Experiments on Quantitative Technology Trends Analysis from the Internet News Resources

Shin-ichi Kobayashi[*], Yasuyuki Shirai[**], Kazuo Hiyane[**],
Fumihiro Kumeno[**], Hiroshi Inujima[*] and Noriyoshi Yamauchi[*]

[*]*kobayash@suou.waseda.jp, {inujima-hiroshi, yamauchi-noriyoshi}@waseda.jp*
Graduate School of Information, Production and Systems, Waseda University,
Kitakyushu, Fukuoka 808-0135, Japan

[**]*{shirai, hiya, kumeno}@mri.co.jp*
Mitsubishi Research Institute, Inc., Tokyo, 100-8141, Japan

Abstract
In business or technology planning, it is strongly required to grasp the overall technology trends and predict what will happen in the near future. A deluge of technology information in recent years, however, makes it quite difficult even by domain experts. We have developed the environment where we can detect and analyze the technology trends from the Internet news resources. In our system, we define several quantitative indicators, such as importance and influence of each keyword in the target field, and relativity and superiority of relationships between keywords. According to these indicators, our system detects key concepts around the target field and provides an environment to analyze the technology trends of the target field. In this article, we report some experimental results using our system based on the content from IT news sites, shows the effectiveness or validity of our method compared with human experts' evaluations, and discuss the feasibility to grasp technology trends from the Internet resources automatically in the future.

Introduction
The Internet is a valuable source of information. In the future, the success or failure of research and development efforts will depend on being able to quickly gather, analyze, and apply the enormous and ever-increasing quantity of information on the Internet. The current Internet search engines, however, cannot be used for systematic information analysis or for recognizing time series changes. We have developed the environment where we can extract the key concepts for the target field and calculate some indicators to analyze the technology trends in time series, based on the co-occurrence analysis from the Internet news resources. In this paper we briefly introduce our method and report some experimental results for the trend detection in the broadband field and the mobile communication field in Japan, based on the content of a Japanese IT news site, ITmedia, and compare with human experts' evaluations.

Related work
In recent years, *web mining* has become popular (Chakrabarti(2003)). Many attempts have been made to use the Internet content and academic papers to organize and present the relationships between technical concepts (Tamura(1992), Aizawa(2000), Watanabe(2001), Boyack(2001,2002), Hassan(2002), Maedche(2002), Morris (2003), Chen (2003), etc.). Among them some efforts, called Emerging Trend Detection (Kontostathis (2004)), have involved time series analysis, which can reveal past and present trends and imminent changes.

Compared with these researches, our method has the following characteristics:

- Leading indicators of changes over time
  News articles on the Internet are newsworthy only at that point in time. However, these articles may also contain information about imminent future trends. In order to grasp the current situation and near future trends of the specified technology field, we provide some new indicators, that is, keyword influence in that field and keyword superiority as well as widely used indicators, that is, keyword importance and keyword relativity.
- Aspect analysis
  In order to make a detailed analysis of technical fields, we sometimes need to focus on a specific aspect of the field (ex. aspect of organizations, technologies, protocols, business

Proceedings of ISSI 2005, pages 544–555.
Edited by P. Ingwersen and B. Larsen.

services, personnel, etc). Our method semi-automatically extracts keywords from patterns that appear frequently in the resources, recognizes the categories to which these keywords belong, and provides an editing and analysis environment in which we can analyze a specified aspect of the field by focussing on the specified categories of keywords.

Technical trend maps and trend indicators

In this section, we briefly introduce the functions of our system whose architecture is shown in Figure 1.

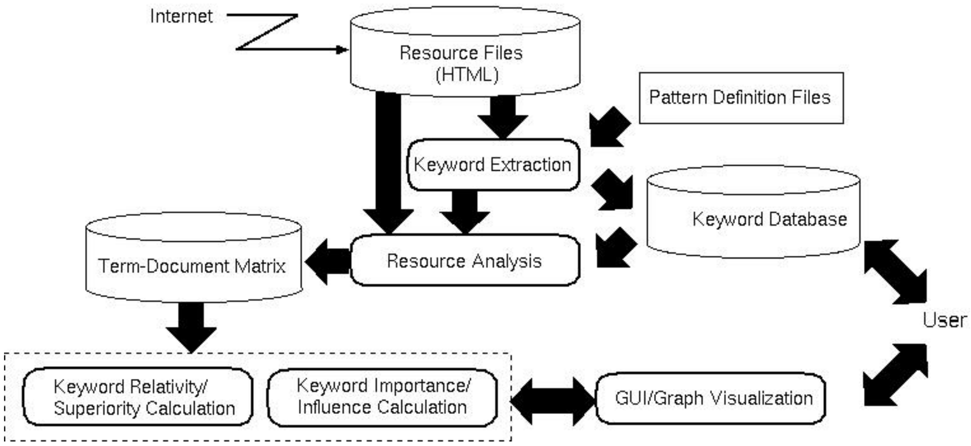

Figure 1 The Architecture of the System

Keyword extraction

This module extracts keywords from resources using pattern matching, including regular expressions, and stores the keywords together with their categories in a keyword database. The patterns depend on the resource type and are defined by the user. In addition, in order to ensure that the appropriate keywords are properly extracted, the dictionaries of stop words and synonyms as well as the method for extracting nouns, etc., must be fine-tuned to match the characteristics of the resource. This module extracts keywords for nine different categories: organization, technical term, system, protocol, business service, person, hardware, social phenomenon/trend, and business term. By considering indicators such as the degree of influence and of importance (discussed later) in terms of individual categories, it is possible to recognize the target field with global perspective (for example, organizationally driven, technologically driven, standards-centred, etc.).

Resource analysis

This module determines whether or not the keywords in the dictionary appear in each resource and generates a term-document matrix. The resources are HTML files obtained from the Internet. Text data are extracted in the appropriate units using the HTML tags.

If we consider *n* documents with *m* keywords, the size of the matrix is $m \times n$. Each element f_{ij} is 1 if keyword *i* is included in document *j* and 0 if not.

$$M = \begin{pmatrix} f_{11} & \cdots & f_{1j} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ f_{i1} & & f_{ij} & & f_{jn} \\ \vdots & & \vdots & \ddots & \vdots \\ f_{m1} & \cdots & f_{mj} & \cdots & f_{mn} \end{pmatrix}$$

In addition, $t_i = (f_{i1}, \ldots, f_{ij}, \ldots, f_{in})$ is the term vector relating to keyword *i*, and $d_j = (f_{1j}, \ldots, f_{ij}, \ldots, f_{mj})$ is the document vector relating to document *j*.

Keyword relativity and superiority

The relativity and superiority are defined between two keywords in a group of specified resources. The relativity *R(i,j)* of keywords *i* and *j* is defined by the similarity of the group of documents in which they appear. *R(i,j)* is defined by the well-known metric, that is, the cosine of the term vectors t_i and t_j, as shown below.

$$R(i,j) = \frac{t_i \bullet t_j}{\|t_i\| \bullet \|t_j\|}$$

The objective of superiority is to extract a hierarchical relationship that is similar to an "is-a" relationship or a "part-of" relationship between two keywords in the group of defined resources. In order to define superiority, the complementary concept i^c for keyword i^c is defined. The complementary concept i^c is a hypothetical concept that bears no similarity (no relationship) to *i*, and is characterized using the following term vector:

$$t_{i^c} = (1 - f_{i1}, \ldots, 1 - f_{ij}, \ldots, 1 - f_{in})$$

For two keywords *i* and *j*, assume *i* is a superior concept to *j*, then in accordance with the relativity between *i* and the complementary concept of *j* would be higher than the relativity between *j* and the complementary concept of *i*. Accordingly, we define the superiority *P(i,j)* of *i* with respect to *j* as follows:

$$P(i,j) = \arccos(R(i,j^c) - \arccos(R(i^c,j))$$

where $-\pi/2 \le P(i,j) \le \pi/2$. If $P(i,j) < 0$, *i* is called to be superior to *j*.

Keyword importance and influence

The importance and influence indicators show the status or positioning of keywords within a specified resource.

Importance is an indicator of relative importance within a specified resource and uses the well-known indicators *tf* and *idf*. The importance indicator $W_R(i)$ of keyword *i* in a resource group *R* (with *n* number of documents) is defined as follows:

$$W_R(i) = tf_R(i) \times \left(\log \frac{n}{df_R(i)} + 1 \right)$$

where $tf_R(i)$ is a total number of appearances of *i* in *R* and $df_R(i)$ is a number of documents in *R* that include *i*.

In order to describe the influence indicator, suppose a keyword group *A* has been defined for a resource group *R*. The influence $I_A(i)$ of keyword $i \in A$ with respect to *A* indicates the degree of influence of *i* for other keywords in *A*. The evaluation considers the degree to which the appearance of *i* in the documents reduces the entropy with respect to the other keywords in *A*, namely (*A-{i}*).

First, the entropy relating to appearance of a keyword in a certain keyword group *S* is indicated by *H(S)* as follows:

$$H(S) = -p(S)\log p(S) - p(\overline{S})\log p(\overline{S})$$

where $p(S)$ ($p(\overline{S})$) is a probability of any of the keywords in *S* being (not being) included in the document. If the total number of documents is *n* and the number of documents containing any keyword in *S* is *m*, then $p(S)=m/n$ and $p(\overline{S}) = (n-m)/n$.

In addition, $H(S|i)$ is the entropy relating to the appearance of *S* under the condition that the appearance of keyword *i* is known. $H(S|i)$ is defined as follows:

$$H(S|i) = -\sum_{x=S,\overline{S}} \sum_{y=i,\overline{i}} p(x,y)\log p(x|y)$$

Where *p(x,y)* is a joint probability of *x* and *y*, and *p(x|y)* is a conditional probability of *x* based on *y*. In this situation, the influence $I_A(i)$ of $i \in A$ with respect to a keyword group *A* is defined as follows:

$$I_A(A) = \frac{H(A-\{i\}) - H(A-\{i\}|i)}{H(A-\{i\})}$$

$I_A(i)$ indicates the proportion of reduction of the entropy for *A-{i}* as a result of knowing *i*. The greater $I_A(i)$ is, the greater the influence of *i* on the appearance of the other keywords of *A*. We consider that a highly influential keyword could induce structural changes in the target field after a certain time. In other words, influence is one of the leading indicators of time change. The nature of the change cannot be determined uniformly, but typically, there is an increase or decrease in the importance of the highly influential keyword itself or in a different keyword that has a high degree of relativity or superiority with the highly influential keyword.

GUI and graph visualization
A graphical user interface provides an interactive interface where we can specify resources by periods, sites, categories, to focus on the target field, find the relative keywords of the user-specified keyword, find the important keywords for the user-specified resources, and calculate the indicators in a displayed graph.

Trend analysis as time series
By continuously calculating these indicators for the specified set of keywords (that are selected by their importance and can be considered as representative keywords that characterize the field) in a time series, it is possible to obtain data on past trends and future developments in that field. The significance of each indicator, as determined by its definition and a time series analysis, is described below.

- Indicators for time series analysis
  We can consider the average values for importance, influence, relativity and superiority of keywords in the target field or each aspect of the field. The average importance means newsworthiness of the field or each aspect as a whole. Increase of the value signifies development/growth of that field or aspect, while decrease of the value signifies stagnation/contraction. The average influence is an indicator of the potential for change. Increase of the value signifies an increase in the signs of change in that field or each aspect. The average relativity is an indicator of the strength of the connection. Increase of the value signifies an increase in the interrelationships in the fields or each aspect of the field. The average superiority is an indicator of structural configuration. Increase of the value signifies a higher hierarchical position for that field or each aspect of the field.
- Indicators for time series analysis of individual keywords
  By investigating the changes of importance and influence for the specified keyword, we can find the changes in time for that keyword. For example, the changes of importance are an indicator that represents growth for that keyword, and the changes of influence are an indicator of the potential for change.

Experiments and evaluation

This section describes experiments that used actual Internet resources in order to verify the effectiveness of the indicators described above. The data used for the experiments comprised ITmedia[1] articles that appeared between 2001 and 2003.

To identify time series trends in the field and each aspect of the field, significant or representative keywords are extracted according to their importance values. We consider that the trends of the field can be revealed by analyzing the trends of the group of keywords. Specifically, the

[1] One of the most famous IT news site in Japan (http://www.itmedia.co.jp).

time series trends for the average values for each indicator for the designated group of important keywords are used to characterize trends in that field.

Test cases and analysis

The ITmedia articles were used to conduct experiments for two cases shown in Table 1, targeting the sub-directories for broadband and mobile communications. These fields were considered appropriate for the experiments because of their technical progress and fast growth in recent years. In each of the broadband and the mobile fields, we have selected 17 and 19 keywords respectively according to their importance, that can be considered as the representative keywords for each field. Figure 2 and Figure 3 show the overall indicators obtained from Case 1 and Case 2 for the broadband and mobile communications fields.

Table 1 Cases of Experiments

Experiments	No.	Period	Fileld
Case 1	1	2001/07-2001/12	Broadband news category of Itmedia
	2	2001/09-2002/02	
	3	2001/11-2002/04	
	4	2002/01-2002/06	
	5	2002/03-2002/08	
	6	2002/05-2002/10	
	7	2002/07-2002/12	
	8	2002/09-2003/02	
	9	2002/11-2003/04	
	10	2003/01-2003/06	
	11	2003/03-2003/08	
	12	2003/05-2003/10	
	13	2003/07-2003/12	
Case 2	3-13	From Period 3 to Period 13 of the Case1	Mobile news category of Itmedia
Case 3	1-13	From Period 1 to 13	Broadband news category of Itmedia (focused on organizations)

Overall trends in the broadband field

The average influence showed high peaks in periods 3 and 4. Many keywords with a high degree of influence appeared during these periods. As if in response, the average importance increased through period 6 and subsequently stabilized. Based on the meanings for the indicators shown in the previous section, this can be interpreted as the high influence of this trend, which first appeared in periods 3 and 4 and became a leading indicator of the broadband "boom" that followed. In actuality, Ministry of Internal Affairs and Communications in Japan (MIC) released its "National Broadband Concept" in October of 2001. At about the same time, the media began reporting widely on the potential for new platforms and services such as WiFi hot spots, IPv6, MPEG-4 video transmission, etc. Because these new possibilities subsequently increased the newsworthiness of the broadband field, the trends in average influence and average importance obtained here can be interpreted as being almost identical to those anticipated in the previous section.

Moreover, the fact that the average superiority gradually increased indicates that some kind of hierarchic structure had developed between the important keywords. In other words, this suggests that the situation has changed from one where new keywords were used just for their trendiness to one where they formed a structured mode for providing services and technologies.

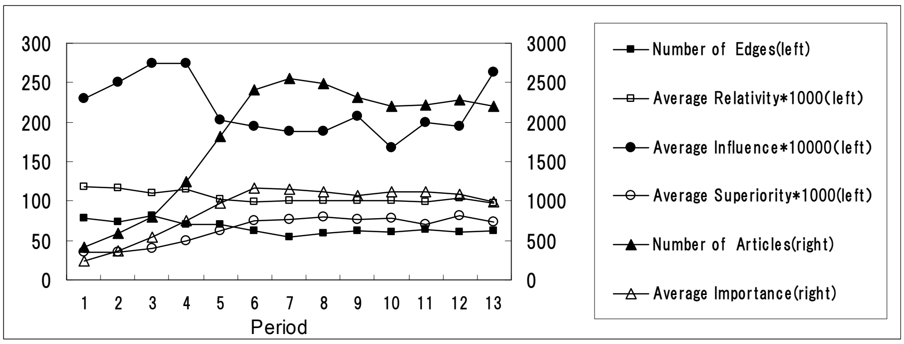

Figure 2 Overall Trends of the Broadband Field

Overall trends in the mobile communications field

The average influence in this field was lower than the average influence in the broadband field and generally decreased throughout all periods. This is a leading indicator of no major changes in average importance. In addition, the average superiority value decreased after period 10, confirming that newsworthy topics in that field tend to dissipate.

In actuality, the newsworthiness of the mobile communications field has tended overall to decrease somewhat as the PDA market shrinks, etc. Moreover, while there are no newsworthy topics galvanizing the industry as a whole, new mobile phone services and terminals continue to appear, so the topics are becoming decentralized. The trends seen in Figure 3 generally seem to agree with the situation of that time.

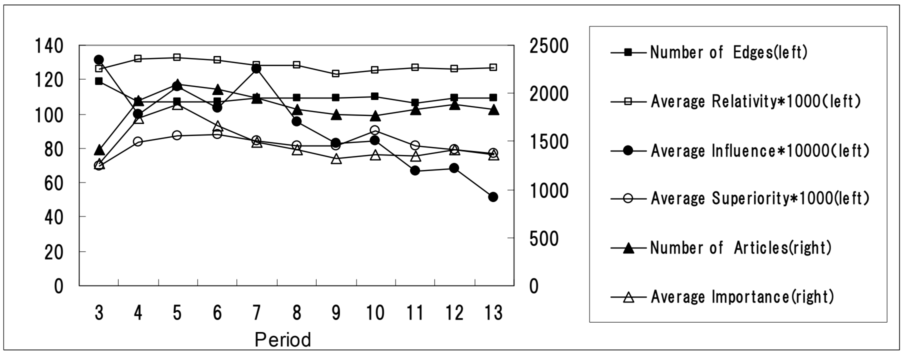

Figure 3 Overall Trends of the Mobile Field

Detailed analysis of the broadband field

Figure 4 and Figure 5 show the trends in importance and influence, respectively, for the keywords obtained for Case 1 in the broadband field.

Noteworthy in the broadband field is the conspicuous growth in the importance of wireless LAN. Up until period 7, this value increased sharply. Wireless LAN was followed in importance by ADSL, mobile phones, IP telephoning and other keywords, all of which showed steady increases in importance, which matched the situation at the time.

The rapid growth of wireless LAN can be predicted from the influence value. Both the absolute value and the increase in the influence of IEEE (IEEE 802.11i) and WiFi hot spots increased rapidly up through period 4, following which the importance of wireless LAN increased.

The rapid growth of wireless LAN can be predicted from the influence value. Both the absolute value and the increase in the influence of IEEE (IEEE 802.11i) and WiFi hot spots increased rapidly up through period 4, following which the importance of wireless LAN increased.

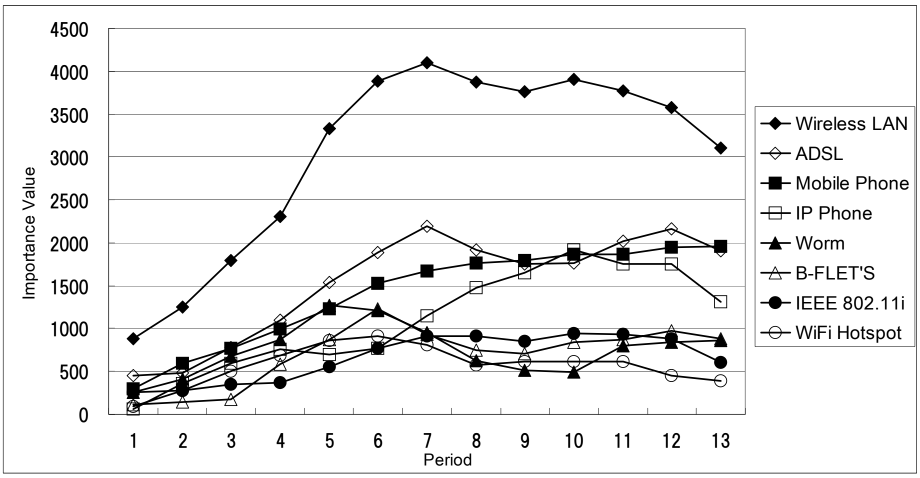

Figure 4 Trends of Importance of the Broadband Field

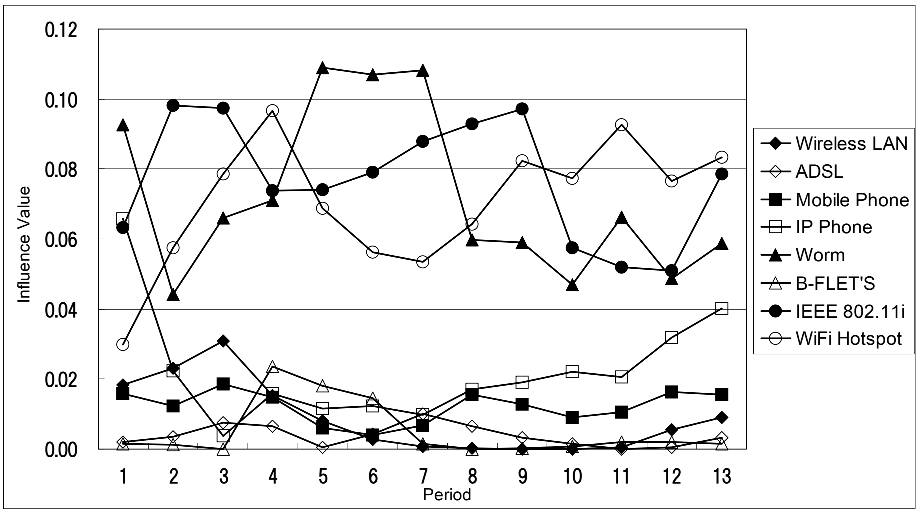

Figure 5 Trends of Influence of the Broadband Field

Figure 6 shows the relationships of the keywords to one another during the specified period. The thickness of the edge indicates a strong relativity, and the orientation of the edge indicates superiority (with the superordinate concept as the starting point). In addition, the size of the node reflects importance. As shown in Figure 6, wireless LAN has a high superiority value and high relativity value with respect to IEEE and WiFi hot spots. In this sense as well, the hypothesis noted in the definition of influence is supported.

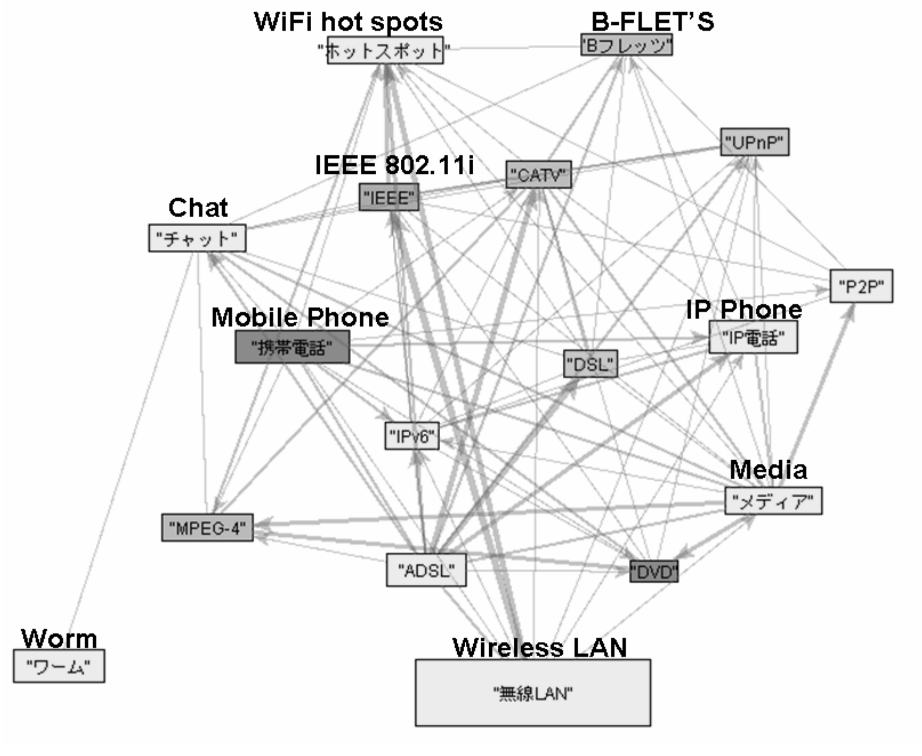

Figure 6 Keywords Map of the Broadband Field at Period 3

Detailed analysis of the mobile communications field

The items that have high importance values in this field are mobile phones and Palm PDAs. However, the importance of mobile phones remained generally constant, while the value for Palm PDAs decreased dramatically, reflecting the sluggish growth in the PDA market.

A comparison of the distributions of influence in this field and the broadband field reveals great differences. Many keywords whose absolute value for influence was low appear within a short period of time. PDC, CDMA, FOMA[2], BREW, i-Appli[3], Java, and other keywords that are mutually competitive or similar continue to increase and decrease in a rather chaotic manner, in general indicating that there was no keyword with a high degree of influence in this field. This shows there is decentralization in the mobile field.

Trends on the organizations in the broadband field

By restricting keywords to the organization category, we can analyze the trends on the aspect of organizations getting into the target technology field. Figure 7 and Figure 8 show the trend on the leading organizations with their importance and influence indicators respectively. In Figure 7, the importance of a communication common carrier F and the competent ministry V increased through all periods. In fact, F leads Japanese broadband field as a platform provider and V presents the national broadband plan through these periods. Conversely, some future service plans proposed by some vendors such as P, L are not necessarily successful in this time. As for influence shown in Figure 8, the trends of R (one of the most popular broadband provider) and J (a security software company) are characteristics. R made a promotion campaign at these periods and had grown into one of the leading provider. After the campaign by R, since other providers also made similar promotion campaign, the influence of R decreased relatively. As for J, the highest value of influence was recorded at the period 7 where some computer worms damaged home computers connected with the Internet via broadband. In these periods, spread of the broadband access exposed the concern about securities for home computers as emerging problems.

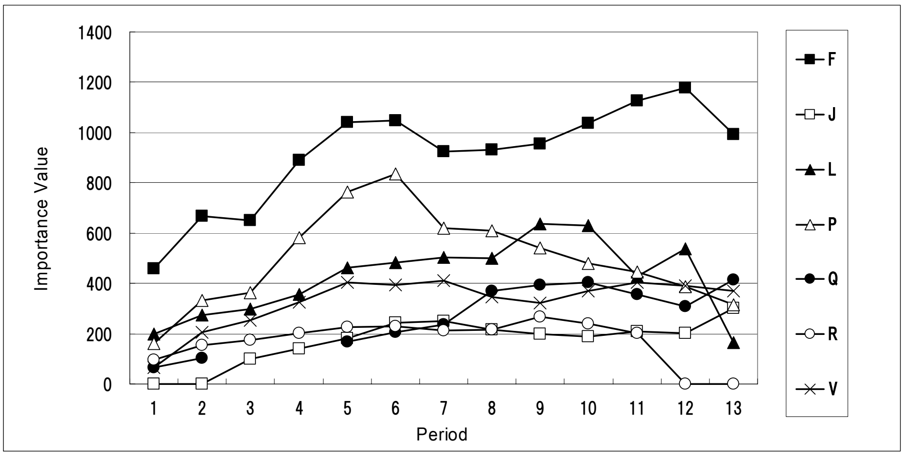

Figure 7 Trends of the Importance w.r.t. Organizations in the Broadband Field

[2] IMT-2000 mobile phone service by NTT-DoCoMo.
[3] Application services on mobile phones by NTT-DoCoMo.

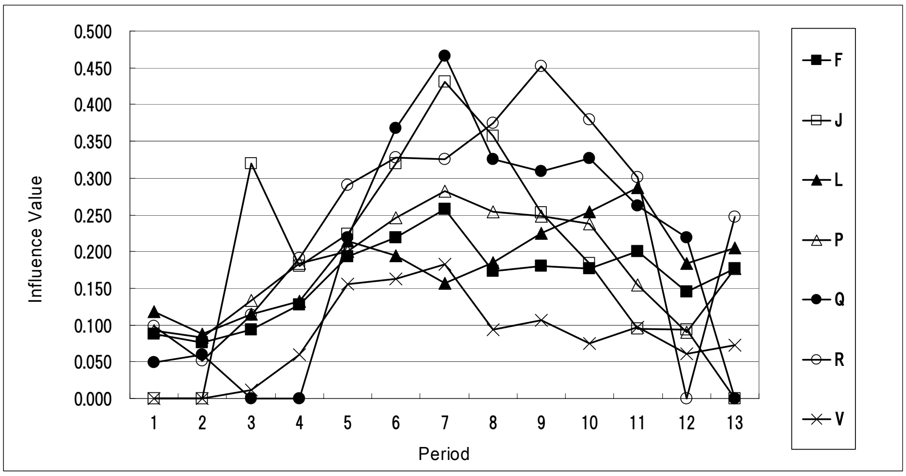

Figure 8 Trends of the Influence w.r.t. Organizations in the Broadband Field

Trends on each keyword in the broadband field

For each keyword, the observation on time-shift of the importance-influence patterns can help us to judge or measure the changes of the maturity/popularity/interests for the keyword. Figure 9 shows the 2-dimensional time-shift pattern for “Mobile Phone”, “IP Phone” and “Worm” in the broadband field. While the influence degrees of “Worm” swing wildly, the importance degrees are generally low. In fact, although the emergence of new viruses sometimes makes “Worm” hot topics in the broadband field, it is not an essential or important stream in this field. The position of “IP Phone” has moved to the importance-high area from the influence-high area. This shows that “IP Phone” has grown up as an important concept starting from a new influential topic. In resent days, “IP Phone” has established one of the major business applications in the field. As for “Mobile Phone”, the influence degrees are generally low and the importance degrees consistently increase. Although “Mobile Phone” could be one of the major platforms in the broadband field, “Mobile Phone” itself is an well-established concept in the field and does not have significant influence to other new broadband technology or services.

Evaluation of influence values as leading indicators

We discuss here that influence values can be regarded as leading indicators for future possible changes. In the broadband field, we chose three keywords “IEEE”, “Worm” and “WiFi Hotspot” as highly influential keywords in the first half of the experimental periods. The

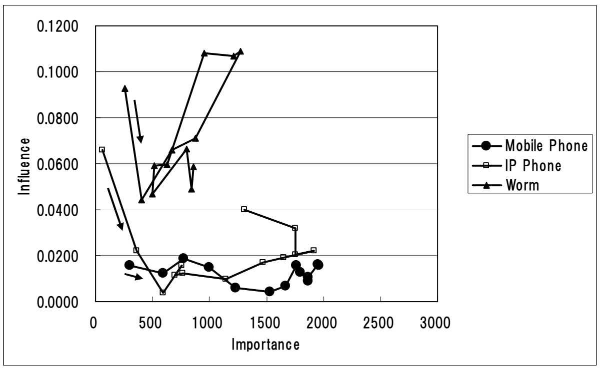

Figure 9 2-dimensional Pattern of Keywords of Broadband Field

average of influence values for three keywords at the period 1-6 are 0.097, 0.097, 0.077 respectively (we set the temporary threshold 0.07 by considering the variance of the influence values of all keywords including other fields). Since determining appropriate threshold value is a hard issue, as first step, it is significant to set the threshold value which detects remarkable changes at least. We assume that if there exist some keywords whose influence values exceed the threshold value, then they show foretaste of some possible changes in the field as noted before. So we expect some changes should occur on the broadband field caused by those three keywords. Here we define the changes as movement of importance values. If a value of each importance or average importance of the field makes substantial change (increase or decrease), we regard this as a change on the field structure. In the broadband field, we can find two major changes: that is, the changes of the importance of "Wireless LAN" and the broadband field itself. As shown in Figure 6, "Wireless LAN", "IEEE" and "WiFi Hotspot" have strong relationships for each. From these facts, we can guess that the changes on "Wireless LAN" (or broadband field) could be suggested by the highly influential keywords.

While in the mobile field, we can find no keywords that satisfy the threshold condition. That means there will be no change in the mobile field. Figure 3 agrees with these suggestions, that is, the average importance value of the mobile field didn't cause significant changes.

Kontostathis (2004) mentioned that there are few projects in this area have used formal evaluation methodologies to determine the effectiveness of the systems. Although the discussions in this section are not fully satisfactory to show the effectiveness of our approach, our study still has significant meaning in terms of the evaluation by the objective method such as precision and recall.

Comparison with human experts' estimation

In order to reinforce the performance of this method, we surveyed the estimations of 10 researchers who are particularly interested in and have minimum 10 years' working experience in these fields so that we can compare the results of this method with human experts' opinions. The evaluation was done from two viewpoints. First evaluation is about the importance and the influence. Keywords can be relatively characterized in terms of the importance and the influence as shown in Figure 10. Keywords in the area A can be considered as important and noteworthy concepts, keywords in B can be considered as not so important but noteworthy hot concepts, keywords in C can be considered as reduced concepts, and keywords in D can be considered as matured concepts for that field. We evaluated this method by comparing the system's and human experts' answers about the quadrants of each keyword.

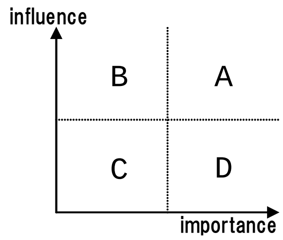

Figure 10 The quadrants of Influence-Importance space

Table 2 shows the results for this comparison. Although the value of concordance rate are not extremely high in themselves, taking into consideration that the answers by human experts hardly agreed with each other, we conclude that our system makes good performance and has ability to judge the status of each technology concept properly. Especially, in case 2 that focuses on the keywords on which over 60% of human experts' answers agreed, the concordance rate increased.

Table 2 The concordance rate of the system and human experts

Field	Case 1		Case 2	
	Concordance Rate	Number of Keywords	Concordance Rate	Number of Keywords
Broadband	53%	Year 2001: 17, Year 2003 : 17	61%	Year 2001: 8, Year 2003 : 6
Mobile	50%	Year 2001: 19, Year 2003: 19	57%	Year 2001: 4, Year 2003 : 11

Case 1: Concordance rate of the method's quadrant and human experts' quadrants regarding all keywords.
Case 2: Concordance rate of the method's quadrant and human experts' quadrants regarding major keywords (over 60% of human experts' answers agreed)

Second evaluation is for the superiority. We surveyed human experts' opinions about the structural configuration of the organizations, at the time of 2001 and 2003. While the results of the system say that higher hierarchization with respect to organizations achieved at the end of 2003 (the average superiority = 0.05) rather than at 2001 (0.03) in the broadband field, eight of ten experts agreed with this trend.

Although we need further verification for practical effectiveness using other fields or resources, as a whole, we recognize that the influence, importance and superiority indicators show good similarity with human experts' estimation.

Discussion

In this paper, we reported the experiments on our quantitative technology analysis system based on the Internet news resources. We have introduced four indicators, that is, importance and influence for each keyword in the target field, and relativity and superiority of relationships between keywords. Although we have not yet established the method to predict future trends systematically, some experimental results including comparison with human experts showed the validity of this method and the use of indicators for technology trend analysis. Future research will extend existing results in several directions including :

- Analysis methods that integrate time series analysis and static analysis
  Time series analysis and static analysis have their own distinct characteristics. Detecting time series characteristics, development of time series animations of keyword maps, etc., are needed to integrate these two methods to enable evaluations and analyze from multiple perspectives.
- Enhancement of the trend detection
  We intend to enhance the method's accuracy and scope so that the method can be applied to new project planning, research and development, and other practical applications.
- Application for text annotation or text classification
  By using the indicators or relationships of keywords, we intend to apply to annotate or classify the Internet news topics. On an experimental basis, we have classified news topics to some classes such as hot news, significant news, newsworthy news, etc. and recognized the effectiveness. We need further theoretical investigation and experiments using enormous amount of news resources.

Furthermore, we have rather essential problem in our approach, that is, how these Internet resources should be selected and how it should be made sure that the selected set of resources and keywords indeed covers technological developments in the target field. Perhaps we should regard the “real trend” of the particular field as composite trends from related various resources and keywords. However, it is not easy to integrate many independent trends since they include diverse trends about the field and different levels of keywords which are thought to characterize the field. There might be discrepancy between trends from different resources. In this research, instead of using many resources and integrating various trends, we selected one resource from the viewpoint of comparison of the system results with human experts. In order to evaluate the performance of the system correctly by comparing with human experts' opinions, human experts should build their opinions based on the same resource as the system uses. For this purpose, we used ITmedia which was the most popular news site among the experts who were involved the evaluation. Therefore, this epistemological problem, how we can choose appropriate resources and keywords for analyzing the real trend, is still remaining unresolved and we need further work for this issue.

Acknowledgments

We would like to thank Softbank IT Media Inc. for granting permission to use the ITmedia news articles for our experiments.

References

Chakrabarti, S. (2003). mining the web - Discovering Knowledge from Hypertext Data. *Morgan Kaufmann*

Tamura, A.(1992). SPRINGS: a system which generates concept map on dynamics between symbols. *Journal of Information Processing, 33(4) (in Japanese)*

Aizawa, A. and Kageura, K.(2000). Calculating Associations between Technical Terms Based on Co-ocurrences in Keyword Lists of Academic Papers. *Trans. IEICE Vol.J83-D-I, No.11, 2000*

Watabe, I. and Kawaoka, T. (2001). The Degree of Association between Concepts using the Chain of Concepts. *Proc. of IEEE International Conference on Systems, Man & Cybernetics*

Boyack, K.W., Wilie, B.N. and Davidson, G.S. (2002). Domain Visualization Using VxInsight for Science and Technology Management. *JASIST, Vol.53, Issue.9*

Boyack, K.W. and Borner, K.(2003). Indicator Assisted Evaluation and Funding of Research. *JASIST, Vol.54, Issue.5*

Hassan, E. (2003). Simultaneous Mapping of Interactions between scientific and technological Knowledge Bases. *JASIST, Vol.54, Issue.5*

Maedche,A., Pekar,V. and Staab,S. (2002). Ontology Learning Part One - On Discovering Taxonomic Relations from the Web. *In Zhong,N., Liu,J. and Yao,Y.Y., editor,Web Intelligence, Springer Verlag*

Morris, S. A., Yen, G., Wu, Z., and Asnake, B.(2003). Timeline visualization of research fronts. *JASIST, Vol.55, Issue.5*

Chen,C, and Kuljis, J.(2003). The Rising Landscape:A Visual Exploration of Superstring Revolution in Physics. *JASIST, Vol.54, Issue.5*

Kontostathis, A., Galitsky, L. M., Pottenger, W. M., Roy, S., Phelps, D.(2004). A Survey of Emerging Trend Detection in Textual Data Mining. *Berry, M. W., editor, Survey of Text Mining, Springer Verlag*