International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 3 (2011) pp. 165 -173

MinEDec: a Decision-Support Model That Combines Text-Mining Technologies with Two Competitive Intelligence Analysis Methods

Yue Dai, Tuomo Kakkonen, Erkki Sutinen

School of Computing, University of Eastern Finland
Joensuu, Finland
{yvedai, tkakkone, sutinen}@cs.joensuu.fi

Abstract: **With the development of information technologies, enterprises are working in a complex and dynamic environment. Monitoring the competitive environment and effectively transforming data into knowledge for decision makers are vital competitive powers. In order to monitor and analyse the competitive environment of businesses, we integrate two well-known competitive intelligence analysis methods, the Five Forces Analysis (FFA) and a SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis with various text-mining technologies in a decision support model Mining Environment for Decisions (MinEDec). We also outline the architecture of a decision support system that is based on the proposed model.**

Keywords: decision support system, competitive intelligence, text mining, the Five Forces framework, SWOT analysis.

I. Introduction

Today enterprises are working in a complex, open and mobilising environment with the ever-inflating information in the modern societies. Changes of environment redefine the way in which business enterprises compete and make decisions, hence affecting their strategies. An enterprise's strategy comes from the need for market share. It consists of all the competitive actions and operational measures used by the leaders and managers [1].

Leaders in modern enterprises deal with decisions that require the integration of both internal and external information from a variety of sources. Several activities need to take place in order to make the right decisions: understanding the external environment, listening to the internal context, analysing and summarising information, and communicating the results in an effective way. Information technologies, such as enterprise-wide systems and data mining can help capture and integrate transactions from a variety of perspectives to support decisions [2].

Competitive intelligence (CI) is a process of monitoring the competitive environment by pulling together data and information from a very large and strategic perspective, to predict or forecast what is going to happen in the competitive environment of an enterprise [3]. Consequently, more and more enterprises are using CI analysis methods, such as competitive positioning analysis, benchmarking analysis, Five Forces Analysis (FFA) and SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis, in support of their decision-making processes [4]. There are various software tools in the market that claim to help the collection and analysis of CI. But none of them can automatically analyse information and generate intelligence through using multiple CI analysis methods and text mining (TM) to support decision making [3, 4, 14, 15].

A decision support system (DSS) helps leaders to make decisions that are unique, rapidly changing, and not easily specified in advance [5]. DSSs have been developed since the middle of the 1970s [15]. There are new frameworks that are based on the traditional DSS. For example, intelligent decision support system (IDSS) adds artificial intelligence (AI) functions to traditional DSS in order to guide users through the decision-making phases and tasks or to supply new capabilities [6]. But most DSSs use data mining to do mathematical and statistical analysis rather than using TM [6, 13, 15].

An emerging set of technologies that promise a great potential for the future of CI and DSS come from the field of TM. TM is closely related to other fields of research on the intersection between computer science and linguistics, namely natural language processing (NLP) and computational linguistics. The aim in TM is to derive useful information from written texts by using methods such as entity extraction, clustering, categorisation and sentiment analysis.

The target of the research described in this article is to establish a decision support model that leverages various TM technologies, SWOT analysis, and the FFA framework to search and analyse unstructured textual data (e.g. newspapers, online sources, customer feedback, reports, and email). By providing the ability of CI analysis based on text sources, the DSS based on our proposed model will be able to seize early warnings of threats and opportunities in the business environment, which are necessary for the proactive strategy of enterprises. In order to create such a system, we need to have a clearly defined model for decision support that functions as the basis of the system design. This paper describes the model as well as the system architecture we have designed.

In Section 2, the background of the current study is described. Section 3 presents the Mining Environment for Decisions (MinEDec) model that combines FFA with SWOT analysis and TM technologies. Section 4 outlines the

CI and DSS system that implements the model. The paper concludes in Section 5 with final remarks.

II. Background

A. The Five Forces framework and MinerVA

It is necessary for leaders in any enterprise to understand the competitive forces in their industry since these will determine the likely successes or failures of particular enterprises within it [1]. The FFA framework is a CI analysis model, which was developed for such business environment analyses (Figure 1) [7].

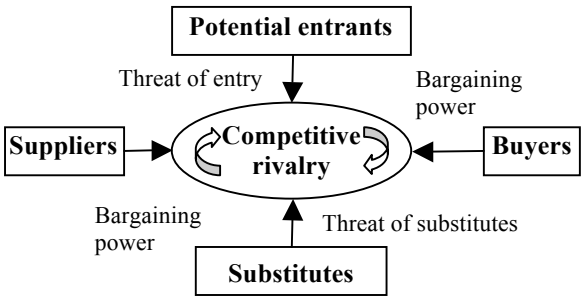

Figure 1. The Five Forces framework [7, 8]

As illustrated in Figure 1, according to the FFA framework rivals, potential entrants, substitute products, suppliers, and buyers are the five basic parties in a competitive environment. The threat of entry means that new entrants will add capacity to the industry and increase the demand and prices, resulting in lower industry profitability. The threat of substitutes describes the risk of market displacement from existing or potential substitutes. The bargaining power of suppliers defines the ability of suppliers to influence the cost, availability, and quality of input materials. The bargaining power of buyers allows the buyers to influence properties such as prices and quality expectations [1, 4, 7]. Understanding and focusing on these five subjects affect both the profit potential and the prospects for achieving competitive advantage.

In [8], we proposed MinerVA a decision-support model based on the FFA framework. The model integrates FFA with three advanced TM technologies – opinion mining, event change detection, and patent trend change mining to monitor the external business environment. The MinEDec model presented in this work is an extension to MinerVA.

B. SWOT analysis

SWOT analysis is a CI method for planning future-oriented strategies [1, 9]. A SWOT analysis is used to evaluate threats and opportunities from a turbulent environment. A SWOT analysis summarises the strengths and weaknesses of a company in order to address the issues that the company is facing or will face, and finds a proper strategic plan. The output of the SWOT analysis, referred to as SWOT matrix, is shown in Table 1.

As illustrated in Table 1, the decision makers need to search the environmental changes and recognise the factors in four categories: strengths, weaknesses, opportunities and threats. The strengths and weaknesses are about the internal environment of the company, compared with the competitors of the company, and generally reflect the company's technology, equipment, personnel, products, markets, management structure, and so on. S1, S2, W1 and W2 represent each factor that belongs to strengths or weaknesses.

The opportunities and threats refer to the external environment factors, which are favourable or unfavourable to the enterprise that is being analysed. In Table 1, O1, O2, T1, and T2 refer to each of these factors. While examples of favourable factors include high technology and a good relationship between buyers, among adverse factors are, for instance, trade policy changes, unexpected events, market changes, and the emergence of competitors.

Four types of strategies can be defined from the SWOT matrix through combining different category factors together. The SO strategy (positive strategy) uses strength points of the business in order to make use of the opportunities. The WO strategy (differentiation strategy) aims at diminishing the weak factors by grasping the opportunities. The ST strategy (gradual strategy) uses strength to reduce the threat factors. The WT strategy (negative or withdrawal strategy) uses defensive approaches to cover the weaknesses and to avoid the threats [10, 11].

Table 1. SWOT matrix [10]

Strategy *SW* / *OT*	**Strengths** S1 S2	**Weaknesses** W1 W2
Opportunities O1 O2	SO strategy S1O1, S1O2 S2O1, S2O2	WO strategy W1O1, W1O2 W2O1, W2O2
Threats T1 T2	ST strategy S1T1, S1T2 S2T1, S2T2	WT strategy W1T1, W1T2 W2T1, W2T2

C. TM technologies

TM refers to the process of deriving high-quality information from texts [12]. It is based on the theoretical foundation of the computational linguistics on one hand and mathematical statistics and data analysis on the other. The technologies used in TM include information retrieval (IR), information extraction (IE), topic tracking, summarisation, categorisation, concept linkage1, information visualisation, and question answering [12, 13].

The main advantages of using TM technologies for CI are the ability to process large amounts of textual data quickly and the objectivity and customisability of the process. The TM process typically includes the following steps [14]:

1) preprocessing input texts into the required format for further analysis (data processing);
2) extraction of important concepts and terms through initial text analysis (concept extraction);
3) identifying patterns and co-occurrences of identified concepts (narrative analysis);
4) developing an automated solution (automatic categorisation); and
5) building a taxonomy of concepts.

In addition to concept extraction, event detection can be performed as part of the extraction process. The aim of this process in a CI system could be, for example, to detect business events such as mergers, acquisitions and product launches.

[1] Concept linkage tools connect related documents by identifying their shared concepts, helping users find information they perhaps would not have found through traditional search methods.

Table 2. The evaluation of existing CI software [3, 4, 14, 23]. Key: DM = data mining, TM = text mining, SD = structured data, UT = unstructured text

Tool name	Vendor	Type of tool	CI methods	Data sources
Innovator	Goldfire	TM	Product lifecycle management, enterprise resource planning	UT from personal data, corporate data, deep web, patent, etc.
Business Objects	SAP	TM	Enterprise information management, performance management	UT from web, internal repositories; preprocess documents to extract meta-data and identify entity types
Knowledge. Works	Cipher	File cabinet	Competitor matrix	SD and UT inside and outside the organisations
WebQL	Caesius	TM	Pricing analytics, customer value analysis	SD and UT from web, internal repositories
STRATEGY!	Strategy Software, Inc.	TM and data visualisation	Benchmarking, SWOT analysis, competitor response profile	SD and UT
Wincite	Wincite Systems LLC	TM and data visualisation	Product/company SWOT, competitor analysis, sales analysis, Porter model	SD and UT from internal repositories
Enterprise Miner	SAS	DM	Modelling and assessment, statistical analysis	SD from web, internal repositories
Wisdom Builder	Wisdom Builder	TM	Knowledge management, relationship management	SD and UT from web, internal repositories
ClearResearch Suite	ClearForest	DM and TM	Knowledge management	SD and UT from web, internal repositories
LUXID®	Temis	TM	Competitor analysis, strategy management, weak signals	SD and UT from web, internal documents, patents, email, etc.

D. Existing TM systems for competitive intelligence

Table 2 summarises the properties of various CI software tools that are available on the market place. It has indicated that while many of the existing software tools claim to be capable of processing texts, their CI capabilities are in reality almost completely focused on numerical data analysis; they have limited or no ability of CI analysis powered by TM. While many of the systems support well-known CI analysis methods, these methods are used independently of each other [3, 4, 14, 15, 23]. Based on the evaluation we concluded that what is currently lacking is an integrated framework that can provide the objectives to analyse and summarise the huge amount of available textual data by using multiple perspectives and models of CI analysis.

III. MinEDec - Integrating SWOT analysis, the Five Forces framework and TM

As discussed in Section 2, integrated CI analysis models and TM technologies have a great potential to support CI and decision making. We believe that integrating CI analysis methods and TM techniques better than has been done thus far will maximise their benefits; the sum of this integrated model will be more than its parts. The main issue to solve is to select the CI and TM methods that are best suited for our purposes and how the integration should be done in order to get the best possible benefits.

A. From data to knowledge

Data is the string of symbols, facts, measurements, statistics, but is not organised to convey any specific meaning: numeric, figures, etc. Information is organised from data in a manner that give it meaning for the recipient; information is data with context and relationships. Intelligence is analysed and value-added information. Knowledge consists of information organised and processed to convey understanding, and experience that are applicable to a current problem or activity (Figure 2) [15, 16].

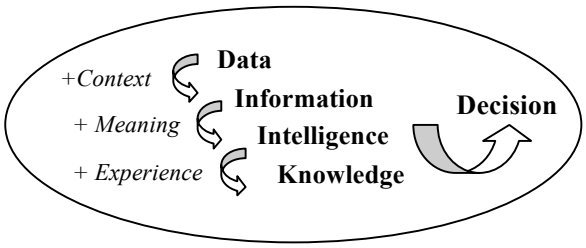

Figure 2. From data to knowledge, to decision [4]

As illustrated in Figure 2, data with context equals information, information with meaning can be intelligence, and intelligence with experience generates knowledge. It reflects the qualitative changes from data to knowledge. Intelligence and/or knowledge are the basis for making decisions, and they must be a useful format to meet strategic needs for enterprises.

Hence, the aim of our decision-support model is to distil unstructured textual data into knowledge that is useful to business decision makers. In order to transfer texts into strategic intelligence, we need to combine CI analysis methods and TM technologies during the transforming process (Figure 3).

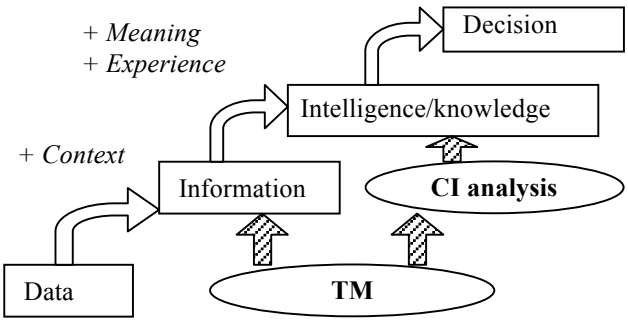

Figure 3. From data to decision supporting by TM and CI

Figure 3 explains the purpose of our model. As illustrated in Figure 3, the technologies of TM, such as IR and IE, can be used to search and summarise unstructured data by adding context, and TM can also support CI analysis models. CI analysis models can extract intelligence/knowledge by adding meaning and experience to information. This can be achieved by combining newly found facts with the knowledge stored in the background knowledge database. Such a model enables leaders to obtain intelligence from countless unstructured data sources, which enables them to make decisions more easily and reliably with the help of TM and CI analysis models.

B. Integrating SWOT analysis and the Five Forces framework

1) Combining SWOT and FFA

FFA is accepted by most researchers of CI as a useful way of analysing the competitive environment. It provides clear objectives and a systematic approach to identify and analyse the relevant business trends and events based on the influence of each of the five factors not only within themselves but also across each of the other forces. The use of FFA in our proposed model provides several advantages. First, it guides information collection, because it helps to identify a set of analytical subjects and makes leaders aware of what data and information to look for. Second, it makes other CI analysis models, such as SWOT analysis, more efficient and focused, because FFA narrows the analysis objectives into five specific objectives [4]. Moreover, these five objectives define the whole framework of an industry, and they are the most important components of the business environment. If leaders can seize the intelligence about these five objectives, they can catch more opportunities for success.

Using SWOT analysis in tandem with FFA can help greatly in understanding the internal environment and the external environment and assessing future courses of action as well as generating strategic options [1, 4].

As illustrated in Figure 4, focusing on the five objectives, we collect major factors that belong to the strengths/weaknesses category or that belong to the opportunities/threats category. These factors can also be used in the SWOT matrix; the integrating matrix can give suggestions about choosing a strategy (SO, WO, ST, or WT) by combining different factors with different objectives.

In order to effectively integrate the two CI analysis models – FFA and SWOT – into one unified decision-support model, we need to consider carefully the properties of both models. We integrate all the properties in Table 2 and the entire factors can be used as the keywords to do information retrieval and IE in MinEDec [17, 18].

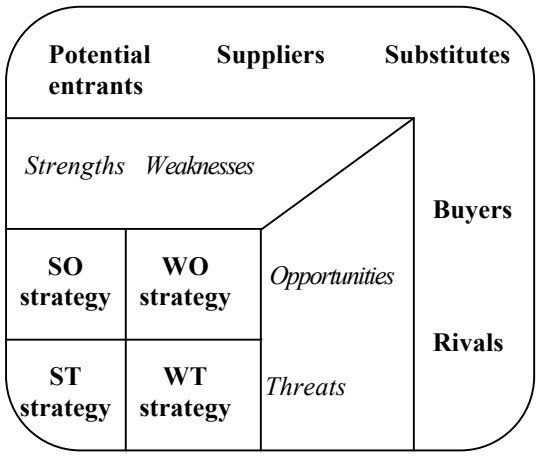

Figure 4. Integrating the Five Forces framework with SWOT matrix

2) Five internal factors and four external factors

As illustrated in Table 3, there are five factors for evaluating internal strengths and weaknesses in our model: *technology*, *price*, *equipment*, *service*, and *attitude*. The leader can obtain the result about each factor for each objective separately. For example, by using an individual objective and factor they can, for example, examine the technology profile of rivals or the attitude of buyers. Alternatively, the leader may decide to use all the factors of an objective to evaluate the whole situation. In addition to analysing their own company, leaders can use the information about technology, price, equipment, service, and buyer's attitude to establish competitors' profiles or suppliers' profiles. Through combining different factors and objectives, new knowledge and intelligence can be generated to support decision making.

Table 3. The major factors used for both SWOT analysis and the Five Forces framework

SWOT / *Objectives*	**Strengths**		**Weaknesses**			**Opportunities**		**Threats**	
	Technology	Price	Equipment	Service	Attitude	Political Shifts	Economic Shifts	Technological Shifts	Social Shifts
Rivals									
Buyers									
Suppliers									
Substitutes									
Potential entrants									

The factors of external opportunity and threat are more general than internal factors. Hence, they indicate trends that are very important for a proactive strategy. Political shifts, for instance, will give a chance to potential entrants, which would be a threat to one's own company. Because the external factors are so general, we can use internal factors to decide that the situation is an opportunity or a threat. Economic shifts, moreover, can increase the power of buyers and suppliers and provide an opportunity. But if the suppliers are more powerful than buyers, it is a threat to one's own company. Technological shifts may be conducted by rivals, and it could be a threat or an opportunity to one's own company also. Social shifts mean the consumers' attention.

3) Detailed and general SWOT analyses

What we want to achieve is to give leaders sufficient support to make a decision. Hence, our model aims at providing detailed and general SWOT models for each objective. For example, when leaders want to know the customers' general information, the model could provide the present and potential customers' profile of the company; when managers want to know the buyers' attitude, they can obtain SWOT intelligence from summarising the buyers' attitudes toward their own products, rivals' products, and substitute products.

Furthermore, considering the fact that the factors that influence each objective are interrelated, the model can help in summarising the whole competitive environment from all the factors. Moreover, the analysis results can be made easier for the leaders to interpret by using various information visualisation methods that enable quick browsing and interpretation of the results.

4) Example

Let us consider mobile communications as an example of how to use the model. The data for this example was taken from [17]. The companies involved in the mobile communication business in Taiwan need to provide real-time news to their customers by using TM technologies to track the buyers' behaviours and attitudes. This example situation could be presented in our model (see Table 3) as follows: when tracking the buyers' attitudes and buyers' attention shifts (social shifts) in the mobile communication market, we might find out that the buyers' attitude seems that they do not pay enough attention to the products of the companies, which is a weakness, and at the same time, buyers' attention is attracted by real-time news. This clearly is an opportunity. Hence, mobile communication companies need to practice WO strategy to diminish the weak factors by grasping the opportunities. Consequently, the result of the analysis is that they need to provide real-time news to their customers.

If using rival/competitor companies, services and technological shifts as major factors, the mobile communication company can find out that the rivals providing a download ringtone service to buyers, which obviously is a threat. Then decision makers move to focus on buyers and attention, and find out their buyers' attention is declining, which is a weakness. In order to avoid the threat, the mobile communication companies need to eliminate the weakness. So they need to provide a download ringtone service to attract and retain their customers according to WT strategy

Also by using rivals, technology, technological shifts and social shifts as major factors, the decision maker can find out that the 3G service has became a hot topic in Taiwan, so the telecommunication company should launch a 3G market, which is an opportunity.

C. Adding TM in the model

Our proposed model analyses the business environment by combining FFA and SWOT based analysis models discussed in the previous subsection with TM technologies. In order to meet the strategic requirements of decision making, leaders need to be effectively guided through the collection and analysis of information. That is where TM technologies step in our model. Figure 5 outlines our decision-support model – MinEDec. As analysed in the previous sections, each of the components (SWOT, FFA, TM) of our model are clear and well established. They all play an important role in supporting strategic decision making in the unified model.

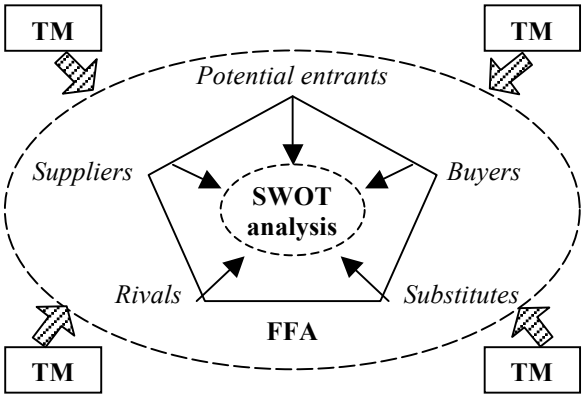

Figure 5. Graphical framework of MinEDec

For the analysis of rivals, TM technologies can be used for tracking the extent of competitive rivalry in the industry, the rivals' products, strategic drift (such as developing an international market in a specific country), and the development of technology and the trends of the whole industry. Furthermore, we can establish a profile for each rival and then use SWOT analysis to show the differences between one's own company and the rival. A detailed SWOT can, for example, be a comparison of own products vs. the rival's products, or own service vs. the rival's service. A general SWOT model can be summarised from detailed SWOTs. A more general model can be used to generate reports about the whole situation of one's own company, including the product, price, service, etc. The decision between the uses of a general or detailed SWOT analysis depends on the strategic needs.

For the analysis of buyers, TM can provide means to detect the attitude of our ultimate customers and immediate customers, and track the behaviours of buyers from these factors: buyer volume tracking from internal data, brand identity, buyers' concentration, price sensitivity tracking from the Internet, the threat of backward integration tracking the buyers' activities from newspapers, analysis of customer opinions from review websites, etc. For the analysis of substitute products, TM can be applied in order to find out developments in related industries and buyer inclination to substitute in order to be aware of the threat in time. TM can be applied in monitoring suppliers to discover the supplier concentration, the impact of materials on cost or

differentiation from opinion mining on the Internet, the presence of substitute material tracking on the Internet, and the threat of forward integration tracking the suppliers' activities from newspapers and other online sources. For potential entrants, the tasks of TM are monitoring the changes in a material's demand and increased price for input [3] [4].

IV. DSS based on MinEDec

This section outlines the architecture of a DSS based on the proposed MinEDec model. MinEDec is being developed as part of a research project entitled "Towards e-leadership: higher profitability through innovative management and leadership systems". The main aim of the project is to develop a TM based DSS system for text documents (both offline and online) that helps business leaders in gathering and analysing CI data for competitor analysis, customer opinion and feedback analysis as well as for aiding decision making. Our system design is based on a reusing and modifying freely available open source Java components, such as Gate [26], Lucene (http://lucene.apache.org/), Jena Semantic Web Framework (http://jena.sourceforge.net/) and YARFRAW (Yet Another RSS Feed Reader And Writer API) (http://yarfraw.sourceforge.net/). This allows us to focus our efforts on the development of CI analysis and decision-support capabilities rather than spending time on "reinventing the wheel".

In order to start a CI analysis process, decision makers first set up the aim of the mining work and define the information sources that are used as inputs. Once the input documents are fetched from offline and online sources (Section 4.1), the system proceeds to applying NLP techniques in order to preprocess the input data before it is passed on to IE and analysis components (Section 4.2). A domain knowledge database (Section 4.3) is needed in order to combine new information with known facts. As a result the system will provide useful intelligence reports about the business environment both in textual and visual format (Section 4.4).

A. Collecting textual data to information

Data collection is an important foundation of a DSS. In order to collect large enough quantities of relevant data, the data sources that the system supports must be diverse. For example, the system can use Internet search engines (for instance, Yahoo and Bing), financial news sites, online newspapers, relevant industry association sites, RSS feeds, and rivals' press releases as the external data sources. For internal text information, the data sources may include emails, reports, and notes, which are generated from other information systems and workflow.

The system utilises search technologies to collect this information automatically from user-defined sources. Because we have already defined the five objectives in previous sections – rivals, buyers, substitutes, suppliers, and potential entrants, some specific data sources will be paid more attention and made sure they are kept up to date in time, for example, some open sources about financial news.

An integral part of the data-collection component of the system is a module that evaluates the quality of the information sources used in order to maximise the quality and reliability of the system inputs. Source quality assessment is based on various factors, such as user ratings and automatic quality indicators.

B. From textual data to knowledge

During the process of TM, four major components are needed: preprocessing, text and data warehouse, domain knowledge base, and information processing. Because the formats of text are varied – ASCII text, MS Word doc, pdf, html, rtf, xml, to name just a few – the system needs to identify and convert them to a unified format that can be processed by all the system components. Furthermore, words are made of multiple parts (for instance, compound words consist of two or more words, words consist of stems and suffixes), words compose sentences, and sentences compose the content of the text according to syntax rules. Performing these initial stages of text analysis is referred to as preprocessing. During preprocessing, NLP methods such as stemming, stop words removal and part-of-speech tagging are applied to decompose text into some meaningful language segments with tags that can be easily extracted.

The next steps in text processing are feature and IE. This step aims to extract relevant information and filter out redundant and irrelevant pieces of information. Feature extraction components extract concepts (such as companies, technologies, products, attitudes) and events (such as launching of a new product, bankruptcy of a competitor). The five objectives (rivals/competitors, buyers, suppliers,

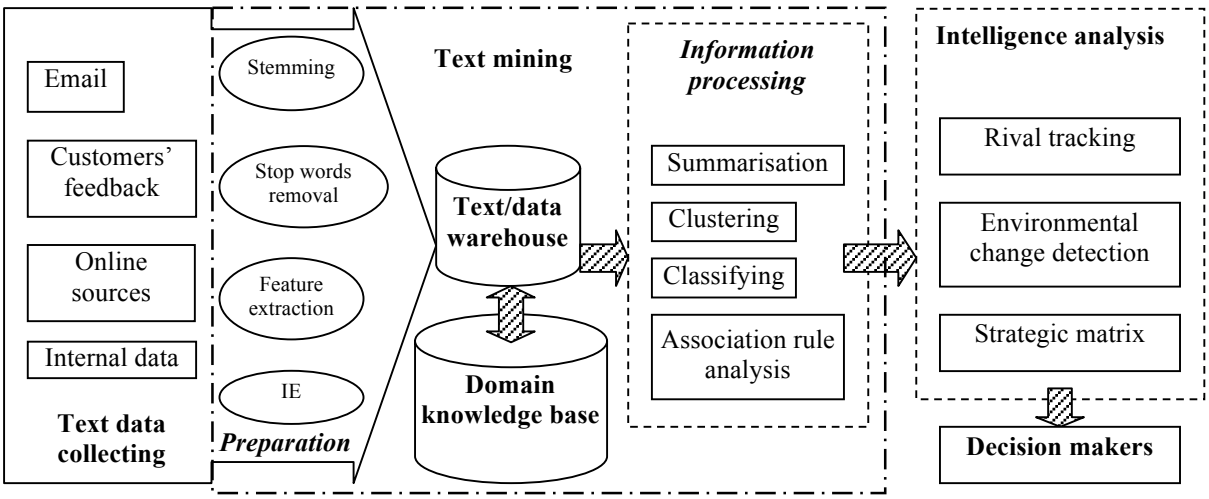

Figure 6. The architecture of a system based on MinEDec, a decision-support model

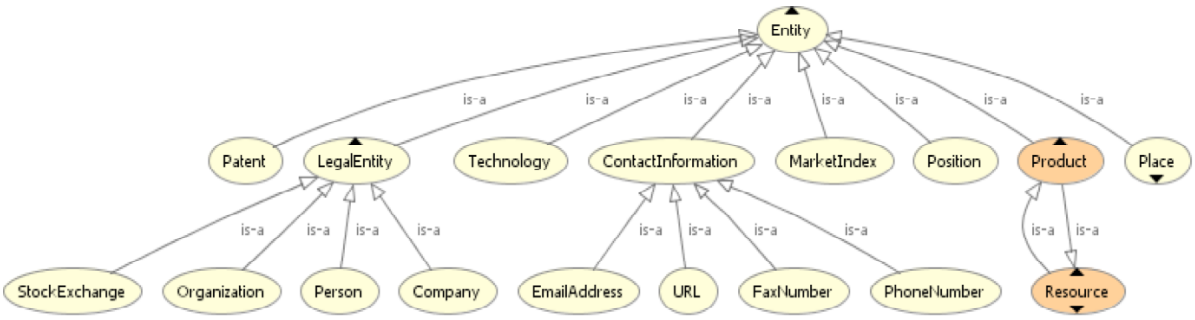

Figure 7. Some of the entity types in CoProE.

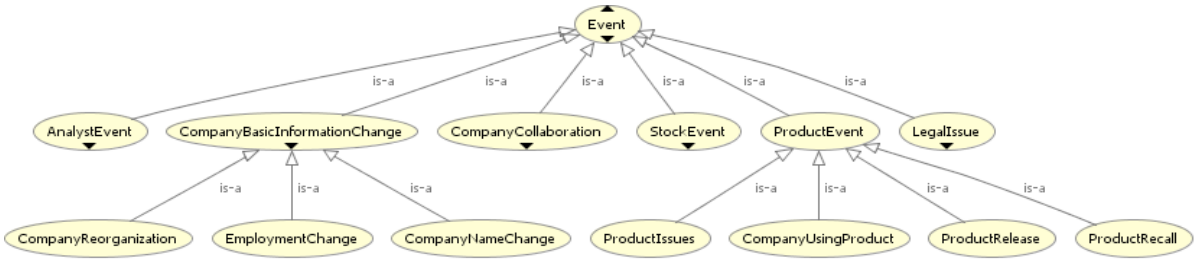

Figure 8. Some of the event types in CoProE.

substitutes, potential entrants) and the nine factors (technology, price, equipment, service, attitude, political change, economic change, technological trend, social attention) defined in MinEDec can guide this process when the system collects and extracts information from data sources.

In our system, we focus on new events, the trend of events, the changes of nine factors, and relationships between events and the five objectives. Using these factors guides data collecting and storing, the system can save time from manual work. The scope of the search, moreover, will be more focused. The results of the preprocessing and extraction are stored into the text and data warehouse.

After the features and information have been stored in the warehouse, information processing methods, such as clustering, classifying, association rule analysis, and summarisation can be used to process the extracted pieces of data into meaningful information.

Clustering and classifying techniques are used for information exploitation and retrieval. For example, the system could use clustering to analyse emails from buyers to find out common keywords that have not been noticed. Association rule analysis can be applied for discovering the relationships and changes between two or more data sets [17, 22]. This can be applied for establishing the relationships and trends between different objectives, technologies, products, events, etc. The aim of summarisation is to extract key information and represent it to the decision maker in a more concise format. Both textual and visual summarisation methods can be utilised.

C. Representing domain knowledge

A key part of the capability of analysing input texts is the existence of knowledge about the domain that is being scrutinised. In the case of our proposed system the domain consist of companies, products and events the user is interested in. Ontologies provide the means for modelling concepts, attributes and relationships in a specific domain. The design of the *company, product, and event* (CoProE) ontology that we are developing arises from the need to have an ontology capable of supporting the domain knowledge needs of our system. Its design reflects the reuse design theme of our system. The ontology is based on a standard product and industry type code set, *United Nations Standard Products and Services Code* (UNSPSC) [24] and an existing ontology, newsEvents, for modelling business events, the affected business entities and the relations between them [25].

The current version of the ontology enables the characterisation of business entities and their key employees, and the classifications in a standardised way of one's own products as well as the modelling of business events and the relations between companies. The ontology currently consists of 16,652 classes, 129 object properties and 26 data properties and is stored in *Web Ontology Language* (OWL) format. Figures 7 and 8 visualise selected parts of the ontology and give examples of types of entities and events we are interested in extracting from the input texts.

Our aim is to leverage the domain knowledge stored in the ontology in multiple ways in MinEDec-driven CI analysis. The TM components of our system use CoProE in ontology-based IE and event detection as well as for knowledge discovery. The user interface can be enriched by combining new information extracted from input documents with known facts that are retrieved from the domain knowledge base.

D. Intelligence analysis and support for decision making

By using TM alone, business leaders gain intelligence about specific objectives. However, our model gives further support to decision making. These additional intelligence/knowledge products are dependent on the IE technology, such as *named entity recognition* (NER), *template relation* (TR), and the *scenario template*.

NER refers to the recognition of proper nouns, time, and number, etc. TR deals with finding out the relationships between the entities. Scenario template forms a description of the whole event or relationship by connecting entities

together [3, 10]. Considering that the expression forms of the same factors are different under different situations, we can cluster and label the keywords, then combine related labels to generate factors, and the factors are displayed in terms of strengths, weaknesses, opportunities, and threats by comparing with rivals or other subjects [11].

After the information-processing phase, information is already processed by summarising, classifying, clustering, and association rule analysis. Also we get the factors belonging to different categories, then we can use change detection to identify the most frequent rules and then compare the most frequent rules over two time periods to observe the subjects' changes or changes of event. These techniques can support rival tracking and environmental change detection.

Furthermore, we can combine factors in a SWOT matrix to give suggestions for planning a strategy. As we reviewed previous research, it does work that uses information technology to support the *analytic hierarchy process* (AHP) technique or the *analytic network process* (ANP) technique to implement SWOT matrix [19, 20, 21]. By introducing these techniques in our system, we can supply a strategy matrix to decision makers to support decision making.

V. Conclusion

We have proposed MinEDec, a decision-support model that combines two well-known and widely-used CI analysis models into a unified model. CI analysis by using this unified model is supported by the use of state-of-the-art TM technologies. We have also outlined the architecture of a DSS that is based on the MinEDec model and applies various TM technologies.

First, we explained that the purpose of our MinEDec model is to transform data into useful knowledge. We then described the functions of SWOT analysis and the FFA framework in a new model for monitoring the business environment. Although there are several CI software tools available, none of them combines TM and several widely accepted CI analysis methods. The proposed model is unique as it analyses the five objectives from the perspective of nine SWOT factors by using TM technologies. Based on this, we have proposed a way of integrating SWOT and FFA models into a unified decision-support model.

We believe that MinEDec can support decision making better than the existing models. Through providing objectives and factors in the business environment, the integrated model can implement the SWOT matrix to give suggestions about strategies by combining different factors. The capability of the proposed DSS in terms of utilising TM, monitoring the five force parties and deeply analysing the competitive environment is worth our attention.

We outlined the design of a DSS system that operationalises the proposed model. Our future and ongoing work on this line of research focuses on developing a working system based on the proposed MinEDec model. After system testing, MinEDec will be evaluated in real business environments.

Acknowledgement

This work was partly supported by grants from the CSC – China Scholarship Council, CIMO – Centre for International Mobility, and the "Towards e-leadership: higher profitability through innovative management and leadership systems" project, which is funded by the European Regional Development Fund and Tekes – the Finnish Funding Agency for Technology and Innovation.

References

[1] G. Johnson, K. Scholes, and R. Whittington, *Exploring Corporate Strategy: Text & Cases*, FT Prentice Hall Financial Times, Harlow, UK, 2008.

[2] V. L. Sauter and D. Free, "Competitive intelligence systems: qualitative DSS for strategic decision making", *The DATA BASE for Advances in Information Systems*, 36(2), pp. 43–57, 2005.

[3] R. Bose, "Competitive intelligence process and tools for intelligence analysis", *Industrial Management and Data Systems*, 108(4), pp. 510–528, 2008.

[4] C. S. Fleisher and B. E. Bensoussan, *Business and Competitive Analysis*, Pearson Education, Inc., Upper Saddle River, New Jersey, USA, 2007.

[5] K. C. Laudon and J. P. Laudon, *Essentials of Management Information Systems: Managing the Digital Firm*, Prentice Hall, London, UK, 2004.

[6] G. Forgionne, M. Mora, J. Gupta, and O. Gelman. "Decision-making support systems", *In Encyclopedia of Information Science and Technology*, Idea Group, Hershey, USA, pp. 759–765, 2005.

[7] M. E. Porter, *On Competition*, Harvard Business Press, Boston, USA, 1998.

[8] Y. Dai, T. Kakkonen, and E. Sutinen, "MinerVA: a decision support model that uses novel text mining technologies", *The 4th International Conference on Management and Service Science (MASS)*, Wuhan, China, pp. 1–4, 2010.

[9] D. H. Kim, B. J. Ahn, J. H. Kim, and J. J. Kim, "The strategic approach using SWOT analysis to develop an intelligent program management information system (iPMIS) for urban renewal projects", *Fourth International Conference on Computer Sciences and Convergence Information Technology*, Seoul, Korea, pp. 320–324, 2009.

[10] H. Y. Shahir, S. Daneshpajouh, and R. Ramsin, "Improvement strategies for agile processes: a SWOT analysis approach", *Sixth International Conference on Software Engineering Research, Management and Applications*, Prague, Czech Republic, pp. 221–228, 2008.

[11] M. Samejima, Y. Shimizu, M. Akiyoshi, and N. Komoda, "SWOT analysis support tool for verification of business strategy", *International Conference on Computational Cybernetics*, Tallinn, Estonia, pp. 1–4, 2006.

[12] J. P. M. D. Oliveira, S. Loh, and L. K. Wives, "Applying text mining on electronic messages for competitive intelligence", *Proceedings of the 5th International Conference on Electronic Commerce and Web Technologies*, Zaragoza, Spain, pp. 277–286, 2004.

[13] D. Sullivan, "The need for text mining in business intelligence", *Information Management Special Reports*, available at: http://www.information-management.com/specialreports/20040210/8100-1.html, 2004.

[14] R. Bose, "Advanced analytics: opportunities and challenges", *Industrial Management & Data Systems*, 109(2), pp. 155–172, 2009.
[15] T. P. Liang, *Decision Support Systems and Intelligent Systems*, Pearson Education International, Upper Saddle River, New Jersey, USA, 2005.
[16] F. Bouthillier and K. Shearer, *Assessing Competitive Intelligence Software: A Guide to Evaluating CI Technology*, Information Today, Inc., Medford, New Jersey, USA, 2003.
[17] D. R. Liu, M. J. Shih, C. J. Liau, and C. H. Lai, "Mining the change of event trends for decision support in environmental scanning", *Expert Systems with Applications*, vol. 36, pp. 972–984, 2009.
[18] M. J. Shih, D. R. Liu, and M. L. Hsu, "Discovering competitive intelligence by mining changes in patent trends", *Expert Systems with Applications*, 37(4), pp. 2882–2890, 2010.
[19] T. Hill and R. Westbrook, "SWOT analysis: it's time for a product recall", *Long Range Planning*, vol. 30, pp. 46–52, 1997.
[20] I. Yüksel and M. Dağdeviren, "Using the analytic network process (ANP) in a SWOT analysis – a case study for a textile firm", *Information Sciences*, vol. 117, pp. 3364–3382, 2007.
[21] M. Kurttila, M. Pesonen, J. Kangas, and M. Kajanus, "Utilizing the analytic hierarchy process (AHP) in SWOT analysis – a hybrid method and its application to a forest – certification case", *Forest Policy and Economics*, vol. 1, pp. 41–52, 2000.
[22] N. Zhou, H. L. Cheng, H. Q. Chen, and S. Xiao, "The framework of text-driven business intelligence", *International Conference of Wireless Communications, Networking and Mobile Computing,* Shanghai, China, pp. 5468–5471, 2007.
[23] Y. Y. Yang, L. Akers, T. Klose, and C. B. Yang, "Text mining and visualization tools – impressions of emerging capabilities", *World Patent Information*, vol. 30, pp. 280–293, 2008.
[24] U. Lösch, N. Nikitina, "The newsEvents ontology – an ontology for describing business events", *8th International Semantic Web Conference, 1st Workshop on Ontology Design Patterns*, Washington DC, USA, 2009.
[25] The United Nations Standard Products and Services Code, http://www.unspsc.org/.
[26] H. Cunningham, D. Maynard, K. Bontcheva, and V. Tablan, "GATE: a framework and graphical development environment for robust NLP tools and applications", *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02),* Philadelphia, USA, 2002.

Author Biographies

Yue Dai is a doctoral student at the EdTech Δ research group, School of Computing, University of Eastern Finland. She received her master's degree of management in the specialty of information science from Nanjing University, China in 2009. Dai's main research interests are strategy management, competitive intelligence, text mining, and sentiment analysis.

Tuomo Kakkonen is a researcher and project manager at School of Computing, University of Eastern Finland. Kakkonen received his PhD in computer science from the University of Joensuu, Finland in 2007. His research interests are related to natural language processing (syntactic parsing, text mining, sentiment analysis) and information retrieval (latent semantic analysis), and their applications in educational technology and business intelligence.

Erkki Sutinen is the leader of EdTech Δ research group at the School of Computing, University of Eastern Finland. His research interests include using technologies for clearly defined needs in complex subject domains, such as programming, in developing countries, and within special education. The applied techniques cover information visualization, information retrieval, data mining, robotics, and design models. He has co-authored and published over 100 academic papers.