of all the input data. Then calculate the mean quantization error (mqe) by the Euclidean distance between the weight vector of the unit and all input vectors.

3) Set the initial map size of the first layer in a small map of, for example, 2x2 units, which is self-organized according to the standard SOM training process.

4) Evaluate the mapping quality by calculating the mean quantization error of each unit in the current layer to determine the error unit according to the largest deviation between its weight vector and the input vectors mapped to it. Then, either a new row or column of units is interpolated between the error unit and its most dissimilar neighbor. The weight vectors of these new units are initialized as the average of their neighbors. Although the training process is very similar to the growing SOM model, it uses a decreasing learning rate and a decreasing neighborhood range, instead of a fixed value. After growing the map, we have calculated the mean *mqe* of all units (*MQE*) in the current map. The map grows until its *MQE* is reduced to a predefined fraction (the growing-stopping criterion) of the *mqe* of the unit in the preceding layer of the hierarchy. In other words, the *MQE* of each map in the current layer should be smaller than a certain fraction value (τ_l) of the unit in the preceding layer. The lower the value of the quantization error, the better the map is trained.

$$MQE_m < \tau_1 \cdot mqe_u,$$

where *m* denotes the units in the current map, and *u* denotes the mapped unit in the preceding layer

5) Determine the depth of each topic in the current layer according to a predefined fraction (τ_2) of the *mqe* of layer 0, i.e., the *mqe* of each unit in the current layer should be smaller than a certain fractional value of the unit in layer 0. The stopping criterion of any unit in the hierarchy is always compared with layer 0.

$$mqe_i < \tau_2 \cdot mqe_0,$$

where *i* denotes the unit in the current layer

The training procedure described in Steps 4 and 5 is used to train the subsequent layers. After training, a hierarchy of the trained maps is generated as the knowledge maps. The labels given to each cluster and determined by LabelSOM [8] are used to represent the topic of the cluster. This completes the knowledge maps.

3.6. Knowledge Maps of Mr. Wang's Managerial Logic

We improve the user interface of GHSOM provided by Dittenback *et al* [4]. The knowledge Maps of Mr. Wang's Managerial Logic includes not only a map view generated from GHSOM but also a tree view generated from the topic name of the map view (Fig 1). Left hand side of the interface is a tree view, the right-top side is a map view, and the lower right hand side is the list view.

Because the label lists suggested by LabelSOM are not all suitable as topic names in this case, we designed a topic selection module to enhance the readability of topic names. We first obtained at most 20 labels for each cluster by the LabelSOM algorithm. The topic name for each cluster was selected from the top two highest-frequency labels based on the statistics of all article titles mapped to the cluster.

4. Results and Discussion

The training results of the model are shown in Fig 1, which illustrate the first layer of the knowledge maps generated by GHSOM. Each unit is assigned a set of up to 5 labels, based on the quantization error vector and the unit's weight vector. The first layer of GHSOM is clustered as the following twelve topics.

- Unit (1,1)¹: Human Resource and Improvement
- Unit (1,2): National Health Insurance System and Bargaining
- Unit (2,1): Tax Return Rules and Petrochemical
- Unit (2,2): Privatization
- Unit (3,1): Steel Industry and Value Chain
- Unit (3,2): Private Enterprise and Privatization
- Unit (4,1): Market Economy and Taiwan Residents
- Unit (4,2): Toyota (Business Administration)²
- Unit (5,1): Reform and Education Material
- Unit (5,2): Hospitals and Medical Level
- Unit (6,1): Aborigine and Students (Education)
- Unit (6,2): Chewing Betel Nuts and Nurses (Social Environments)



Fig 1: The top layer of knowledge maps

Users can be further guided to browse more detailed topics in the first layer by clicking the “next level” hyperlink of each cluster. For example, in Unit (1,2) users can be guided to the second layer (Fig 2), which represents the following topics:

¹ We use the notation (x,y) to refer to the unit in row x and column y, starting with (1,1) in the upper left corner.

² The topic name in brackets may be more appropriate for this cluster in terms of its contents clustered in this unit.

- Unit (1,1): Fraudulent Bidding and Construction Firms
- Unit (1,2): Engineering and Audit Committee
- Unit (1,3): Audit Committee and Fundamental Design
- Unit (1,4): Private Schools and Public Schools
- Unit (2,1): Bargaining and Good Firms
- Unit (2,2): Quoting Firms and Transparency
- Unit (2,3): National Health Insurance System and Payment
- Unit (2,4): Salary and Information



Fig2: The second layer of Unit (1,2) on the top map

If users are interested in the “Construction Firms” topic, they can click on the hyperlink on the topic name of Unit (1,1) in Fig 2. In the last step, users can see the content of the speech document (Fig 3) by clicking on the hyperlink in “Table of Contents” frame. For example, Unit (1,1) of the second layer contains the “How to prevent fraudulent bidding by good engineering management”. Both the content of the article (Fig 3) and the features (labels) used to describe the article are presented to users. The labels selected as features of Unit (1,1) are “construction firms”, “contractor”, “quality of construction”, “construction” and “fraudulent bidding”. These labels seem to relate to this speech document.

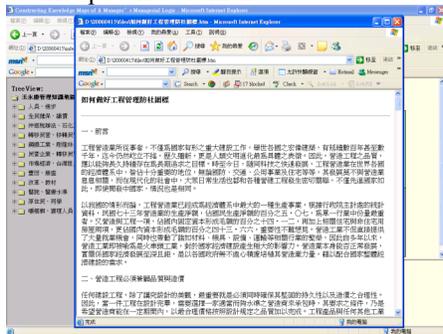


Fig 3: The content of Unit (1,1) of the second map

We will evaluate the performance of the knowledge maps by the managers who work for FPG and are familiar with Mr. Wang’s thinking.

5. Conclusion

In this paper, we apply some text mining techniques in constructing knowledge maps of Mr. Wang’s managerial logic. It’s a new trial using a text mining

approach in exploring a top manager’s managerial logic. We think it is an alternative choice for identifying, understanding, and analyzing managerial logics. In future research, researchers can apply other text mining tools for this research issue. For example, case-based reasoning that is often applied to document classification application may be useful in this issue. The data set can be the logics of other entrepreneurs, successful managers or renowned politicians. Besides, the use of the knowledge maps in resolving problems about businesses or helping decision makers may be another future research direction.

6. References

- [1] Afuah, A., *Business models- A strategic management approaches*, 2004, McGraw-Hill/Irwin.
- [2] Chen, C., Kuljis, J., and Paul, R. J. “Visualizing latent domain knowledge,” *IEEE Transactions on Systems, Man, and Cybernetics*, 31(4), 2001, pp.518-529.
- [3] Chen, K. J., and Liu, S. H. “Word identification for Mandarin Chinese sentences,” *Proceedings of the Fifteenth International Conference on Computational Linguistics*, Nantes, 1992.
- [4] Dittenbach, M., Rauber A., and Merkl, D. “Uncovering hierarchical structure in data using the growing hierarchical self-organizing map,” *Neurocomputing*, 48, 2002, pp.199-216.
- [5] Hambrick, D. C., and Mason, P. A., “Upper Echelons: The Organization as a Reflection of Its Top Managers,” *The Academy of Management Review*, Apr. 1984, 9, pp.193-206.
- [6] Krogh, G. V. “Care in Knowledge Creation,” *California Management Review*, 40(3), 1998, pp.133-153.
- [7] Ong, T.-H., and Chen, H., “Updateable PAT-Tree Approach to Chinese Key Phrase Extraction using Mutual Information: A Linguistic Foundation for Knowledge Management,” *Proceedings of the Second Asian Digital Library Conference*, Nov. 8-9, 1999, pp. 63-84.
- [8] Rauber, A., and Merkl, D., Automatic labeling of self-organizing maps: making a treasure map reveal its secrets. In: N. Zhong, L. Zhou (Eds.), *Proceedings of the Third Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 1999, number LNCS/LNAI 1574 in Lecture Notes in Artificial Intelligence, pp. 228-237. Beijing, China.
- [9] Salton, G., *Automatic text processing*, 1999, New York: Addison-Wesley.
- [10] Zack, M. H., “Managing codified knowledge,” *Sloan Management Review*, Summer, pp. 45-58, 1999.