Text Data Mining of Care Life Log by the Level of Care Required Using KeyGraph

Muneo Kushima, Kenji Araki, Tomoyoshi Yamazaki, Sanae Araki, Taisuke Ogawa, Noboru Sonehara

***Abstract*—In the present study, to classify the vast amount of Care Life Log data that occurs in nursing in one Miyazaki Hospital Long-term Health Care Facility by level of care required, data mining was carried out. The characteristic vocabulary from the Long-term Health Care Facility's Care Life Log was used to integrate and analyze the level of care required.**

There are five levels of care, with Level 1 vocabulary including recreation, toilet, morning, afternoon, etc. The level of care gradually increases from Level 1 to Level 5, which has vocabulary that includes tube, danger, treatment, removal, and discovery. The higher the level, the worse the health condition and therefore the greater care required. These levels allow for a clear analysis of a patient's condition. This analysis has led to an improvement in Quality of Life as well as a decrease in mismatches between the level of care required for patients and the level of care given by caretakers.

The nursing field requires efficiency in health care services. Because of this, improvement and continuous data collection are important. There is a need for the collection of data as a whole in the long-term building of health care services as well as large-scale data collection.

In the future, we aim to develop an EMR(Electronic Medical Record) that can be created semi-automatically in accordance with the level of care required.

***Index Terms*—text data mining, care life log, KeyGraph, Electronic Medical Record**

I. Introduction

Data mining searches for correlations among items by analyzing a great deal of such accumulated data as sales data and telephone call histories. Text data mining resembles data mining because it extracts useful knowledge and information by analyzing the diversified viewpoints of written data [1].

Recently, interest has risen in text data mining because it uncovers useful knowledge buried in a large amount of accumulated documents [2]. Research has started to apply text data mining to medicine and healing [3]. In addition, the speed of electronic medical treatment data is accelerating because of the rapid informationization of medical systems, including EMRs.

Recently, research on data mining in medical treatment that aims for knowledge and pattern extraction from a huge accumulated database is increasing. However, many medical documents, including EMRs that describe the treatment information of patients, are text information. Moreover, mining such information is complicated.

The data arrangement and retrieval of such text parts become difficult because they are often described in a free format; the words, phrases, and expressions are too subjective and reflect each writer [4]. Perhaps in the future, the text data mining of documents will be used for lateral retrieval, even in the medical treatment world, not only by the numerical values of the inspection data but also by computerizing documents [5].

In this present study, to classify the vast amount of Care Life Log data that occurs in nursing in one Miyazaki Hospital Long-term Health Care Facility by level of care required, text data mining was carried out using KeyGraph. Then we visualized this information.

II. Care Life Log and and the Level of Care Required

Care Life log records a period of 24 hours of the caregiver's activity. It is also utilized as a long-term service content record. The recording itself is not the main purpose, but it transmits information to others, accumulates and analyzes data, and aims to connect the service to better care.

The level of care required is categorized as Standard Support 1 and 2, and Essential Support 1,2,3,4, and 5. Essential Support 1 indicates that a person can eat and use the restroom by themselves.

Essential Support 5 indicates a person is mostly unable to do these things by themselves.

Essential Support Levels are outlined below:

(1)Level 1: He or she needs care by others when performing complex actions or moving. There is a noticed decrease in physical and mental capabilities.

(2)Level 2: The same conditions as Level 1 with the addition of needing some assistance when eating or using the restroom.

(3)Level 3: The patient cannot use the restroom by themselves and needs assistance performing any action indicated by Level 2.

(4)Level 4: The patient can hardly use the restroom or perform any action indicated in Level 3.

Manuscript received Nov. 28, 2016 ; revised Dec. 31, 2016.

M. Kushima, K. Araki, T. Yamazaki, S. Araki, and T. Ogawa are with the Faculty of Medicine, University of Miyazaki Hospital. (e-mail:muneo_kushima@med.miyazaki-u.ac.jp, http://mit.med.miyazaki-u.ac.jp/.
5200, Kihara, Kiyotake-cho, Miyazaki-shi 889-1692 Japan.
Tel: +81-985-85-9057, Fax: +81-985-84-2549.
N. Sonehara is with the Information and Society Research Division, National Institute of Informatics.
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan.
Tel: +81-3-4212-2521, Fax: +81-3-4212-2120.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(5)Level 5: The patient can hardly eat or use the restroom and needs assistance with almost all actions.

III. EMR

When the medical information system was updated in May, 2006, the University of Miyazaki Hospital introduced a package version of the EMR system, which was developed in collaboration with a local IT company. The recorded main data include a patient's symptoms, laboratory results, prescribed medicines, and the tracking of the changed data. Cases that make both the images of X-rays and the appended material electronic are not infrequent either. If a network is used, EMR can be shared not only in one hospital but also among two or more hospitals.

The text data in EMR consist of paper notations about inspection reports, in-patient care plans, nutrition management plans, bedsore-prevention plans, fall checks, operation notes, and summaries. The doctor fills in the passage record and the nurses fill in the nursing records, which include the life and inspection history of a patient.

The care life log also has small notes about reservations etc. Since no guidelines exist about recording text, ambiguous feelings or impressions are sometimes included. Care workers remember or take notes about what their patients say while working and later input them into the EMR.

There are four recording categories: subjective data (S), objective data (O), assessment (A), and plans (P).

(1)S: information directly gleaned from patients.

(2)O: objective facts and observations about the patients appearance or state by co-medicals.

(3)A: evaluations and judgments derived from this information.

(4)P: future plans and care actually taken.

The care life log, which records the care activities practiced by nurses, contains many notes about nursing processes. It helps ensure high quality nursing and evaluates nursing practices.

IV. TEXT DATA MINING APPLICATION TO MEDICINE

Text data mining is often used to analyze information hidden in the text of a document and to extract key words, phrases, and even concepts from written documents. Text data mining or data mining, which is roughly equivalent to text analytics, refers to the process of deriving high-quality information from texts.

Text data mining usually structures the input text (often by parsing, adding derived linguistic features, removing others and insertion into a database), deriving patterns within the structured data and finally evaluating and interpreting the output.

Fig. 1 shows the process of text data mining. Two particular aspects should be considered when applying text data mining to a medical context. Second, final decisions can be obtained regarding courses of treatment.

One difficulty with applying text data mining to medicine is the entire process of identifying symptoms for understanding the associated risks while taking appropriate action.

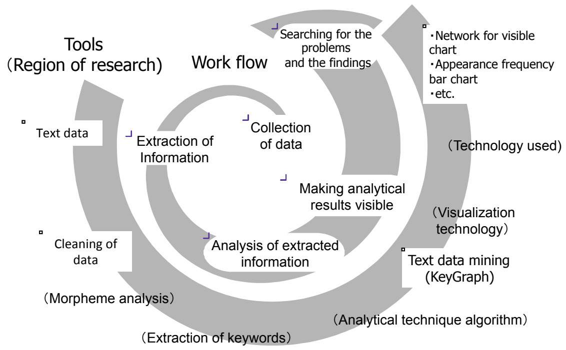

Fig.1 process of text data mining

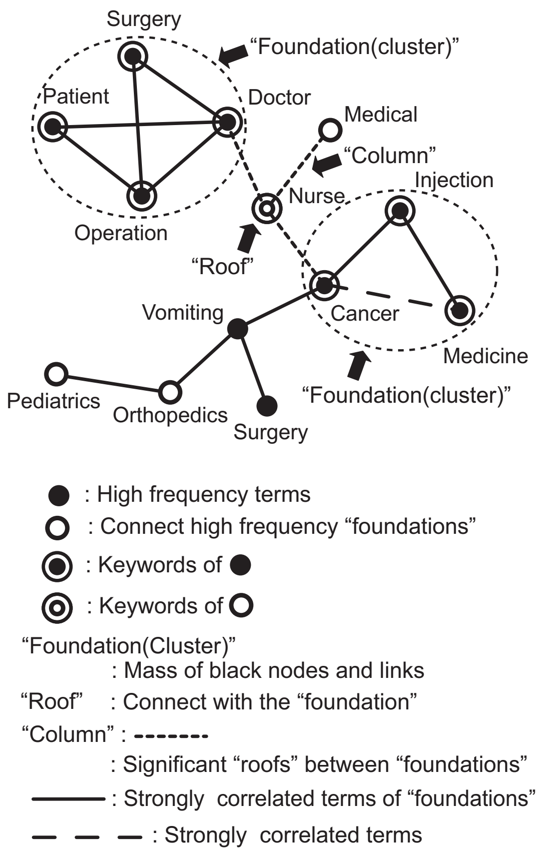

Fig.2 KeyGraph example when applied to text data

V. KEYGRAPH

We applied KeyGraph to the text data mining technique [6],[7]. We also applied it for extracting key words.

A. Example of KeyGraph Performance

Figure 2 shows an example when it is applied to text data.

(1)Black nodes indicate items that frequently occur in a data set.

(2)White nodes indicate the items that occur less frequently overall but frequently occur with black nodes in a data set.

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(3)Double-circled nodes indicate items whose co-occurrence frequency with black nodes is especially high. Double-circled nodes are considered keywords.
(4)Links indicate that the connected item pair frequently co-occurs in a data set.
(5)Solid lines form a foundation, which dotted lines connect.

Foundations, which are circles of dotted lines, are obtained from the text data. In Fig. 3, two foundations have strong linkages with event-sets: {doctor, surgery, patient, operation}, and {cancer, medicine, injection}.

B. Outline of KeyGraph

Instead of giving a detailed explanation of KeyGraph, we briefly outline it here. KeyGraph consists of three major components derived from construction metaphors. Each component is described as follows:

1)Foundations: sub-graphs of highly associated and frequent terms that represent basic concepts in the data. A foundation is defined as a cluster that consists of black nodes linked by solid lines. The foundations are underlying common contexts because they are formed by a set of items that frequently co-occur in the data set.

2)Roofs: terms that are highly associated with foundations.

3)Columns: associations between foundations and roofs that are used for extracting keywords, i.e., the main concepts in the data. A column is a dotted line that connects foundations. Although the common context represented by a foundation is widely known, the context represented by a column is not. Columns are important because they connect two common contexts in items that do not frequently occur.

VI. Analysis Results

In the present study, to classify the vast amount of Care Life Log data that occurs in nursing in one Miyazaki Hospital Long-term Health Care Facility by level of care required, text data mining was carried out using KeyGraph. The following analysis results are shown:
(1) care life log by the level of care 1 required : Fig.3
1)Foundation 1: The caregivers have enjoyable conversations with others while using a coloring book.
2)Foundation 2: The cared person has been actively participating in morning and afternoon recreation.
3)Foundation 3: The cared person requires assistance putting on and taking off trousers as well as getting in and out of a wheelchair.
The foundations are obtained from the text data with event-sets 1 { progress, conversation, coloring, use, story, easy, time }, 2 { progress, participation, a.m., p.m., lively }, and 3:{ urination, assistance, inner, boarded, up and down, same room }.
(2) care life log by the level of care 2 required : Fig.4
1)Foundation 1: Cared person was admitted in a wheelchair from the hospital.
2)Foundation 2: They require assistance putting on and taking off clothing, but exhibit an attitude and willingness to perform the task by themselves.
The foundations are obtained from the text data with event-sets 1 { progress, car chair, proficient, establishment, hospital, admission, explanation }, 2 { progress, assistance, symptom, abdomen, behavior, nursing care, personally, willingness, removable, dependent }.
(3) care life log by the level of care 3 required : Fig.5
1)Foundation 1: When caregivers visited the room to change a diaper, the patient had spilled water.
2)Foundation 2: Patient is able to move from a wheelchair to the bed with only minimal assistance and a handrail.
The foundations are obtained from the text data with event-sets 1 { progress, appearance, exchange, diapers, correspondence, daytime }, 2 { progress, assistance, hand, handrail, car chair, behavior, motion }.
(4) care life log by the level of care 4 required : Fig.6
1)Foundation 1: Patient spends time in bed after meals, They cannot talk when being moved to and from a wheelchair.
2)Foundation 2: Patient is confined to a wheelchair, but is able to move the chair using their feet.
The foundations are obtained from the text data with event-sets 1 { progress, correspondence, lunch, morning, car chair, family, sleep, stomach, credit }, 2 { progress, a.m., leg, south, drive, round trip, boarded }.
(5) care life log by the level of care 5 required : Fig.7
1)Foundation 1: The doctor conducted treatment for a pressure ulcer after a patient has bathed.
2)Foundation 2: The doctor transported the patient to the bed after the examination.
The foundations are obtained from the text data with event-sets 1 { implementation, treatment, bathing, sacrum, acne, decubitus }, 2 { progress, doctor, transport, duty, things, medical examination, through }.

VII. Consideration

The following is an overall evaluation of this research:
(1)Due to KeyGraph analysis, characteristic keywords were extracted from nursing and passage records and the formation of foundations was confirmed.
(2)Grasping the characteristics of foundations makes it possible to classify keywords and medical practices in each department.
(3)It may be possible to visualize the work practiced by nurses and doctors.

VIII. Conclusion

In this present study, to classify the vast amount of Care Life Log data that occurs in nursing in one Miyazaki Hospital Long-term Health Care Facility by level of care required, text data mining was carried out using KeyGraph.

Sentences were analyzed into morphemes, and the relations between feature vocabularies were analyzed by a text data mining technique to visualize this information.

In addition, this result identified vocabularies relating to the proper methods of treatment, resulting in a concise summary of the vocabularies extracted from the care life log. Text data mining is expected to become a valuable technique in the analysis of care documents in the future.

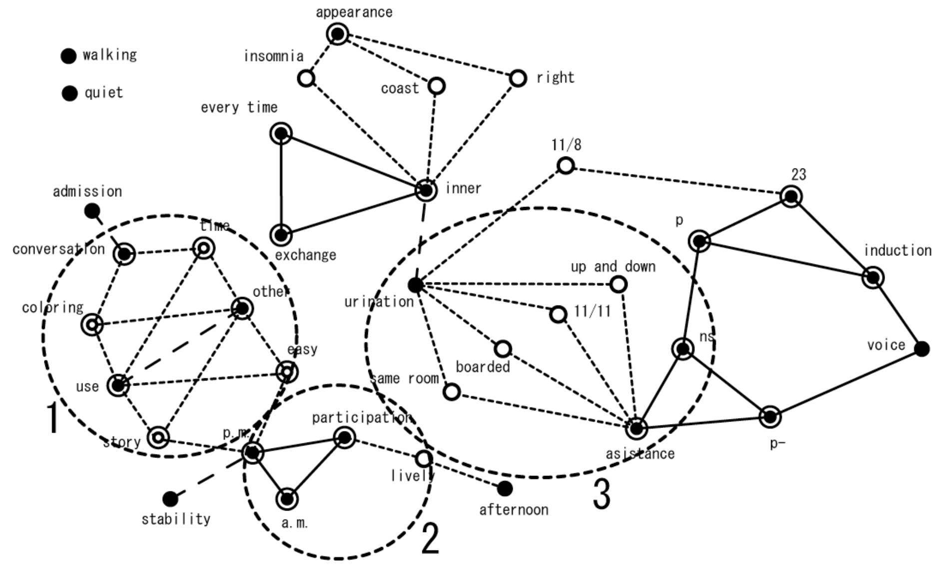

Fig.3 Care life log by the level of care 1 required

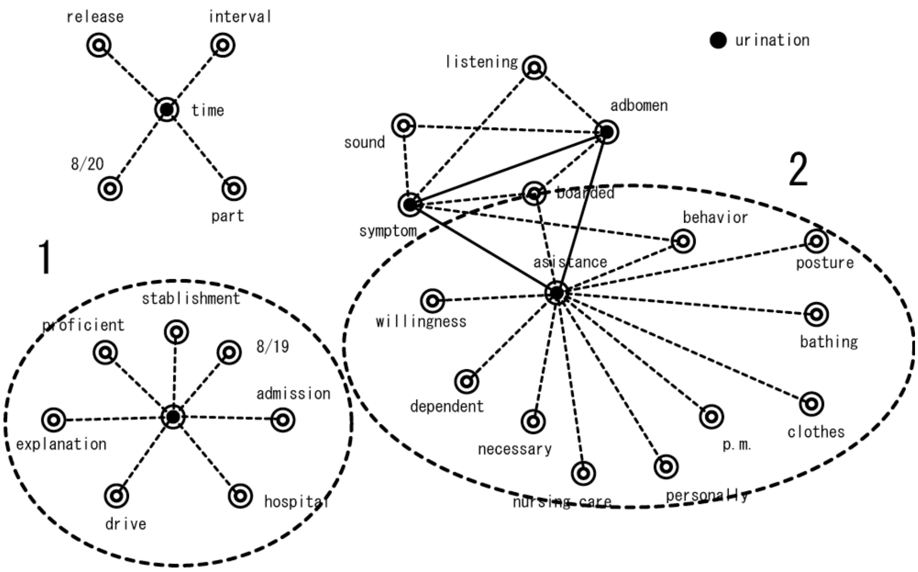

Fig.4 Care life log by the level of care 2 required

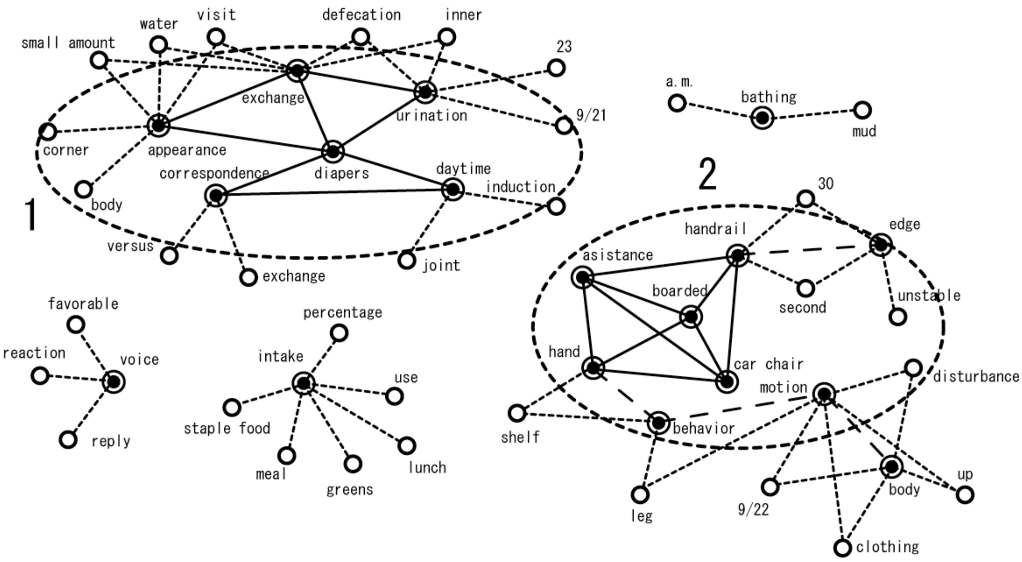

Fig.5 Care life log by the level of care 3 required

ISBN: 978-988-14047-3-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

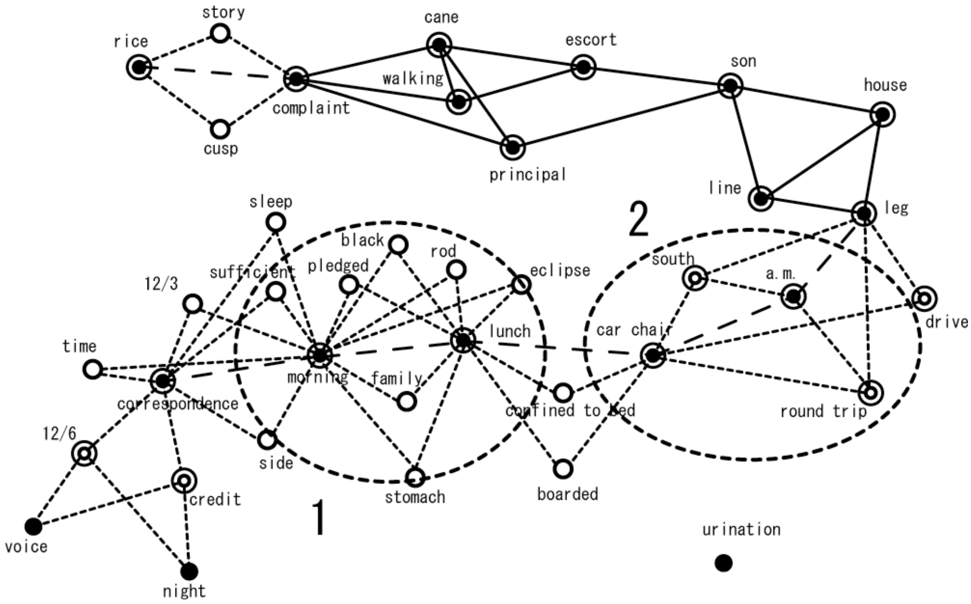

Fig.6 Care life log by the level of care 4 required

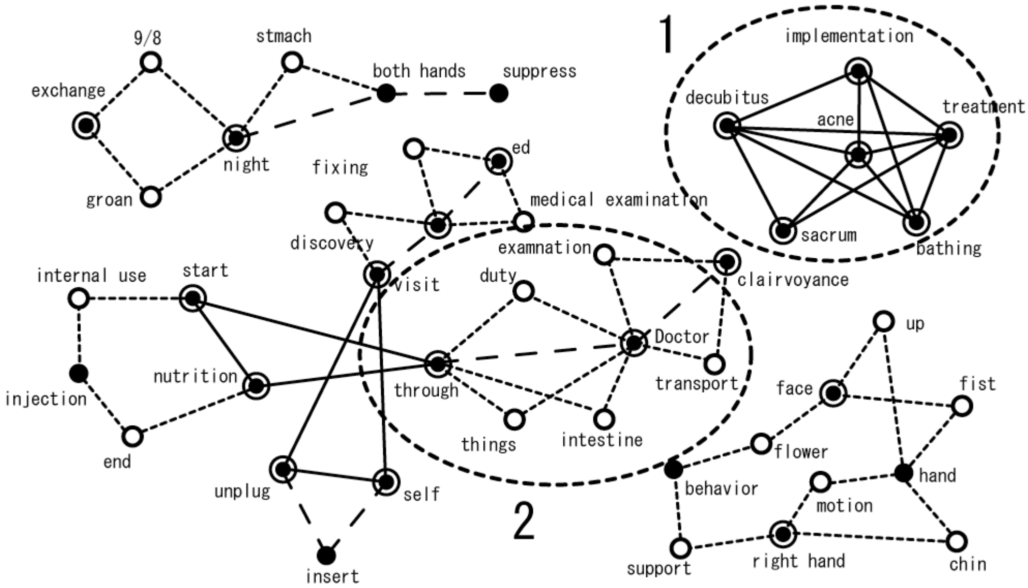

Fig.7 Care life log by the level of care 5 required

ACKNOWLEDGMENT

The authors thank the members of the Medical Information Department at the University of Miyazaki Hospital.

This work was supported by National Institute of Informatics and the Japanese Society for Artificial Intelligence.

REFERENCES

[1] Y. Matsumura, H. Nakano, H. Kusuoka, K. Park, M. Matsuoka, H. Oshima, M. Hayakawa and H. Takeda, “Clinic Hospital Cooperation System Based on The Network Type Electronic Patient Record,” *Japan Association for medical Informatics*, vol. 22, no. 1, pp. 19-26, 2002.

[2] S. Murayama, K. Okuhara and H. Ishii, “Innovation in Manufacturing Premise by New Finding Obtained from Accident Relapse Prevention Report,” *Proceedings of The 13th Asia Pacific Management Conference*, Melbourne, Australia, vol. 13, pp. 1124-1129, 2007.

[3] M. Usui and Y. Ohsawa, “Chance Discovery for Decision Consensus in Company - Touchable, Visible, and Sharable for a Textile Manufacture Company -,” *Japan Society for Fuzzy Theory and Intelligent Informatics*, vol. 15, no. 3, pp. 275-285, 2004.

[4] H. Abe, S. Hirano and S. Tsumoto, “Evaluation Temporal Models Based on Association Mining From Medical Documents,” *Japan Association for Medical Informatics*, vol. 27, no. 1, pp. 33-38, 2007.

[5] K. Izumi, K. Goto and M. Matsui, “Analysis of Financial Markets' Fluctuation by Textual Information,” *The 23rd Annual Conference of the Japanese Society for Artificial Intelligence*, 2009.

[6] Kyung-Joong. Kim, Myung-Chul. Jung and Sung-Bae. Cho, “KeyGraph-based chance discovery for mobil contents management system," *International Journal of Knowledge and Intelligent Engineering System*, Vol.11, pp. 313-320, 2007.

[7] Y. Ohsawa, Benson Nels E and M. Yachida, “KeyGraph: Automatic Indexing by Co-occurrence Graph based on Building Construction Metaphor,” *Fifth International Forum on Research and Technology Advances in Digital Libraries (ADL'98)*, 1998.