Information Extraction and Visualization from Internet Documents

By Dai KUSUI,* Kenji TATEISHI* and Toshikazu FUKUSHIMA*

ABSTRACT The Internet displays a large number of Web pages, and many email messages are sent and received. Such Internet documents are important information sources in daily life and company activities. However, since they are so large and varied it is very difficult to extract useful information for specific purposes when retrieving them. Individuals and companies themselves must acquire new added value by analyzing available Internet documents. For such purposes, automatic information extraction, analysis, and visualization technologies are needed. As concrete examples of such technologies this paper describes two systems that extract and visualize product reputation information from Internet Web pages and “who knows what” information from email messages.

KEYWORDS **Information extraction, Visualization, Reputation, Opinion, “Who knows what”, Knowledge management**

1. INTRODUCTION

The Internet displays several billions of Web pages, and many email messages are sent and received. If an average of thirty emails is received every day, in a year one person will receive more than 10,000 of them.

Internet documents are indispensable information sources in daily life and company activities. However, since they are so large and various, it is very difficult to extract useful information for specific purpose from them. Individuals and companies themselves must acquire new added value by analyzing available Internet documents.

For example, such search engines as Google and Yahoo are useful for locating desired information from Web pages. However, they are not necessarily suitable under the following situations when an individual:

- Wants to examine what kind of detail has become a new topic of global interest.
- Wants to investigate the reputation of a product or service before purchasing it.
- Wants to examine the reputation of the products of his own company or competitors.
- Wants to investigate the effects of advertisements and marketing campaigns, etc.

Email databases can be effectively used not only for retrieving specific mails as an information source but also other purposes. For example, useful new information can be acquired by analyzing relationships during email messages.

- Grasping the thread of email contents quickly by visualizing the quotation relationship of email messages.
- Examining who communicates with whom about what and analyzing the flow of information about a business.
- Discovering who knows a lot about what kind of things and who is a key individual for certain themes.

For such purposes, automatic information extraction, analysis, and visualization technologies are needed. As concrete examples of such technologies, this paper describes two systems. The first extracts and visualizes product reputation information from Internet Web pages, and the second extracts and visualizes “who knows what” information from email messages.

*Internet Systems Research Laboratories

†As the products introduced in this paper are sold for the domestic market, some sections and figures feature explanations by the Japanese language.

2. REPUTATION SEARCH ENGINES

2.1 Importance of Opinions on the Internet

The dramatic spread of the Internet enables us to deliver our own message to the public and to communicate with many people. Since the Internet is a special space where everyone has the opportunity to make and disseminate his/her own messages, we can expect a large amount of diverse opinions. A tool that is able to collect and analyze these opinions efficiently can be used for the following purposes.

(1) Pre-Purchasing Product Surveys

Before making purchases, consumers can benefit from searching for and acquiring other opinions from Web or Blog sites that discuss many products. However, general-purpose search engines such as Google sometimes provide much wrong information unrelated to product opinions. According to our research, when a product name is entered as a keyword into a general-purpose search engine, the proportion of Web pages that include opinions in the search results averages only 16%[1]. Therefore, an opinion-specialized search engine that efficiently collects subjective views is important.

(2) Market Research

It is important for corporate activities at any step of product proposal, development, and improvement to acquire and use feedback from consumers. Recently, questionnaire analysis through the Internet is becoming popular because the Internet can gather such data more effectively. However, it is still expensive to collect opinions in fields where products have short life-cycles or on all products including those of competitors. Moreover, gathering open answers is more expensive than closed answers. Therefore, if a system can gather opinions more cheaply and speedily, we expect that it will be able to replace questionnaire data.

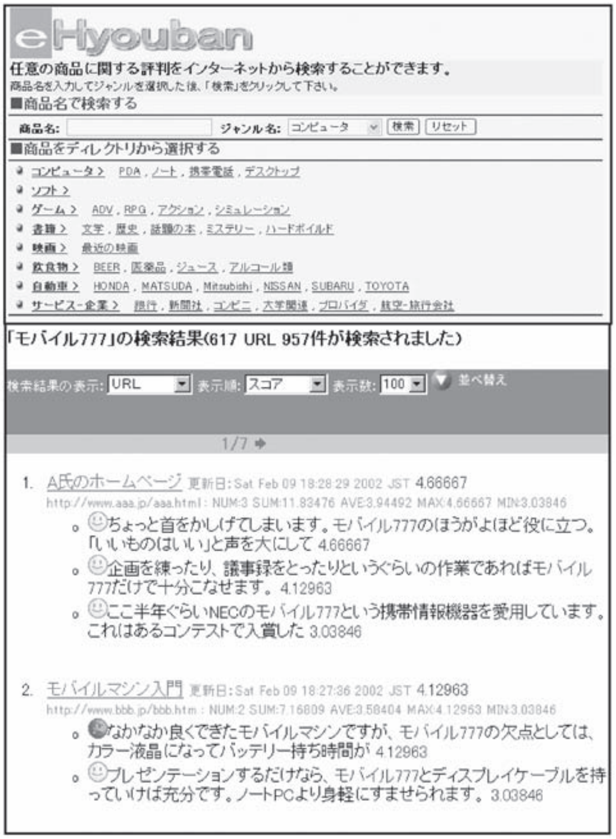
(3) Risk Management for Enterprises

There are many Internet communities such as BBS or Weblog. Since such communities are often anonymous, some messages may denigrate enterprises. Rapid detection of such harmful information is needed for risk management. Such Internet monitoring services as Gala (http://gala.jp) and eWatch (http://www.ewatch.com) regularly watch Web sites and BBS selected by their clients, informing them of injurious information on their products as soon as it is found. However, such services are expensive and take a long time to produce results since they are usually performed by humans rather than by software. Therefore, if automation can be realized, it will become a more valuable tool.

In this paper, we describe an opinion-specialized search engine called “Reputation Search Engine” (RSE)[2,6] that extracts and classifies opinions from many kinds of Web sources. RSE has two functions. One extracts opinions on a specific product from Web documents. The other classifies extracted opinions into positives or negatives. Using such functions, RSE is available for the three purposes described above.

2.2 RSE Examples

Figure 1 shows screenshots of RSE. When a user gives such keywords as a product name on an initial Web page, RSE lists Web documents including opinions of the keyword as search results. The opinions are clipped from the original Web pages, grouped by URL, and then sorted in the order of opinion-likeliness scores (See Section 2.3). An icon that represents positive (smiling) or negative (crying) opinion is shown at the side of each opinion. Users can effectively read the clipped opinions and their



任意の商品に関する評判をインターネットから検索することができます。

商品名を入力してジャンルを選択した後、「検索」をクリックして下さい。

■商品名で検索する

商品名: ジャンル名: コンピュータ

■商品をディレクトリから選択する

- コンピュータ > PDA , ノート , 携帯電話 , デスクトップ
- ソフト >
- ゲーム > ADV , RPG , アクション , シミュレーション
- 書籍 > 文学 , 歴史 , 話題の本 , ミステリー , ハードボイルド
- 映画 > 最近の映画
- 飲食物 > BEER , 医薬品 , ジュース , アルコール類
- 自動車 > HONDA , MATSUDA , Mitsubishi , NISSAN , SUBARU , TOYOTA
- サービス-企業 > 銀行 , 新聞社 , コンビニ , 大学関連 , プロバイダ , 航空-旅行会社

「モバイル777」の検索結果(617 URL 957件が検索されました)

検索結果の表示: URL 表示順: スコア 表示数: 100

1. A氏のホームページ 更新日:Sat Feb 09 18:28:29 2002 JST 4.66667
   http://www.aaa.jp/aaa.html : NUM:3 SUM:11.83476 AVE:3.94492 MAX:4.66667 MIN:3.03846
   - ☺ちょっと首をかしげてしまいます。モバイル777のほうがよほど役に立つ。「いいものはいい」と声を大にして 4.66667
   - ☺企画を練ったり、議事録をとったりというぐらいの作業であればモバイル777だけで十分こなせます。 4.12963
   - ☺ここ半年ぐらいNECのモバイル777という携帯情報機器を愛用しています。これはあるコンテストで入賞した 3.03846
2. モバイルマシン入門 更新日:Sat Feb 09 18:27:36 2002 JST 4.12963
   http://www.bbb.jp/bbb.htm : NUM:2 SUM:7.16809 AVE:3.58404 MAX:4.12963 MIN:3.03846
   - ☹なかなか良くできたモバイルマシンですが、モバイル777の欠点としては、カラー液晶になってバッテリー持ち時間が 4.12963
   - ☺プレゼンテーションするだけなら、モバイル777とディスプレイケーブルを持っていけば充分です。ノートPCより身軽にすませられます。 3.03846

Fig. 1 Screenshots of RSE.

positive/negative icons. In addition, each opinion has time information, enabling temporal analysis, as shown in **Fig. 2**.

For example, the results from Figs. 1 and 2 can be used in the following way. First, Fig. 2 shows an example where opinions on a certain product have been counted for a five-day period. From this chart, it can be observed that the proportion of negative opinions after the product release date increased drastically, whereas they used to be low. This observation means that this product did not meet consumer expectations. The results of Fig. 1 reflect what consumers have specifically said about the product.

2.3 System Architecture

The RSE consists of the following three modules: Opinion Extraction Module, Opinion Classification Module, and Web Documents Collection Module.

2.3.1 Opinion Extraction Module

To develop RSE, we first must clarify our definition of the term "opinion."

1) An opinion is composed of an object name and an evaluative expression.
2) An opinion has a semantic relationship between the two expressions.
3) An opinion reflects a subjective judgment.

As above defined, RSE extracts opinions using an evaluative expression dictionary and pattern-matching rules. This method is often employed in the field of information extraction, especially in Named Entity Extraction[3].

First, RSE finds opinion candidates that fulfill the above first condition. An opinion candidate is a sentence that includes an object name and an evaluative expression, which indicates either a positive or negative evaluation of the object name. Evaluative expressions are prepared in advance as a dictionary on a domain basis. This dictionary is created manually from Web sites such as 'Yahoo! Message Boards,' where opinions on objects are often discussed. Each evaluative expression in the dictionary is given a positive or negative label. For example, in computer domains, "良い(good)," "使いやすい(handy)" and "満足(satisfactory)" are used as positive expressions and "遅い(slow)," "不満(dissatisfied)" and "悪い(bad)" are used as negative ones. Similarly, in alcoholic beverage domains, "おいしい(delicious)," "良い(good)," "好き(like)" are positive, "いまいち(unsatisfied)," "悪い(bad)," and "強すぎる(too strong)" are negative.

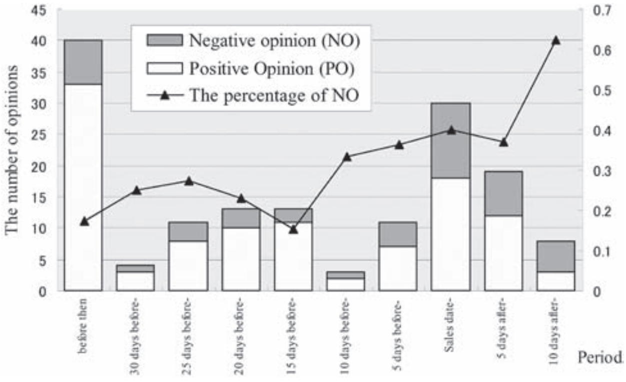


Fig. 2 Examples of temporal analysis.

RSE next calculates opinion-likeliness scores using pattern-matching rules that fulfill the above second and third conditions. The following shows examples of opinion candidates, where Mobile777 indicates an object name and "良い(good)" denotes an evaluative expression.

(a) Mobile777、これは良い!! (Mobile777 is a good product!)
(b) Mobile777を持っております。ICQを使いたいのですがどうすれば良いでしょうか？ (I have a Mobile777, and I want to use ICQ on it. What is a good way to do that?)
(c) Mobile777が良いという人もいるでしょうが . . . (Some say that Mobile777 is a good product, but...)
(d) PCの調子が悪いため、Mobile777を使用していますが . . . (Since there is something wrong with my PC, I've used a Mobile777...)

Here, (a) is a correct example, but (b)-(d) are incorrect because these sentences do not satisfy the second condition; (b) does not have a relationship between the two expressions, and (c) and (d) do not convey subjective judgments. Opinion-likeliness calculation gives opinion (a) high scores and opinions (b) to (d) low scores by using pattern-matching rules. These rules are learned manually from training examples (See Reference [2] for detail). Opinion candidates whose scores are above the threshold are regarded as opinions.

Evaluation of the opinion extraction method was conducted in two domains: computers and alcoholic beverages. The results showed precision of 72% and 84% in the computer and alcoholic beverage domains, respectively, and 78% overall. In addition, these two domains revealed the advantage of our method over an SVM text classification method.

2.3.2 Opinion Classification Module

The system classifies opinions by using labels attached to evaluative and negative expressions. The system first looks up the label of the evaluative expression and then counts the frequency of negative expressions located close before and after the evaluative expressions. If the frequency is an even number, the system classifies the opinion as the label. If it is an odd number, it classifies the opinion as the reverse of the label.

In the experiment, we used the grammatically negative expressions "ない(not or no)" and fixed the scope for finding negative expressions to 12 bytes after the evaluative expression. Precision was 87% (119/137) for the computer domain, 93% (154/166) for the alcoholic beverage domain, and 90% (273/303) in total. The overall accuracy was found to be high.

2.3.3 Web Documents Collection Module

RSE collects Web documents in three ways. First, it gathers them from Web sites selected in advance. Next, it utilizes the results of general-purpose search engines in which the system first gathers URLs by throwing an object name at them and collecting their Web documents. Finally, using crawlers it checks message boards that general search engines do not reach. These crawlers are designed to suit the format of each message board. Since each crawler is attuned to the format of a message board, their posted dates can be added to extracted opinions.

3. EMAIL BASED KNOWLEDGE MANAGEMENT SYSTEM CALLED "Interaction Viewer"

3.1 Motivation and Problems

The motivation behind this system was the reuse of business knowledge and expertise possessed by employees. There are two problems with effectively reusing knowledge.

(1) Knowledge Extraction, Management and Maintenance Costs

It is necessary to be able to use knowledge with the usual tools (email and Web browsers). Automatic generation and update of knowledge bases by log analysis of emails, schedules, room reservation systems and so on, are important.

(2) Useless Knowledge

Most knowledge is effective only under specific conditions. The accumulation of knowledge that easily utilizes "whom should it ask?" is important.

Since this system directly uses email folders as knowledge or case bases, it is especially advantageous in such busy organizations as high-tech product customer support organizations. For example, References [4] and [5] state the importance of case-based reasoning technology in customer support and help desk services.

3.2 Design Decisions

3.2.1 Model of an Email Message

An email message consists of header and body sections, as shown in **Fig. 3**. The header section has several attributes including a subject, a sender, a receiver, date, time, and so on. Each attribute is an index of the email. The body section consists of several parts: a quotation part, a sentence part and a signature part at the end.

There are two types of references. One is inserted in the header section. The subject often includes "re," which indicates that the message is a reply to a previous email. The in-reply-to and reference attributes indicate the previous email message's message-id. The other reference type includes quotation parts included in the email body copied from the previous email message, which is also an important email index. The keywords included in the quotation parts also become indices of the email.

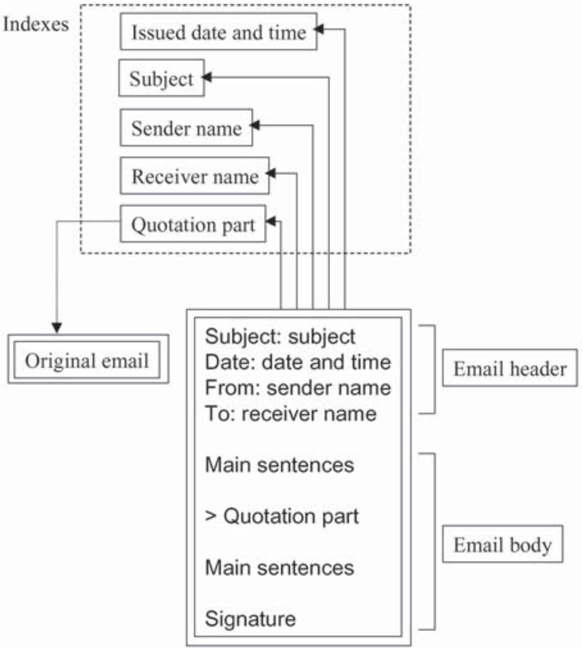


Fig. 3 An email structure.

3.2.2 Analyzing Email Message Sequences

First, a person generates a new email message including an initial question with a new subject in the mail header. The initial question is called the root question. The replies of subsequent email messages are called answers. If an answer satisfies the root question, then the question and the answer email pair constructs knowledge. An answer email may include not only a partial answer but also a related question, such as "what was the status of the device manager when you encountered the problem?" Such a question clarifies the root question and is called a subsequent question. The sequence of subsequent questions automatically constructs a discrimination tree.

Figure 4 shows relationships among emails. A subsequent email sometimes includes a description that raises a new question related to the root question. Subsequent email messages might quote the new question instead of replying to the root question, meaning that the main topic has been changed from the root to the new question. The email starting the new question can be regarded as a new root question, although the SUBJECT attribute includes the same description as the former root question email. It is difficult for a computer to automatically find such a topic change in an email list. A system must be made to comprehend meaning and to handle such topic changes.

3.3 System Development

Based on the results examined in Section 3.2, we designed an email-based knowledge management system called "Interaction Viewer" that has two modes: email search and "who knows what" search. A user can visualize the context of email messages in email search mode and the relationship among users or among users and topics in "who knows what" search mode.

3.3.1 Context Visualization

The email search mode loads a list of email messages. Interaction Viewer analyzes the relationships among the email messages and displays a tree structure.

Figure 5 shows the Interaction Viewer screen in email search mode. A user can search email threads by specifying keywords or the time period. A new window opens by selecting the thread of the search result. The tree structure among email messages is arranged at the top part of the window. When a subject is selected, the text in the mail is displayed downward. When the triangle on the side of the subject is selected, the referential relationships among email messages can be displayed downward, as shown in **Fig. 6**.

Figure 6 uses indentation to show messages with the email addresses of the senders. The user

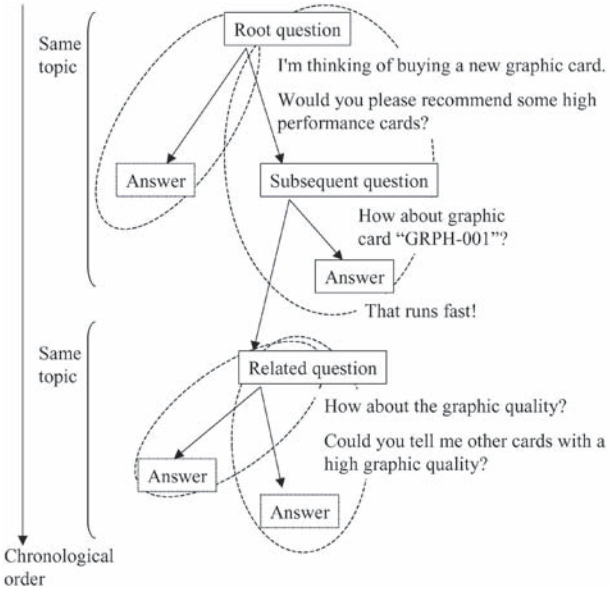
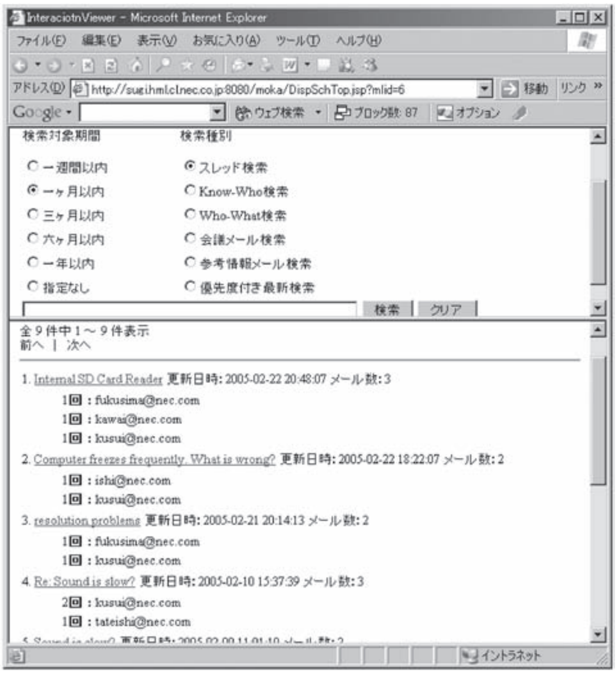

Same topic
Root question
I'm thinking of buying a new graphic card.
Would you please recommend some high performance cards?
Answer
Subsequent question
How about graphic card "GRPH-001"?
Answer
That runs fast!
Same topic
Related question
How about the graphic quality?
Could you tell me other cards with a high graphic quality?
Answer
Answer
Chronological order

Fig. 4 Typical sequence in an email list.


InteraciotnViewer – Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
アドレス(D) http://sugi.hml.cl.nec.co.jp:8080/moka/DispSchTop.jsp?mlid=6
移動 リンク »
Google ウェブ検索 ブロック数: 87 オプション
イントラネット

検索対象期間

○ 一週間以内
● 一ヶ月以内
○ 三ヶ月以内
○ 六ヶ月以内
○ 一年以内
○ 指定なし

検索種別

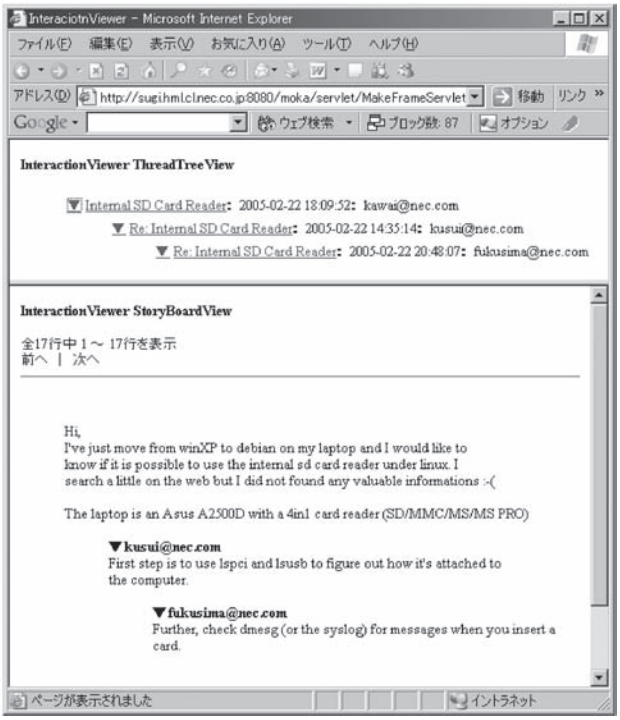
● スレッド検索
○ Know-Who検索
○ Who-What検索
○ 会議メール検索
○ 参考情報メール検索
○ 優先度付き最新検索

検索 クリア

全９件中１～９件表示
前へ ｜ 次へ

1. Internal SD Card Reader 更新日時: 2005-02-22 20:48:07 メール数: 3
  1回 : fukusima@nec.com
  1回 : kawai@nec.com
  1回 : kusui@nec.com
2. Computer freezes frequently. What is wrong? 更新日時: 2005-02-22 18:22:07 メール数: 2
  1回 : ishi@nec.com
  1回 : kusui@nec.com
3. resolution problems 更新日時: 2005-02-21 20:14:13 メール数: 2
  1回 : fukusima@nec.com
  1回 : kusui@nec.com
4. Re: Sound is slow? 更新日時: 2005-02-10 15:37:39 メール数: 3
  2回 : kusui@nec.com
  1回 : tateishi@nec.com
5. Sound is slow? 更新日時: 2005-02-09 11:01:10 メール数: 2

Fig. 5 IV screen in email search mode.


InteraciotnViewer - Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
アドレス(D) http://sugi.hml.cl.nec.co.jp:8080/moka/servlet/MakeFrameServlet
移動 リンク »
Google ウェブ検索 ブロック数: 87 オプション
ページが表示されました
イントラネット

InteractionViewer ThreadTreeView

▼ Internal SD Card Reader: 2005-02-22 18:09:52: kawai@nec.com
▼ Re: Internal SD Card Reader: 2005-02-22 14:35:14: kusui@nec.com
▼ Re: Internal SD Card Reader: 2005-02-22 20:48:07: fukusima@nec.com

InteractionViewer StoryBoardView

全17行中 1 ～ 17行を表示
前へ | 次へ

Hi,
I've just move from winXP to debian on my laptop and I would like to know if it is possible to use the internal sd card reader under linux. I search a little on the web but I did not found any valuable informations :-(

The laptop is an Asus A2500D with a 4in1 card reader (SD/MMC/MS/MS PRO)

▼kusui@nec.com
First step is to use lspci and lsusb to figure out how it's attached to the computer.

▼fukusima@nec.com
Further, check dmesg (or the syslog) for messages when you insert a card.

Fig. 6 Visualizing email flows.

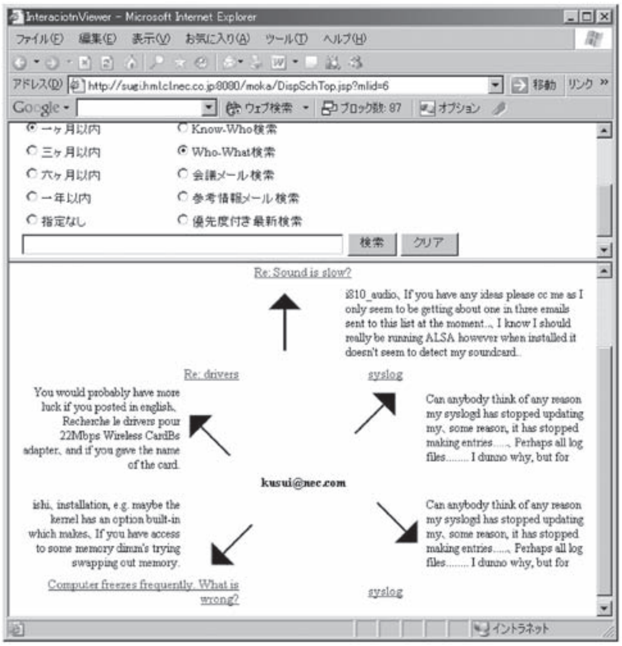
understands message contexts through referential relationships.

3.3.2 Who Knows What Knowledge Base

Interaction Viewer generates and automatically updates the who knows what knowledge base by analyzing the relationships among email messages. A user can search "who knows what" for human skills and "who knows whom" for human relations in the "who knows what" search mode.

Figure 7 shows the human skills search screen ("who knows what") results in "who knows what" mode. The network visualizes human skills, interests, and so on. The center of the network is the searched user, around which there are five topics. The arrow width between the searched user and topic shows the number of messages the user sent on the topic. The topic's keywords are added to the arrow. A user can change five search results in sequence.

Figure 8 shows the screen of human relations search ("who knows whom") results in "who knows what" mode. The network visualizes human relationships. The center of the network is the searched user, around which there are five other users. The arrow width between the searched and other users shows the number of messages between the users. Message keywords between users are added to the arrow. A user can change five search results in sequence.


InteraciotnViewer - Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
アドレス(D) http://sugi.hml.cl.nec.co.jp:8080/moka/DispSchTop.jsp?mlid=6
移動 リンク »
Google ウェブ検索 ブロック数: 87 オプション
検索
クリア
イントラネット

● 一ヶ月以内 ○ Know-Who検索
○ 三ヶ月以内 ● Who-What検索
○ 六ヶ月以内 ○ 会議メール検索
○ 一年以内 ○ 参考情報メール検索
○ 指定なし ○ 優先度付き最新検索

Re: Sound is slow?
i810_audio., If you have any ideas please cc me as I only seem to be getting about one in three emails sent to this list at the moment..., I know I should really be running ALSA however when installed it doesn't seem to detect my soundcard..

Re: drivers
You would probably have more luck if you posted in english., Recherche le drivers pour 22Mbps Wireless CardBs adapter., and if you gave the name of the card.

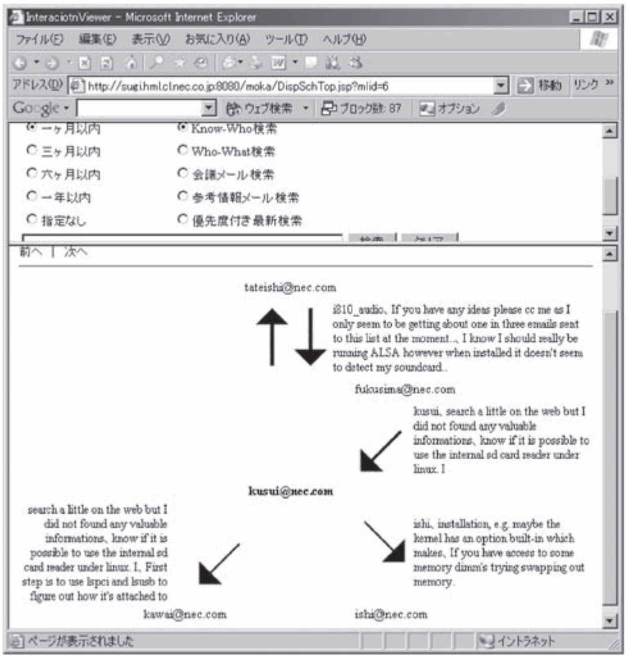
syslog
Can anybody think of any reason my syslogd has stopped updating my., some reason, it has stopped making entries......, Perhaps all log files......... I dunno why, but for

kusui@nec.com

ishi., installation, e.g. maybe the kernel has an option built-in which makes., If you have access to some memory dimm's trying swapping out memory.
Computer freezes frequently. What is wrong?

Can anybody think of any reason my syslogd has stopped updating my., some reason, it has stopped making entries......, Perhaps all log files......... I dunno why, but for
syslog

Fig. 7 Screen of Human Skills Search.


InteraciotnViewer - Microsoft Internet Explorer
ファイル(F) 編集(E) 表示(V) お気に入り(A) ツール(T) ヘルプ(H)
アドレス(D) http://sugi.hml.cl.nec.co.jp:8080/moka/DispSchTop.jsp?mlid=6
移動 リンク »
Google ウェブ検索 ブロック数: 87 オプション
ページが表示されました
イントラネット

● 一ヶ月以内 ● Know-Who検索
○ 三ヶ月以内 ○ Who-What検索
○ 六ヶ月以内 ○ 会議メール検索
○ 一年以内 ○ 参考情報メール検索
○ 指定なし ○ 優先度付き最新検索

前へ | 次へ

tateishi@nec.com
i810_audio., If you have any ideas please cc me as I only seem to be getting about one in three emails sent to this list at the moment..., I know I should really be running ALSA however when installed it doesn't seem to detect my soundcard..

fukusima@nec.com
kusui., search a little on the web but I did not found any valuable informations., know if it is possible to use the internal sd card reader under linux. I

kusui@nec.com

search a little on the web but I did not found any valuable informations., know if it is possible to use the internal sd card reader under linux. I., First step is to use lspci and lsusb to figure out how it's attached to
kawai@nec.com

ishi., installation, e.g. maybe the kernel has an option built-in which makes., If you have access to some memory dimm's trying swapping out memory.
ishi@nec.com

Fig. 8 Screen of Human's Relations Search.

4. CONCLUSION

In this paper, we described how to acquire new added value by analyzing such Internet documents as Web pages and email messages. We introduce two systems as concrete examples. The first is a reputation search engine that extracts and visualizes reputation information from Internet Web pages. The second is an email-based knowledge management system that extracts and visualizes "who knows what" information from email messages. We will increase the number of kinds of targeted Internet documents and develop more varied analytical methods in the future.

REFERENCES

[1] K. Tateishi, Y. Ishiguro and T. Fukushima, "A reputation search engine that gathers people's opinions from the Internet," *Technical Report NL-144-11, Information Processing Society of Japan*, pp.75-82, 2001 (in Japanese).

[2] K. Tateishi, Y. Ishiguro and T. Fukushima, "A Reputation Search Engine that Collects People's Opinions by Information Extraction Technology," *IPSJ Transactions on Databases*, **22**, 2004.

[3] D. Appelt and D. Israel, "Introduction to Information Extraction Technology," Tutorial for IJCAI-99, Stockholm, 1999.

[4] T. M. Goker and Roth-Berghofer, "Development and Utilization of a Case-Based Help-Desk Support System in a Corporate Environment," *Case-Based Reasoning Research and Development, Proceedings of the ICCBR99*, pp.132-146, 1999.

[5] H. Thomas, R. Foil and J. Dacus, "New Technology Bliss and Pain in a Large Customer Service Center," *Case-Based Reasoning Research and Development, Proceedings of the ICCBR97*, pp.166-177, 1997.

[6] K. Tateishi, Y. Ishiguro and T. Fukushima, "A Reputation Search Engine from the Internet", *J. of the Japanese Society for Artificial Intelligence*, **19**, 3, pp. 317-323, 2004 (In Japanese).

Received February 28, 2005

* * * * * * * * * * * * * * *

Dai KUSUI received his B.S degree in applied mathematics and physics and an M.S degree in applied systems science from Kyoto University in 1990 and 1992, respectively. He joined NEC Corporation in 1992 and now works in the Internet Systems Research Laboratories. He is engaged in the research and development of intelligent interactive systems.

Mr. Kusui is a member of IPSJ.

Kenji TATEISHI received his B.S degree in physics from the Science University of Tokyo in 1997 and M.S degree in information engineering from Kyushu University in 1999. He joined the Internet Systems Research Laboratories, NEC Corporation, in 1999 and has been engaged in the field of Information Extraction and Information Retrieval on the Internet. He received a Best Paper Award for the 64th IPSJ National Convention in 2002.

Mr. Tateishi is a member of IPSJ.

Toshikazu FUKUSHIMA received his B.S. degree in physics from the University of Tokyo and joined NEC Corporation in 1982. He is currently a Senior Manager of the Ubiquitous Intelligence Technology Group in the Internet Systems Research Laboratories, NEC. He received his Ph.D from Kyushu University in 1998, a Best Paper Award for Young Researchers of the 45th IPSJ National Convention in 1992, Best Paper Award of the 53rd IPSJ National Convention in 1996, the 23rd IPSJ Best Paper Award in 1992, the 6th IPSJ Sakai Special Researcher Award in 1997, the 51st OHM Technology Award in 2003, and the Best Paper Award of the 15th IEICE Data Engineering Workshop in 2004.

Dr. Fukushima is a member of IPSJ, JSAI, ANLP, JSIK, INFOSTA, and ACM.

* * * * * * * * * * * * * * *