Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Chance Discovery with Data Crystallization for Scenario Formation in Solar Cell Technology

Tzu-Fu Chiu[1], Chao-Fu Hong[2], Ming-Yeu Wang[3], and Yu-Ting Chiu[4]
[1]Department of Industrial Management and Enterprise Information, Aletheia University, Taipei, Taiwan, R.O.C.
Email: chiu@email.au.edu.tw
[2]Department of Information Management, Aletheia University, Taipei, Taiwan, R.O.C.
Email: cfhong@email.au.edu.tw
[3] Department of Bio-industry and Agribusiness Administration, National Chiayi University, Chiayi, Taiwan, R.O.C.
Email: mywang@mail.ncyu.edu.tw
[4]Department of Information Management, National Central University, Taoyuan, Taiwan, R.O.C.
Email: gloria@mgt.ncu.edu.tw

***Abstract*—As the solar cell, one of renewable energies, is growing at a fast pace, the recognition of its technological situation becomes necessary for a company and stakeholders nowadays. Meanwhile, the patent data contains plentiful technological information from which is worth exploring to extract further knowledge. Therefore, a graph-based approach, chance discovery, is employed so as to analyze the patent data, to form the technological scenarios, and to explain the tendency of solar cell technology. Finally, several topics of solar cell technology have been identified, the directions of each topic have been depicted, and the relations between topics have been also observed.**

***Keywords*—scenario formation, chance discovery, data crystallization, solar cell, patent data.**

I. Introduction

It is essential for a company or stakeholders to realize the situation of a certain technology so that the company can review its development directions of that technology and the stakeholders can examine the suitability of their relevant investments. Within technological information, up to 80% of the disclosures in patents are never published in any other form [1]. Therefore, patent analysis has been recognized as an important task at the government and company levels. Through appropriate analysis, technological details and relations, business trends, novel industrial solutions, or making investment policy can be achieved [2]. Apart from those existing methods, a graph-based approach, Chance Discovery, will be employed in this area in order to visualize the technological scenarios of solar cell in U.S.

II. Related Work

This section will be divided into three subsections to introduce the related work. They are: scenario formation, solar cell, and chance discovery.

A. Scenario Formation

A scenario is a product that describes some possible future state and/or that tells the story about how such a state might come about [3]. The main categories of scenarios are: predictive (forecasts, what-if), explorative (external, strategic), and normative (preserving, transforming) [4]. The scenario approach (i.e., scenario analysis, scenario planning, or scenario building) can be comprised of five stages: scenario preparation, scenario-field analysis, scenario prognostics, scenario development, and scenario transfer [5], and had been applied in perceiving, framing minds, thinking of (top) managers, decision support, environmental study, and some other areas [6].

In this study, a new approach, chance discovery, will be applied in the scenario formation to visualize the overview of solar cell via patent data.

B. Solar Cell

A solar cell or photovoltaic (PV) cell is a device which converts sunlight into electricity by the photovoltaic effect [7]. Solar cell, a sort of green energy, is clean, renewable, sustainable, and good for protecting our environment. In recent years (2003-2006), total PV production grew in average by almost 50% worldwide, whereas the thin film segment grew by almost 80% (from a very low level) and reached 196 MW or 8% of total PV production in 2006 [8].

In order to understand the development situation of solar cell in U.S., this study utilizes chance discovery with data crystallization to analyze the patent data of year 2007 from USPTO (the United States Patent and Trademark Office) [9].

C. Chance Discovery

A chance means an event or a situation with significant impact on decision making. Chance discovery is to become aware of a chance and to explain its significance, especially if the chance is rare and its significance is unnoticed [10]. In addition, a chance can be also conceived either as an opportunity or as a risk; where desirable effects from an opportunity should be promoted, and undesirable effects from a risk should be prevented [11]. Chance discovery has been applied for designing product via patent data [12, 13].

1) KeyGraph: In chance discovery, KeyGraph is a tool for visualizing relations among data items, corresponding to events in the real world [14]. KeyGraph is generalized from a document-indexing method in order to extract essential events and the causal structures among them from an event sequence

[15]. A KeyGraph diagram mainly consists of clusters and hubs, where clusters are composed of co-occurring frequent items (words in a document, or events in a sequence) and hubs are items not so frequent as the ones in clusters but co-occurring with multiple clusters [15]. Among hubs, the ones with higher key values are chances. The algorithms for generating a KeyGraph and the formula for calculating key values can refer to [16].

2) Data crystallization: In order to detect the unobservable significant events, data crystallization aims at presenting the hidden structure among events, which inserts dummy items representing the potential existence of unobservable events to the given data set [14]. Unobservable events and their relations with other events are to be visualized by applying KeyGraph [12]. A generic data crystallization algorithm can be summarized as follows [17]: (1) event identification; (2) clustering; (3) dummy event insertion; (4) co-occurrence calculation; and (5) topology analysis. Identifying dummy events via data crystallization has been applied for designing new products of a surface inspection system (i.e., a machine for detecting defects on couple charged devices) [12].

III. Research Design for Scenario Formation

The research design for scenario formation is organized as in Fig. 1. It will be divided into eight phases and described in the following subsections. They are: data preprocessing, association diagram generation, user-guided association diagram generation, KeyGraph generation, data crystallization, diagram integration, pattern identification, and scenario formation.

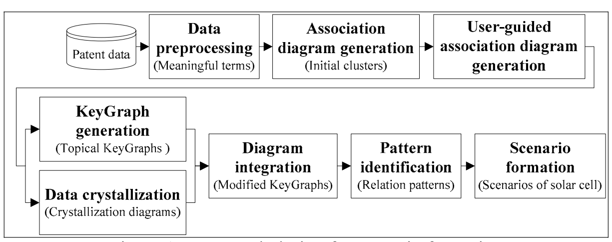

Figure 1. Research design for scenario formation

A. Data Prepocessing

In the first phase, the patent data of solar cell (during a certain period of time) will be downloaded from the USPTO. For considering an essential part to represent a complex patent data, the abstract field is selected as the object for this study. An English POS tagger (i.e., a Part-Of-Speech tagger for English) from the University of Tokyo [18] will be employed to perform word segmenting and labeling on the patent data (i.e., the abstract field). Additionally, a file of stop words and a file of synonyms will be built so as to facilitate the data clean-up processing. After data preprocessing, the meaningful terms will be obtained and then passed onto the following phases.

B. Association Diagram Generation

Second phase is designed to draw an association diagram using the term frequency and term association of the meaningful terms. Term frequency $F(n_i)$ is defined by Equation (1) using a Boolean function $B(s)$ for a proposition s in Equation (2), where b_j is the jth data subset of terms. Term association is defined by the Jaccard coefficient in Equation (3) as a measure of the co-occurrence [19].

$$F(n_i) \equiv \sum_{j=0}^{|b|-1} B(n_i \in b_j) \quad (1)$$

$$B(s) = \begin{cases} 1 & \text{if } s \text{ is true} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$J(n_i, n_j) \equiv \frac{F(n_i \cap n_j)}{F(n_i \cup n_j)} \quad (3)$$

Through the proper thresholds setting of frequency and association, a number of initial clusters will be generated in the association diagram. Furthermore, each initial cluster can be viewed as a technological topic according to the domain knowledge.

C. User-guided Association Diagram Generation

In order to explore the technological topic of every initial cluster in the association diagram, several key terms will be selected from each cluster (based on the user's knowledge) to represent that topic. Subsequently, the selected key terms will be used to sift out a data subset from the whole patent data set. Afterward, this sifted data subset will be utilized to draw an association diagram, that is a user-guided association diagram [20], for every initial cluster.

D. KeyGraph Generation

According to the user-guided association diagrams, a topical KeyGraph will be drawn for visualizing the focused clusters (i.e., subtopics) and chances within each technological topic. This topical KeyGraph will be then integrated with a crystallization diagram to construct a modified KeyGraph.

E. Data Crystallization

Similar to the above phase, this phase is designed to employ the data crystallization technique to generate dummy events via the user-guided association diagrams. Firstly, a dummy event will be added as an extra item into each patent of the data subset. For example, "DUMMY-102" is a dummy event for the 2nd patent of the 1st data subset. Secondly, according to each new data subset, the user-guided association diagram will be redrawn with dummy events included. Lastly, the associations (i.e., links) from each dummy event to other events will be recorded and then inserted into the previous user-guided association diagram respectively to form a crystallization diagram.

F. Diagram Integration

As the same focused clusters (i.e., subtopics) are existing both in the above topical KeyGraph and crystallization diagram of each topic, it becomes possible for combining these two graphs to build a new integrated diagram, namely "modified KeyGraph", so as to display the chances, dummy events, as well as focused clusters at the same time in the same graph.

G. Pattern Identification

According to the modified KeyGraphs, the focused clusters, chances, and dummy events will be applied to recognize the relation patterns by observing the relationships between clusters and chances, between clusters and dummy events, and between chances and dummy events. Firstly, the chance linked between clusters or among clusters will be marked and used to construct the "cluster A (i.e., technique A) relates to cluster B (i.e., technique B) via chance C" pattern. Secondly, the dummy event linked between clusters, or among clusters will be marked and used to construct the "cluster A relates to cluster B via dummy event D" pattern, where the meaning of dummy event D can refer to the title of a specific patent which includes that dummy event. Lastly, the links between or among chances and dummy events will be marked and used to construct the "group X connects with group Y using chance C or dummy event D" pattern.

H. Scenario Formation

According to the relation patterns from each modified KeyGraph and based on the domain knowledge, a sub-scenario of solar cell will be formed so that a certain aspect of the technology can be depicted successively. Consequently, the overall scenario of solar cell can be formed by integrating the sub-scenarios of various aspects together. The sub-scenarios and overall scenario can demonstrate the situation of technology, features of subtopic, directions of topic, and tendencies of connected topics.

IV. Experiment and Explanation

The experiment has been performed according to the research design. The experimental results would be illustrated in the following seven subsections: result of data preprocessing, result of association diagram generation, result of user-guided association diagram generation, result of KeyGraph generation and data crystallization, result of diagram integration, result of pattern identification, and result of scenario formation.

A. Result of Data Prepocessing

As the aim of this study is to visualize the scenarios of solar cell, the patent documents of solar cell were the target data for this experiment. Mainly, the abstract field of patent documents was used in this study. Therefore, 81 patent items of year 2007 were collected from USPTO, using key words: "solar cell, or photovoltaic cell, or PV cell" on "title field or abstract field". The POS tagger was then triggered to do the data preprocessing upon the collected 81 patent items. Consequently, the patent data of year 2007 was cleaned up readily for the further processing phases.

B. Result of Association Diagram Generation

Using the meaningful terms from data preprocessing, the association diagram was created as in Fig. 2. In the diagram, twenty initial clusters were found while the number of comprising nodes of a cluster was set to no less than four. According to the domain knowledge, these clusters were named as: metal-oxide, nanocrystal, telemetering-system, LCD (i.e., liquid crystal display), dye-sensitized, active-semiconductor-structure, LED (i.e., light emitting diode), monitoring-terminal, burette, TCO (i.e., transparent conductive oxide), collection-region, thin-panel-enclosure, adjacent-row, thin-film-1, thin-film-2, intruder-detector, hammock, lantern, light-intensity, and spectrum from cluster ① to ⑳ respectively.

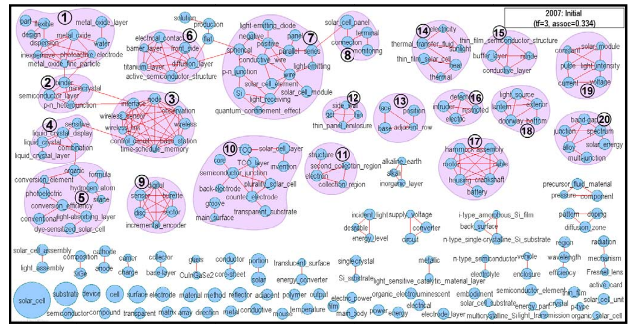

Figure 2. Association diagram with initial clusters

Within the above 20 initial clusters, four technological topics and one application topic had been identified based on the comparison among clusters and the domain knowledge. These topics were: cluster thin-film (including thin-film-1 ⑭ and thin-film-2 ⑮), dye-sensitized ⑤, LED ⑦, TCO ⑩, as well as cluster application (including telemetering-system ③, monitoring-terminal ⑧, burette ⑨, intruder detector ⑯, hammock ⑰, and lantern ⑱). These topics were then used to generate the user-guided association diagrams.

C. Result of User-guided Association Diagram Generation

Before drawing a user-guided association diagram of a specific topic: thin-film as an example, five key terms which are thin-film-solar-cell, thin-film-semiconductor-structure, thermal-transfer-fluid, conductive-layer, and nitride, were selected for filtering out the irrelevant data, and the remaining data (i.e., 13 patents) were used to generate the user-guided association diagram (in Fig. 3). After the data-subset-sifted execution, there were 13 patents for thin-film, 11 patents for dye-sensitized, 8 patents for LED, 8 patents for TCO, and 8 patents for application topic. The other user-guided association diagrams for dye-sensitized, LED, TCO, and application topics were generated successively.

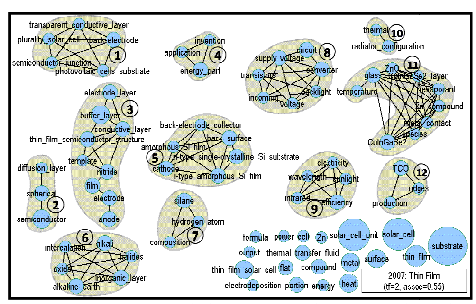

Figure 3. A user-guided association diagrem of thin-film

D. Result of KeyGraph Generation and Data Crystallization

Based on the data of a user-guided association diagram, a KeyGraph of each topic (e.g., thin-film) was drawn so as to illustrate the containing focused clusters (12 subtopics) and chances (9 chances) within the topic. Five KeyGraphs for thin-film, dye-sensitized, LED, TCO, and application topics were generated successively.

Subsequently, also based on the data of a user-guided association diagram and by adding dummy events to the sifted data subset, a crystallization diagram of each topic (e.g., thin-film) was drawn so as to demonstrate the containing focused clusters and dummy events (13 dummy events, from DUMMY-101 to DUMMY-113) within the topic. Five crystallization diagrams for thin-film, dye-sensitized, LED, TCO, and application topics were generated successively. Both the KeyGraph and crystallization diagram were prepared for the following diagram integration process.

E. Result of Diagram Integration

By integrating the KeyGraph and crystallization diagram, a modified KeyGraph of thin-film was generated as an example in Fig. 4. The other modified KeyGraphs for dye-sensitized, LED, TCO, and application topics would be produced successively and used for conducting the pattern identification process.

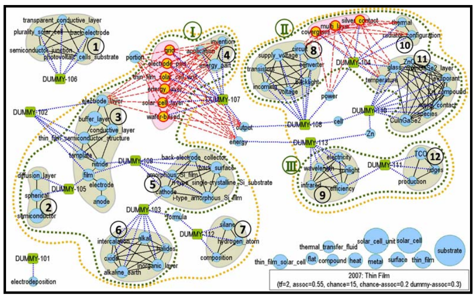

Figure 4. A modified KeyGraph of thin-film

F. Result of Pattern Identification

Based on the modified KeyGraph of each topic, the links radiating from chances or dummy events to focused clusters (i.e., subtopics) were used to recognize the relation patterns between or among clusters for each topic. The executions of relation pattern identification were completed and explained in the following subsections.

1) Relation patterns of thin-film

The modified KeyGraph of thin-film (i.e., Fig. 4), formed by 12 subtopics, 9 chances, and 13 dummy events, was applied for finding out the relation patterns for thin-film. The outcome of relation patterns was summarized in Table I, where subtopic-1 related to subtopic-2 (or 3) via chances (i.e., Cs) or dummy events (i.e., DEs).

TABLE I. RELATION PATTERNS OF THIN-FILM TOPIC

no.	*subtopic-1*	*via*	*subtopic-2 or 3*
1	nitride-thin-film ③	thin-film-energy-layer (8 Cs)	energy-part ④
2	backlight-converter ⑧	silver-contact-multi-layer (3 Cs)	thermal-radiator ⑩, Zn-compound ⑪
3	diffusion-layer ②	diffusion-layer-pn-junction (DE-105)	nitride-thin-film ③
4	nitride-thin-film ③	i-type-amorphous-Si-film (DE-109)	amorphous-Si-film ⑤
5	amorphous-Si-film ⑤	opto-electronic-device (DE-103)	alkali-halides ⑥
6	thermal-radiator ⑩	thin-film- thermal-radiator (DE-104)	Zn-compound ⑪
7	infrared-wavelength ⑨	transparent-conductive-oxide-film (DE-111)	TCO-ridge ⑫

2) Relation patterns of dye-sensitized

Similarly, the modified KeyGraph of dye-sensitized (in Fig. 5), formed by 10 subtopics, 9 chances, and 11 dummy events, was applied for finding out the relation patterns.

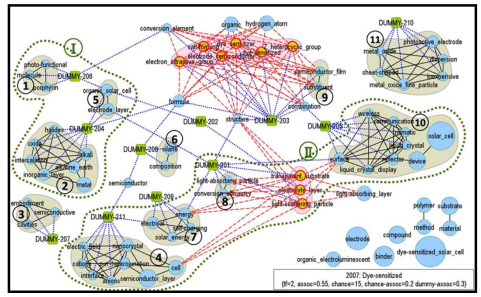

Figure 5. A modified KeyGraph of dye-sensitized

The outcome of relation patterns of dye-sensitized was summarized below (in Table II).

TABLE II. RELATION PATTERNS OF DYE-SENSITIZED

no.	*subtopic-1*	*via*	*subtopic-2 or 3*
1	nanocrystal ④, self-charging ⑦	light-scattering (3 Cs)	light-absorbing ⑧, nematic-LCD ⑩
2	porphyrin-molecule ①	porphyrin-polymer (DE-206)	organic-solar-cell ⑤
3	alkali-halides ②	opto-electronic-device (DE-204)	organic-solar-cell ⑤
4	nanocrystal ④	i-type-amorphous-Si-film (DE-211)	self-charging ⑦
5	organic-solar-cell ⑤	bi-layer-PV-cell (DE-201)	self-charging ⑦, light-absorbing ⑧, nematic-LCD ⑩

3) Relation patterns of LED

Likewise, the modified KeyGraph of LED (in Fig. 6), formed by 9 subtopics, 10 chances, and 8 dummy events, was applied for finding out the relation patterns.

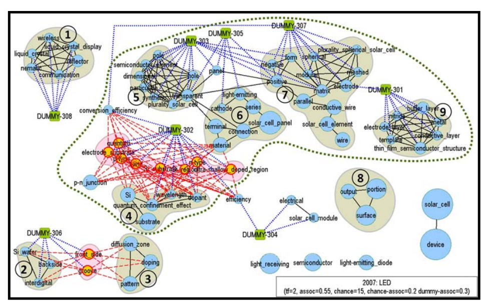

Figure 6. A modified KeyGraph of LED

The outcome of relation patterns of LED was summarized below (in Table III).

TABLE III. RELATION PATTERNS OF LED

no.	*subtopic-1*	*via*	*subtopic-2 or 3*
1	quantum-confinement ④	Si-substrate (8 Cs)	light-emitting ⑥
2	interdigital-semiconductor ②	groove (2 Cs)	diffusion-zone ③
3	transparent-synthetic ⑤	light-emitting-device (DE-305)	light-emitting ⑥, plurality-spherical ⑦
4	plurality-spherical	conductive-layer	nitride-thin-film ⑨

no.	subtopic-1	via	subtopic-2 or 3
	⑦	(DE-301)	

4) Relation patterns of TCO

Continuously, the modified KeyGraph of TCO (in Fig. 7), formed by 11 subtopics, 8 chances, and 8 dummy events, was applied for finding out the relation patterns.

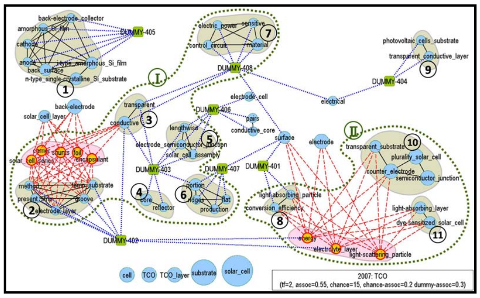

Figure 7. A modified KeyGraph of TCO

The outcome of relation patterns of TCO was summarized below (in Table IV).

TABLE IV. RELATION PATTERNS OF TCO

no.	subtopic-1	via	subtopic-2 or 3
1	temporary-substrate ②	encapsulant-shunt (5 Cs)	transparent-conductive ③
2	light-absorbing-particle ⑧	electrolyte-layer (3 Cs)	transparent-substrate ⑩, light-absorbing-layer ⑪
3	transparent-conductive ③, reflector ④	elongated-solar-cell (DE-403)	semiconductor-junction ⑤, flat-ridge ⑥
4	transparent-conductive ③	electro-optical-cell (DE-408)	sensitive-material ⑦

5) Relation patterns of application

Repeatedly, the modified KeyGraph of application (in Fig. 8), formed by 7 subtopics, 12 chances, and 8 dummy events, was applied for finding out the relation patterns.

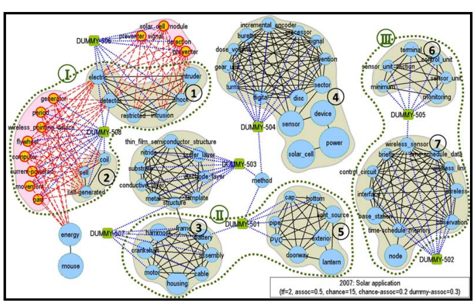

Figure 8. A modified KeyGraph of application

The outcome of relation patterns of application was summarized below (in Table V).

TABLE V. RELATION PATTERNS OF APPLICATION TOPIC

no.	subtopic-1	via	subtopic-2 or 3
1	intruder-detedtor ①	wireless-pointing (8 Cs), computer-mouse (DE-508)	computer-mouse ②
2	hammock ③	lantern (DE-501)	lantern ⑤
3	monitoring-terminal ⑥	monitoring-terminal (DE-505)	telemetering-system ⑦

G. Result of Scenario formation

According to the relation patterns derived from the above pattern identification process, the individual sub-scenario of each topic and an overall scenario of all topics could be formed to depict the situation and directions of solar cell technology upon the patent data from USPTO in 2007.

1) Sub-scenario of thin-film

According to Table I and based on domain knowledge, the relation patterns were applied to form a sub-scenario of thin-film for the solar cell technology. Basically, nitride-thin-film might combine with energy-part via thin-film-energy-layer (i.e., no. 1); backlight-converter might combine with thermal-radiator and Zn-compound via silver-contact-multi-layer (no. 2); and so on. Furthermore, the individual relation patterns could be interrelated to form an interrelation pattern, for example (Group I): nitride-thin-film might combine with energy-part, amorphous-Si-film, diffusion-layer, and alkali-halides via energy-layer, pn-junction, and anti-reflection-film (no. 1, 3, 4, 5), which seemed to indicate an amorphous-thin-film direction. Additionally, Group I could connect to subtopic "silane-hydrogen ⑦" via node "formula" to form a slightly larger group; Group II (no. 2, 6) could connect to Group III (no. 7) via node "Zn" to form a new larger group, which seemed to point out a Zn-compound-thermal-radiator direction.

2) Sub-scenario of dye-sensitized

According to Table II, the sub-scenario of dye-sensitized could be described as follows. Nanocrystal, self-charging, light-absorbing, and nematic-LCD might be combined together to form Group II via light-scattering (no. 1), which seemed to indicate a nanocrystal-light-absorbing direction; porphyrin-molecule, organic-solar-cell, and alkali-halides might be combined together to form Group I via porphyrin-polymer and opto-electronic-device (no. 2, 3), which seemed to denote a porphyrin-organic-solar-cell direction. Moreover, Group I could connect to Group II via bi-layer-PV-cell (DE-201) to form a new larger group.

3) Sub-scenario of LED

According to Table III, the sub-scenario of LED could be described as follows. Basically, quantum-confinement might combine with light-emitting via Si-substrate (no. 1); interdigital-semiconductor might combine with diffusion-zone via groove (no. 2); and so on. Furthermore, quantum-confinement, light-emitting, transparent-synthetic, plurality-spherical, and nitride-thin-film might be combined together (no. 1, 3, 4) to form a group, which seemed to imply a quantum-confinement & light-emitting-device direction.

4) Sub-scenario of TCO

According to Table IV, the sub-scenario of TCO could be described as follows. Temporary-substrate, transparent-conductive, sensitive-material, reflector, semiconductor-junction, and flat-ridge might be combined together via encapsulant-shunt, elongated-solar-cell, and electro-optical-cell (no. 1, 3, 4) to form Group I, which seemed to indicate a temporary-substrate & sensitive-material direction; light-absorbing-particle might combine with transparent-substrate and light-absorbing-layer via electrolyte-layer (no. 2) to form

Group II, which seemed to denote a light-absorbing & transparent-substrate direction. Moreover, Group I could connect to Group II via temporary-substrate (DE-402) to form a new larger group.

5) Sub-scenario of application

According to Table V, the sub-scenario of application could be described as follows. Intruder-detector and computer-mouse might be combined together to form Group I via wireless-pointing (no. 1), which seemed to indicate a detector-type product; hammock and lantern might be combined together to form Group II via lantern (no. 2) which seemed to denote a home-appliance-type product; monitoring-terminal and telemetering-system might be combined together to form Group III via monitoring-terminal (no. 3) which seemed to imply a system-type product. Additionally, Group II might connect to subtopic "burette ④" via node "method".

6) Overall scenario of solar cell

By summarizing the above five sub-scenarios, the overall scenario of solar cell could be depicted as follows. The popular topics in 2007 were thin-film, dye-sensitized, LED, TCO, and application. Among topics, thin-film might connect to dye-sensitized via opto-electronic-device (a common dummy event, i.e., DE-103 and DE-204), which seemed to indicate a optoelectronics tendency (in thin-film and dye-sensitized); thin-film might connect to LED via nitride-thin-film (a common subtopic, i.e., ③ in Fig. 4 and ⑨ in Fig. 6), which seemed to denote a nitride-thin-film-LED tendency (in thin-film and LED); dye-sensitized might connect to TCO via light-absorbing (a common subtopic, i.e., ⑧ in Fig. 5 and ⑧ ⑪ in Fig. 7), which seemed to imply a light-absorbing tendency (in dye-sensitized and TCO).

V. Conclusion

According to the experiment, the technological topics of solar cell for year 2007 via patent data have been found: thin-film, dye-sensitized, LED, TCO, and application. Each topic was examined to explore the relation patterns via chances and dummy events.

Based on the identified patterns, every topic was explored to find out the potential development directions. For examples, topic thin-film seemed to indicate an amorphous-thin-film and a Zn-compound-thermal-radiator directions; topic dye-sensitized seemed to denote a nanocrystal-light-absorbing and a porphyrin-organic-solar-cell directions. Furthermore, using the summarized information, the overall trend of solar cell technology might be a optoelectronics tendency (in thin-film and dye-sensitized), a nitride-thin-film-LED tendency (in thin-film and LED), and a light-absorbing tendency (in dye-sensitized and TCO). Consequently, the scenario formation of solar cell has been attained via chance discovery using the patent data of USPTO in 2007.

In the future research, the relation patterns of patent data can be analyzed by some other methods such as association rules, social network analysis, or social computing so as to insure the validity of the experiment result. In addition, the data source can be expanded from USPTO to some other kinds like Japan, EPO, or WIPO in order to observe the technological situation globally.

Acknowledgment

This research was supported by the National Science Council of the Republic of China under the Grants NSC 97-2410-H-415-043.

References

[1] M. Blackman, "Provision of patent information: a national patent office perspective", World Patent Information, vol. 17 (2), 1995, pp. 115-123.

[2] Y. Tseng, C. Lin, and Y. Lin, "Text mining techniques for patent analysis", Information Processing and Management, no. 43, pp. 1216-1247, 2007.

[3] P. Bishop, A. Hines, and T. Collins, "The current state of scenario development an overview of techniques", Foresight, vol. 9, no. 1, pp. 5-25, 2007.

[4] L. Borjeson, M. Hojer, K. Dreborg, T. Ekvall, and G. Finnveden, "Scenario types and techniques: towards a user's guide", Future, no. 38, pp. 723-739, 2006

[5] J. Gausemeier, A. Fink, and O. Schlake, "Scenario management: an approach to develop future potentials", Technological Forecasting and Social Change, no. 59, pp. 111-130, 1998.

[6] T. Postma and F. Liebl "How to improve scenario analysis as a strategic management tool", Technological Forecasting & Social Change, no. 72, pp. 161-173, 2005.

[7] Solar Cell. http://en.wikipedia.org/wiki/Solar_cell [2009/03/12]

[8] A. Jager-Waldau, "PV status report 2007: research, solar cell production and market implementation of photovoltaics", JRC Technical Notes, 2007.

[9] USPTO: the United States Patent and Trademark Office. http://www.uspto.gov/ [2008/10/10]

[10] Y. Ohsawa and H. Fukuda, "Chance discovery by stimulated groups of people: application to understanding consumption of rare food", Journal of Contingencies and Crisis Management, vol. 10, no. 3, pp. 129-138, 2002.

[11] Y. Maeno and Y. Ohsawa, "Human-computer interactive annealing for discovering invisible dark events", IEEE Transactions on Industrial Electronics, vol. 54, no. 2, pp. 1184-1192, April 2007.

[12] K. Horie, Y. Maeno, and Y. Ohsawa, "Data crystallizatio applied for designing new products", Journal of Systems Science and Systems Engineering, vol. 16, no. 1, pp. 34-49, March 2007.

[13] K. Horie, Y. Maeno, and Y. Ohsawa, "Human-interactive annealing process with pictogram for extracting new scenarios for patent technology", Data Science Journal, vol. 6, supplement, pp. s132-s136, March 2007.

[14] Y. Ohsawa, "Data Crystallization: Chance Discovery Extended for Dealing with Unobservable Events", New Mathematics and Natural Computation, vol. 1, no. 3, pp. 373-392, 2005.

[15] Y. Ohsawa, "KeyGraph: visualized structure among event clusters", in: Y. Ohsawa and P. McBurney (eds.) Chance Discovery, pp. 262-275, Springer-Verlag, 2003.

[16] Y. Ohsawa, N. E. Benson, and M. Yachida, "KeyGraph: automatic indexing by cooccurrence graph based on building construction metaphor", Proceedings of the Advanced Digital Library Conference (IEEE ADL 1998), pp. 12-18, 1998.

[17] Y. Maeno and Y. Ohsawa, "Stable deterministic crystallization for discovering hidden hubs", Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, vol. 2, pp. 1393-1398, 2006.

[18] An English POS tagger. http://www-tsujii.is.s.u-tokyo.ac.jp/~tsuruoka/postagger/ [2008/07/28]

[19] Y. Maeno and Y. Ohsawa, "Predicting relevant empty spots in social interaction", Journal of Systems Science & Complexity, vol. 21, no. 2, pp. 161-171, 2008.

[20] A. Vailaya, P. Bluvas, R.t Kincaid, A. Kuchinsky, M. Creech, and A. Adler, "An architecture for biological information extraction and representation", Bioinformatics, vol. 21, no. 4, pp. 430-438, 2005.