# Requirement-Oriented Core Technological Components' Identification Based on SAO Analysis

Chao Yang[1, 2] · Donghua Zhu[1] · Xuefeng Wang[1] · Yi Zhang[1, 2] · Guangquan Zhang[2] · Jie Lu[2]

[1] School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China

[2] Decision Systems and e-Service Intelligence Lab, Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, NSW 2007, Australia

E-mail addresses: yc_2009@hotmail.com, zhudh111@bit.edu.cn, wxf5122@bit.edu.cn, yizhangbit@gmail.com, guangquan.zhang@uts.edu.au, jie.lu@uts.edu.au

**Abstract:** Technologies play an important role in the survival and development of enterprises. Understanding and monitoring the core technological components (e.g., technology process, operation method, function) of a technology is an important issue for researchers to develop R&D policy and manage product competitiveness. However, it is difficult to identify core technological components from a mass of terms, and we may experience some difficulties with describing complete technical details and understanding the terms-based results. This paper proposes a Subject-Action-Object (SAO)-based method, in which (1) a syntax-based approach is constructed to extract the SAO structures describing the function, relationship and operation in specified topics; (2) a systematic method is built to extract and screen technological components from SAOs; and (3) we propose a "relevance indicator" to calculate the relevance of the technological components to requirements, and finally identify core technological components based on this indicator. Based on the considerations for requirements and novelty, the core technological components identified have great market potential and can be useful in monitoring and forecasting new technologies. An empirical study of graphene is performed to demonstrate the proposed method. The resulting knowledge may hold interest for R&D management and corporate technology strategies in practice.

**Keywords:** Subject-Action-Object (SAO); Patent Analysis; Text Mining; Technological Components Identification.

## 1. Introduction

Today's global economy depends on technological innovation (Porter and Cunningham 2004). Emerging technology is a key element of competitive advantage for firms and countries (Ronald N. Kostoff et al. 2004). It represents progressive (sometimes explosive) developments for industries (Zhang et al. 2014a), and for this reason engineers and scientists focus on the identification of technological components for a technology of interest (Porter and Cunningham 2004). Technological components can also serve as the basis for further research (e.g., technological forecasting (Zhu and Porter 2002; J. Guo et al. 2016), identify technology opportunities (B. Yoon and Park 2005; J. Yoon and Kim 2012), patent map analysis (Tseng et al. 2005), extract technological intelligence (Zhu and Porter 2002; R. N. Kostoff et al. 2008), and explore innovation trajectory (lo Storto and Ieee 2008)).

There are three methods for identifying technological components: qualitative analysis, indicators' analysis, and citation analysis. The qualitative and expert based method is an important part of contemporary technology analysis. Combining empirical analyses with a diverse set of expertise is often indispensable in interdisciplinary research. Indicators' analysis focuses on the design of the evaluation indicator and usually uses existing keywords to identify technological components. Citation analysis focuses on the citation relationship and usually builds the network of technology to identify core technological components.

However, there are two challenges in identifying core technological components for a technology of interest: (1) technology is in rapid development—it has complexity and uncertainty—and sometimes, it lacks uniform industry standards and technical specifications, especially for emerging technology. These characteristics make it difficult for a person who is not an expert to understand and describe the complete details of technological components; (2) Words-based methods ignore verb-related phrases, pay more attention to the "system components" themselves but not to the relationships between topics. Thus, words-based methods can be problematic in describing complete technical details (e.g., process, method, material treatment, operation and requirements process) and misunderstanding would likely occur if we simply focused on these isolated terms. For example, the terms *chemical vapor deposition* and *graphene film* are retrieved and they are key terms in the field of Graphene. However, analysts like us, who lack strong technical knowledge in this domain, do not know how the two terms are used in the field. At this time, if we take note of the sentence containing these terms and then analyze the SAO structures in it, the meaning becomes clearer. We present an example: "Chemical vapor deposition method can be used in preparing high quality graphene film." The subject *chemical vapor deposition* and the object *graphene film* are easily connected by the action *prepare*. In this instance, we derive the idea to extend the term analysis to SAO structures.

Considering these concerns, this paper attempts to build up a method that combines SAO structures with bibliometric analysis, for identifying core technological components and minimizing the use of expert knowledge. We introduce SAO structures (a sequence of verbs and nouns) that can describe the function, relationship, operation and requirements with little human intervention. The main contribution of this paper is (1) an SAO-based method that is constructed to extract and screen technological components. The SAO semantic structure can combine verbs and nouns to present a detailed description of technology "function" (Choi et al. 2011), in this case, it reduce the need for using expert knowledge to summarize technological information from mass data. (2) identifying core technological components based on the relevance of the technological components to requirements. We first construct an SAO-based requirement identification method. Because SAO structure contains verb, it is good at identify the description of requirements. For example, “improve”, “stabilize”, “enhance” usually express the meaning of requirements. With the help of verbs in SAOs, we can reduce the use of expert knowledge in requirements identification. At last, we identify core technological components based on the relevance with requirements.

The core technological components identification process emphasizes significance, novelty, and requirements relevance, which ensure that the core technological components identified have great market potential. Meanwhile, With the help of SAO, we can arrive at more complete and clear technical details of new technologies.

The rest of this paper is organized as follows: In Section 2, we summarize the key literature of core technological components’ identification and SAO-based semantic analysis. Section 3 describes our data and elaborates on our methodology. In Section 4, we present a case study of patents related to graphene technology. We draw conclusions in Section 5.

## 2. Literature review

### 2.1. Identifying core technological components

The methods for identifying core technological components can be grouped into three main categories: qualitative analysis, indicators analysis, and citation analysis.

The most common methods are qualitative and expert based (Boon and Moors 2008; Simpson et al. 2008). It is one of the most important parts of contemporary technology analysis. However, with the development of interdisciplinary research, it is getting more and more costly to combine the strengths of various experts effectively.

Indicators analysis focuses on the design of the evaluation indicator (e.g., number of patents in a specific year (Bengisu 2003), and the similarity among conference sessions (Furukawa et al. 2015). Researchers usually design indicators based on existing categories, keywords or indexing terms to identify technological components and explore the newness, growth and market potential of technology (Cozzens et al. 2010; Vidal-Espana et al. 2007; Seymour 2008; H. Guo et al. 2011; Tseng et al. 2005; B. Yoon and Park 2005). However, the results of the indicators’ analysis may vary with the length of time windows (Rotolo et al. 2015). Indicators’ analysis also shows less focus on assessing relevance between technological components and requirement.

Citation analysis focuses on the citation and co-citation relationship inherent in the data. The most common way is to build the network of technology to identify core technological components (Erdi et al. 2013; Kajikawa et al. 2008; Cho and Shih 2011). We can also construct citation networks based on subject categories (Rafols et al. 2010). Based on the citation analysis, we can track research domains, identify technological components, and detect research fronts. However, there are some limitations in citation analysis. One of the problems is that there will be a time lag between the birth of a technological component and its appearance in the databases. Another limitation is that citation analysis cannot reflect the influences of public policy, patent laws, and the pace of economic growth.

There are also some hybrid approaches. Researchers try to combine patent citation, technology cycle, opinions of specialists, co-word analysis, and various quantitative indicators to identify technological components (Ju and Sohn 2015; Cozzens et al. 2010; Abercrombie et al. 2012). However, compared to describing high level concepts, identifying core technological components (e.g., process, method, operation) is always a challenge.

### 2.2. SAO analysis

Traditional keyword-based approaches ignore the role verbs play in the analysis of technological documents and deliver an understanding of technology information that is too shallow (Liu and Singh 2004; Choi et al. 2011; J. Guo et al. 2016). Potential relationships and value-added information are overlooked or unexplored. SAO is a triple structure extracted from text corpus. Subjects and objects are terms. Actions are verbs that represent the operation by which, or relationship between, those terms.

SAO has the potential to describe the detailed technical information (Wang et al. 2015). Cascini et al. (2004) believed that subjects and objects may refer to components of the system, and actions may refer to functions performed by the technology. Bergmann et al. (2008) believed that SAO structures can be organized in problem-solution formats. A number of researchers of Semantic TRIZ (Theory of Inventive Problem Solving) use the concept of SAO structure (Verbitsky 2004). They believe that SAO can be used to represent the Problem & Solution pattern, and to understand “what problems occurred” and “what solutions were used to solve these problems” (Zhang et al. 2014b). It is easy to map the “subject/object” to the “problem,” while transferring the whole SAO model to the “solution” with its “action” or “function” (Zhang et al. 2014b).

SAO is a useful tool that has been used to support technology mining. Choi et al. (2011) presented a method that formulates an SAO network and applied actor network theory to analyze technology implications. Some studies calculated the similarity of technologies or patents based on SAO structures, and then detected the risks of patent infringement (Bergmann et al. 2008; Park et al. 2012), produced an inventor competence map (Moehrle et al. 2005), and identified technological opportunities (J. Yoon and Kim 2012). Kim et al. (2009) argued that by extracting SAO structures, it is possible to identify a manifestation of technology and develop a technological trend discovery system.

In summary, SAO has the potential to identify core technological components. With the development of the SAO semantic analysis, scholars can put their creativity to full use and apply it to other fields.

## 3. Methodology

This paper proposes an SAO-based method to identify core technological components for technologies of interest. Technological components are a series of technology processes, operation methods, functions, and material treatments (e.g., “chemical vapor deposition is used in graphene preparation”, “adding water into a graphene mixture”, and “introducing microwave radiation assistance”). SAO structure can express detailed semantic information, which makes it easy to identify technological components. Fig. 1 shows the process of core technological components identification.

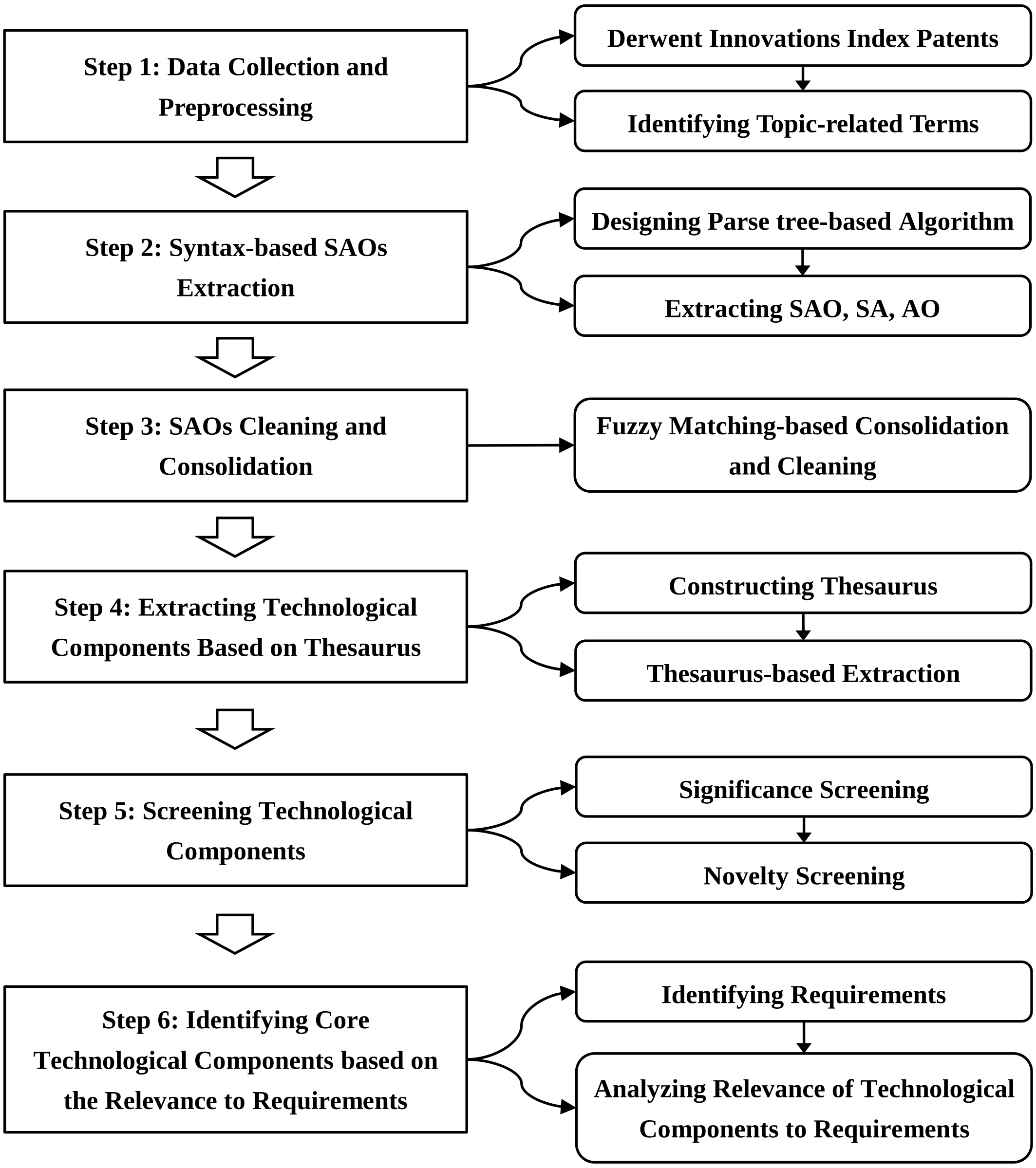
Step 1: Data Collection and Preprocessing
Derwent Innovations Index Patents
Identifying Topic-related Terms
Step 2: Syntax-based SAOs Extraction
Designing Parse tree-based Algorithm
Extracting SAO, SA, AO
Step 3: SAOs Cleaning and Consolidation
Fuzzy Matching-based Consolidation and Cleaning
Step 4: Extracting Technological Components Based on Thesaurus
Constructing Thesaurus
Thesaurus-based Extraction
Step 5: Screening Technological Components
Significance Screening
Novelty Screening
Step 6: Identifying Core Technological Components based on the Relevance to Requirements
Identifying Requirements
Analyzing Relevance of Technological Components to Requirements

**Fig. 1.** Steps of core technological components' identification.

### 3.1. Step 1: Data collection and preprocessing

Derwent Innovations Index (DII) is chosen as our patent data source. There are two main reasons for this: (1) The DII collects extensive patents from more than 40 patent organizations all over the world, which makes it an especially comprehensive patent database; and (2) the DII affords value-added patent information with its 60 years of patent indexing experience, and patents are rewritten into English for the purpose of clarifying obscure and legalistic terminology.

One challenge in SAOs extraction is how to ensure that the SAOs have a close relevance to specific technology topics. To solve this challenge, we introduced a set of preprocessing approaches that included identifying topic-related terms and extending topic-related terms. Topic-related terms are a collection of words/phrases that can indicate the important content of a specific technology field. These terms will be used as the candidate subject and object of SAO. In the chemistry or material field, we applied ChEBI to acquire the initial topic-related terms. But terms acquired from ChEBI cannot cover all the content of the technology field. In order to achieve all topic-related terms (1) we identified the sentences containing topic-related terms as core sentences; (2) we identified nouns in the core sentences and annotated them as core words; and (3) we extended core words to core terms based on natural language processing. The core terms will be used as topic-related terms.

### 3.2. Step 2: Syntax-based SAOs extraction

A syntax-based approach was constructed to extract the SAOs that described the function, relationship, and operation in specified topics. The SAO extracted in this paper is a bit different from the general one extracted with natural language processing. For performing statistical analysis, we had to ensure that the document-SAO matrix was not too sparse. To solve this question, (1) we extracted SAO structures from various tenses, voices and sentence elements; and (2) because sometimes there is no subject/object in the sentence or the subject/object is a pronoun, we extracted the broader SAO structure that includes SAO, SA, and AO structures. Following the two principles, based on parse tree, we designed a set of algorithms that applied the syntax-based extraction rule and topic-related terms to perform SAO Extraction (C. Yang et al. 2015; Chao Yang et al. 2017). There are seven kinds of rules according to the modes of Action (shown in Table 1). We implemented the algorithms with GATE (Cunningham et al. 2013).

**Table 1.** SAO structure extraction rules.

| Number | Extraction Rule | Example[b] |
|---|---|---|
| 1 | Extracting the SAO of Simple Present Tense[a] | A does B |
| 2 | Extracting the SAO of Passive Voice | A is done by B |
| 3 | Extracting the SAO of Infinitive | A (does) to do B |
| 4 | Extracting the SAO of Gerund | A (does) doing B |
| 5 | Extracting the SAO of Present Participle | A doing B |
| 6 | Extracting the SAO of Past Participle | A done by B |
| 7 | Additional rules | A involves/comprises B |

[a] Patent records (e.g. DII) are usually simple present tense.

[b] A and B are topic-related terms (taken from Step 1).

### 3.3. Step 3: SAO Cleaning and Consolidation

After the extraction of SAOs, these SAOs should be cleaned and consolidated as some similar concepts are presented by different SAOs. This step removed all general terms and consolidated synonyms, ambiguities, and different variant forms of SAO (C. Yang et al. 2015; Chao Yang et al. 2017).

The SAO is cleaned and consolidated using thesaurus and fuzzy matching: (1) different variant forms of words, such as singular/plural and synonyms, have been combined; (2) a stop word list is used to remove common SAOs; (3) a thesaurus of synonyms (including verbs and nouns) is constructed to combine similar SAO components; (4) we use fuzzy matching to combine similar SAOs. This step is fulfilled with VantagePoint.

### 3.4. Step 4: Extracting technological components based on thesauri

A systematic method is designed to extract technological components from SAOs. This method contains four steps: (1) keywords of papers whose keywords contain graphenes in WEB OF SCIENCE (2015–2016) were obtained; (2) removing common words and irrelevant words based on VantagePoint (VantagePoint); (3) based on the results of (1) and (2), constructing thesauri that contain the core technology terms in the graphene field; and (4) identifying SAOs which contain the terms in thesauri above, and these SAOs is the initial technological components (will be further screened in next step). For instance, “graphene preparation” is a term in the thesaurus and we use “graphene preparation” to search the SAOs set. We then obtain initial technological components: “copper foil used in graphene preparation” and “chemical vapor deposition performed in graphene preparation”. The use of SAO semantic structure can combine verbs and nouns to present a detailed description of technology “function”, which reduce the need for using expert knowledge to summarize technological information from mass data.

### 3.5. Step 5: Screening technological components

With the extraction of massive technological components, one challenge is how to identify the most critical technological components. A screening method is designed based on the characteristics of core technological components: significance and novelty. The proposed method explores significance and novelty of technological components. The analysis methods are list in Table 2.

**Table 2.** Screening framework.

| **Two aspects of screening** | **Methods** |
|---|---|
| Significance | Frequency statistics analysis<br>Technological components correlation analysis |
| Novelty | Technology cycle |

(1) Significance screening

Significance screening contains frequency statistics’ analysis and technological components’ correlation analysis: (1) technology is a convergence of previously separated research streams (Day and Schoemaker 2000). That means the importance of a technological component can be expressed by the accumulation of patents (or publications) containing this technological component. Thus, frequency statistics can be used to perform significance screening; (2) critical technological components have the potential to exert a considerable impact on the other technologies (Rotolo et al. 2015). As a result, if we put the core technological components into the technology interaction network, we will see that these core technological components are highly connected with other technological components. Thus, technological components’ correlation analysis is used to perform significance screening.

Technological components' correlation analysis generates a matrix based on the co-occurrence frequency of technological components. If two technological components occur frequently in the same patents, then we can conclude that they have a strong correlation. If there are m technological components $\{Tc_1, Tc_2, \cdots, Tc_m\}$ in all patents, then these elements make up an m × m matrix M. We call this matrix the symbiosis-based correlation matrix. Element $m_{ij}$ in matrix M is the co-occurrence frequency of two technical components $Tc_i$ and $Tc_j$. By normalizing matrix M, we obtain a correlation matrix $M'$. Element $m_{ij}$ represents the degree of correlation between $Tc_i$ and $Tc_j$.

To maximize the presentation of the effectiveness of the technological components' correlation analysis results, we generate visual technology maps (Figure 6) based on the matrix $M'$ via VantagePoint (VantagePoint). In the visual map, each node represents one technological component. The size of the node reflects the number of patents associated with the technological components. The line between two nodes indicates the degree of correlation between them. The absence of a line between two nodes means the correlation between those two nodes is lower than the cut-off value specified for that map.

(2) Novelty screening

A frequently seen trajectory (technology cycle) of technological components is the 'S' shaped curve (Cozzens et al. 2010; Carrillo and Gonz ález 2002), where in the early stages the technological component shows poor performance. In the following portion of the curve, the technological component takes off since some of the problems encountered in the first phase have been solved and customer acceptance has increased (shown in Fig. 2). Novelty screening chooses technological components that are at the stage of "ascent stage" in the technology cycle. They are usually the key parts of new technology.

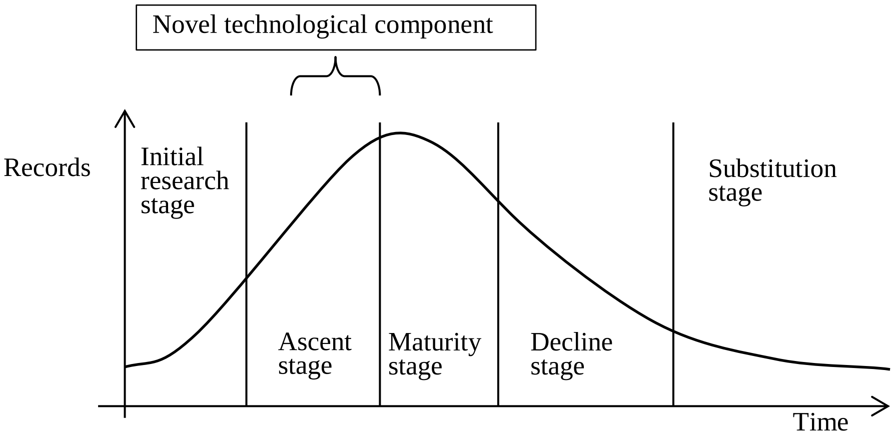

**Fig. 2.** The trajectory of technological components.

### 3.6. Step 6: Identifying core technological components based on the relevance to requirements

Requirements are series of technical requirements, standards, and customer expectations (e.g., material has high electrical conductivity, improves stability, improves surface area, and increases production efficiency). The value of a technology is embodied in the satisfaction of requirements.

Core technological components are finally identified based on the relevance of technological components to requirements. We propose a “relevance indicator” to calculate the relevance of the technological components to requirements. The introduction of requirements-based analysis makes the core technological components’ identification take into account external social factors (e.g., customer requirements, market standards, and market potential), and more accuracy. The technological components that satisfy key requirements usually have great market potential, and we choose them as core technological components. There are three steps:

(1) Identifying requirements via SAOs. Thesauri (terms and verbs, e.g., improve, quality, stability) that have close relationships with requirements are built based on keyword statistics and literature reviews, and then SAO sets are searched with this thesaurus to identify the SAOs that describe requirements.

(2) Screening core requirements according to the requirement frequency and requirement cycle. Firstly, we use frequency statistics of patents containing specific requirement to evaluate the significance of requirement. Secondly, core requirement is usually grows rapidly in recent years, and therefore, a frequently seen trajectory of core requirement development is the 'S' shaped curve. Finally, we combine similar requirements.

(3) Core technological components are identified based on the relevance of the technological components to requirements. We propose a “relevance indicator” to calculate the relevance of the technological component to requirements. Co-occurrence algorithm is the basis of the “relevance indicator,” that is, if a technological component and a requirement occur frequently in same patents, we can conclude that there is a strong correspondence between them. This co-occurrence-based measure calculates the relevance between a technological component to requirements set. There are n technological components $\{tc_1, tc_2, \cdots, tc_n\}$ and m requirements $\{rc_1, rc_2, \cdots, rc_m\}$. These elements make up an $\mathrm{n} \times \mathrm{m}$ matrix M where element $m_{ij}$ is the co-occurrence frequency of $tc_i$ and $rc_j$. “Relevance indicator” TRI is the sum of co-occurrence frequencies of a technological component with all requirements. TRI is used to rank these technological components and identify core technological components. The formula (the relevance indicator of $tc_i$) is:

$$TRI_i = m_{i1} + m_{i2} + \cdots + m_{ij}$$

For example, there are technological component “ $tc_1 = 'hydrogen\ peroxide\ added\ reaction\ system'$ ” and five requirements “ $rc_1 = 'improve\ efficiency'$ ”, “ $rc_2 = 'increase\ production'$ ”, “ $rc_3 = 'have\ high\ quality'$ ”, “ $rc_4 = 'improve\ stability'$ ”, “$rc_5 = 'improve\ electric\ conductance'$”. The co-occurrence frequencies of "$tc_1$" with “$rc_1$”, “$rc_2$”, “$rc_3$”, “$rc_4$”, “$rc_5$” are “$m_{11} = 3$”, “$m_{12} = 1$”, “$m_{13} = 2$”, “$m_{14} = 0$”, “$m_{15} = 3$”. So the relevance indicator of $tc_1$ is $TRI_1 = m_{11} + m_{12} + m_{13} + m_{14} + m_{15} = 3 + 1 + 2 + 0 + 3 = 9$. Higher values indicate a higher ranking, and this means that the technological components satisfy greater requirements and can be identified as a core technological component.

We used Gephi to perform the visualization of relevance between technological components and requirements. In the visual map, nodes represent technological components and requirements. Technological components have the same color and are located in the outer ring. Requirements have different colors and are located in the center of the circle. The bigger the node, the more nodes connect with it. The line between two nodes indicates the degree of relevance between them. It is a directed map in which the line is always directed from technological components to requirements.

## 4. Case study in graphene

### 4.1. Data collection and SAOs' identification

Graphene is a two-dimensional material and has shown great potential in the field of semiconductor, electronics, battery energy, and composites industries. Up to now, a lot of patents have been published. It is invaluable to identify the core technological components of graphene for that will bring us great benefits to a country's technological position.

In the case study, we chose the Derwent Innovations Index (DII) as our patent data source. The search strategy is that all DII patents from 1963 to November 2014 whose title contained the word "graphene" were downloaded (Shapira et al. 2010). This strategy resulted in a total of 7,413 patent family records spanning 30 countries, 1,803 institutional affiliations, and 7,299 inventors (C. Yang et al. 2015). The data is downloaded on 20.11.2014 (19.3MB, contact author at yc_2009@hotmail.com for getting data set).

After "patents preprocessing," "syntax-based SAOs extraction," and "SAOs Cleaning and Consolidation," we finally achieved 54,947 SAOs.

### 4.2. Extracting and screening technological components

Based on the method of Step 4, we obtained two thesauri: verbs and terms (shown in Table 3). With the thesauri, we finally arrived at 19,956 technological components.

**Table 3.** The examples of thesauri for technological components.

| **Verbs** | **Terms** |
|---|---|
| Involves | electron beam lithography |
| Include | chemical synthesis |
| Mix | electrochemical preparation |
| Dissolve | graphene oxide reduction |
| Add | catalytic transformation |
| Dope | microwave assisted hydrothermal method |
| Provide | ultrasonic exfoliation method |
| Use | laser |
| Carry | spin coating |
| Introduce | supersonic spray |

Then, technological components' screening was performed in the following step:

(1) Significance screening

Technological components are ranked based on the frequency of records containing these components. Fig. 3 shows part of the technological components. The horizontal axis of the map represents the numbers of patent records. The most frequently occurring technical component is "perform ultrasonic treatment."

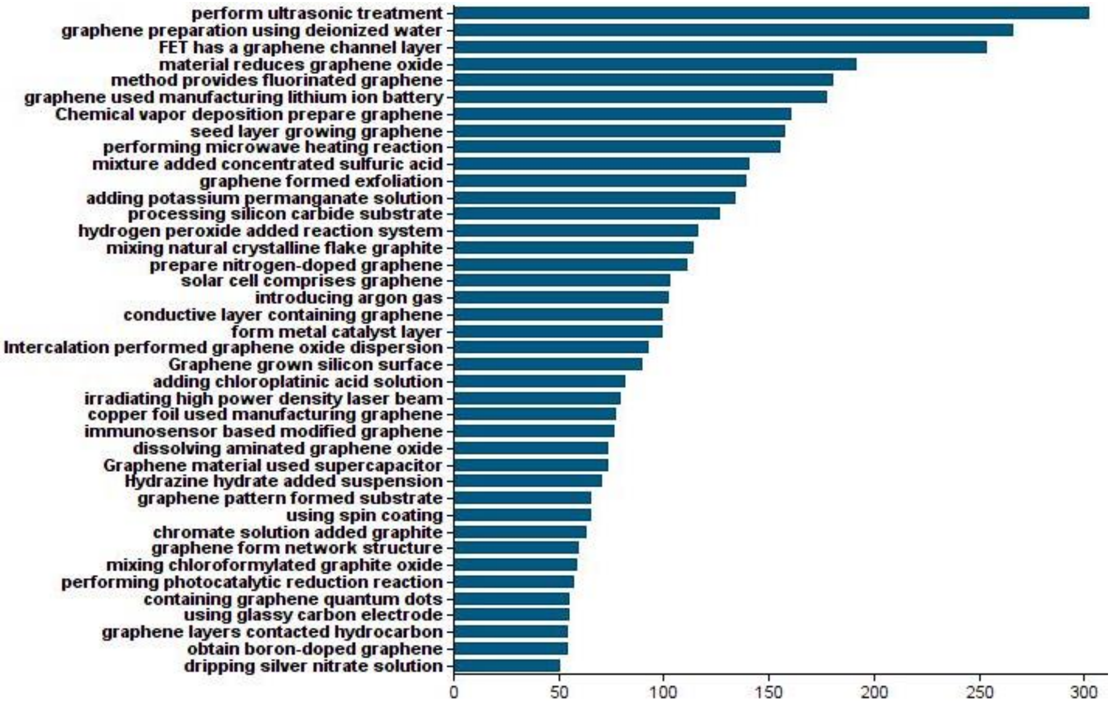

**Fig. 3.** Example of technological components' frequency statistics.

We chose the top 50 technical components (account for 42.8% of all patents) to generate the Auto-Correlation Map using VantagePoint (Fig. 4). We focused on the nodes that have many connections with other nodes. The more central the node is, the more likely this node presents an important technological component. The nodes with more than 3 connections are listed in Table 4.

**Table 4.** The nodes with more than 3 connections.

| The technological components (SAOs) satisfying significance |
|---|
| "irradiating high power density laser beam," "conductive layer containing graphene," "adding ethylenediamine," "using spin coating," "graphene pattern formed substrate," "adding polyacrylonitrile carbon fiber," "carrying hydrothermal synthesis method," "Chemical vapor deposition prepare graphene," "adding chloroplatinic acid solution," "graphene used manufacturing lithium ion battery," "performing microwave heating reaction," "dissolving aminated graphene oxide," "performing photocatalytic reduction reaction," "Intercalation performed graphene oxide dispersion," "dripping silver nitrate solution," "adding sodium borohydride powder," "adding sodium hydroxide solution," "mixing chloroformylated graphite oxide," "adding aqueous ammonia," "Hydrazine hydrate added suspension," "graphene preparation using deionized water," "introducing argon gas," "adding potassium permanganate |

solution,” “hydrogen peroxide added reaction system,” “perform ultrasonic treatment,” “mixture added concentrated sulfuric acid,” “mixing natural crystalline flake graphite,” and “chromate solution added graphite”

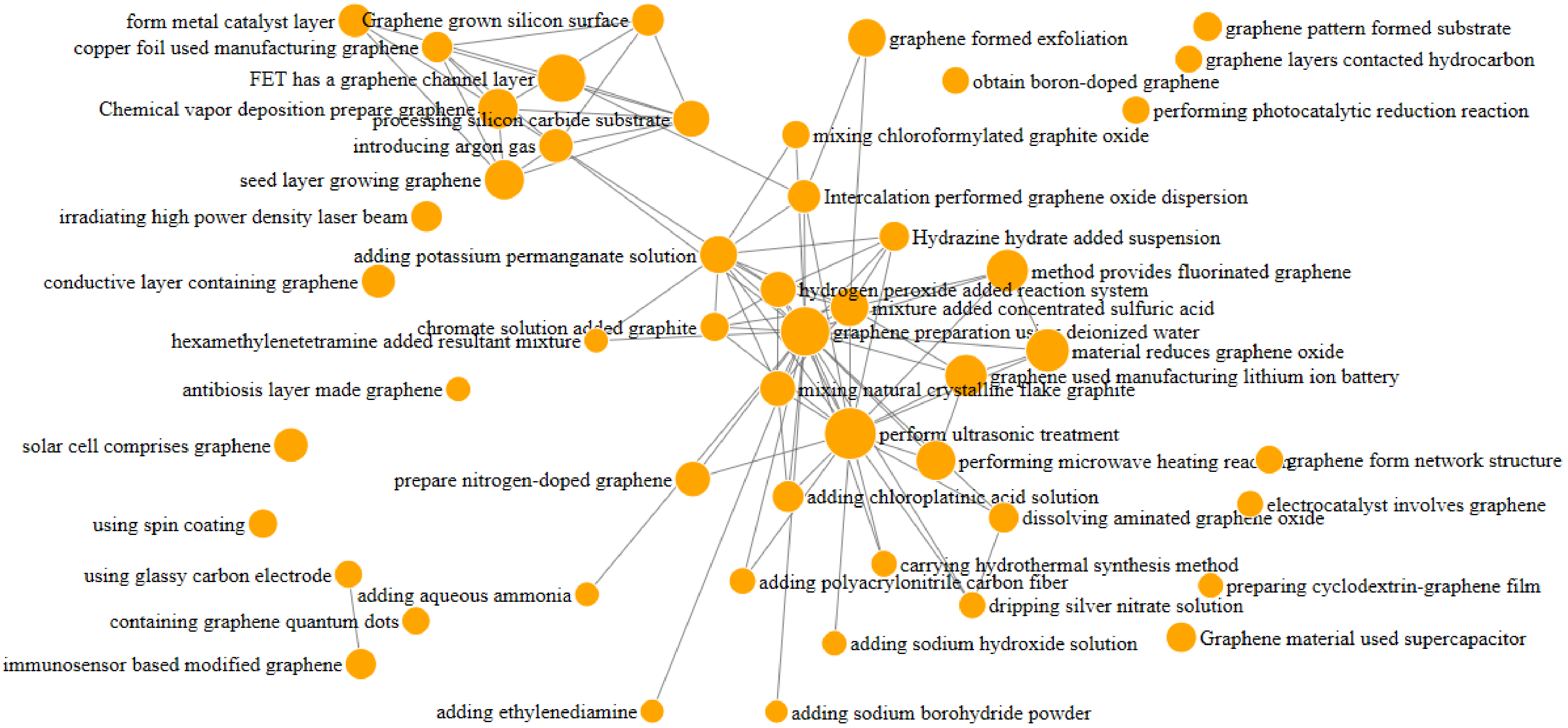
form metal catalyst layer
Graphene grown silicon surface
copper foil used manufacturing graphene
FET has a graphene channel layer
Chemical vapor deposition prepare graphene
processing silicon carbide substrate
introducing argon gas
seed layer growing graphene
irradiating high power density laser beam
adding potassium permanganate solution
conductive layer containing graphene
chromate solution added graphite
hexamethylenetetramine added resultant mixture
antibiosis layer made graphene
solar cell comprises graphene
prepare nitrogen-doped graphene
using spin coating
using glassy carbon electrode
adding aqueous ammonia
containing graphene quantum dots
immunosensor based modified graphene
adding ethylenediamine
graphene formed exfoliation
graphene pattern formed substrate
graphene layers contacted hydrocarbon
obtain boron-doped graphene
performing photocatalytic reduction reaction
mixing chloroformylated graphite oxide
Intercalation performed graphene oxide dispersion
Hydrazine hydrate added suspension
method provides fluorinated graphene
hydrogen peroxide added reaction system
mixture added concentrated sulfuric acid
graphene preparation using deionized water
material reduces graphene oxide
graphene used manufacturing lithium ion battery
mixing natural crystalline flake graphite
perform ultrasonic treatment
performing microwave heating reac
graphene form network structure
electrocatalyst involves graphene
adding chloroplatinic acid solution
dissolving aminated graphene oxide
carrying hydrothermal synthesis method
adding polyacrylonitrile carbon fiber
preparing cyclodextrin-graphene film
dripping silver nitrate solution
Graphene material used supercapacitor
adding sodium hydroxide solution
adding sodium borohydride powder

**Fig. 4.** Correlation map of technological components.

(2) Novelty screening

Fig. 5 shows the development of technological components. We identified the technological components that experienced early poor performance stages and are now taking off (the “ascent stage”). Considering the significance of the screen results, we finally achieved 16 technical components (shown in Table 5).

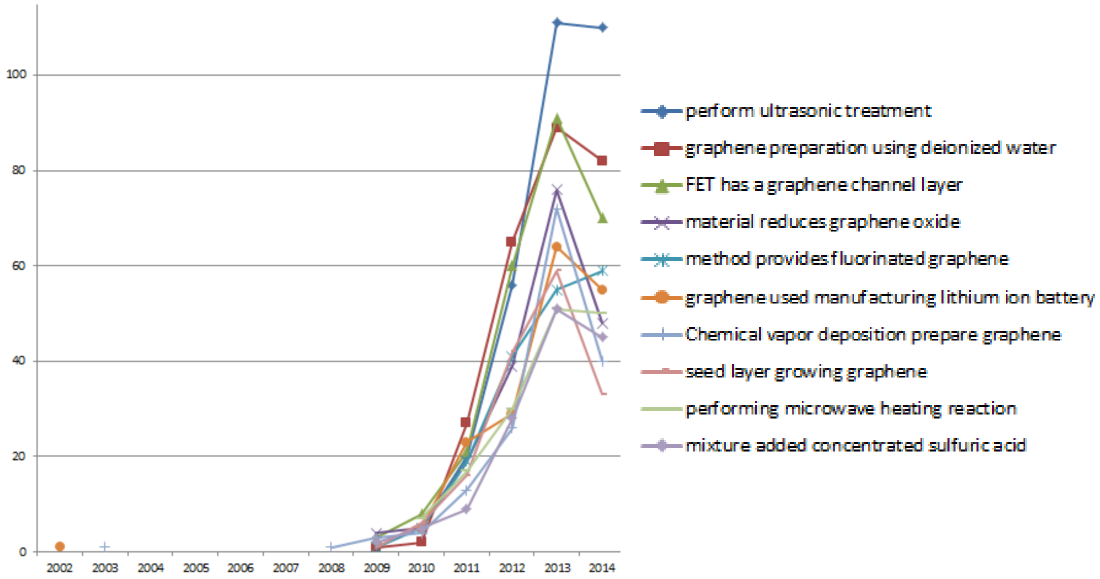

**Fig. 5.** Example of development of technological components.

**Table 5.** Novel technological components.

| The technological components satisfying novelty |
|---|
| “graphene pattern formed substrate,” “adding polyacrylonitrile carbon fiber,” “adding chloroplatinic acid solution,” “graphene used manufacturing lithium ion battery,” “performing microwave heating reaction,” “dissolving aminated graphene oxide,” “Intercalation performed graphene oxide dispersion,” “dripping silver nitrate solution,” “mixing chloroformylated graphite oxide,” “adding aqueous ammonia,” “Hydrazine hydrate added suspension,” “adding potassium permanganate solution,” “hydrogen peroxide added reaction system,” “perform ultrasonic treatment,” “mixing natural crystalline flake graphite,” and “chromate solution added graphite” |

## 4.3. Identifying core technological components based on its relevance to requirements

Firstly, we identify requirements in the graphene field with the thesauri (shown in Table 6) that is built based on keyword statistics and literature reviews. The 7413 patents yielded a total of 1554 requirements. Secondly, we use frequency statistics of patents containing specific requirement to evaluate the significance of requirements. Figure 6 displays the top 30. Thirdly, similarly to technological components, requirements also shows a 'S' shaped trajectory. Figure 7 shows the development of graphene related requirements. Fourthly, we identify the intersection of frequency statistics and requirement cycle above, and combine similar requirements. Finally, six core requirements were achieved: “has high energy storage capacity,” “improves stability,” “graphene manufactured high specific surface area,” “improves electrochemical performance,” “remove impurity component” and “graphene film has improved uniformity.”

**Table 6.** The examples of thesauri for requirements.

| **Verbs** | **Terms** |
|---|---|
| improve | conductivity |
| prepare | performance |
| increase | production efficiency |
| intensify | Electric conductance |
| stabilize | Mechanical property |
| measure | quality |
| reduce | stability |
| perform | capacitance |
| detect | dispersibility |
| enhance | purity |

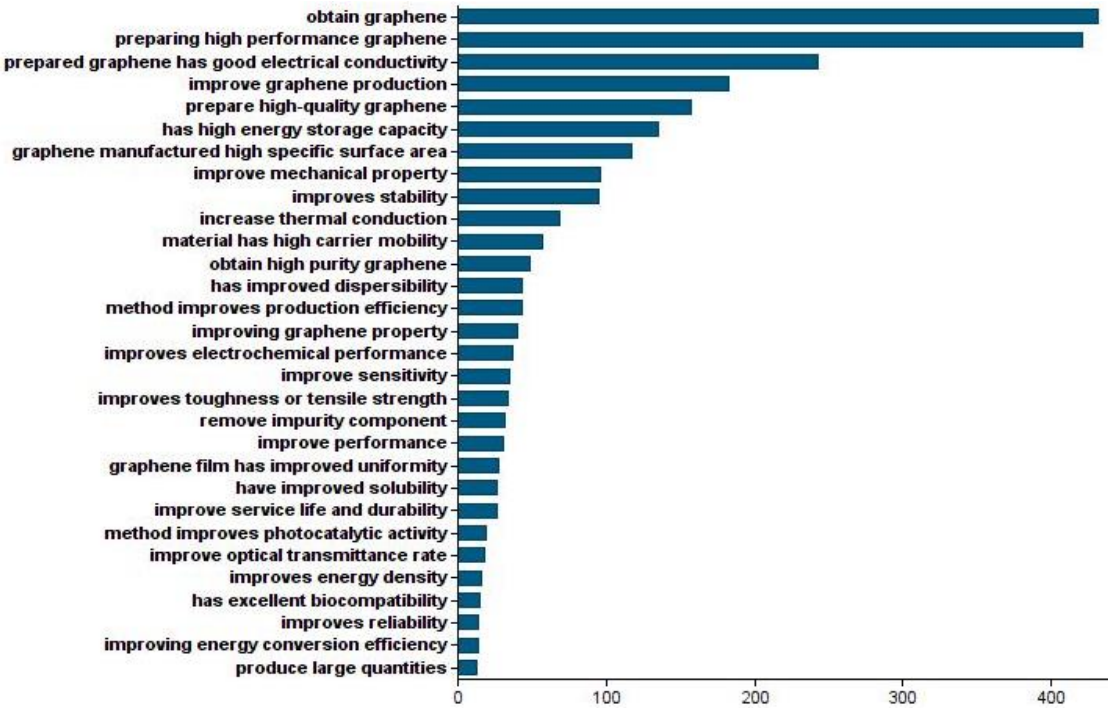

**Fig. 6** Example of requirements (top 30)

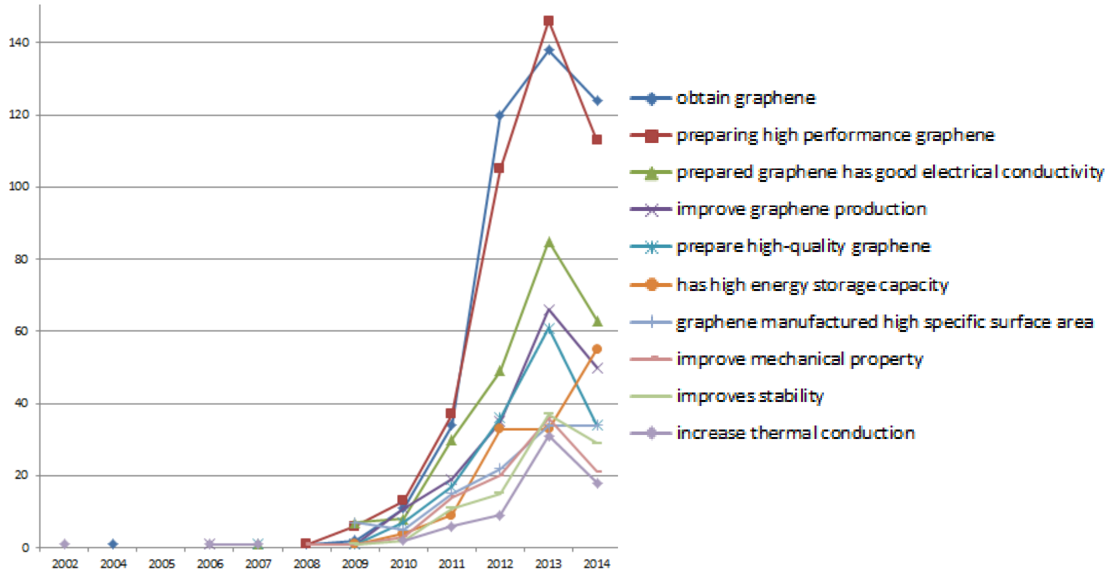

**Fig. 7** Example of requirements time distribution

We calculated the "relevance indicator" of all technological components and performed the visualization of relevance between technological components and requirement (shown in Fig. 8). Technological components are located in the outer ring. Requirements are located in the center of circle. The thickness of the edge expresses the strength of the relevance between them.

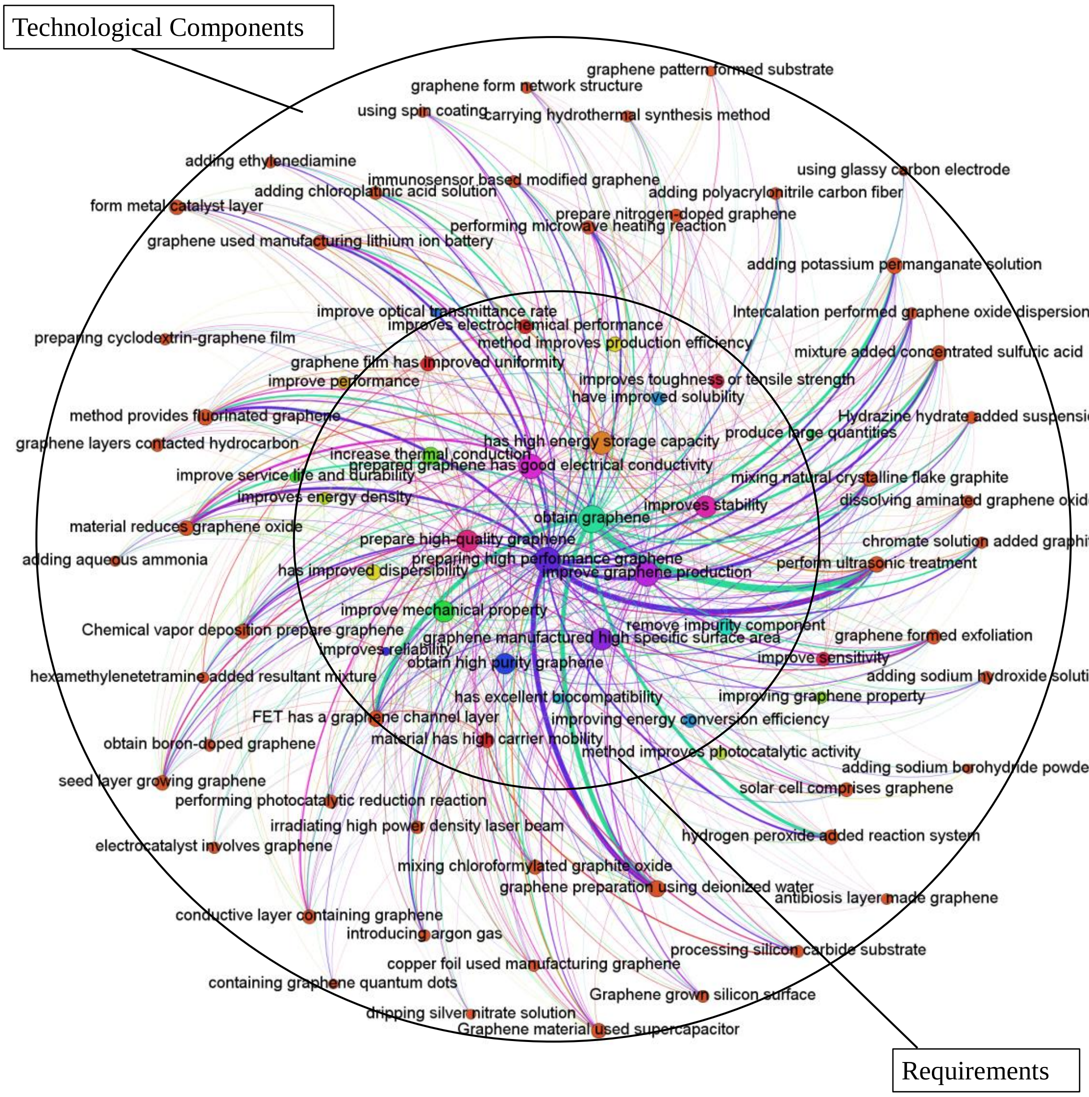
Technological Components
graphene pattern formed substrate
graphene form network structure
using spin coating
carrying hydrothermal synthesis method
adding ethylenediamine
immunosensor based modified graphene
using glassy carbon electrode
adding chloroplatinic acid solution
adding polyacrylonitrile carbon fiber
form metal catalyst layer
prepare nitrogen-doped graphene
performing microwave heating reaction
graphene used manufacturing lithium ion battery
adding potassium permanganate solution
improve optical transmittance rate
improves electrochemical performance
Intercalation performed graphene oxide dispersion
preparing cyclodextrin-graphene film
method improves production efficiency
mixture added concentrated sulfuric acid
graphene film has improved uniformity
improve performance
improves toughness or tensile strength
have improved solubility
method provides fluorinated graphene
Hydrazine hydrate added suspension
graphene layers contacted hydrocarbon
has high energy storage capacity
produce large quantities
increase thermal conduction
prepared graphene has good electrical conductivity
improve service life and durability
mixing natural crystalline flake graphite
improves energy density
improves stability
dissolving aminated graphene oxide
obtain graphene
material reduces graphene oxide
prepare high-quality graphene
chromate solution added graphite
adding aqueous ammonia
preparing high performance graphene
perform ultrasonic treatment
has improved dispersibility
improve graphene production
improve mechanical property
remove impurity component
Chemical vapor deposition prepare graphene
graphene manufactured high specific surface area
graphene formed exfoliation
improves reliability
improve sensitivity
obtain high purity graphene
hexamethylenetetramine added resultant mixture
adding sodium hydroxide solution
has excellent biocompatibility
improving graphene property
FET has a graphene channel layer
improving energy conversion efficiency
obtain boron-doped graphene
material has high carrier mobility
method improves photocatalytic activity
adding sodium borohydride powder
seed layer growing graphene
solar cell comprises graphene
performing photocatalytic reduction reaction
irradiating high power density laser beam
hydrogen peroxide added reaction system
electrocatalyst involves graphene
mixing chloroformylated graphite oxide
graphene preparation using deionized water
antibiosis layer made graphene
conductive layer containing graphene
introducing argon gas
processing silicon carbide substrate
copper foil used manufacturing graphene
containing graphene quantum dots
Graphene grown silicon surface
dripping silver nitrate solution
Graphene material used supercapacitor
Requirements

**Fig. 8.** Relevance map of technological components and requirements.

Based on Fig. 8, we chose the technological components whose relevance indicators are higher than ten as core technological components (e.g., the co-occurrence frequencies of "perform ultrasonic treatment" with the six core requirements are 11, 2, 7, 3, 4 and 2, so the "relevance indicator" is 29. Then "perform ultrasonic treatment" is chosen as a core technological component). The reason for the threshold "10" is that we can cover all the 6 core requirements and remove most of the unimportant technological components with this number. Core technological components and corresponding key requirements are presented in Table 7.

**Table 7.** Core technological components.

| Core Requirements | Core Technological Components |
|---|---|
| high energy storage capacity | **1** graphene used manufacturing lithium ion battery<br>**2** perform ultrasonic treatment<br>**3** chromate solution added graphite |

| | |
|---|---|
| | **4** adding polyacrylonitrile carbon fiber<br>**5** adding potassium permanganate solution<br>**6** performing microwave heating reaction<br>**7** Intercalation performed graphene oxide dispersion<br>**8** dissolving aminated graphene oxide |
| improving stability | **1** performing microwave heating reaction<br>**2** graphene used manufacturing lithium ion battery<br>**3** dissolving aminated graphene oxide<br>**4** mixing natural crystalline flake graphite<br>**5** hydrazine hydrate added suspension<br>**6** adding polyacrylonitrile carbon fiber<br>**7** adding chloroplatinic acid solution<br>**8** mixing chloroformylated graphite oxide |
| manufacturing high specific surface area graphene | **1** perform ultrasonic treatment<br>**2** hydrogen peroxide added reaction system<br>**3** Intercalation performed graphene oxide dispersion<br>**4** performing microwave heating reaction<br>**5** mixing natural crystalline flake graphite<br>**6** adding potassium permanganate solution<br>**7** adding aqueous ammonia, |
| improving electrochemical performance | **1** perform ultrasonic treatment<br>**2** adding chloroplatinic acid solution<br>**3** graphene used manufacturing lithium ion battery<br>**4** mixing chloroformylated graphite oxide |
| removing impurity component | **1** perform ultrasonic treatment<br>**2** hydrazine hydrate added suspension<br>**3** mixing natural crystalline flake graphite<br>**4** hydrogen peroxide added reaction system |
| improving uniformity of graphene film | **1** hydrazine hydrate added suspension<br>**2** perform ultrasonic treatment<br>**3** dripping silver nitrate solution |

### 4.4. Comparison with LDA model

Core technological component are technology process, operation method and function of a technology. It contains more specific information than keyword. There is no direct way of identifying core technological components in recent studies. To demonstrate the advantages of the proposed method, we compare our method with LDA model which can produce similar results and has a potential in identifying core technological components.

The data set used in LDA model is the same with the proposed method above. To improve the effect of LDA model, we introduce term clumping to identify the topic words of patents. Based on the results of term clumping, we fulfill the LDA model. The setting of LDA model are: maximum number of iterations is 10000, document-topic associations is 2.0, topic-term associations is 0.5. We set 10 topics in the application of LDA model. The top 5 words and their probability distribution are listed in Table 8.

**Table 8.** Top 5 words and probability distribution of 10 topics.

| Topic 1 | | Topic 6 | |
|---|---|---|---|
| Graphene | 0.0421 | composite | 0.0671 |
| Device | 0.0414 | polymer | 0.0442 |
| Light | 0.0253 | material | 0.0387 |
| electric | 0.0211 | fiber | 0.0361 |
| system | 0.0173 | conductivity | 0.0347 |
| Topic 2 | | Topic 7 | |
| material | 0.2086 | electrode | 0.1449 |
| composite | 0.0988 | Graphene | 0.0675 |
| battery | 0.0420 | cell | 0.0391 |
| Ion | 0.0337 | sensor | 0.0263 |
| lithium | 0.0331 | membrane | 0.0244 |
| Topic 3 | | Topic 8 | |
| solution | 0.1092 | carbon | 0.1365 |
| water | 0.0686 | gas | 0.0514 |
| degrees | 0.0443 | catalyst | 0.0450 |
| Acid | 0.0441 | Heating | 0.0418 |
| mixture | 0.0257 | nanotube | 0.0402 |
| Topic 4 | | Topic 9 | |
| layer | 0.2045 | oxide | 0.1863 |
| Graphene | 0.0888 | Graphene | 0.1862 |
| device | 0.0648 | graphite | 0.1211 |
| electrode | 0.0368 | liquid | 0.0406 |
| Semiconductor | 0.0317 | dispersion | 0.0339 |
| Topic 5 | | Topic 10 | |
| Graphene | 0.1944 | film | 0.1414 |
| sheet | 0.0670 | substrate | 0.1245 |
| structure | 0.0564 | Graphene | 0.1121 |
| surface | 0.0399 | metal | 0.0835 |
| Form | 0.0286 | surface | 0.0537 |

LDA model can present topic distribution and topic words, and has a significant potential in topic (technological components and requirements) identification. We compare the result of proposed method (requirement-oriented core technological components' identification method) with LDA model in two aspects:

(1) Interpretation and information richness

Compared with LDA model, the result of proposed method is much better in the interpretation of technological components. Based on table 7 and table 8, we can see that LDA model ignore the role verbs play in the analysis of technological documents and deliver an shallow understanding of technology information. Potential relationships and value-added information are overlooked or unexplored. SAO is a triple structure extracted from text corpus and can present problem-solution patterns. In this case, SAO has the potential to describe the detailed technical information, e.g., “what problems occurred” and “what solutions were used to solve these problems". Subjects and objects can refer to components of the system, and actions can refer to functions performed by the technology.

(2) Semantic disambiguation.

In LDA model, homonyms and synonyms of words result in ambiguous interpretations. But SAO is a triple structure extracted from a text corpus. Subjects and objects are terms or phrases that are closely related to the topic. Actions are verbs that represent the operation by which, or the relationship between, those terms and phrases. The development of natural language processing techniques has allowed SAO structures to express rich semantic information and gained recognition as a powerful tool for identifying concepts in a corpus. So SAO structure has the ability to solve the problem of ambiguous interpretations resulted by homonyms and synonyms of words.

## 5. Conclusions and further studies

In the current age of Big Data, it is common sense to combine traditional bibliometric analysis with semantic analysis, and this paper could be considered as this kind of an attempt. We proposed an SAO-based approach for semantic information retrieval, and then extract candidate technological components. A systematic method that considered significance and novelty was built to screen and select technological components. At last, a requirement relevance analysis was used to identify core technological components.

The main advantages of the proposed method are:

(1) SAO structure is used to achieve core technological components, which can be helpful in identifying new, complex, and uncertain concepts in fast growing technologies. With the lack of uniform industry standards and technical specifications in growing technology, terms can be complex, uncertain and changing over time, but SAOs are relatively stable in the evolution process and can address more complete semantic understandings. The reason is that SAOs have verbs to describe action, and have Subject/Object to present more than one concept.

(2) The proposed method introduces requirements-oriented analysis which makes the core technological components’ identification take into account the relevance of technological components to requirements, and therefore becomes more accurate.

The proposed method served to identify core technological components, describe specific technical details of a technology (e.g., process, method, material treatment, operation, and requirements’ process). This method presents capabilities for R&D planning and generates Competitive Technical Intelligence (CTI) to inform strategic management.

The process of core technological components’ identification emphasizes significance, novelty, and requirements' relevance, which ensure that they have great market potential and can support the forecasting of new technologies. The proposed method can be helpful in solving general challenges of forecasting new technologies: (1) SAOs are relatively stable and can describe more complete technical details, which is helpful in solving the problem of technological forecasting—the complexity and uncertainty of the emerging concept caused by the rapid development of new technologies. (2) SAO is helpful for solving the problem of ambiguous interpretations resulted by homonyms and synonyms of words, especially in multidisciplinary and interdisciplinary research fields.

There are also several limitations to this paper. We emphasized recall more than the integrity of SAO. Compared to SAO, SA and AO lost a part of the information. We engaged experts for setting indictor thresholds, but a systematic setting process would be able to improve the efficiency of qualitative approaches. We anticipate further studies in four directions: (1) to continue to improve the SAO extraction algorithm to consolidate similar SAOs; (2) to introduce network-based techniques for relationship identification among S (Subject) and O (Object); (3) to introduce a systematic approach to weigh/rank the SAO structures for supporting bibliometric analysis in further steps; and (4) to extend the empirical study to address multiple ST&I data sources.

**Acknowledgments:** This work is supported by the General Program of National Natural Science Foundation of China (Grant Nos. 71673024, 71373019), the Australian Research Council (ARC) under discovery grants DP140101366 and DP150101645. This paper was also funded by the International Graduate Exchange Program of Beijing Institute of Technology.

## References

Abercrombie, R. K., Udoeyop, A. W., & Schlicher, B. G. (2012). A study of scientometric methods to identify emerging technologies via modeling of milestones. *Scientometrics, 91*(2), 327-342, doi:10.1007/s11192-011-0614-4.

Bengisu, M. (2003). Critical and emerging technologies in Materials, Manufacturing, and Industrial Engineering: A study for priority setting. *Scientometrics, 58*(3), 473-487, doi:10.1023/B:SCIE.0000006875.61813.f6.

Bergmann, I., Butzke, D., Walter, L., Fuerste, J. P., Moehrle, M. G., & Erdmann, V. A. (2008). Evaluating the risk of patent infringement by means of semantic patent analysis: the case of DNA chips. *R&D Management, 38*(5), 550-562, doi:10.1111/j.1467-9310.2008.00533.x.

Boon, W., & Moors, E. (2008). Exploring emerging technologies using metaphors – A study of orphan drugs and pharmacogenomics. *Social Science & Medicine, 66*(9), 1915-1927, doi:http://dx.doi.org/10.1016/j.socscimed.2008.01.012.

Carrillo, M., & González, J. M. (2002). A new approach to modelling sigmoidal curves. *Technological Forecasting and Social Change, 69*(3), 233-241, doi:http://dx.doi.org/10.1016/S0040-1625(01)00150-0.

Cascini, G., Fantechi, A., & Spinicci, E. (2004). Natural language processing of patents and technical documentation. In S. Marinai, & A. Dengel (Eds.), *Document Analysis Systems VI* (Vol. 3163, pp. 508-520, Lecture Notes in Computer Science). Berlin: Springer Berlin Heidelberg.

Cho, T. S., & Shih, H. Y. (2011). Patent citation network analysis of core and emerging technologies in Taiwan: 1997-2008. *Scientometrics, 89*(3), 795-811, doi:10.1007/s11192-011-0457-z.

Choi, S., Yoon, J., Kim, K., Lee, J. Y., & Kim, C. H. (2011). SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells. *Scientometrics, 88*(3), 863-883, doi:10.1007/s11192-011-0420-z.

Cozzens, S., Gatchair, S., Kang, J., Kim, K.-S., Lee, H. J., Ord óñez, G., et al. (2010). Emerging technologies: quantitative identification and measurement. *Technology Analysis & Strategic Management*, *22*(3), 361-376, doi:10.1080/09537321003647396.

Cunningham, H., Tablan, V., Roberts, A., & Bontcheva, K. (2013). Getting More Out of Biomedical Documents with GATE's Full Lifecycle Open Source Text Analytics. *PLoS Comput Biol, 9*(2), e1002854, doi:10.1371/journal.pcbi.1002854.

Day, G. S., & Schoemaker, P. J. H. (2000). Avoiding the Pitfalls of Emerging Technologies. *California Management Review, 42*(2), 8-33, doi:10.2307/41166030.

Erdi, P., Makovi, K., Somogyvari, Z., Strandburg, K., Tobochnik, J., Volf, P., et al. (2013). Prediction of emerging technologies based on analysis of the US patent citation network. *Scientometrics, 95*(1), 225-242, doi:10.1007/s11192-012-0796-4.

Furukawa, T., Mori, K., Arino, K., Hayashi, K., & Shirakawa, N. (2015). Identifying the evolutionary process of emerging technologies: A chronological network analysis of World Wide Web conference sessions. *Technological Forecasting and Social Change*, *91*, 280-294, doi:http://dx.doi.org/10.1016/j.techfore.2014.03.013.

Guo, H., Weingart, S., & B örner, K. (2011). Mixed-indicators model for identifying emerging research areas. *Scientometrics, 89*(1), 421-435, doi:10.1007/s11192-011-0433-7.

Guo, J., Wang, X., Li, Q., & Zhu, D. (2016). Subject–action–object-based morphology analysis for determining the direction of technological change. *Technological Forecasting and Social Change, 105*, 27-40, doi:http://dx.doi.org/10.1016/j.techfore.2016.01.028.

Ju, Y., & Sohn, Y. (2015). Patent-based QFD framework development for identification of emerging technologies and related business models: A case of robot technology in Korea. *Technological Forecasting and Social Change, 94*, 44-64, doi:10.1016/j.techfore.2014.04.015.

Kajikawa, Y., Yoshikawa, J., Takeda, Y., & Matsushima, K. (2008). Tracking emerging technologies in energy research: Toward a roadmap for sustainable energy. *Technological Forecasting and Social Change, 75*(6), 771-782, doi:http://dx.doi.org/10.1016/j.techfore.2007.05.005.

Kim, Y., Tian, Y., Jeong, Y., Jihee, R., & Myaeng, S.-H. (2009). Automatic discovery of technology trends from patent text. In *2009 ACM Symposium on Applied Computing* (pp. 1480-1487). Honolulu, Hawaii: ACM.

Kostoff, R. N., Boylan, R., & Simons, G. R. (2004). Disruptive technology roadmaps. *Technological Forecasting and Social Change, 71*(1–2), 141-159, doi:http://dx.doi.org/10.1016/S0040-1625(03)00048-9.

Kostoff, R. N., Solka, J. L., Rushenberg, R. L., & Wyatt, J. A. (2008). Literature-related discovery (LRD): Water purification. *Technological Forecasting and Social Change, 75*(2), 256-275, doi:10.1016/j.techfore.2007.11.009.

Liu, H., & Singh, P. (2004). ConceptNet - a practical commonsense reasoning tool-kit. *Bt Technology Journal, 22*(4), 211-226.

lo Storto, C., & Ieee (2008). *Exploring innovation trajectories in high-tech industries through patent analysis: the case of the optical memories industry* (Iemc - Europe 2008: International Engineering Management Conference, Europe, Conference Proceedings: Managing Engineering, Technology and Innovation for Growth). New York: Ieee.

Moehrle, M. G., Walter, L., Geritz, A., & Muller, S. (2005). Patent-based inventor profiles as a basis for human resource decisions in research and development. *R & D Management, 35*(5), 513-524, doi:10.1111/j.1467-9310.2005.00408.x.

Park, H., Yoon, J., & Kim, K. (2012). Identifying patent infringement using SAO based semantic technological similarities. *Scientometrics, 90*(2), 515-529, doi:10.1007/s11192-011-0522-7.

Porter, A. L., & Cunningham, S. W. (2004). *Tech mining: Exploiting new technologies for competitive advantage* (Vol. 29). New York: John Wiley & Sons.

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology, 61*(9), 1871-1887, doi:10.1002/asi.21368.

Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology? *Research Policy, 44*(10), 1827-1843, doi:10.1016/j.respol.2015.06.006.

Seymour, R. (2008). Platinum Group Metals Patent Analysis and Mapping A REVIEW OF PATENTING TRENDS AND IDENTIFICATION OF EMERGING TECHNOLOGIES. *Platinum Metals Review, 52*(4), 231-240, doi:10.1595/147106708x362735.

Shapira, P., Youtie, J., & Carley, S. (2010). Graphene research profile: UK and US publications, 2000-2010. Program on Nanotechnology Research and Innovation System Assessment: Georgia Institute of Technology Atlanta.

Simpson, S., Packer, C., Carlsson, P., Sanders, J. M., Ibarluzea, I. G., Fay, A. F., et al. (2008). Early identification and assessment of new and emerging health technologies: Actions, progress, and the future direction of an international collaboration-EuroScan. *International Journal of Technology Assessment in Health Care, 24*(4), 518-524, doi:10.1017/s0266462308080689.

Tseng, Y. H., Lin, C. J., & Lin, Y. I. (2005). Text mining for patent map analysis. *Information Processing & Management, 43*(5), 1216–1247.

VantagePoint. www.theVantagePoint.com. Accessed 19 November 2016.

Verbitsky, M. (2004). Semantic TRIZ. http://www.triz-journal.com/archives/2004/. Accessed 5 January 2015.

Vidal-Espana, F., Leiva-Fernandez, F., Prados-Torres, J. D., Perea-Milla, E., Gallo-Garcia, C., Irastorza-Aldasoro, A., et al. (2007). Identification of new and emerging technologies. *Atencion Primaria, 39*(12), 641-646, doi:10.1157/13113954.

Wang, X., Qiu, P., Zhu, D., Mitkova, L., Lei, M., & Porter, A. L. (2015). Identification of technology development trends based on subject–action–object analysis: The case of dye-sensitized solar cells. *Technological Forecasting and Social Change, 98*, 24-46, doi:10.1016/j.techfore.2015.05.014.

Yang, C., Zhu, D., & Wang, X. (2017). SAO Semantic Information Identification for Text Mining. *International Journal of Computational Intelligence Systems, 10*(1), 593 - 604, doi:10.2991/ijcis.2017.10.1.40.

Yang, C., Zhu, D., & Zhang, G. Semantic-Based Technology Trend Analysis. In *2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 24-27 Nov. 2015 2015* (pp. 222-228). doi:10.1109/ISKE.2015.43.

Yoon, B., & Park, Y. (2005). A systematic approach for identifying technology opportunities: Keyword-based morphology analysis. *Technological Forecasting and Social Change, 72*(2), 145-160, doi:http://dx.doi.org/10.1016/j.techfore.2004.08.011.

Yoon, J., & Kim, K. (2012). Detecting signals of new technological opportunities using semantic patent analysis and outlier detection. *Scientometrics, 90*(2), 445-461, doi:10.1007/s11192-011-0543-2.

Zhang, Y., Zhou, X., Porter, A. L., Gomila, J. M. V., & Yan, A. (2014a). Triple Helix innovation in China’s dye-sensitized solar cell industry: hybrid methods with semantic TRIZ and technology roadmapping. *Scientometrics, 99*(1), 55-75, doi:10.1007/s11192-013-1090-9.

Zhang, Y., Zhou, X., Porter, A. L., & Vicente Gomila, J. M. (2014b). How to combine term clumping and technology roadmapping for newly emerging science & technology competitive intelligence: “problem & solution” pattern based semantic TRIZ tool and case study. *Scientometrics, 101*(2), 1375-1389, doi:10.1007/s11192-014-1262-2.

Zhu, D., & Porter, A. L. (2002). Automated extraction and visualization of information for technological intelligence and forecasting. *Technological Forecasting and Social Change, 69*(5), 495-506, doi:http://dx.doi.org/10.1016/S0040-1625(01)00157-3.