Expert Systems with Applications 39 (2012) 2927–2938

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

TrendPerceptor: A property–function based technology intelligence system for identifying technology trends from patents

Janghyeok Yoon, Kwangsoo Kim *

Department of Industrial and Management Engineering, Pohang University of Science and Technology, San 31, Hyoja-dong, Nam-gu, Pohang, Kyungbuk 790-784, Republic of Korea

ARTICLE INFO

Keywords:
Technology intelligence
Property
Function
Patent mining
Patent analysis
Social network analysis
TRIZ trend
Natural language processing

ABSTRACT

Technology intelligence systems are vital components for planning of technology development and formulation of technology strategies. Although such systems provide computation supports for technology analysis, much effort and intervention of experts, who may be expensive or unavailable, is required in gathering processes of information for analysis. As a remedy, this paper proposes *TrendPerceptor*, a system that uses a property–function based approach. The proposed system assists experts (1) to identify trends in invention concepts from patents, and (2) to perform evolution trend analysis of patents for technology forecasting. For this purpose, a module of the system uses grammatical analysis of textual information to automatically extract properties and functions, which show innovation directions in a given technology. Using the identified properties and functions, a module for invention concept analysis based on network analysis and a module for evolution trend analysis based on TRIZ (Russian acronym of the Theory of Inventive Problem Solving) trends are suggested. This paper describes the architecture of a system composed of these three modules, and illustrates two case studies using the system.

1. Introduction

Modern economies emphasize the role of research and development (R&D) that promotes the creation, diffusion and accumulation of intellectual properties within economic systems. In general terms, R&D refers to “creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of knowledge to devise new applications” (OECD, 2008). As a method to improve the effectiveness of R&D activities, technology intelligence has been introduced to plan technology development and formulate technology strategies. Technology intelligence is “the capture and delivery of technological information as part of the process whereby an organization develops an awareness of technological threats and opportunities” (Kerr, Mortara, Phaal, & Probert, 2006). Technology intelligence includes technology monitoring, technology assessment, and technology forecasting (Lichtenthaler, 2004). As technology lifecycles shorten and business environments become more globalized, technology intelligence capabilities have become increasingly important for experts such as researchers, practitioners and R&D policy makers.

Technology intelligence tools have several advantages over expert-based approaches (Yoon, 2008), including:

1. Ability to analyze large quantities of information which cannot be analyzed by humans alone.
2. Ability to generate much useful information which humans cannot produce. For example, they can visualize the relationship between technology and companies, and analyze the characteristics of technology using statistical analysis.
3. Ability to support decision making processes with relevant information including technology assessment and technology forecasting.

Technology intelligence tools assist experts to make strategic technology plans. With respect to content analysis of technical documents, many technology intelligence tools have been actively developed to identify technology trends and give technological insights. The Theory of Inventive Problem Solving (Russian acronym: TRIZ) (Altschuller, 1984) has been used as a tool for technology intelligence (Schuh & Grawatsch, 2004). Among its many tools, TRIZ trend analysis identifies the evolutionary status of systems to predict further improvement of technology by using defined system evolution patterns called TRIZ trends (Mann, 2002, 2003). Thus, TRIZ trend analysis gives decisive information to determine the threats and opportunities presented by competing technologies (Schuh & Grawatsch, 2004). Although tools exist to facilitate identification of TRIZ trends, they rely heavily on skills and knowledge of TRIZ experts in gathering TRIZ trend-related information and determining specific TRIZ trends and trend phases (Verhaegen, D’Hondt, Vertommen, Dewulf, & Duflou, 2009). Diva (Morris,

---

* Corresponding author. Tel.: +82 54 279 8231.
E-mail addresses: janghyoon@postech.ac.kr (J. Yoon), kskim@postech.ac.kr (K. Kim).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.08.154

DeYong, Wu, Salman, & Yemenu, 2002), a bibliometric analyzer of collections of scientific literature and patents, is a computer program that visualizes patent clusters and gives insight into trends in the technological field of interest. Patinformatrics tools (Trippe, 2003) such as VantagePoint and Aureka can cluster data, mapping document clusters and generate co-occurrence matrices so that analysts can discover relationships or trends in a given technology. In addition, Techpioneer (Yoon, 2008), an expert system for excavating potential technology opportunities, uses text mining and morphology analysis; the system outputs valuable information such as the trends of keywords and the morphology of existing technology, and identifies promising opportunities for technology development. However, in the aspect of content analysis of technical documents, these tools require that the set of keyword or key phrase patterns be defined in advance; this task relies heavily on the knowledge and effort of domain experts, who may be expensive or unavailable. Furthermore defining keyword or key phrase patterns of adjacent or emerging technology areas may be a difficult task even for experts, because their knowledge cannot extend over all technology areas. Fundamentally, these experts need to more concentrate on their knowledge activities that identify technology trends and promising technologies, but too much of their effort and time should not be consumed in gathering processes of information for technology analysis.

To eliminate these dependencies on effort of experts, this paper proposes *TrendPerceptor*, a system that uses a property–function based approach. The proposed system assists experts (1) to identify trends in invention concepts from patents, and (2) to perform evolution trend analysis of patents for technology forecasting.

Properties and functions imply invention concepts of a system (Dewulf, 2005, 2006; Verhaegen et al., 2009). Properties, which are the specific characteristics of a system or its sub-systems, are usually expressed using adjectives; functions, which are useful actions of a system or its sub-systems, are usually expressed using verbs (Dewulf, 2006). By exploiting these grammatical relationships, properties and function can be obtained using grammatical analysis of textual information. To this end, *TrendPerceptor* uses natural language processing (NLP) to automatically extract properties and functions from patent documents. This system facilitates experts to analyze technological trends of a given technology by providing a network composed of the identified properties and functions. Additionally, the system assists experts to predict further improvements of a system by automating the processes of TRIZ trend analysis. The final output of TRIZ trend analysis consists of evolutionary potential radar plots which show the evolution status of patents related to a technology; they can be used as valuable input for technology forecasting based on TRIZ trends.

Chapter 2 presents the groundwork of this research. On the basis of the groundwork, chapter 3 describes the system architecture of *TrendPerceptor* in detail, and chapter 4 introduces two case studies using the proposed system. Finally chapter 6 presents conclusions and future work.

2. Groundwork

2.1. Property–function analysis

Dewulf (2006) presented a property–function based approach to identify connections among products, processes and systems in different domains: a property addresses 'what a system is or has' and expresses a specific characteristic of a system or its subsystems; a function addresses 'what a system does or undergoes' and expresses a useful action of a system or its subsystems. On the basis of investigation of about 16,000 patents from the U.S. Patent and Trademark Office (USPTO), he found that properties are generally expressed using adjectives, whereas functions are generally expressed using verbs (Dewulf, 2006). This means that properties and functions can be identified using grammatical analysis of textual information (Dewulf, 2006). However, using only adjectives and verbs does not sufficiently represent concreteness of technology domains because this research aims at identifying technological trends from specific technology areas. Therefore this paper proposes representing properties and functions in the form of binary relations, which are simple but concrete two-word expressions: properties are expressed as 'adjective + noun' forms, and functions are expressed as 'verb + noun' forms. Using the extracted properties and functions, *TrendPerceptor* facilitates technology trend analysis such as invention concept analysis of patents and evolution trend analysis of patents.

2.2. Stanford dependencies representation and Stanford parser

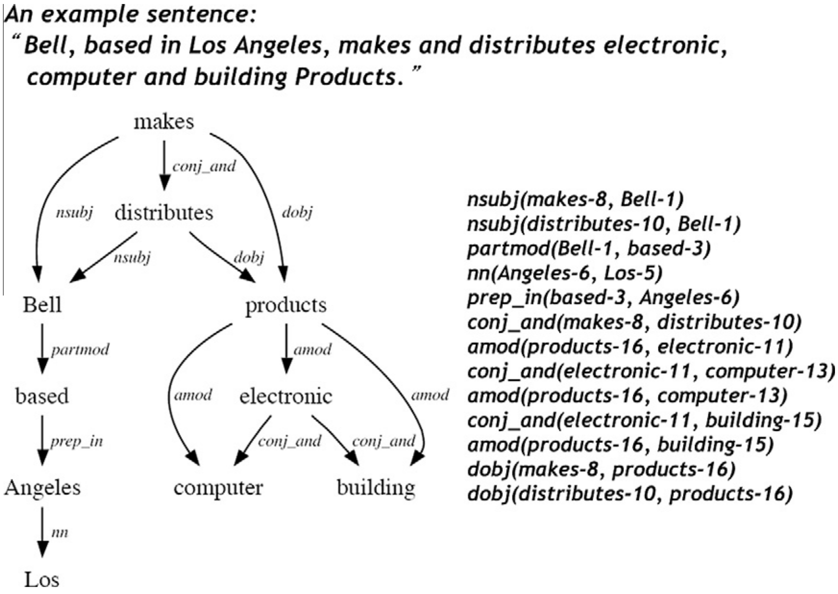
Because properties and functions can be identified by analysis of the grammatical structure of textual information, this paper uses the Stanford parser to automatically extract information concerning properties and functions. The Stanford parser is an NLP parser that provides a JAVA-application programming interface and is based on the Stanford dependencies representation. The Stanford dependencies representation for English was designed to provide a simple description of the grammatical relationships in a sentence that can be easily understood and effectively used by people who have no linguistic expertise (de Marneffe & Manning, 2008B). Currently, the dependencies are categorized into 55 grammatical relations, which are all binary: a grammatical relation holds between a governor and a dependent (de Marneffe & Manning, 2008A). Any grammatically correct sentence can be represented by the Stanford dependencies (Fig. 1) and, conversely, the Stanford parser can extract specific binary relations from the sentence. In this paper, the proposed system uses the Stanford parser to preprocess texts of patents for further analysis.

2.3. Network analysis

Network analysis is a popular method for identifying technology trends in bibliometrics (Lee & Jeong, 2008), which is a set of tools used to study or measure texts and information. Network analysis using technology keywords encompasses (1) defining a set of keyword patterns or key phrase patterns, (2) generating a network that codifies the relationship between occurrences of the patterns, and (3) identifying technological trends by interpreting the network (Lee & Jeong, 2008; Yoon, 2008; Chang, Wu, & Leu, 2010). The network allows identification of priorities and relationships among the patterns by applying social network analysis (SNA), which maps and measures relationships and interactions among people, groups, organizations, computers or other connected entities (Hanneman & Riddle, 2005). For invention concepts analysis of patents, the system proposed in this paper organizes a network by using co-occurrences of properties and functions, and outputs files that are formatted appropriately for further analysis using external SNA tools such as NetMiner, UCINET and Pajek.

2.4. TRIZ trend analysis

Originally, eight laws of technical system evolution were identified (Altschuller, 1984), but recently an updated list of 35 TRIZ trends was presented, incorporating new domains and reflecting new innovative solutions (Mann, 2002). These trends are grouped into space (12 trends), interface (17 trends) and time (6 trends). These evolution laws show specific evolution sequences of systems (Petrov, 2002). TRIZ philosophy follows the concept that systems evolve toward increasing ideality, where ideality = (conceived


An example sentence:
"Bell, based in Los Angeles, makes and distributes electronic, computer and building Products."

nsubj(makes-8, Bell-1)
nsubj(distributes-10, Bell-1)
partmod(Bell-1, based-3)
nn(Angeles-6, Los-5)
prep_in(based-3, Angeles-6)
conj_and(makes-8, distributes-10)
amod(products-16, electronic-11)
conj_and(electronic-11, computer-13)
amod(products-16, computer-13)
conj_and(electronic-11, building-15)
amod(products-16, building-15)
dobj(makes-8, products-16)
dobj(distributes-10, products-16)

Fig. 1. An example of Stanford dependencies representation.

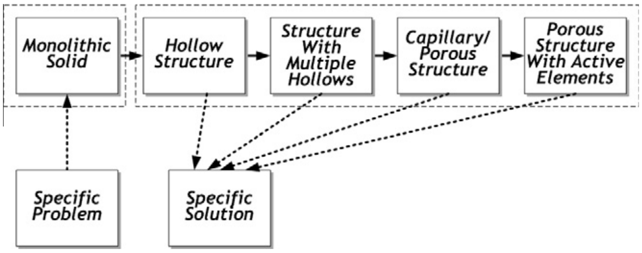


Fig. 2. An example of evolution trend (the space segmentation).

benefit)/(cost + harm) (Mann, 2003). The direction to increase ideality of a system is generally represented as a left-to-right trend; for example, the ideality of a system in the 'capillary/porous structure' phase is larger than the ideality of a system in the 'monolithic solid' (Fig. 2). Evolutionary potential is the difference between the evolutionary limit and the maturity of the current system with respect to each trend (Mann & Dewulf, 2002). Therefore, evolutionary potential radar plots, which show the multi-dimensional evolutionary status of a system, can be used as valuable input for TRIZ trend based technology forecasting. In this paper, *TrendPerceptor* automates the processes of evolution trend analysis.

3. System architecture

3.1. Basic concepts

TrendPerceptor: is a technology intelligence system that supports invention concepts analysis of patents (Yoon, Choi, & Kim, 2011) and system evolution trend analysis of patents (Yoon & Kim, 2011) (Fig. 3). Although many tools provide processes for technology intelligence such as identifying technology trends and capturing technology opportunities (Yoon, 2008), at many points much effort of experts is consumed in gathering of information for technology analysis. For these reasons, the system uses grammatical analysis with NLP to automatically extract information concerning properties and functions from patent documents. In turn, the system provides experts with outputs for invention concepts analysis of patents and evolution trend analysis of patents.

3.2. Architecture

The system architecture (Fig. 4) of *TrendPerceptor* is composed of three modules: the property–function mining module (PFMM), the invention concept analysis module (ICAM) and the evolution trend analysis module (ETAM). Modules that consist of two or three sub-modules utilize external modules including the Stanford parser and semantic similarity checker, and interface with specific databases to retrieve and store processed data. PFMM is a preprocessing module; it extracts properties and functions from texts of patents using NLP. ICAM identifies innovation concepts of patents. ICAM generates a network using properties, functions and their co-occurrences, then provides technological implications by interpreting the network. ETAM automates identification of the evolutionary potential of patents related to a given technology using the properties and functions, and gives insights for further improvement of the technology.

3.3. Property–function mining module

Most tools using text mining require a set of predefined keyword patterns or key phrase patterns. However, PFMM employs a property–function based approach to automatically extract information related to invention concepts of patents. Although various technical documents exist, this paper considers only patents because they contain up-to-date and reliable information about inventions. With little effort of experts, PFMM outputs the preprocessed data such as properties, functions and their co-occurrences. PFMM consists of a patent search engine and a property–function extractor.

3.3.1. Patent search engine

To identify technology trends, patent documents should be collected in advance. To this end, the patent search engine has an interface with public patent databases. Using patent retrieval queries composed of textual information related to a target technology, and bibliographic information such as international patent code, applicants and application date, the engine outputs a list of retrieved patents into a Microsoft Excel file or a text file.

3.3.2. Property–function extractor

The property–function extractor identifies properties and functions by grammatical analysis of abstracts of the retrieved patents

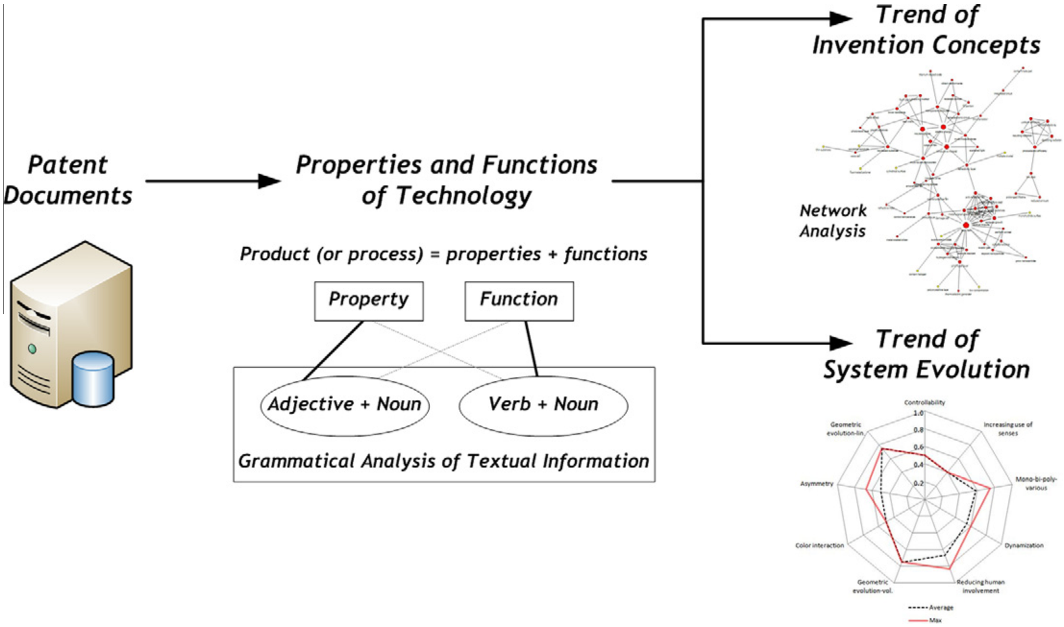


Fig. 3. Basic concepts of *TrendPerceptor*.

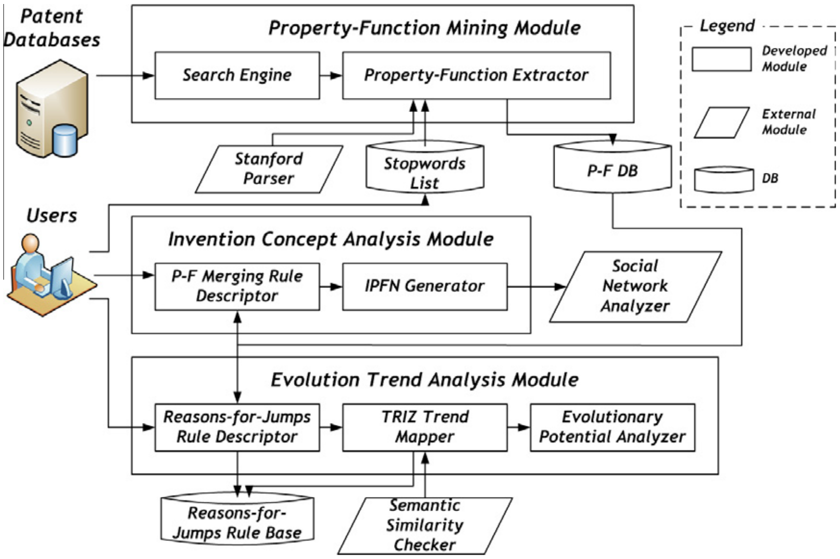


Fig. 4. Architecture of *TrendPerceptor*.

because the human-generated abstracts of patents can be considered as the most important part (Chen, Tokuda, & Adachi, 2003; Tong, Cong, & Lixiang, 2006; Verhaegen et al., 2009). Dewulf (2006) and Verhaegen et al. (2009) identified properties from textual information of patents using adjectives. Hirtz, Stone, McAdams, Szykman, & Wood (2002) used 'verb + noun' pairs to define functions of mechanical systems. This paper defines the properties and functions as binary relations, which are simple but domain-specific two-word expressions. Properties are usually binary relations that can be expressed as 'adjective + noun' forms; functions are usually the binary relations that can be represented as 'verb + noun' forms. To automate gathering of properties and functions, the extractor uses the Stanford dependencies representation and the Stanford parser.

Among the 55 Stanford dependencies, five Stanford typed dependencies such as adjectival modifiers (amod), direct objects (dobj), infinitival modifiers (infmod), participial modifiers (partmod) and relative clause modifiers (rcmod) can be grammatically transformed into 'adjective + noun' or 'verb + noun' forms (Table 1).

Many binary relations are automatically extracted from a patent using the Stanford parser, but many of them can be 'noise' binary relations that are irrelevant to properties and functions of the patent. For this reason, a list of extended English stopwords (STOPWORDS, 2010) is used to filter out unintended or too-general binary relations; for example, binary relations containing 'comprise', 'invention', 'apparatus', 'have' and 'make' are eliminated.

A patent document includes many sentences and each sentence may contain properties and functions. To perform further analysis in Subsections 3.4 and 3.5, the property–function extractor stores properties and functions identified by each sentence into the property–function database (Fig. 5).

Table 1
The five Stanford dependencies related to properties and functions.

Stanford dependency (SD)	Description and example: SD name(dependent, governor)
Amod (adjectival modifier)	An adjectival modifier of a noun phrase (NP) is any adjectival phrase that serves to modify the meaning of the NP “Sam eats red meat” amod(meat, red)
Dobj (direct object)	The direct object of a verb phrase (VP) is the noun phrase which is the (accusative) object of the verb; the direct object of a clause is the direct of the VP which is the predicate of that clause “She gave me a raise” dobj(gave, raise)
Infmod (infinitival modifier)	An infinitival modifier of an NP is an infinitive that serves to modify the meaning of the NP “I do not have anything to say” infmod(anything, say)
Partmod (participial modifier)	A participial modifier of an NP or VP is a participial verb form that serves to modify the meaning of the NP or VP “Truffles picked during the spring are tasty” partmod(truffles, picked)
Rcmod (relative clause modifier)	A relative clause modifier of an NP is a relative clause modifying the NP. The relation points from the head noun of the NP to the head of the relative clause, normally a verb “I say the man you love” rcmod(man, love)

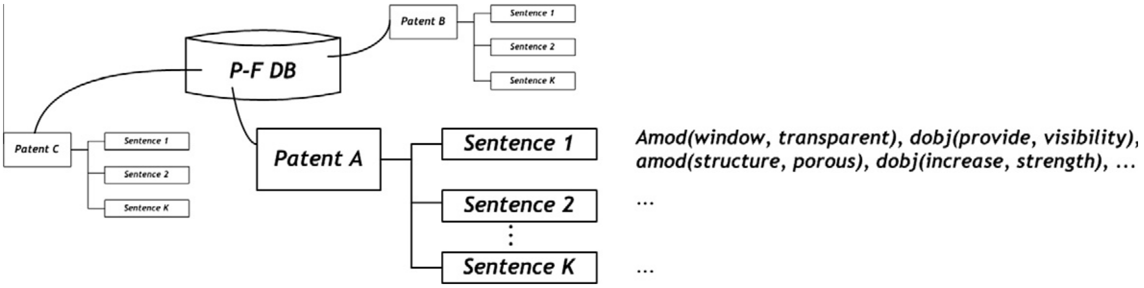


Fig. 5. Structure of the property–function database.

3.4. Invention concept analysis module

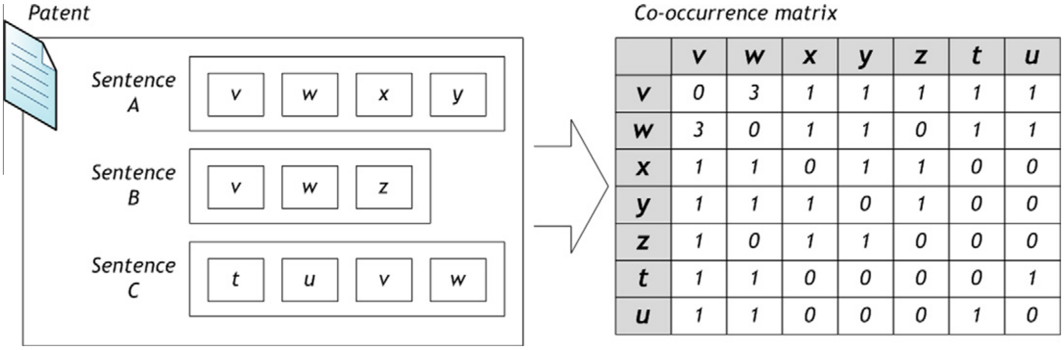
ICAM provides invention concept analysis. In a patent document, the occurrence of properties and functions in the same sentence implies that they are technologically connected to each other. Consider a sentence ‘the transparent window provides visibility’. In this sentence, ‘transparent window’ is a property and ‘provide visibility’ is a function. In the technological aspect, the property ‘transparent window’ is a method or a material to achieve the function ‘provide visibility’, and conversely the function ‘provide visibility’ is a use or an objective that the property ‘transparent window’ provides. ICAM uses the identified properties and functions as nodes and their co-occurrences in the same sentence as links. Because co-occurrences do not have a direction, ICAM organizes an undirected valued network and produces the appropriate data formats that can be used by external SNA tools, which can identify technological implications of properties and functions using analysis indicators of SNA such as degree, centrality and density.

3.4.1. Property–function merging rule descriptor

Because many of the properties and function may be semantically identical, the property–function merging rule descriptor categorizes them into groups. The descriptor provides several functions including creation of a new group, addition of a property or a function to an existing group, and deletion of irrelevant properties and functions which were not filtered out by stopwords. In this system, a file defining the groups and their elements is called ‘mapping rule base’. For network analysis, properties in a group can be merged into a representative property; functions in a group can be merged into a representative function. Instead of gathering defined keywords or key phrases for text mining, this system uses grammatical analysis to collect binary relations concerning innovation concepts, then the relations into groups.

3.4.2. Invention property–function network generator

Using the mapping rule base, all binary relations are changed into their representatives. Then a co-occurrence matrix for each patent can be organized. The co-occurrence method is used



Co-occurrence matrix

	v	w	x	y	z	t	u
v	0	3	1	1	1	1	1
w	3	0	1	1	0	1	1
x	1	1	0	1	1	0	0
y	1	1	1	0	1	0	0
z	1	0	1	1	0	0	0
t	1	1	0	0	0	0	1
u	1	1	0	0	0	1	0

Fig. 6. Co-occurrence matrix of a patent.

because the properties and functions that appear in the same sentence are related to each other. Consider patent P: sentence A has a set of binary relations <v, w, x, y>, sentence B has <v, w, z>, and sentence C has <t, u, v, w>. Row headers and column headers are labeled as v, w, x, y, z, t and u, which are the representatives of defined groups in the mapping rule base. The co-occurrence frequency of v and w is 3, and that of other pairs is 0 or 1, so a co-occurrence matrix for each patent can be defined (Fig. 6).

If identical nodes are merged, then identical links can be merged automatically and the frequency of merged links is the sum of frequencies of the identical links; for example, if t and v co-occur in other patents three times, then in the merged co-occurrence matrix, t and v co-occur four times. Likewise, a co-occurrence matrix for the whole patent set, which is called the Invention Property–Function Network (IPFN), can be built by merging all patents' co-occurrence matrices. Because several external SNA tools such as NetMiner, UCINET and Pajek can be used to analyze an IPFN, the IPFN generator has a function that outputs the appropriate data formats for these tools.

3.4.3. Social network analyzer

Using the output files of the IPFN generator, SNA tools identifies technological implications of properties and functions. Three analysis indicators are used: degree, centrality and density. The degree of a node is defined as the number of links or the sum of values of links incident to the node (Diestel, 2005); it shows the importance of the node in a network. The centrality of a node can be used to determine its relative importance, based on how well it connects the network (Freeman, 1979). The density of a network indicates how closely all nodes in the network are related, and indicates the completeness of a network.

Applying the concepts of degree, centrality and density to the IPFN, technological implications of properties and functions can be directly formulated (Table 2). For example, if the degree of a property is high, the property is very likely to be a method or a material that can be used for various purposes or applications, or to be a technological verified, cost-effective or dominant method or material. Likewise, if the degree of a function is high, it has a strong possibility of being a usually-required or necessary use or objective in a given technology.

3.5. Evolution trend analysis module

ETAM provides evolutionary status of a technology as decisive information for technology forecasting based on TRIZ trends. To this end, a 'reasons for jumps' rule base which arranges the properties and functions according to TRIZ trends and trend phases is first defined, and in turn specific TRIZ trends and trend phases are detected by measuring semantic sentence similarity between binary relations from patents and binary relations from the rule base. After normalizing the identified trends and trend phases, this module visualizes the evolution status of patents related to given technology using evolutionary potential radar plots.

3.5.1. 'Reasons for jumps' rule descriptor

The 'reasons for jumps' rule descriptor supports defining a 'reasons for jumps' rule base, a set of properties and functions in the form of a binary relation which can determine the specific trends and trend phases. They are expressed in the 'adjective + noun' or 'verb + noun' forms. The properties and functions in the rule base can be categorized into two types: generic and domain-specific.

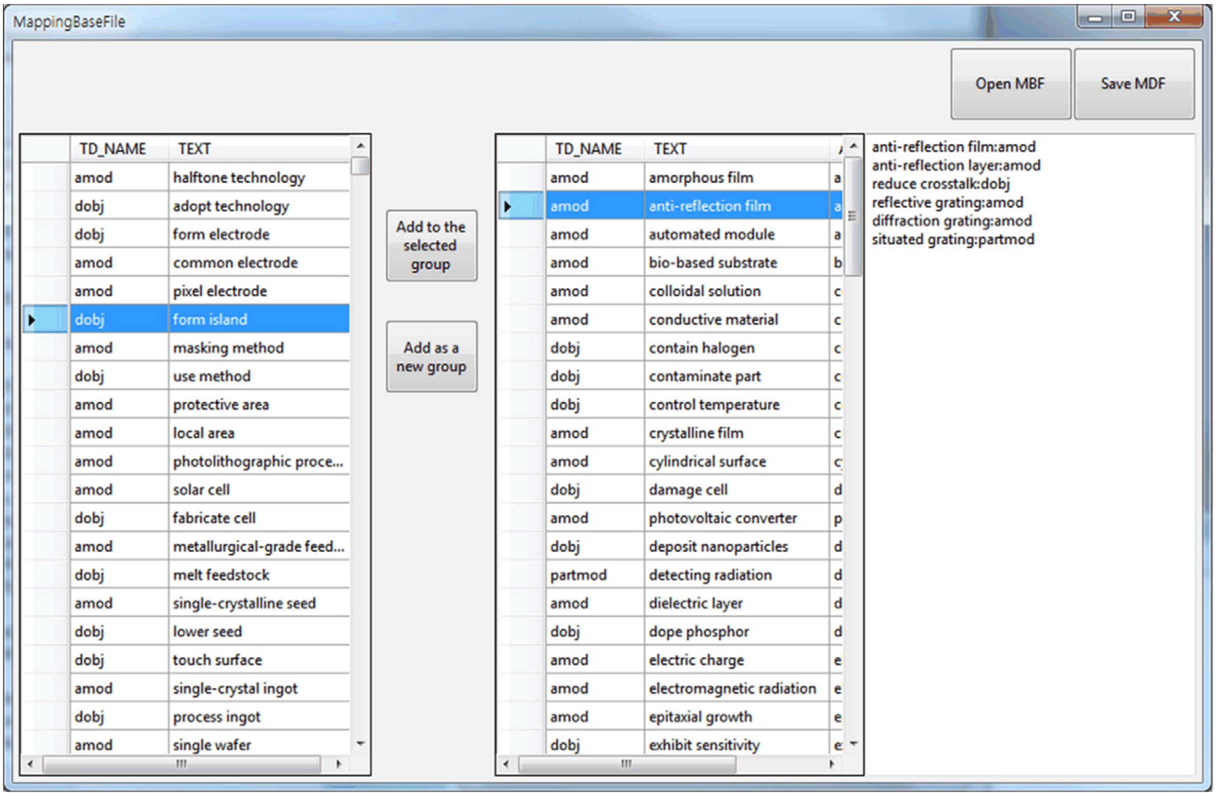
Generic properties and functions can be identified from the TRIZ trends, which capture the specific sequences of the evolution of a system. The trends are followed by a list of examples of the particular trend and a list of reasons distilled from other solutions to suggest why the jumps might offer a benefit (Mann, 2002). Generic properties and functions collectable from these TRIZ trends are, for example, 'smooth surface', 'easy grip', 'roughened surface', 'increase noise' and 'active pores'. Domain-specific binary relations can be identified using the training sets of randomly selected patents. Using repeated training to update the rule base can increase its reliability. In this way, a 'reasons for jumps' rule base for the surface segmentation trend can be defined (Table 3).

Table 2
Technological implications of properties and functions (Yoon et al., 2011).

		Technological implications	
		Property	Function
Degree	High	Dominant, cost-effective, useful, technologically verified method or material	Usually required or necessary uses or objectives
Centrality	High betweenness	Methods or materials applicable to various purposes. Methods or materials used with other methods or materials	Uses or applications using various methods or materials. Uses or objectives used with other uses or applications
	High Closeness	Widely-used or leading methods or materials over a product. Technologically verified key methods or materials	Widely-required or necessary uses over a product. Key objectives of the patents related to a product
Density	Large size; high density	Highly coupled to other properties or functions. Frequently used, dominant, technologically verified methods or materials	Highly connected to other properties or functions. Frequently used, dominant, technologically verified uses or objectives
	Large size; low density	Prospective methods or materials, but not yet dominant and technologically verified	Prospective uses or objectives, but not yet dominant

Table 3
An example of 'reasons for jumps' rule base of the surface segmentation trend (umbrella) (Yoon & Kim, 2011).

Evolution stage	Binary relations (properties and functions related to trend jump)
Smooth	Smooth surface, slick surface, soft surface, smoothen surface, soften surface, ...
Ribbed	Rib protrusions, easy grip, aerodynamic drag, reduce drag, improve traction, improve drainage, increase surface, increase space, aesthetic appearance, improve design, increase noise, deliberate weak-point, improve join, improve coupling, ...
3D roughened	Aerodynamic controllability, self cleaning, identification marks, 3D surface, roughened surface, low-drag surface, high friction, ...
Active	Active surface, active pores, bio-active surface, enhanced pores, ...

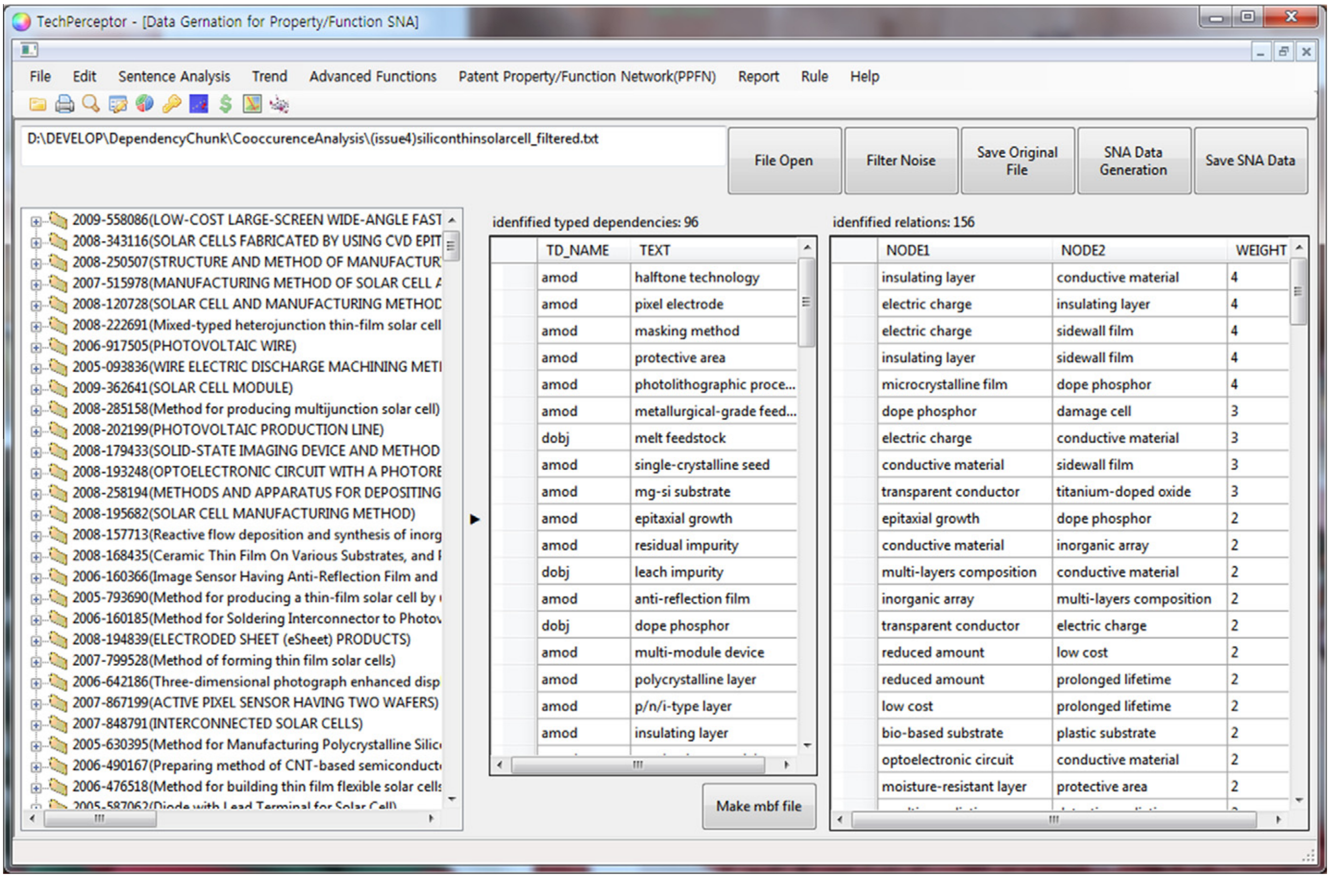


TD_NAME	TEXT
amod	halftone technology
dobj	adopt technology
dobj	form electrode
amod	common electrode
amod	pixel electrode
dobj	form island
amod	masking method
dobj	use method
amod	protective area
amod	local area
amod	photolithographic proce...
amod	solar cell
dobj	fabricate cell
amod	metallurgical-grade feed...
dobj	melt feedstock
amod	single-crystalline seed
dobj	lower seed
dobj	touch surface
amod	single-crystal ingot
dobj	process ingot
amod	single wafer

TD_NAME	TEXT
amod	amorphous film
amod	anti-reflection film
amod	automated module
amod	bio-based substrate
amod	colloidal solution
amod	conductive material
dobj	contain halogen
dobj	contaminate part
dobj	control temperature
amod	crystalline film
amod	cylindrical surface
dobj	damage cell
amod	photovoltaic converter
dobj	deposit nanoparticles
partmod	detecting radiation
amod	dielectric layer
dobj	dope phosphor
amod	electric charge
amod	electromagnetic radiation
amod	epitaxial growth
dobj	exhibit sensitivity

anti-reflection film:amod
anti-reflection layer:amod
reduce crosstalk:dobj
reflective grating:amod
diffraction grating:amod
situated grating:partmod

Fig. 7. Defining a mapping base file.



D:\DEVELOP\DependencyChunk\CooccurenceAnalysis\(issue4)siliconthinsolarcell_filtered.txt

2009-558086(LOW-COST LARGE-SCREEN WIDE-ANGLE FAST
2008-343116(SOLAR CELLS FABRICATED BY USING CVD EPIT
2008-250507(STRUCTURE AND METHOD OF MANUFACTUR
2007-515978(MANUFACTURING METHOD OF SOLAR CELL
2008-120728(SOLAR CELL AND MANUFACTURING METHOD
2008-222691(Mixed-typed heterojunction thin-film solar cell
2006-917505(PHOTOVOLTAIC WIRE)
2005-093836(WIRE ELECTRIC DISCHARGE MACHINING MET
2009-362641(SOLAR CELL MODULE)
2008-285158(Method for producing multijunction solar cell)
2008-202199(PHOTOVOLTAIC PRODUCTION LINE)
2008-179433(SOLID-STATE IMAGING DEVICE AND METHOD
2008-193248(OPTOELECTRONIC CIRCUIT WITH A PHOTORE
2008-258194(METHODS AND APPARATUS FOR DEPOSITING
2008-195682(SOLAR CELL MANUFACTURING METHOD)
2008-157713(Reactive flow deposition and synthesis of inorg
2008-168435(Ceramic Thin Film On Various Substrates, and
2006-160366(Image Sensor Having Anti-Reflection Film and
2005-793690(Method for producing a thin-film solar cell by
2006-160185(Method for Soldering Interconnector to Photo
2008-194839(ELECTRODED SHEET (eSheet) PRODUCTS)
2007-799528(Method of forming thin film solar cells)
2006-642186(Three-dimensional photograph enhanced disp
2007-867199(ACTIVE PIXEL SENSOR HAVING TWO WAFERS)
2007-848791(INTERCONNECTED SOLAR CELLS)
2005-630395(Method for Manufacturing Polycrystalline Silic
2006-490167(Preparing method of CNT-based semiconduct
2006-476518(Method for building thin film flexible solar cell
2005-587062(Diode with Lead Terminal for Solar Cell)

idenfified typed dependencies: 96

TD_NAME	TEXT
amod	halftone technology
amod	pixel electrode
amod	masking method
amod	protective area
amod	photolithographic proce...
amod	metallurgical-grade feed...
dobj	melt feedstock
amod	single-crystalline seed
amod	mg-si substrate
amod	epitaxial growth
amod	residual impurity
dobj	leach impurity
amod	anti-reflection film
dobj	dope phosphor
amod	multi-module device
amod	polycrystalline layer
amod	p/n/i-type layer
amod	insulating layer

idenfified relations: 156

NODE1	NODE2	WEIGHT
insulating layer	conductive material	4
electric charge	insulating layer	4
electric charge	sidewall film	4
insulating layer	sidewall film	4
microcrystalline film	dope phosphor	4
dope phosphor	damage cell	3
electric charge	conductive material	3
conductive material	sidewall film	3
transparent conductor	titanium-doped oxide	3
epitaxial growth	dope phosphor	2
conductive material	inorganic array	2
multi-layers composition	conductive material	2
inorganic array	multi-layers composition	2
transparent conductor	electric charge	2
reduced amount	low cost	2
reduced amount	prolonged lifetime	2
low cost	prolonged lifetime	2
bio-based substrate	plastic substrate	2
optoelectronic circuit	conductive material	2
moisture-resistant layer	protective area	2

Fig. 8. Identified nodes and links of the IPFN.

3.5.2. TRIZ trend mapper

The TRIZ trend mapper aims at automatically identifying TRIZ trends and trend phases of patents. Although the ‘reasons for jumps’ rule bases may include various binary relations for determination of specific trends and trend phases, it cannot be extended to include all possible binary relations. Therefore, the mapper automatically identifies the specific TRIZ trends and trend phases by measuring semantic sentence similarity between binary relations in ‘reasons for jumps’ rule bases and binary relations extracted from patents.

In general, processes of the semantic sentence similarity measurement between two sentences follows five steps: (1) tokenizing the sentences; (2) word stemming; (3) part-of-speech tagging; (4) determining the most likely sense for every word in each sentence; (5) computing the similarity of the sentences based on the similarity between pairs of corresponding words. Resnik (1999) formulated a measure of similarity between two concepts as follows:

$$sim(c_1, c_2) = \frac{2 * depth(lcs(c_1, c_2))}{depth(c_1) + depth(c_2)},$$

where lcs is the lowest common subsumer of two concepts c_1 and c_2, and $depth$ is the distance from a concept node c_i to the root of a concept hierarchy. The similarity score of two concepts is $0 < sim(c_1,c_2) \leqslant 1$. The matching average (Simpson & Dao, 2005) to compute similarity between two sentences is:

$$MatAvg(X, Y) = \frac{2 * Match(X, Y)}{|X| + |Y|},$$

where $|\cdot| = X$ or Y is the number of set tokens in the sentence, and $Match(X, Y)$ is the sum of similarity of the matching word tokens between sentences X and Y. The matching average score between two sentences is $0 < Matavg(X, Y) \leqslant 1$. Using the WordNet semantic dictionary (Miller, 1995) as the concept hierarchy, the measure for similarity measurement between two words (Resnik, 1999), and

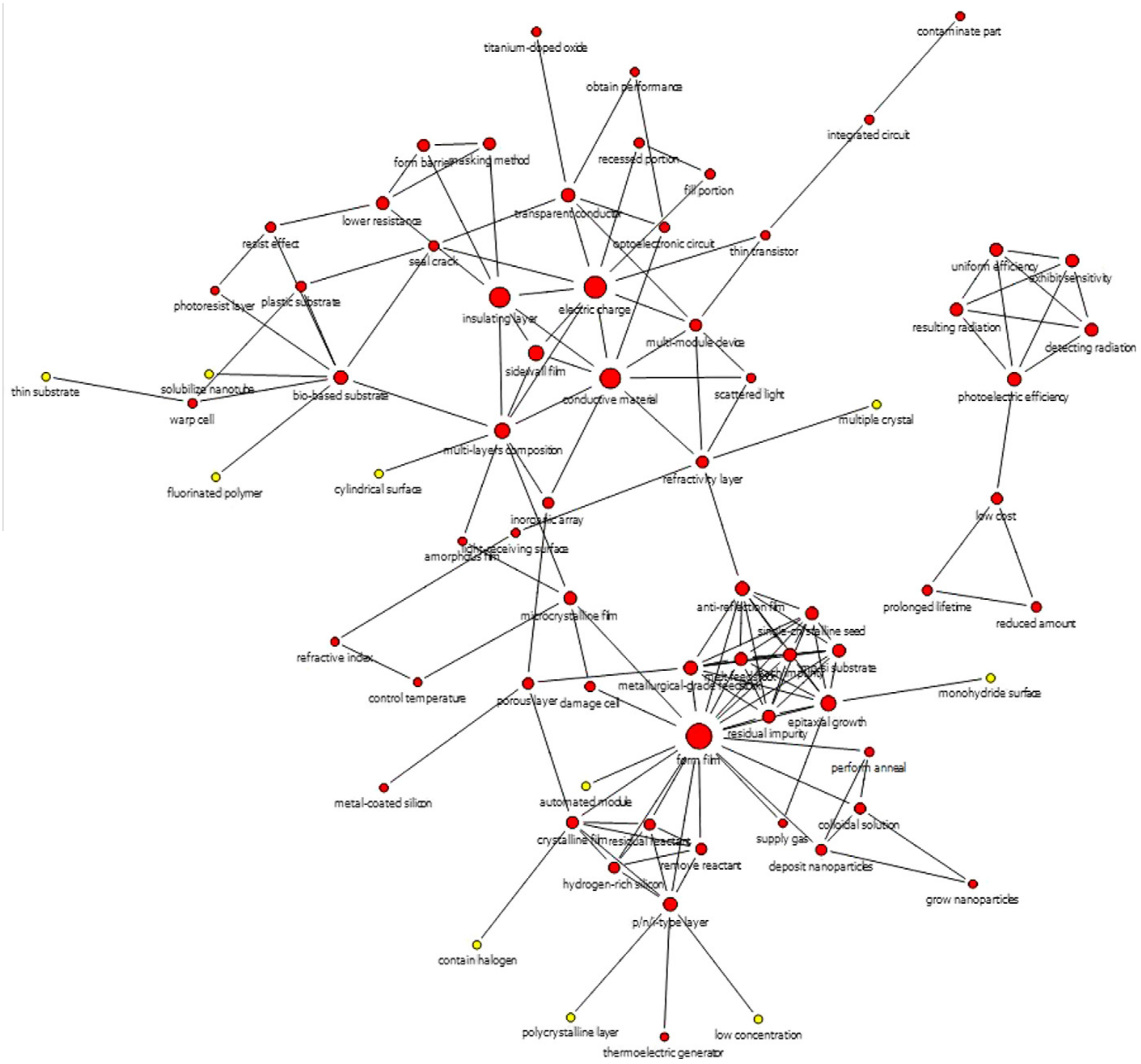


Fig. 9. Visualization of the IPFN (Yoon et al., 2011).

the matching average for similarity measurement between two sentences, a .NET based semantic similarity checker has been implemented as a C# library (Simpson & Dao, 2005). The mapper employs the semantic similarity checker. If a property or a function extracted from a patent is related to two or more phases of a trend, the mapper chooses the more evolved one. The TRIZ trend mapper also provides a similarity threshold: the mapper performs the trend mapping increasingly tightly as the acceptable similarity confidence score approaches 1. Many TRIZ trends can be identified using this approach except numerical trends such as the ‘mono-bi-poly’ trends (similar, various and increasing differences), the ‘macro-to-nano’ trend and the ‘decreasing density’ trend.

3.5.3. Evolutionary potential analyzer

The evolutionary potential analyzer uses evolutionary potential radar plots to visualize the evolutionary status of patents. To this end, the analyzer normalizes each identified trend phase with respect to the number of total trend phases in that trend. Next, the analyzer identifies untapped evolution trends by comparing average trend phases and maximum trend phases of patents, and suggests which trends the system should evolve to develop discontinuous technologies. Because the visualization and suggestion of the analyzer can be used as valuable input for technology forecasting based on TRIZ trends, they can be incorporated in design processes and managerial decision making processes.

4. Case study with prototype system

Using the proposed system, two case studies are introduced. The case of IPFN uses patents related to silicon-based thin film solar cells and the case of ETA uses patents related to umbrellas.

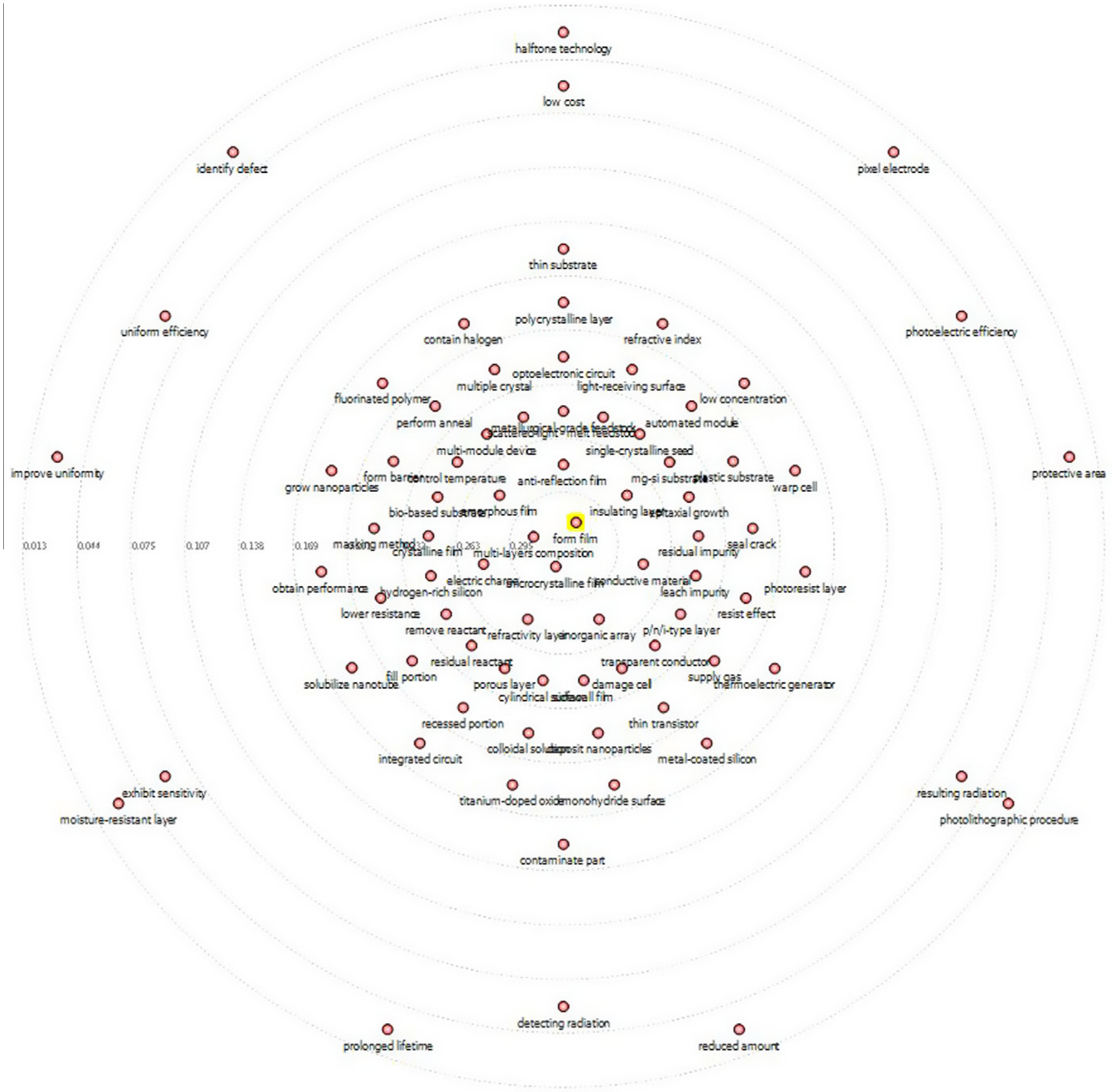


Fig. 10. Closeness centrality analysis of the IPFN.

4.1. Case 1: silicon-based thin film solar cells

4.1.1. Data collection

Solar energy has become the most promising renewable energy source because it is unlimited and non-polluting. A solar cell is a device that exploits the photovoltaic effect to convert sunlight directly into electricity. Silicon-based thin film solar cells are among the most advanced types of solar cells. Using a patent search query, this research collected patents from USPTO; 50 patents were used for analysis. Using PFMM, binary relations concerning properties and functions were extracted, and binary relations that contain the English stopwords were eliminated.

4.1.2. Organizing IPFN

To merge the co-occurrence matrices of all patents, first we defined a mapping base file to recognize identical nodes by grouping the identified binary relations (Fig. 7). The property–function merging rule descriptor can create new groups, add a property or a function to existing groups, and delete irrelevant properties or functions. Finally the binary relations are classified into 96 types of nodes (properties and functions). Then the IPFN generator merged the co-occurrence matrices of all patents into an IPFN, and 156 links (co-occurrences) were identified from the IPFN (Fig. 8). The generated IPFN can be analyzed using external SNA tools, so the IPFN generator in this case study transformed the IPFN into the data format appropriate to NetMiner and created a file.

4.1.3. Analysis of IPFN

NetMiner, an SNA tool, visualizes the IPFN and analyzes SNA indicators such as degree, centrality and density. In this case study, degree analysis and closeness centrality analysis are introduced. First, the nodes with high degree could be identified; larger nodes

Table 4
The six TRIZ trends and trend phases.

TRIZ trend	Trend phases (normalized phase)
Space segmentation	Monolithic solid (0.20) → hollow structure (0.40) → structure with multiple hollow (0.60) → capiliary/porous structure (0.80) → porous structure with active elements (1.00)
Surface segmentation	Smooth surface (0.25) → surface with rib protrusions (0.50) → 3D roughened surface (0.75) → roughened surface + active pores (1.00)
Asymmetry	Symmetrical system (0.33) → partial asymmetry (0.67) → matched asymmetry (1.00)
Dynamization	Immobile system (0.20) → jointed system (0.40) → fully flexible system (0.60) → fluid or pneumatic system (0.80) → field based system (1.00)
Controllability	Direct control action (0.25) → action through intermediary (0.50) → addition of feedback (0.75) → intelligent feedback (1.00)
Human involvement	Human (0.17) → human + tool (0.33) → human + powered tool (0.50) → human + semi automated tool (0.67) → human + automated tool (0.83) → automated tool (1.00)

Table 5
Identified trends and trend phases (Yoon & Kim, 2011).

Patent no.	Identified binary relations	TRIZ trends (normalized phases)
2009476098	Hold umbrella(dobj), connected spring(partmod), move pole(dobj), pole angle(amod), adjust angle(dobj), allows pivot(rcmod), allow motion(dobj), rotary bearing(amod), outdoor table(amod), tilt umbrella(dobj)	DYN (0.6)
2008155952	Cover relationship(dobj), open umbrella(partmod), wider area(amod), provide area(dobj), reduce weight(dobj)	SPS (0.6)
2008157729	Ventilating umbrella(amod), provide umbrella(dobj), protects umbrella(rcmod), contain trunk(dobj), multi-layered canopy(amod), contain plurality(dobj), make garlands(partmod), synthetic leaves(amod), distal end(amod), middle end(amod), hinged end(amod)	SPS (0.6)
2007224570	Self-inflating shield(amod), particular umbrella(amod), inflatable envelope(amod), flexible material(amod), chemical reaction(amod), produced reaction(rcmod), dissolve granulate(dobj), citric acid(amod), applied pressure(partmod), predetermined point(amod), trigger reaction(dobj), self-inflating balloon(amod)	DYN (0.8) CON (0.75)
2008155351	Connecting device(amod), designed device(partmod), pivoted part(amod), lengthwise ribs(amod), transverse ribs(amod), hollow sleeve(amod), tubular sleeve(amod), sheathed sleev(partmod), symmetrical wings(rcmod), elongated wings(amod), same line(amod), central bore(amod), use rivet(dobj), enhance strength(dobj)	SPS (0.4) ASY (0.33) DYN (0.6)
2009453398	Capable structure(amod), cool wind(amod), produce wind(dobj), hollow rod(amod), circumferential portion(amod), ventilating member(amod), conduct member(dobj), inhale air(dobj), provide wind(dobj), comfortable effect(amod), provide effect(dobj)	SPS (0.4)
2008082994	Provide visibility(dobj), crossing individual(partmod), cross intersection(dobj), dark weather(mod), upper end(amod), gripping end(amod), enclosed light(partmod), transparent casing(amod), automotive vehicle(amod), avoid accidents(dobj)	SPS (0.4) SUP (0.5) HUI (0.67)
2008099457	Motorized fan(amod), drives fan(rcmod), drive rotation(dobj), leading edge(amod), unshaped portion(amod), trailing edge(admod), preceding blade(amod), vertical spacing(amod), create spacing(dobj), adjacent blades(amod), create slope(dobj), flows channel(rcmod), stale air(amod), hot air(amod), push air(dobj), cooling breeze(amod), provide breeze(dobj)	SPS (0.4) DYN (0.6) HUI (0.50)
2008071077	Automatic umbrella(amod), longitudinal axis(amod), mirror sets(partmod), mirror images(dobj), equal numbers(amod), distal ends(amod), respective rib(amod), mutual connection(amod), shorter links(amod), proximal ends(amod), fixed ring(partmod), intermediate rings(amod), biased spring(partmod), automatic opening(amod), geometric relation(amod)	ASY (0.67) HUI (0.67)
2006589427	Attached umbrella(amod), insert umbrella(dobj), pole assembly(amod), cylindrical pole(amod), extending pole(patmod), rotatable chuck(amod), separate tool(amod), portable told(amod), handheld tool(amod), rotate assembly(dobj), cutting section(amod), anchoring section(amod), attached section(rcmod)	DYN (0.6)
2007809236	Solar umbrella(amod), automatic device(amod), heat device(dobj), solar panel(amod), momentary switch(amod), anchor device(dobj), malleable surfaces(amod), rechargeable batteries(amod), electric motor(amod), tubular member(amod), waterproof switch(amod), 3-function switch(amod), additional methods(amod)	SPS (0.4) HUI (0.67) CON (0.5)

SPS: Space Segmenation, SUS: Surface Segmentation, ASY: Asymmetry, DYN: Dynamization, CON: Controllability, HUI: Human Involvement.

indicate a higher degree. (Fig. 9). They include ‘form film’, ‘multi-layers composition’, ‘anti-reflection film’, ‘*p*-*i*-*n*’ type layer’, ‘leach impurity’, ‘multi-module device’ and ‘refractive layer’. These high-ranked properties and functions are strongly related to the dominant methods or objectives of silicon-based thin film solar cells. The reason that the technology trends of thin film solar cells moved from bulk type to thin film was the increase in the price of silicon. Therefore, patents to reduce the use of silicon and increase electricity conversion efficiency have been actively developed. Actually, many patents have presented methods of forming thin films such as *n*-layer, *i*-layer, *p*-layer, substrate and transparent conductive oxide, methods of constructing multi-layered thin film structure such as tandem cells and triple cells, and methods of assembling solar cell modules. In this way, degree analysis identifies key concepts related to a target technology.

Closeness of nodes can be visualized using concentric circles (Fig. 10); nodes closest to the center have highest closeness centrality. Properties and functions with high centrality in the IPFN were ‘form film’, ‘microcrystalline film’ and ‘multi-layered composition’. Actually, these are the most dominant and necessary methods or objectives prerequisite to development of thin film solar cells. Especially, many patents about microcrystalline silicon manufacture have been suggested to overcome amorphous silicon’s limitation that its photoconductivity can be reduced by prolonged exposure to intense light. The high centrality of ‘refractivity layers’ and ‘anti-reflection film’ imply that light trapping is a widely-used method in silicon-based thin film solar cells. In fact, they are strongly related to improvement of electricity conversion efficiency because they reduce reflection of incident rays and increase the light path within light trapping layers.

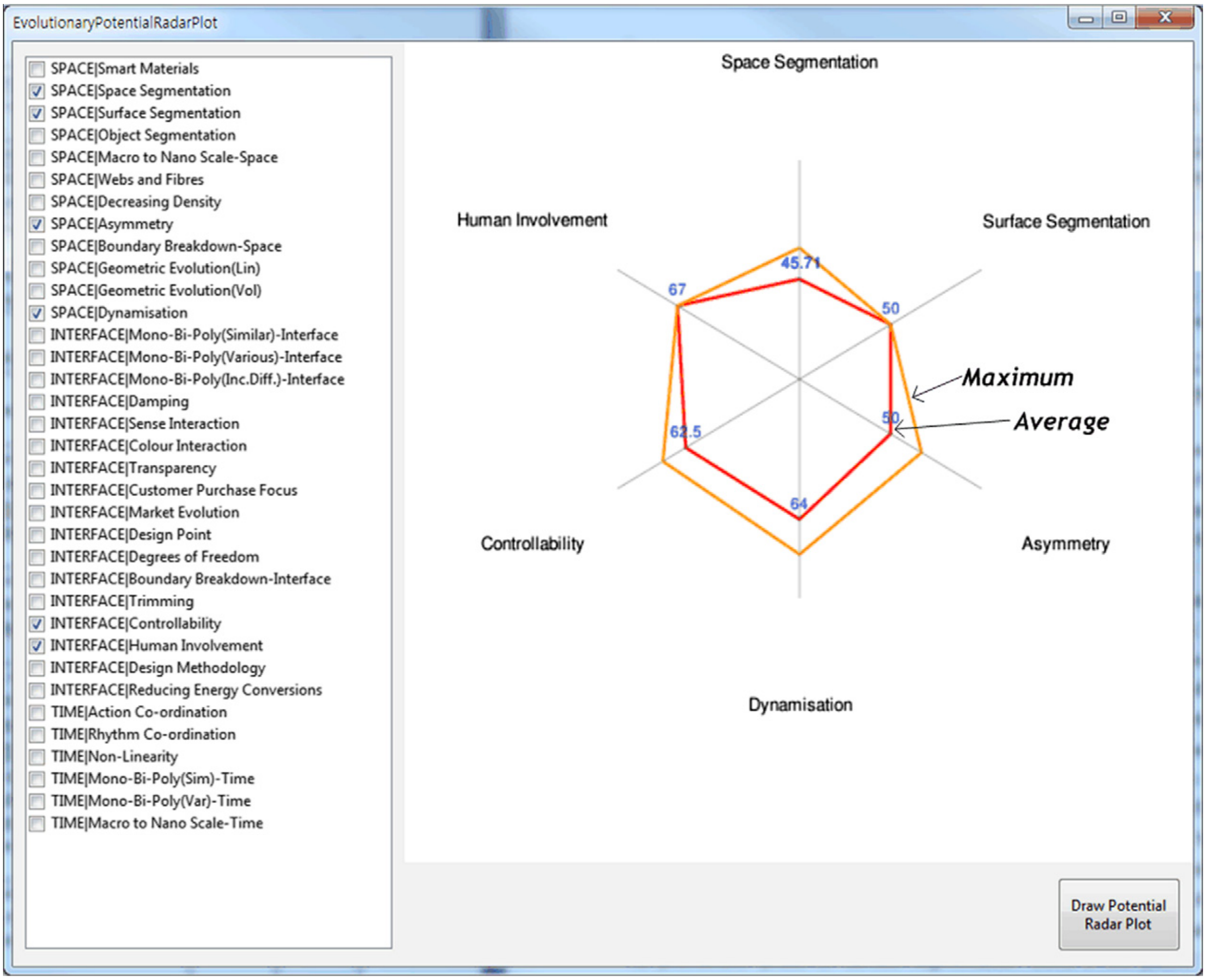
4.2. Case 2: umbrellas

4.2.1. Data collection

To predict further improvements of a system, this case study uses the patents of umbrella product family. About 4200 patents categorized into the international patent category code ‘A45B’ and having the word ‘umbrella’ in their titles were collected from USPTO. With respect to the six TRIZ trends (Table 4) including space segmentation, surface segmentation and dynamization, a ‘reasons for jumps’ rule base was defined by referring to updated TRIZ trends (Mann, 2002) and by conducting a training process that updates the rule bases gathering the domain-specific binary relations from 20 randomly selected patents. To ensure that the rule base was correct, this training process was repeated ten times. Then 11 randomly selected patents were prepared for evolution trend analysis of the umbrella product family.

4.2.2. Automated TRIZ trend identification

Properties and functions obtained from patents were related to specific trends and trend phases by the TRIZ trend mapper (Table 5). To this end, the mapper measured the semantic sentence similarity between the binary relations obtained from patents and the binary relations in the ‘reasons for jumps’ rule bases, then



☐ SPACE|Smart Materials
☑ SPACE|Space Segmentation
☑ SPACE|Surface Segmentation
☐ SPACE|Object Segmentation
☐ SPACE|Macro to Nano Scale-Space
☐ SPACE|Webs and Fibres
☐ SPACE|Decreasing Density
☑ SPACE|Asymmetry
☐ SPACE|Boundary Breakdown-Space
☐ SPACE|Geometric Evolution(Lin)
☐ SPACE|Geometric Evolution(Vol)
☑ SPACE|Dynamisation
☐ INTERFACE|Mono-Bi-Poly(Similar)-Interface
☐ INTERFACE|Mono-Bi-Poly(Various)-Interface
☐ INTERFACE|Mono-Bi-Poly(Inc.Diff.)-Interface
☐ INTERFACE|Damping
☐ INTERFACE|Sense Interaction
☐ INTERFACE|Colour Interaction
☐ INTERFACE|Transparency
☐ INTERFACE|Customer Purchase Focus
☐ INTERFACE|Market Evolution
☐ INTERFACE|Design Point
☐ INTERFACE|Degrees of Freedom
☐ INTERFACE|Boundary Breakdown-Interface
☐ INTERFACE|Trimming
☑ INTERFACE|Controllability
☑ INTERFACE|Human Involvement
☐ INTERFACE|Design Methodology
☐ INTERFACE|Reducing Energy Conversions
☐ TIME|Action Co-ordination
☐ TIME|Rhythm Co-ordination
☐ TIME|Non-Linearity
☐ TIME|Mono-Bi-Poly(Sim)-Time
☐ TIME|Mono-Bi-Poly(Var)-Time
☐ TIME|Macro to Nano Scale-Time

Fig. 11. Evolutionary potential radar plot (umbrella product family).

related the properties and functions of patents to specific trends and trend phases. For TRIZ trend mapping, the threshold value of acceptable similarity score of the TRIZ trend mapper was set to 0.90. In patent 2007224570, ‘flexible material’ was directly related to the ‘fully flexible system’ phase of the dynamization trend, and ‘self-inflating shield’ and ‘trigger reaction’ could be directly related to the ‘addition of feedback’ phase of the controllability trend. Likewise, trends and trend phases of other patents could be identified.

4.2.3. Analysis of evolutionary potential

Using the identified evolution trends and trend phases, the evolutionary potential analyzer depicts the normalized evolutionary potential radar plots (Fig. 11). The average evolutionary potential shows the overall evolutionary status of a product related to the collected patents. If the average evolution phase of a trend is low and the difference between average evolution phase and maximum evolution phase in the trend is zero, then the trend is an untapped area (Mann, 2002; Verhaegen et al., 2009). In this case, the normalized average evolution phase of the ‘surface segmentation’ trend is 0.50 and the difference between average potential and maximum potential is zero, so this indicates the trend has room for further improvements.

5. Conclusion and future research

As a vital component for technology development planning and technology strategy formulation, technology intelligence tools have been actively developed. Although many tools using textual information of patents include automated processes of technology analysis, at many point much effort of experts, who may be expensive or unavailable, is consumed in gathering of information for technology analysis. Fundamentally, these experts need to more concentrate on their knowledge activities such as technology trend identification and technology forecasting. To eliminate this dependency on experts, this paper proposed *TrendPerceptor*, a property–function based technology intelligence system that assists experts such as researchers, practitioners and R&D policy makers (1) to identify trends in invention concepts from patents, and (2) to perform evolution trend analysis of patents for technology forecasting. Properties and functions show invention concepts of systems: properties are related to materials and methods of a system; functions are related to uses and objectives of a system. In the PFMM, properties and functions are automatically extracted using grammatical analysis of textual information in patent documents. ICAM assists experts to identify invention concepts of a technology by providing a network composed of the extracted properties, functions and their co-occurrences. ETAM detects specific TRIZ trends and trend phases of a technology, and depicts the evolutionary potential as a normalized radar plot, which can be used valuable input for technology forecasting based on TRIZ trends.

Although this paper presented a system to help experts to identify technology trends in the aspect of invention concepts and evolution trends, more work is required. First, to support analysis of IPFN, the system used NetMiner, which is a commercial SNA tool, as a network analyzer. To increase applicability and convenience of the system, further work will develop an integrated system using open source libraries such as Java Universal Network/Graph Framework (JUNG, 2010). Second, the accuracy of TRIZ trend identification in ETAM should improved. Currently the module identifies most likely TRIZ trends and trend phases by referring only to binary relations defined in ‘reasons for jumps’ rule bases. In future work, a more reliable TRIZ trend identification module should be developed that integrates semantic sentence similarity measurement and classifiers based on machine learning, including support vector machines and *k*-nearest neighbors.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea Government (MEST) (No. 2009-0088379).

References

Altschuller, G. (1984). *Creativity as an exact science: The theory of the solution of inventive problems*.

Chang, P., Wu, C., & Leu, H. (2010). Using patent analyses to monitor the technological trends in an emerging field of technology: A case of carbon nanotube field emission display. *Scientometrics, 82*(1), 5–19.

Chen, L., Tokuda, N., & Adachi, H. (2003). A patent document retrieval system addressing both semantic and syntactic properties: Association for Computational Linguistics.

de Marneffe, M., & Manning, C. (2008A). Stanford typed dependencies manual, Technical report, 2008.

de Marneffe, M., & Manning, C. (2008B). The Stanford typed dependencies representation: Association for Computational Linguistics.

Dewulf, S. (2005). DIVA® - directed variation: Solving conflicts in TRIZ (Part 1). *The TRIZ Journal*.

Dewulf, S. (2006). Directed variation: Variation of properties for new or improved function product DNA, a base for ‘connect and develop’. *ETRIA TRIZ Futures*.

Diestel, R. (2005). *Graph theory*. *Graduate Texts in Mathematics*. Heidelberg: Springer.

Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social networks, 1*(3), 215–239.

Hanneman, R., & Riddle, M. (2005). *Introduction to social network methods*. CA: University of California, Riverside.

Hirtz, J., Stone, R., McAdams, D., Szykman, S., & Wood, K. (2002). A functional basis for engineering design: Reconciling and evolving previous efforts. *Research in Engineering Design, 13*(2), 65–82.

JUNG. (2010). Java Universal Network/Graph Famework (JUNG). online available from <http://www.jung.sourceforge.net/index.html>.

Kerr, C., Mortara, L., Phaal, R., & Probert, D. (2006). A conceptual model for technology intelligence. *International Journal of Technology Intelligence and Planning, 2*(1), 73–93.

Lee, B., & Jeong, Y. (2008). Mapping Korea’s national R&D domain of robot technology by using the co-word analysis. *Scientometrics, 77*(1), 3–19.

Lichtenthaler, E. (2004). Technological change and the technology intelligence process: A case study. *Journal of Engineering and technology Management, 21*(4), 331–348.

Mann, D. (2002). *Hands-on systematic innovation*. Creax press.

Mann, D. (2003). Better technology forecasting using systematic innovation methods. *Technological Forecasting and Social Change, 70*(8), 779–795.

Mann, D., & Dewulf, S. (2002). Evolutionary potential in technical and business systems. *TRIZ Journal*.

Miller, G. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*(11), 41.

Morris, S., DeYong, C., Wu, Z., Salman, S., & Yemenu, D. (2002). DIVA: A visualization system for exploring document databases for technology forecasting. *Computers and Industrial Engineering, 43*(4), 841.

OECD. (2008). OECD Factbook 2008: Economic, Environmental and Social Statistics. Online available from <http://www.titania.sourceoecd.org/vl=6549685/cl=18/nw=1/rpsv/factbook/070101.htm>.

Petrov, V. (2002). The laws of system evolution. *The TRIZ Journal, 3*, 9–17.

Resnik, P. (1999). Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research, 11*(95), 130.

Schuh, G., & Grawatsch, M. (2004). TRIZ-based technology intelligence.

Simpson, T., & Dao, T. (2005). WordNet-based semantic similarity measurement. Online available from <http://www.codeproject.com/KB/string/semanticsimilaritywordnet.aspx>.

STOPWORDS. (2010). English Stopwords. Online available from <http://www.ranks.nl/resources/stopwords.html>.

Tong, L., Cong, H., & Lixiang, S. (2006). Automatic classification of patent documents for TRIZ users. *World Patent Information, 28*(1), 6–13.

Trippe, A. (2003). Patinformatics: Tasks to tools. *World Patent Information, 25*(3), 211–221.

Verhaegen, P. A., D’Hondt, J., Vertommen, J., Dewulf, S., & Duflou, J. R. (2009). Relating properties and functions from patents to TRIZ trends. *CIRP Journal of Manufacturing Science and Technology, 1*(3), 126–130.

Yoon, B. (2008). On the development of a technology intelligence tool for identifying technology opportunity. *Expert Systems with Applications, 35*(1-2), 124–135.

Yoon, J., & Kim, K. (2011). An automated method for identifying TRIZ evolution trends from patents. *Expert Systems with Applications, 38*(12), 15540–15548.

Yoon, J., Choi, S., & Kim, K. (2011). Invention property-function network analysis of patents: a case of silicon-based thin film solar cells. *Scientometrics, 86*(3), 687–703.