

Chapter 2

Foundations of Intelligent Knowledge Management

Knowledge or hidden patterns discovered by data mining from large databases has great novelty, which is often unavailable from experts' experience. Its unique irreplaceability and complementarity has brought new opportunities for decision-making and it has become important means of expanding knowledge bases to derive business intelligence in the Big Data era. Instead of considering how domain knowledge can play a role in each stage of data mining process, this chapter concentrates on a core problem: whether the results of data mining can be really regarded as "knowledge". The reason is that if the domain knowledge is quantitatively presented, then the theoretical foundation can be explored for finding automatic mechanisms (algorithms) to use domain knowledge to evaluate the hidden patterns of data mining. The results will be useful or actionable knowledge for decision makers. To address this issue, the theory of knowledge management should be applied. Unfortunately, there appears little work in the cross-field between data mining and knowledge management. In data mining, researchers focus on how to explore algorithms to extract patterns that are non-trivial, implicit, previously unknown and potentially useful, but overlook the knowledge components of these patterns. In knowledge management, most scholars investigate methodologies or frameworks of using existing knowledge (either implicit or explicit ones) support business decisions while the detailed technical process of uncovering knowledge from databases is ignored.

This chapter aims to bridge the gap between these two fields by establishing a foundation of intelligent knowledge management over large databases or Big Data. Section 2.1 addresses the challenging problems to data mining. Section 2.2 enables to generate "special" knowledge, called intelligent knowledge base on the hidden patterns created by data mining. Section 2.3 systematically analyzes the process of intelligent knowledge management—a new proposition from original data, *rough* knowledge, *intelligent* knowledge, and *actionable* knowledge as well as the four transformations (4 T) of these items. This study not only promotes more significant research beyond data mining, but also enhances the quantitative analysis of knowledge management on hidden patterns from data mining. Section 2.4 will outline some interesting research directions that will be elaborated in the rest of chapters.

2.1 Challenges to Data Mining

Since 1970s, researchers began systematically exploring various problems in knowledge management (Rickson 1976). However, people have been interested in how to collect, expand and disseminate knowledge for a long time. For example, thousands of years ago, Western philosophers studied the awareness and understanding of the motivation of knowledge (Wiig 1997). The ancient Greek simply believed that personal experience forms all the knowledge. Researchers at present time pay more attention to management of tacit knowledge and emphasize on management of people as focusing on people's skills, behaviors and thinking patterns (Wang 2004; Zhang et al. 2005).

Thanks to the rapid development of information technology, many western companies began to widely apply technology-based tools to organize the internal knowledge innovation activities. Thus it drove a group of researchers belonging to technical schools to explore how to derive knowledge from data or information. For instance, Beckman (1997) believes that knowledge is a kind of humans' logical reasoning on data and information, which can enhance their working, decision-making, problem-solving and learning performance. Knowledge and information are different since knowledge can be formed after processing, interpretation, selection and transformation of information (Feigenbaum 1977).

In deriving knowledge by technical means, data mining becomes popular for the process of extracting knowledge, which is previously unknown to humans, but potentially useful from a large amount of incomplete, noisy, fuzzy and random data (Han and Kamber 2006). Knowledge discovered from algorithms of data mining from large-scale databases has great novelty, which is often beyond the experience of experts. Its unique irreplaceability and complementarity has brought new opportunities for decision-making. Access to knowledge through data mining has been of great concern for business applications, such as business intelligence (Olson and Shi 2007).

However, from the perspective of knowledge management, knowledge discovery by data mining from large-scale databases face the following challenging problems.

First, the main purpose of data mining is to find hidden patterns as decision-making support. Most scholars in the field focus on how to obtain accurate models. They halt immediately after obtaining rules through data mining from data and rarely go further to evaluate or formalize the result of mining to support business decisions (Mcgarry 2005). Specially speaking, a large quantity of patterns or rules may be resulted from data mining. For a given user, these results may not be of interest and lack of novelty of knowledge. For example, a data mining project that classifies users as "current users, freezing users and lost users" through the use of decision tree classification algorithm produced 245 rules (Shi and Li 2007). Except for their big surprise, business personnel cannot get right knowledge from these rules (Shi and Li 2007). The expression of knowledge should not be limited to numbers or symbols, but also in a more understandable manner, such as graphics, natural languages and visualization techniques. Knowledge expressions and qualities from different data mining algorithms differ greatly, and there are inconsistencies,

even conflicts, between the knowledge so that the expression can be difficult. The current data mining research in expressing knowledge is not advanced. Furthermore due to the diversification of data storages in any organizations, a perfect data warehouse may not exist. It is difficult for data mining results based on databases or data warehouses to reflect the integration of all aspects of data sources. These issues lead to the situation that the data mining results may not be genuinely interesting to users and can not be used in the real world. Therefore, a “second-order” digging based on data mining results is needed to meet actual decision-making needs.

Second, many known data mining techniques ignore domain knowledge, expertise, users’ intentions and situational factors (Peng 2007). Note that there are several differences between knowledge and information. Knowledge is closely related to belief and commitment and it reflects a specific position, perspective or intention. Knowledge is a concept about operations and it always exists for “certain purposes”. Although both knowledge and information are related to meaning, knowledge is in accordance with the specific situation and acquires associated attributes (Nonaka et al. 2000; Zeleny 2007). From the culture backgrounds of knowledge, Westerners tend to emphasize on formal knowledge, while Easterners prefer obscure knowledge. It is also believed that these different kinds of knowledge are not totally separated but complementary to each other. In particular, they are closely linked in terms of how human and computer are interacted in obtaining knowledge. Because of the complexity of knowledge structure and the incrementality of cognitive process, a realistic knowledge discovery needs to explore interactively different abstraction levels through human-computer interaction and then repeat many times. Keeping the necessary intermediate results in data mining process, guiding role of human-computer interaction, dynamic adjusting mining target, and users’ background knowledge, domain knowledge can speed up the process of knowledge excavation and ensure the effectiveness of acquired knowledge. Current data mining tools are unable to allow users to participate in excavation processes actually, especially for second-order excavation. In addition, both information and knowledge depend on specific scenarios, and they are relevant with the dynamic creation in humans’ social interaction. Berger and Luckman (1966) argued that interacting people in certain historical and social scenario share information derived from social knowledge. Patterns or rules generated from data mining must be combined with specific business context in order to use in the enterprise. The context here includes relevant physics, business and other externally environmental and contextual factors, which also covers cognition, experience, psychology and other internal factors of the subject. It is the key element to a complete understanding of knowledge, affecting people’s evaluation about knowledge. A rule may be useful to enterprises in a certain context, for a decision maker, at a certain time, but in another context it might be of no value. Therefore, context is critical for data mining and the process of the data mining results. In the literature, the importance of context to knowledge and knowledge management has been recognized by a number of researchers (Dieng 1999; Brezillion 1999; Despres 2000; Goldkuhl 2001; Cap 2002). Though people rely on precise mathematical expressions for scientific findings, many scientific issues cannot be interpreted by mathematical forms. In fact in

the real world, the results of data mining should be interacted effectively with the company reality and some non-quantitative factors before they are implemented as actionable knowledge and business decision support. These factors include the bound of specific context, expertise (tacit knowledge), users' specific intentions, domain knowledge and business scenarios (Zhang et al. 2008).

Third, common data mining process stops at the beginning of knowledge acquisition. The organizations' knowledge creation process derived from data should use different strategies to accelerate the transformation of knowledge in different stages of the knowledge creation, under the guidance of organizational objectives. Then a spiral of knowledge creation is formed, which creates conditions for the use of organizational knowledge and the accumulation of knowledge assets. At present, data mining process only covers knowledge creation part in this spiral, but does not involve how to conduct a second-order treatment to apply the knowledge to practical business, so as to create value and make it as a new starting point for a new knowledge creation spiral. Therefore, it cannot really explain the complete knowledge creation process derived from data. There is currently very little work in this area. In the ontology of data mining process, the discovered patterns are viewed as the end of the work. Little or no work involving the explanation of knowledge creation process at organizational level is studied in terms of implementation, authentication, internal process of knowledge, organizational knowledge assets and knowledge recreation. From the epistemological dimension, it lacks a deep study about the process of data - information - knowledge - wisdom, and the cycle of knowledge accumulation and creation is not revealed. A combination of organizational guides and strategies needs to decide how to proceed with the knowledge guide at the organizational level so that a knowledge creation process derived from data (beyond data mining process) and organizational strategies and demands can be closely integrated.

Based on the above analysis, in the rest of this book, the knowledge or hidden patterns discovered from data mining will be called "rough knowledge." Such knowledge has to be examined at a "second-order" in order to derive the knowledge accepted by users or organizations. In this book, the new knowledge shall be called "intelligent knowledge" and the management process of intelligent knowledge is called intelligent knowledge management. Therefore, the focus of the study has the following dimensions:

- The object of concern is "rough knowledge".
- The stage of concern is the process from generation to decision support of rough knowledge as well as the "second-order" analysis of organizational knowledge assets or deep-level mining process so as to get better decision support.
- Not only technical factors but also non-technical factors such as expertise, user preferences and domain knowledge are considered. Both qualitative and quantitative integration have to be considered.
- Systematic discussion and application structure are derived for the perspective of knowledge creation.

The purposes of proposing intelligent knowledge management are:

- Re-define rough knowledge generated from data mining for the field of knowledge management explicitly as a special kind of knowledge. This will enrich the connotation of knowledge management research, promote integration of data mining and knowledge management disciplines, and further improve the system of knowledge management theory in the information age.
- The introduction of expertise, domain knowledge, user intentions and situational factors and the others into “second-order” treatment of rough knowledge may help deal with the drawbacks of data mining that usually pays too much emphasis on technical factors while ignoring non-technical factors. This will develop new methods and ideas of knowledge discovery derived from massive data.
- From the organizational aspect, systematic discussion and application framework derived from knowledge creation based on massive data in this paper will further strengthen and complement organizational knowledge creation theory.

2.2 Definitions and Theoretical Framework of Intelligent Knowledge

In order to better understand intelligent knowledge management, basic concepts and definitions are introduced in this subsection.

The research of intelligent knowledge management relates to many basic concepts such as original data, information, knowledge, intelligent knowledge and intelligent knowledge management. It is also associated with several relevant concepts such as congenital knowledge, experience, common sense, situational knowledge etc. In order to make the proposed research fairly standard and rigorous from the beginning, it is necessary to give the definition of these basic concepts. Moreover, the interpretation of these concepts may provide a better understanding of intrinsic meanings of data, information, knowledge, and intelligent knowledge.

Definition 2.1 Data is a certain form of the representation of facts.

The above definition that is used in this paper has a general meaning of “data.” There are numerous definitions of data from different disciplines. For example, in computing, data is referred to distinct pieces of information which can be translated into a different form to move or process; in computer component or network environment, data can be digital bits and bytes stored in electronic memory; and in telecommunications, data is digital-encoded information (Webopedia 2003; Whatis.com 2005). In information theory, data is abstractly defined as an object (thing) that has the self-knowledge representation of its state and the state’s changing mode over time (Zhong 2007). When it is a discrete, data can be expressed mathematically a vector of n-dimensional possible attributes with random occurrences. Without any physical or analytic processing to be done, given data will be treated as “original” in this paper. Therefore, original data is the source of processing other forms (such as information, rough knowledge, intelligent knowledge and others).

From the perspective of forms, the data here includes: text, multimedia, network, space, time-series etc.

From the perspective of structure, the data includes: structured, unstructured and semi-structured data; as well as more structured data which current data mining or knowledge discovery can deal with.

From the perspective of quantity, the data includes: huge amounts of data, general data and small amounts of data etc.

Data, judging from its nature, is only the direct or indirect statements of facts. It is raw materials for people to understand the world.

Therefore, the characteristics of the original data here include: roughness (original, roughness, specific, localized, isolated, superficial, scattered, or even chaotic), extensive (covering a wide range), authenticity and manipulability (process through data technology). After access to original data, appropriate processing is needed to convert it into abstract and universal applicable information. Thus, the definition of information is given as:

Definition 2.2 Information is any data that has been pre-processed to all aspects of human's interests.

Traditionally, information is the data that has been interpreted by human using certain means. Both scientific notation and common sense share the similar concepts of information. If the information has a numerical form, it may be measured through the uncertainty of an experimental outcome (The American Heritage Dictionary of the English Language 2003), while if it cannot be represented by numerical form, it is assigned for an interpretation through human (Dictionary of Military and Associated Terms 2005). Information can be studied in terms of information overload. Shi (2000) classified information overload by exploring the relationships between relevant, important and useful information. However, definition 2 used in this paper is directly for describing how to get knowledge from data where information is an intermediate step between these two. It is assumed that the pre-processed data by either quantitative or qualitative means can be regarded as information. Based on the concepts of data and information, the definition of rough knowledge is presented as follows:

Definition 2.3 Rough Knowledge is the hidden pattern or “knowledge” discovered from information that has been analyzed by the known data mining algorithms or tools.

This definition is specifically made for the results of data mining. The data mining algorithms in the definition means any analytic process of using artificial intelligence, statistics, optimization and other mathematics algorithms to carry out more advanced data analysis than data pre-processing. The data mining tools are any commercial or non-commercial software packages performing data mining methods. Note that data pre-processing normally cannot bring a qualitative change of the nature of data and results in information by definition 2, while data mining is advanced data analysis that discovers the qualitative changes of data and turns information into knowledge that has been hidden from human due to the massive data. The representation of rough knowledge changes with a data mining method.

For example, rough knowledge from association method is rules, while it is a confusion matrix for the accuracy rates by using a classification method.

The purpose of defining data, information and rough knowledge is to view a general expression of the data mining process. This paper will call the process and other processes of knowledge evolution as “transformations.”

The transformation from data (or original data) to rough knowledge via information is called the first transformation, denoted as T_1 . Let K_R stand for the rough knowledge and D denote as data. Then the first type of transformation can be expressed as:

$$T_1 : D \rightarrow K_R \quad \text{or} \quad K_R = T_1(D)$$

As it stands, T_1 contains any data mining process that consists of both data preprocessing (from data to information) and data mining analysis (from information to rough knowledge). Here the main tasks of T_1 can include: characterization, distinction, relevance, classification, clustering, outlier analysis (abnormal data), evolution analysis, deviation analysis, similarity, timing pattern and so on. Technologies of T_1 include extensively: statistical analysis, optimization, machine learning, visualization theory, data warehousing, etc. Types of rough knowledge are potential rules, classification tags, outlier labels, clustering tags and so on.

Characteristics of rough knowledge can be viewed as:

- (i) Determined source: from results of data mining analysis;
- (ii) Part usability: the possibility of direct support for business may exist, but much may not be used directly.
- (iii) Rough: without further refinement, rough knowledge contains much redundant, one-sided or even wrong knowledge. For example, the knowledge generated from over-training has high prediction accuracy rate about the test set, but the effect is very poor;
- (iv) Diversity: knowledge needs to be shown by a certain model for decision-making reference. There are many forms of rough knowledge, for instance, summary description, association rules, classification rules (including decision trees, network weights, discriminant equations, probability map, etc.), clusters, formulas and cases and so on. Some representations are easy to understand, such as decision trees, while some manifestations have poor interpretability, such as neural networks.
- (v) Timeliness: compared with humans' experience, rough knowledge is derived from data mining process in a certain time period, resulting in short cycle. It may degrade in the short term with environmental changes. In addition, there are conflicts sometimes between the knowledge generated from different periods. As a result, as the environment changes the dynamic adaptability can be poor.

While rough knowledge is a specific knowledge derived from the analytic data mining process, the human knowledge has extensively been studied in the field of knowledge management. The item knowledge has been defined in many different

ways. It is generally regarded as individual's expertise or skills acquired through learning or experience (Wikipedia 2008). In the following, knowledge is divided as five categories in terms of its contents. Then, these can be incorporated into rough knowledge from data mining results for our further discussion on intelligent knowledge.

Definition 2.4 Knowledge is called Specific Knowledge, denoted by K_S if it contains the certain state and rules of an object expressed by human.

Specific knowledge is a cognitive understanding of certain objects and can be presented by its form, content and value (Zhong 2007). Specific knowledge has a strict boundary in defining its meanings. Within the boundary, it is knowledge; otherwise, it is not (Zeleny 2002).

Definition 2.5 Knowledge is called Empirical Knowledge, denoted by K_E if it directly comes from human's experience gained from empirical testing.

Note that the empirical testing in definition 5 is referred to specifically non-technical, but practical learning process from which human can gain experience. If it is derived from statistical learning or mathematical learning, knowledge is already defined as rough knowledge of definition 2.2. Empirical testing here can be also referred to as intermediate learning, such as reading from facts, reports or learning from others' experiences. When these experiences are verified through a scientific learning, they will become "knowledge". Otherwise, they are still "experiences" (Zhong 2007).

Definition 2.6 Knowledge is called Common Sense Knowledge, denoted as K_C if it is well known and does not need to be proved.

Common sense is the facts and rules widely accepted by most of humans. Some knowledge, such as specific knowledge or empirical knowledge can become common sense as they are gradually popularized. Therefore, it is also called "post-knowledge" (Zhong 2007).

Definition 2.7 Knowledge is called Instinct Knowledge, denoted by K_H if it is innate as given functions of humans.

Instinct knowledge is heritage of humans through the biological evolution and genetic process. It does not need to be studied and proved. If instinct knowledge is viewed as a "root" of the knowledge mentioned above, then a "knowledge ecosystem" can be formed. In the system, instinct knowledge first can be changed into empirical knowledge after training and studying. Then, if empirical knowledge is scientifically tested and confirmed, it becomes specific knowledge. As the popularity of specific knowledge develops, it is common sense knowledge. However, the system is premature since the creation of human knowledge is quite complex and could not be interpreted as one system (Zhong 2007).

Definition 2.8 Knowledge is called Situational Knowledge, denoted as K_U if it is context.

The term *context* used in this paper, associated with knowledge and knowledge activities, is relevant to conditions, background and environment. It includes not

only physical, social, business factors, but also the humans' cognitive knowledge, experience, psychological factors (Pan 2005).

Situational knowledge or context has the following characteristics:

- (i) It is an objective phenomenon which exists widely;
- (ii) It is independent of knowledge and knowledge process, but keeps a close interaction with knowledge and knowledge process;
- (iii) It describes situational characteristics of knowledge and knowledge activities. Its function is to recognize and distinguish different knowledge and knowledge activities. To humans, their contexts depict personal characteristics of one engaging in intellectual activities (Pan 2005).

Based on the above definitions of different categories of knowledge, a key definition of this paper is given as:

Definition 2.9 Knowledge is called Intelligent Knowledge, denoted as K_I if it is generated from rough knowledge and/or specific, empirical, common sense and situational knowledge, by using a “second-order” analytic process.

If data mining is said as the “first-order” analytic process, then the “second-order” analytic process here means quantitative or qualitative studies are applied to the collection of knowledge for the pre-determined objectives. It can create knowledge, now intelligent knowledge, as decision support for problem-solving. The “second-order” analytic process is a deep study beyond the usual data mining process. While data mining process is mainly driven by a series of procedures and algorithms, the “second-order” analytic process emphasizes the combinations of technical methods, human and machine interaction and knowledge management.

Some researchers in the field of data mining have realized its importance of handling the massive rules or hidden patterns from data mining (Ramamohanarao 2008; Wong 2008; Webb 2008). However, they did not connect the necessary concepts from the field of knowledge management in order to solve such a problem for practical usage. Conversely, researchers in knowledge management often ignore rough knowledge created outside humans as a valuable knowledge base. Therefore, to bridge the gap between data mining and knowledge management, the proposed study on intelligent knowledge in the paper is new.

As discussed above, the transformation from information to rough knowledge T_1 is essentially trying to find some existing phenomenological associations among specific data. T_1 is some distance away from the knowledge which can support decision-making in practice. The “second-order” analytic process to create intelligent knowledge from available knowledge, including rough knowledge, can be realized in general by transformation T_2 , defined as follows:

$$T_2 : K_R \cup K \rightarrow K_I \quad \text{or} \quad K_I = T_2(K_R \cup K),$$

where $K = \rho(K_S, K_E, K_C, K_H, K_U)$ is a power set.

K_S Specific Knowledge;

K_E Empirical Knowledge;

K_C Common Sense Knowledge;
 K_H Instinct Knowledge;
 K_U Situational Knowledge.

The above transformation is an abstract form. If the results of the transformation are written in terms of the components of intelligent knowledge, then the following mathematical notations can be used:

- (i) Replacement transformation: $K_I = K_R$
- (ii) Scalability transformation: $K_I = \alpha K_R$, where $-\infty < \alpha < +\infty$
- (iii) Addition transformation: $K_I = K_R + K_I$
- (iv) Deletion transformation: $K_I = K_R - K_I$
- (v) Decomposition transformation:

$$K_I = \alpha_1 K_{R1} + \alpha_2 K_{R2} + \alpha_3 K_{R3} + \dots, \text{ where } -\infty < \alpha_i < +\infty$$

In the above, replacement transformation is a special case of scalability transformation, and they, together with addition and deletion transformations are parts of decomposition transformation.

The coefficients of $\{\alpha_1, \alpha_2, \alpha_3 \dots\}$ in the decomposition represent the components of $K = \rho(K_S, K_E, K_C, K_H, K_U)$ distributed in the knowledge creation process.

The intelligent knowledge has the following characteristics:

- (i) The process of intelligent knowledge creation fully integrates specific context, expertise, domain knowledge, user preferences and other specification knowledge, and makes use of relevant quantitative algorithms, embodying human-machine integration principle;
- (ii) Since intelligent knowledge is generated from the “second-order” analytic process, it is more valuable than rough knowledge;
- (iii) It provides knowledge to people who need them at the right time, under appropriate conditions.
- (iv) The objective of intelligent knowledge is to provide significant inputs for problem-solving and support strategic action more accurately.

To explore more advanced issues in the meaning of knowledge management, intelligent knowledge can be further employed to construct a strategy of problem-solving by considering goal setting, specific problem and problem environment.

Restricted by the given problem and its environmental constraints, aiming at the specific objectives, a strategy of solving the problem can be formed based on related intelligent knowledge. To distinguish the strategy that has been used in different fields, the strategy associated with intelligent knowledge is called intelligent strategy.

Intelligent Knowledge

A Study beyond Data Mining

Shi, Y.; Zhang, L.; Tian, Y.; Li, X.

2015, XVI, 150 p. 47 illus., 24 illus. in color., Softcover

ISBN: 978-3-662-46192-1