

Social Chance Discovery

Chao-Fu Hong
Department of Information
Management
Aletheia University
Taipei, Taiwan (R.O.C.)
cfhong@email.au.edu.tw

Chien-Jen Huang
Department of Information
Management
Aletheia University
Taipei, Taiwan (R.O.C.)
norman@mail.au.edu.tw

Jheng-Long Wu
Graduate School of Management
Sciences
Aletheia University
Taipei, Taiwan (R.O.C.)
arlon1024@gmail.com

Abstract—Information technology such as data mining or text mining, can extract useful data structure from data. But extracted data structure was hardly used in explaining the meaning of data itself. In reality, a laborer can hardly catch any meaningful job information from members of his work group (strong tie) when it comes to changing jobs. However, some of his friends (weak tie), although not so close to him, associated to other work groups, can bring him information about new jobs. This information is very important in job hunting. This kind of personal central base social networks can be used to realize a person's motivation and enrich his social capital, extends his resources to the resources of his linked friends. This process is called social chance discovery. Base on this concept, in this study, we build model of Qualitative Chance Discovery and apply it on DVD news. Experiment results show this model can be used as an interactive tool to help company understands the way of breaking entry barrier to compete with existing companies in the harsh environment of DVD industry.

Keywords—qualitative analysis, quantitative analysis, chance discovery, social computing, social capital

I. WHAT IS THE CHANCE?

Yoshikawa [1] surveyed what the chance meant and found that it had some characteristic meanings: opportunity, probability, fortune and chance involving a certain degree of risk. Furthermore, KeyGraph, developed by Ohsawa [2], utilized connection probability to define the confidence of a term connecting with associative clusters and searches for key links and key terms. These key terms are used to detect weak but important events, which in turn are used to predict future chances. Starting from one scenario, this process connects to other scenario through bridge, and creates an innovative scenario.

In reality, the weak sign may work as weak tie (or chance) illustrated by Granovetter [3] in the following story: a laborer wanted to change his job, he and his close friends (strong tie) shared the same job information. However, some of his friends, who were associated to other work groups, were not so close to him. One day, friends from other groups came to see him and bring him new job information. In this case, friends who were not so close served as the weak tie (bridge), helped him in connecting to other clusters and gave him useful information. This information is very important in job hunting. How does this type of process actually work depends on the organization of different clusters, where user finds short cuts (bridge, weak

tie or friends of other groups) between different clusters and bring potential innovation. This phenomenon of finding future chance (bridge) by using social relations is defined as Chance Discovery [2].

II. SOCIAL CHANCE DISCOVERY

In practice, there was no enough evidences to explain the weak tie effect. For example, in the ceiling effect, top person tries to use the weak tie relation, but the low position person only has little resource, in this case link to him only get few help [4]. Similarly, Marsden and Hurlbert [5] compared many cases and found that the person in the lower position is in a worse position to catch resource than the person in a higher position. In addition to using the useful path presented by the weak tie effect, user has to work hard to achieve the goal. Furthermore, to measure the social effect of a person, both the location of network and connectivity of this person in the network are used in social chance study [6].

The results of previous discussion indicate that the society may be clearly described by some meaningful social relations, such as the relation of a person's family or his friend. A person has to maintain his social relations for the following two reasons:

- 1) To ensure a person can reciprocally and stably get the social capital to maintain his resource. This kind of social relation can be used to help him catching the social resources to maintain his benefit.
- 2) Create some extra benefit.

As previous discussion, the personal central base social network is created with personal social relations, such as the person's family relation and friend networks. Information technology, like text mining, plays an important role in visualizing this motivational based social network. Besides helping people understand his immediate resources, resources of his linked persons have also become part of his social capital. In human society when a person needs to advance his social position to be a leader, he plans how to get larger social capital for catching the leader position. This kind of process is called the social chance discovery. In practice, user can use social chance discovery process to analyze and evaluate which strategy is good for a firm, and find an innovative path to help

the firm climbs to the top of hierarchical social structure. This path is called social chance of the firm.

III. METHODOLOGE AND CASE STUDY:

Optical player, including CD and DVD player, is one of the necessary peripherals of PC industry. On 2001, about 75 million optical players sold worldwide were made in Taiwan, over 50% of global share. This tremendous amount made Taiwan the most important optical payer manufacturing base in the world. Due to the competitive nature of this industry, in order to maintain market share, all of the leading firms need to gain competitive edge by either developing future standards and/or applying new technologies. This phenomenon makes the study of their social chances an intriguing one.

Because there was no Taiwanese company in the leading group of this industry, we decided to choose manufacturers that appeared in Taiwan stock market as our research subject. Under this criterion, Liteon, Benq, Btc, Aopen, Qsi and Sanyo were set as topic word and were inputted into QCD system (Qualitative Chance Discovery) [7] to extract and visualize the valid documents. From the initial output of QCD system, it was clear that the relation of the terms of DVD-ROM, DVD-RAM, CD-ROM, CD-RW, DVD-RW, DVD+RW, DVD-Dual and DVD-Multi were needed in constructing the industry environment map. These terms are determined as value words (w_{v1}, w_{v2}, \dots). The algorithm is described as following:

1) *Recognizing interesting topic*: Because optical player industry is an important industry in Taiwan. We chose Liteon, Benq, Btc, Aopen, Qsi and Sanyo as topic word (key term) (w_{f1}, \dots).

2) *Sifting out data from meaningful but small documents*: Selected key terms are inputted into information retrieve system to sift out relevant data concerning specific topic. This mechanism is defined as following.

$$\text{If } T_{wf} \in T_{Top50} \text{ of document then} \quad (1)$$

$$\text{document} \in \text{Meaningful document subset}$$

Where T_{wf} indicates a term in key term set (w_{f1}, \dots); T_{Top50} is the set of terms whose frequency is larger than or equal to top 50% in one document. This mechanism specifies that when a key term in a document belongs to T_{Top50} , this document becomes part of the meaningful document subset.

In order to extract the relevant sentences for constructing the topic hierarchical social network, relevant information (meaningful document set) was inputted into QCD system again. Fig. 1 and Fig. 2 depict the results of QCD system concerning Benq and Sanyo respectively.

1) *Discovering relevant decision variables*: Once the Decision Making Variables have been identified from the individual topic associative graph, these variables are used as input parameters of sentence retrieve system as shown in Fig. 1 and Fig. 2.

2) *Sifting out the meaningful but small sentences data*: Beside serving the purpose of removing non-relevant sentences, these variables also narrowing down data size and sift out the valid sentences, where each valid sentence contains keyword w_{f1} and at least one value word (w_{v1}, w_{v2}, \dots) as shown in (2).

$$\text{If } (w_1, w_2, \dots)_s \in (\text{one}(w_{v1}, w_{v2}, \dots)) \text{ and } (w_{f1}) \text{ then} \quad (2)$$

$$\text{sentence} \in \text{Valid Sentences subset}$$

Where $(w_1, w_2, \dots)_s$ consist of all words in one sentence, and $(\text{one}(w_{v1}, w_{v2}, \dots))$ means at least one word belongs to value words (w_{v1}, w_{v2}, \dots).

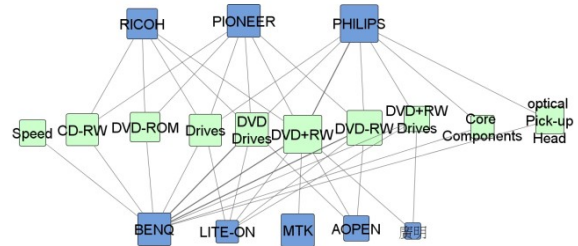


Figure 1. The social hierarchical network of Benq

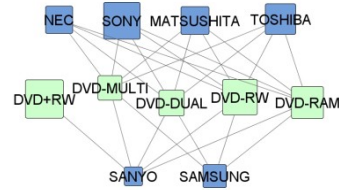


Figure 2. The social hierarchical network of Sanyo

Because user needed to know the big picture of the industry of optical player, our interactive system integrated seven social hierarchical networks to emerge the environment of optical player market as shown in Fig.3.

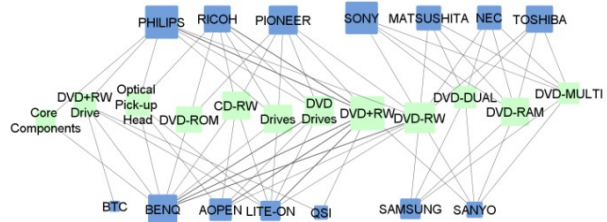


Figure 3. Use trailing group spot to catch the industrial environment of DVD player by QCD.

Using trailing group to extract valid information, experimental result indicates that the hierarchical social network is a two layers structure; one is the leading group (top layer) and the

other is the trailing group (bottom layer). By common sense, the position of leading group is more important than the position of trailing group. Besides, in text mining field, more important word has higher frequency. So the firms' frequency is used to see which company belongs to the member of leading/trailing group. Experimental result is shown in table one evidences that the positional importance makes significant difference between the leading group and the trailing group. Therefore, in this case, these firms can be separated into two layers, the leading and the trailing group in constructing the hierarchical social society.

TABLE I. COMPARE THE POSITION IMPORTANCE BETWEEN THE LEADING GROUP AND FOLLOWING GROUP

Variable	t-test for Equality of Means			
	t	Sig	Mean	Std.
Position	2.768	0.017*	38.14286	13.78158

Note * p<0.05; Two-tailed test.

Besides computing word's frequency for determining word's position, our QCD method also capable of computing word pair frequency as their associative value. Based on the definition of social chance, in this study, QCD system could help user discover the promising candidate who has chance to climb from trailing position to leading position. Focusing on Fig.3, user could find two separate trailing groups, one was Taiwanese maker and the other was the foreign maker (Sanyo and Samsung). The next intriguing question raise in user's mind would be: are these two groups different in connecting with all related technologies.

TABLE II. COMPARE THE CONNECTION IMPORTANCE OF THE TAIWAN'S MAKERS WITH THE FOREIGN MAKERS

Variable	t-test for Equality of Means			
	t	Sig	Mean	Std.
DVD-ROM	0.598	0.576	0.80000	1.33866
DVD+RW	1.266	0.261	7.50000	5.92199
DVD-RW	0.000	1.000	0.00000	2.59230
DVD-DUAL	-10.614	0.000***	-2.60000	0.24495
DVD-RAM	-5.715	0.005**	-1.40000	0.24495

Note ** p<0.01; *** p<0.001; Two-tailed test.

The connection differences between these two groups are shown in Table 2. Only on two technologies, DVD-RAM and DVD-Dual, these two groups have significant differences. From Fig.1 and Fig.2, user could observe that Sanyo and Samsung not only connect with Phillip, the most important supplier for DVD+RW technology, but also connect with Toshiba, NEC and Sony, the most important suppliers of DVD-RW technology. Meanwhile, Taiwan's makers only connect with the Phillip for getting DVD+RW technology. That is to say, Sanyo and Samsung have extra social capital for creating social chance to enter the blue ocean in the future. Surprisingly, the experiment results in Table 2 shows no significant difference on DVD+RW and DVD-RW between the above two groups (i.e., for DVD+RW 0.261>0.1 and for DVD-RW

1.0>0.1), because in 2002, these technologies are absolutely necessary for manufacturing DVD player for all makers. Therefore, all firms have to compete in the red ocean.

If user only focused on technological innovation, we collected the DVD industrial news from 1997 to 2002. The primary date of the value words of DVD is as determined as following:

- 1) After preprocessing document, the date of the value word is set as the date of the news.
- 2) For the same value word, sort all dates. The earliest date is defined as the primary date of the value word, or the primary year of a new technology.
- 3) Based on word to word associative analysis, create the technology road map (TRM).

Fig. 4 describes the process of generating road map for DVD technologies.

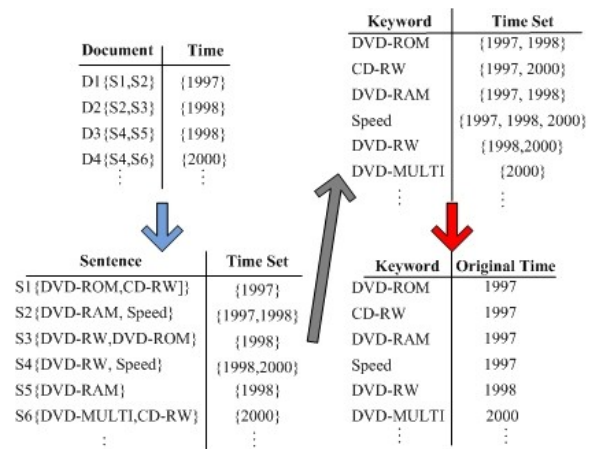


Figure 4. Process of generating technology road map

According to the above process, QCD system sketches the technology road map of DVD as shown in Fig.5. It is clear in Fig.5 that the evolution of DVD technology started from DVD-ROM and DVD-RAM, passing through DVD-RW and DVD+RW, and finally reached DVD-Dual. This result evidenced that the integration of two technologies, DVD-RW and DVD+RW, formatted DVD-Dual technology.

Previous experiment results show that DVD-Dual is a new technology integrates two DVD specifications, DVD-RW and DVD+RW. From company point of view, we believe a company can enter the market quickly if the company can offer cheaper and friendlier machine for consumer. Meanwhile, from consumer point of view, we also believed if the price was close consumer would not care much about the specifications of DVD in enjoying his leisure life. He would just buy a new DVD-Dual player. That is why foreign companies were more actively moving up to the leading grouping this competition.

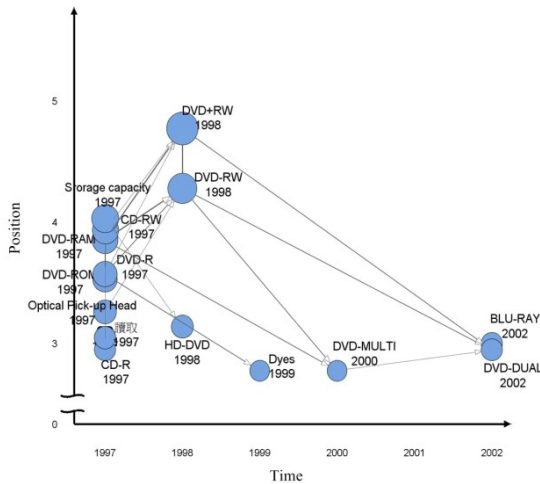


Figure 5. Technology Road Map (TRM)

The yearly frequency of different technologies is very important. It could be used to assess which technology belongs to hot market, which other technology falls into the category of decreasing market, and most importantly, which technology has potential of future innovative market. Our QCD system first calculated yearly frequency for all technologies to find out the year when the technology had the highest appearing frequency. Furthermore, to indicate increasing/decreasing effect, frequencies of present year were also added to compare with the maximum frequency. In the following chart, red colored ring represents decreasing effect, and blue colored ring represents increasing effect.

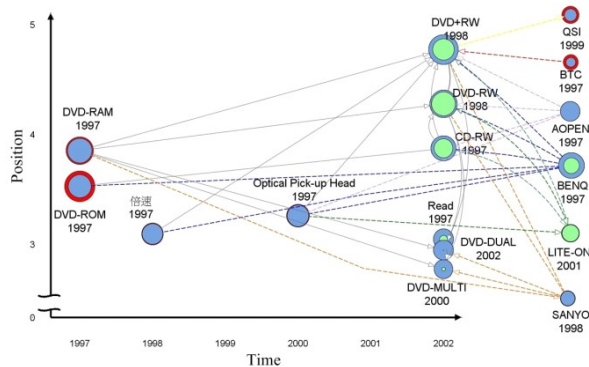


Figure 6. Time depended technology increasing and decreasing effect

In Fig.6, the markers are all relocated to its highest frequency year (i.e., the year shown by the marker is its primary year.) The markers located on 2002 means these technologies are still increasing; the markers located before

2002 means these technologies are decreasing. Furthermore, on the location the year 2002, hot technologies are presented on higher position, and some new technologies are found on lower position. Finally, QCD system puts firm names to the right side and connects them with these technologies accordingly. It is worth noticing that Sanyo connects to both hot technologies and innovative technologies.

In today's harsh consumer product industry, if a firm can find some innovative path to climb to a higher position in the hierarchical social structure, that path is called its social chance. The experimental results (Fig.5 and Fig.6) of this study indicate that Sanyo tried to enter the blue ocean of the competition. Sanyo accomplished this difficult task by using social relations (social capital) to connect with Phillip and Toshiba (Fig.3), and get help in developing DVD-Dual technology, a promising technology that can help Sanyo enter the new market quickly.

IV. CONCLUSION

In this study we developed a two step QCD process model. First, personal social network was created to help visualizing what resource is useful. Second, by using this model, user can integrate all relevant subjects to present the overall social network. By applying QCD model on DVD industry, the experiment results indicate how Sanyo uses its social relations to cooperate with Phillip and Toshiba to create a social chance in developing DVD-Dual technology. This social chance helped Sanyo entered the new market quickly. The contributions in this paper are twofold: first, starting from qualitative analysis our model of social chance discovery process is capable of doing quantitative analysis to fulfill user's need. Second, the visualized social environment generated by can help users visualize some weak path in the need of creating social chance. In the future, more experiments are yet to be done to verify the usefulness of QCD system.

REFERENCES

- [1] F. Yoshikawa, "Decisions by Chance and on Chance: Meanings of Chance in Recent News Stories," Chance Discovery, Springer Verlag, Berlin Heidelberg New York, 2003.
- [2] Y. Ohsawa, KeyGraph, "Visualized Structure Among Event Clusters," Chance Discovery, Springer Verlag, Berlin Heidelberg New York, 2003.
- [3] M. Granovetter, "The strength of weak ties: a network theory revised," Sociological Theory, vol. 1, 1983, pp. 201-233.
- [4] N. Lin, W. M. Ensel, and J. Vaughn, "Social resources and the strength of ties," American Sociological Review, vol. 46, 1981, pp. 393-405.
- [5] P. Marsden, and j. Hurlbert, "Social resources and mobility outcomes: A replication and extension," Social Forces, vol. 67, 1988, pp. 1038-1059
- [6] P. Marsden, and Campell, "Measuring tie strength," Social Forces, vol. 83, 1984, pp. 482-501.
- [7] C. Hong, "Qualitative Chance Discovery-Extracting competitive Advantage," Information Sciences, vol. 179, 2009, pp.1570-1583.