

Literature Recommendation to Support Clinical Diagnostic Decision — An Approach Based on Link Analysis in the Weighted Keywords Network

Ying Sun

Abstract—We present an approach of selecting medical literature for the purpose of supporting clinical diagnostic decision. The proposed approach is based on link analysis in the weighted keywords network. The keywords in the network are extracted from the keywords lists in the articles, the Mesh lexicon and acronyms, with which the full relation of keywords and articles (the link network) is built. In this paper we introduced in details how to extract the collection of valid sentences and diagnosis-related keywords based on the keywords co-occurrence analysis and existing description of symptoms. The paper also lays out the detailed process of measuring the relevance between keywords and articles with link analysis in the weighted keywords network. Finally, we report some related experiments and the results of user evaluations.

Index Terms—Clinical decision support, link analysis, weighted keyword network, medical information retrieval.

I. INTRODUCTION

In recent years, deciding how to retrieve the most relevant research articles in the huge collection has become a challenging task of information service and attracted more interests in the research communities as a means to decrease ambiguity and return results which are more likely to meet user requirements effectively. The task of TREC clinical decision support track 2014 is aimed to retrieve biomedical articles useful for answering generic questions of a specified type about a medical case report. One challenge in improving searching medical literature comes from the medical terms. Since medical science is a relatively special research area and its terms have their own distinct characteristics such as using so many special medical terms, acronyms, and even Greek letters. It is also obvious that the task for the clinical diagnosis is always more complicated than the normal query task. We cannot use just a few simple terms to denote symptoms in the diagnosis because even the same symptom may indicate many different possible diseases when it occurs with other different symptoms. Meanwhile, what we need is not any link or text having matched terms but useful articles which can provide more information of detailed treatment according to the existing description of symptoms.

The TREC clinical diagnosis decision support track has three types of tasks, which includes diagnosis, test and

treatment. Accurate test suggestion and further treatment fully relies on effective and accurate diagnosis, which constitutes the key content of this paper and we will make exploration in this research area based on our previous research work.

This paper is organized as follows. In the next section, we will discuss some related background work and our motivation for this research. Principles and detailed approaches for building the collection of standard keywords are described in Section III. In Section IV, we will discuss the detailed process of extracting the relation of keywords and documents with the introduction of keywords normalization. Section V is the most important part in this paper which discusses how to automatically recommend the clinical diagnosis literature based on our collection of relevant sentences and diagnosis-related keywords. The experimental evaluation and results are presented in Section VI. In the final section, we share our conclusions and more plans for future work.

II. RELATED WORKS

The common applications of information technology in the clinical decision support service include tools for information management [1], tools for focusing attention [2], and tools for diagnosis decision support [3]. Medical information management is a wide concept which includes many subtypes such as patient health record management, seamless connection to other information systems, image viewing support, and proper user authentication and data encryption [4]. The most advanced method in this research area is to use mobile platforms for providing service anywhere and anytime [5]. Tools for focusing attention can alert users about clinically important drug–drug interactions [6]. Some tools can give useful suggestions for computerized provider orders such as drug selection and dosage recommendations [7], [8]. Clinical diagnosis means an analysis on the available data collections to determine the pathophysiologic explanation for patient symptoms. Sometimes clinical diagnosis needs to decide which questions to ask, tests to order, or procedures to perform, and it also needs to determine the value of the results relative to associated risks or financial costs [9]. This is a typical evidence-based medicine [10] and related research has also been greatly improved recently.

Effective clinical diagnosis needs three indispensable requirements such as accurate data, adequate knowledge base, and efficient solution which is always an important research point in the recommendation service of medical literatures [11]. Some systems in the early days can conduct diagnostic

Manuscript received April 20, 2015; revised June 9, 2015.

Ying Sun is with the University at Buffalo, the State University of New York, %34 Baldy Hall, Buffalo, NY, 14260 USA (e-mail: sun3@buffalo.edu).

differentiation and give the useful tips about what information is needed for enhancing precision of judgment, such as DXplain [12], [13], QMR [14], [15] and etc. Some other systems can summarize patient records and present physicians with the results in an easy-understanding format [16].

The success of these systems mainly depends on the highly standardized writing format in the patient health records since it will be easier to abstract the customized results from them [17]. Jaspers, Smeulers, Vermeulen and Peute [18] also indicate that the clinical diagnosis decision support system cannot provide enough useful patient outcomes. Although it can help physicians more significantly and has greater practitioner performance, it is essential for physicians to understand these results and also necessary for users to have the relevant knowledge and experiences. Meantime, there are some scholars who point out that the effectiveness of clinical diagnosis decision support systems depends greatly on the degree of acceptance by physicians [19]. Wright, Sittig, Ash, Bates, Feblowitz, Fraser and Middleton [20], [21] think it should be worth more concerns that the corresponding management and governance pay a great role in the application and development of technology of the clinical diagnosis decision support.

By now, it has been concluded that the intelligent diagnosis decision support systems have not been used widely, instead they often focused on the drug selection and prescription support [20], [21]. This current dilemma also contributes tremendously to the further development in this scholarly field. However, in recent years most great achievements have been made in the applications of diagnosis decision support service for some specific diseases since many specific diseases have their own standard diagnostic processes and therapeutic criteria so it will be easier to make these traditionally manual processes automated. For example, Hoeksema, Bazy-Asaad, Lomotan, Edmonds, Ramírez-Garnica, Shiffman and Horwitz [22] point out that the clinical decision support service has been proposed as a promising approach to improve guideline-based care for the treatment of pediatric asthma. This system can provide many useful assessments of impairment, risk, control, severity, and generated treatment recommendations for new patients according to the guidelines for the treatment of pediatric asthma created by the National Asthma Education and Prevention Program in the United States. Other experiments have proved the implementation of evidence-based clinical decision support systems in the emergency department was associated with a significant decrease (20%) in use, and increase (69%) in yield, of CT pulmonary angiography for the evaluation of acute pulmonary embolism [23]. But Romano and Stafford [24] think many quality indicators show no significant difference in the quality between visits with and without clinical decision support systems although they also admit that higher quality has been noted for diet counseling in the high risk adults with the support of these systems. All of these researches give us an inspiration for further work on it.

While there are many different techniques that have been used for implementing these applications, here we examine three techniques which play a critical role in the clinical decision support service: automatic question answering,

artificial intelligence based on the rule bank, and link-based analysis.

One increasingly popular method for the clinical decision support service is automatic question answering. This approach usually decomposes the complex problems into fact-seeking questions or maps them into other similar simple questions [25]. The method combined with the semantic domain model is also a potential solution such as PICO framework. In PICO framework, all diagnoses can be divided into four types: problem, intervention, comparison and outcome. All of the patient symptoms can be reformulated with more normalized expression so that the diagnosis can be achieved by automatic answer extraction and refining from the query results based on existing medical knowledge base and clustering methods of natural language processing [26]. However, this method is significantly dependent on describing symptoms with PICO correctly and effectively [27].

In recent years, many techniques of artificial intelligence based on rule bank have attracted more interests in the research communities as a means to implement the automatic diagnosis. For example, the automated diagnosis of sea cucumber diseases has been implemented with reasoning machine using BP neural network based on rule bank constructed from typical disease cases [28]. There are also many other diagnosis simulation systems of respiration diseases based on BP neural network model [29]. However, it is very crucial that these systems must have a typical and effective training dataset. As a result, the achievements of these approaches have been made mostly in the diagnosis of specific diseases not general diseases.

As opposed to these methods, link-based analysis has wider application and been proved an effective means to deal with the big data in many research areas. There are also many specific types of link-based analysis in the research of clinical diagnosis support service. One type is bipartite network analysis [30]. All diseases and symptoms will be classified into two collections by which their relevance can be calculated with some iterative algorithms. Based on this method, Chen, Lu and Liang [31] explore the pathogenesis of hereditary diseases with the measure of relevance between genetic diseases and virulence genes. However, our experiments have shown that this method cannot get satisfied results if only using keywords as the semantic unit since the information in one keyword is so limited and corresponding resolution power is also lower, while this method is advantageous to the analysis of gene fragment since they have more information and higher degree of differentiation. Another type is hyperlink-based algorithm which has been widely used in the fields such as Web page recommendation. It provides us a feasible way to measure the relevance of articles based on the semantic information of keywords, especially for the recommendation service of medical literature in the clinical diagnosis support service. Since this approach does not need pre-existing normalized expressing of symptoms and many rule banks, it will significantly benefit our research which explores how to retrieve the recommended articles based on the direct descriptions of patient symptoms, which also constitutes the main topic of this paper.

III. COLLECTION OF STANDARD KEYWORDS

To build a keyword network, the first decision is where and how to collection the keywords. We decide to use the keywords provided by the authors of articles.

A. Keyword List

All keywords provided by the authors in the collection are extracted to form the original keyword list. There are some difficulties in extract these keywords correctly. They are: 1). Different medical databases often have different requirements of writing rules and some authors use many unusual formats which makes it very difficult to extract keywords correctly with just one approach. Although all articles are in XML format, however within the element, keywords are separated by different separators. We can only parse with those commonly used ones, e.g. comma, colons, semicolons. 2) Different element names are used by different journals in the collection, such as “keywords”, “key terms”, “key indexing terms”, etc.

In order to solve these problems, we designed a simple yet effective algorithm to extract keywords, which is based on the statistics analysis of term frequency. Its main steps include: Retrieval all the keywords from all articles using delimiter of XML tags;

- 1) Delete the keywords in the stop list and remove invalid ones such as pure digital numbers, field names and keywords that are too long;
- 2) Aggregate respective term frequencies of keywords in each field (such as title, abstract, body text) of each article again;
- 3) Delete the keywords with the lowest document frequency (lower than a specific threshold).

With step 3, we don't need to identify all possible separators and element names. This approach has two other advantages. One is to expand current keyword list for each article by identifying keywords appear in other fields of the article, but not listed as keywords, and included them in the keyword list of the article. Another advantage is to reformulate more accurate keywords list. If some articles have spelling mistakes in their keywords lists, we can still retrieve all the valid keywords from their titles and bodies after scanning all the fields for all the valid keywords and removing ones with lower term frequencies.

B. Enrich the Keyword List with Medical Lexicon

We choose MeSH as a supplementary source of keywords. All the keywords in the MeSH are divided into three hierarchical levels. The top level is Descriptor, the second one is Concept and the lowest level is Term. Each concept belongs to one descriptor and each term belongs to one concept. Concept is the most formal and most useful content in the MeSH since the role of descriptor is mainly about grouping related concepts into the existing classes and each concept always has many terms which usually are some corresponding synonyms with different spellings. Concepts in the MeSH include all descriptors literally.

It is more helpful that MeSH also provides the mapping between different concepts. We can obtain more synonyms and related keywords combined with this mapping of concepts and the hierarchy of MeSH so that all the keywords

can be expanded in each article and recall of retrieval can be improved enormously. For example, if there is a correlative mapping between two concepts, we can add all the terms belonging to one concept into the keywords list of another concept respectively and vice versa.

C. Acronyms

Acronyms have the great importance in describing symptoms and diseases in the medical literature. Although they seldom occur in keywords lists, they are widely used in the body text. However, there are not so many available lexicons of acronyms available and most of them are incomplete. It is necessary to explore a novel approach to obtain the lexicon of acronyms automatically for our research.

The ideal scenario would be that each acronym occurs in the articles with corresponding original term as well. But it is not true in most of the cases since the writing style and abbreviation rules have greater diversities. For example, some acronyms do not use all of first capital letters of all terms and some of them even use lowercase letters. Some articles only use acronyms without referring to their corresponding original terms at all. Based on these observations, we designed a two-step extracting algorithm. The first step is to find out all the pairs of acronyms and their corresponding original terms from which we can extract all the valid mapping relations between them. The second step is to make judgment for each acronym and decide what its true meaning is in one article. We will discuss the second step in the next section and here only the first step is introduced.

The main idea of our heuristics algorithm in the first step is the way in which the pairs of acronyms and corresponding original terms always follow a basic occurring rule of format: $A*B*(AB*)$. A and B here denote two capital letters of the first two sequent keywords in the original term since nearly all the acronyms do use these two capital letters as the beginning. Asterisk means any letters, digitals or Greek letters. This algorithm will scan all the body texts and retrieve the strings having this feature. The detailed steps include:

- 1) Retrieve all the opening curly braces in each article;
- 2) Check whether the first two letters after each opening brace are capital;
- 3) If yes, retrieve the position of these two letters before opening brace. If their order is same as the one after brace, all the keywords from the first capital letter to brace can be extracted as the original term and then be further checked whether it includes some invalid characters such as colons, periods and et al. Acronym can be also extracted from the brace to the first non-alphabetical or non-numeric character after the brace. Then a pair of acronym and original term can be generated;
- 4) Validate and filter again after extracting all the pairs of acronym and original term. For example, delete the pairs which length of acronym is equal to or longer than original term, delete the pairs which acronym is part of other acronyms and both of them have the same original terms, delete the pairs which letters in the acronym occur in the corresponding original term with different order;
- 5) Aggregate the document frequencies of all the acronyms and delete ones with lower document frequency since they may be wrong-spelt or have no sense at all even if

they are literally correct.

IV. BUILDING KEYWORDS AND DOCUMENTS NETWORK

A. Relationship Extraction Strategies

Different strategies of extraction are used according to the different types of keywords.

For the keywords in the combined keyword list, we can scanned each document and extract all the keywords from the body texts directly.

For the keywords in the MeSH lexicon, the mapping relation of concepts can be also used for the expansion of current keyword list for each article. That is to say, if a MeSH term is a keyword of one article, terms within the same concept with the term will be added into the keyword list of this article with same term frequency.

For acronyms, after the basic extraction finished using method introduced in previous section, the second step is to scan all the acronyms case-sensitively again in each article. Since one acronym may refer to many different original terms, it is necessary to decide which one is most likely correct. The algorithm based on the analysis of term frequency is used once more. The basic procedure is to scan all the original terms of each acronym in each article and choose the original term with the highest term frequency as the candidate. If there are many candidates with the same term frequency, one with the longest length will be chosen as the final corresponding term. All the acronyms will be replaced with the corresponding term after processing so that there is no acronym any more in the final relation of keywords and documents.

B. Normalization of Keywords

To improve effectiveness, all keywords are normalized using existing dictionaries and simple stemming heuristic rules such as deleting the last letter "s" if the keyword will be the same as another keyword in the list.

We selected three levels of analysis: sentence level, segment level and the whole article level. Segment level analysis is based on the XML tag. We aggregate all the term frequencies of each keyword in these three levels.

V. AUTOMATIC RECOMMENDATION SUPPORT CLINICAL DIAGNOSTIC DECISION

The information of patient symptoms can be collected and summarized by the clinicians, which often includes healthy conditions, current status, existing diseases, related symptoms and et al. The proper diagnosis and necessary tests will be suggested by the clinicians with their own experiences and knowledge in the past. It is obviously useful that these medical literatures full of diagnoses and case descriptions can provide a profuse and sound foundation for the clinical decision support.

This case description is usually the source of query terms in medical information retrieval. However, it usually includes a relatively large number of terms and the term frequencies in the query are mostly no more than 1. It will increase false positive with the retrieval of so many non-relevant articles if

considering all the keywords in the description. It is necessary to measure the importance of each keyword in the description and those with the highest weight should be selected as the final query terms. There are two steps needed. The first step is to sort all the keywords on the reversed order of document frequency and those with lower document frequency should be removed. The second step is to retrieve the pairs of keywords occurring in the same sentence since the combination of keywords will be more useful to reveal the possible diseases. The sentences are better units for this co-occurrence analysis because the descriptions of symptoms are usually limited in one segment. The numbers of keywords in the co-occurring pairs are usually 2 while the fewer are 1 or 3. The manual selection and expert decision by the clinicians should be introduced into this process for the enhancement of describing the necessary and significant symptoms. This is the only work in our research which needs the extra intelligent help manually while we also think it is natural and necessary for achieving the more precise recommendation.

So far we have constructed two important data sets: one is the term frequencies of keywords in three levels of occurrence unit: the sentences, the segments and the body texts. The second is the selected pairs of keywords from the description of patient's symptoms. It has been proved only utilizing one type of occurrence unit is not effective enough to make the decision correctly. The sentences can provide a more precise and high-related base for the occurrence analysis but many potentially useful occurring keywords may be ignored inevitably since many articles use a longer text to describe the symptoms of patient. However, the segments and the body texts also induce so many insignificant co-occurrences of keywords.

We design a novel method which synthesizes respective advantages of these analyses based on the different levels of units. For each selected pair of keywords from the description of patient, each sentence will be chosen as a basic analyzing unit firstly and checked whether it consists of at least one pair of keywords. Those sentences with more pairs above the specific threshold will be selected as the valid sentences and those segments having more valid sentences will be also selected as the valid segments. The basic idea of this process is that we think most relevant articles can introduce the possible diseases or some other useful information when these valid segments appear in them. In order to improve the precision of retrieval, each segment will be filtered again. The segments will be grouped by their occurring articles and those articles with less valid segments will be removed with included segments. At last, all the sentences in the final collection of valid segments constitute a collection of the valid sentences.

A. Retrieval of the Diagnosis-Related Keywords

The diagnosis-related keywords can be helpful to reveal the possible diseases and corresponding treatment or give more clues and suggestions for the clinicians. From the collection of valid sentences, the diagnosis-related keywords can be extracted by the link-based algorithm combined with weighting measurement.

The typical PageRank is an efficient link-based algorithm of measuring weight of each Web page in network. The core

Web pages often are linked by more other Web pages so that they have more in-degrees. Just as the requirement of hyperlink analysis and PageRank algorithm, each keyword in the collection of valid sentences can be seen as a node and the links between these nodes can be got by the co-occurrence of corresponding keywords. Since the link-based algorithm needs directed links, the direction of each link will be set from the keyword with the higher document frequency to one with lower document frequency. This design is based on the assumption that some parts of one article also discuss the relevant information of diagnosis if they introduce many important symptoms and those keywords with the lower document frequency are usually more important and have the higher resolution power. Based on this design of algorithm, these keywords can gain more inflowing weights in the iterative computation.

The weight of each node of keyword can be calculated as:

$$weight_{n+1}(keyword_k) = (1 - \alpha) + \alpha \times \sum \frac{idf(inKeyword_i) \times weight_n(inKeyword_i)}{C(inKeyword_i)} \quad (1)$$

The parameter n means the count of iteration and C means the out-degree of a node. Here the parameter inKeyword means the keywords at the beginning of the link. This algorithm pays more attention to the weight of each node of keyword based on the document frequency so the document frequency is combined in this calculation. That is to say, only the keywords with higher document frequencies and more in-degrees can get higher weights.

B. Measuring the Relevance of Articles

The relevance of all the related articles in the collection of valid sentences can be calculated by the average weights of corresponding keywords. More keywords with higher weights an article has, more relevant it is. The basic measure of one article's relevance is shown as:

$$weight(article_k) = avg(weight(keyword_i^{articlek})) \quad (2)$$

VI. EXPERIMENTS

A. Data Set

We used the TREC clinical decision support track 2014 collection. It includes 733,328 articles in the Open Access Subset of PubMed Central (PMC).

B. The Experiment of Extracting the Collection of Keywords

Using the extraction method introduced in Section III, the total number of keywords from the keywords lists is 347,999. The stop word list we used is from Reuters Corpus Volume 1 which has 25 common terms. In the MeSH lexicon, the total number of descriptions is 27,149, 51,525 concepts and 218,985 terms. The number of mapping relations between concepts and terms is 49,412. The total number of pairs of

acronyms and corresponding original terms is 289,670 before validation and 139888 after validation.

After duplication removal and aggregation, we can get 311,379 keywords in the final collection. The total number of occurrence pairs of keywords and sentences is 881,930,831 in which there are 148,757,694 sentences and 24,054,364 segments.

C. The Recommending Results of Medical Articles

We test 10 standard topics and 1 demo topic. Each result only displays at most 5 articles with the highest relevance. Here two typical test topics and their results are introduced below:

1) Topic description: A woman in her mid-30s presented with dyspnea and hemoptysis. CT scan revealed a cystic mass in the right lower lobe. Before she received treatment, she developed right arm weakness and aphasia. She was treated, but four years later suffered another stroke. Follow-up CT scan showed multiple new cystic lesions. Possible relevant articles are shown in Table I:

TABLE I: THE RESULTS OF DEMO TOPIC

DocID	Title
3025345	A Case of a Pulmonary Arteriovenous Malformation With Ebstein's Anomaly
3148967	Stroke in hereditary hemorrhagic telangiectasia patients. New evidence for repeated screening and early treatment of pulmonary vascular malformations: two case reports
3180463	Infarctions in the vascular territory of the posterior cerebral artery: clinical features in 232 patients
3287708	Maximising adherence to study protocol within pharmaco-rehabilitation clinical trials
3420796	Sporadic Hemiplegic Migraine with Seizures and Transient MRI Abnormalities

The pairs of keywords we choose include: dyspnea and hemoptysis, mass and lobe, weakness and aphasia, stroke. After scanning, 88512 valid segments are obtained and the number of corresponding articles is 30622. After filtering, 58 valid segments are confirmed which include 1728 keywords with 38456 directed links, within 625 sentences and 52 articles. We can conclude from the Table I that the first two articles are mainly about the pulmonary arteriovenous malformation and highly relevant to the diagnosis result. Although the third one discusses migraine and the fourth one is mainly about the infarction in posterior cerebral artery, they also can provide a useful comparison for the further diagnosis. The last one discusses the rehabilitation in stroke with less useful information and lower relevance. Compared with the existing suggested three results, our result only has the first one of them (docid: 3148967). The reason is that other two articles in the suggested three results have not been chosen as the candidates of valid articles according to our algorithm. For example, the second article (docid: 2987927) has no sentence with more than two valid keywords since it discusses those symptoms dispersedly within a so long text. The third one (docid: 3082226) has no description of symptom at all. This result also reveals some possible shortcomings in our approach since it mainly use the keywords of symptoms as the query terms.

2) Topic description: A 58-year-old African-American woman presents to the ER with episodic pressing/burning

anterior chest pain that began two days earlier for the first time in her life. The pain started while she was walking, radiates to the back, and is accompanied by nausea, diaphoresis and mild dyspnea, but is not increased on inspiration. The latest episode of pain ended half an hour prior to her arrival. She is known to have hypertension and obesity. She denies smoking, diabetes, hypercholesterolemia, or a family history of heart disease. She currently takes no medications. Physical examination is normal. The EKG shows nonspecific changes. Possible relevant articles are shown in Table II:

TABLE II: THE RESULTS OF TOPIC 1

DocID	Title
2994533	Primary pericardial malignant mesothelioma and response to radiation therapy
3258729	Epipericardial fat necrosis – a rare cause of pleuritic chest pain: case report and review of the literature
3490454	Does the patient with chest pain have a coronary heart disease? Diagnostic value of single symptoms and signs – a meta-analysis
3681367	Medical emergencies on board commercial airlines: is documentation as expected?

The first article is highly relevant since it mainly discusses the primary pericardial malignant mesothelioma and has the same description of symptoms in the topic case. The second one is a literature review about in-flight medical emergencies. It is not a relevant article but has so many valid keywords coming from review text so that it gains a higher weight. This result also reveals the same shortcoming of purely using the keywords of symptoms as the query terms. The third one is relevant since it discusses the epipericardial fat necrosis. The last one is mainly about the chest pain and coronary heart disease which provides many useful information and comparison cases.

D. User Evaluation

All the 10 topic cases without suggested answer and 1 demo topic case have been tested by the approach of this paper. We invite 11 users to evaluate the results of experiment. 5 users are clinicians with the specific medical science knowledge and 6 users major in library and information science with the rich comprehension of information retrieval. Each user is asked to evaluate 2 queries and mark his/her satisfaction for the results with the top five highest weights. We use a Likert scale of 5 in which 5 is satisfied and 1 is not satisfied. The results of evaluation are shown in Table III:

TABLE III: THE USER EVALUATION RESULTS

The order of article	1 st	2 nd	3 rd	4 th	5 th
User 1	5	5	4	5	3
User 2	5	5	2	1	4
User 3	5	4	5	5	4
User 4	5	5	5	4	4
User 5	5	5	2	3	1
User 6	4	4	5	5	
User 7	4	3	3	4	3
User 8	3	4	3	2	5
User 9	5	5	5	3	4
User 10	5	5	5	5	4
User 11	1	2	3	4	5

Since some results of queries are less than 5, there are some blank in Table III. We can conclude that the total satisfaction

of all users is 3.98 which is 33% higher than the average satisfaction. The results also tell us that the fewer numbers of results provided, the higher average satisfaction is. The detailed results are shown in Table IV:

TABLE IV: THE AVERAGE SATISFACTION AT TOP 5 RANKS

The top n articles	Average satisfactory
1	4.27
2	4.07
3	3.94
4	3.95
5	3.98

VII. CONCLUSIONS AND OUTLOOK

We have presented an approach of medical literature recommendation for the clinical diagnostic decision support and demonstrated that the semantic analysis based on link-based algorithm combined with the keyword co-occurrence analysis can be used to effectively retrieve the relevant articles according to the description of patient symptoms. In our future work, we plan to continue evaluating the stability and expand our data collection in other areas for conducting a wider and more user evaluation. As mentioned above, purely utilization of the semantic information of keywords can cause some difficulties when the articles are too long or their contents are too different from the description of information need although these articles are really relevant ones. Our future work will also involve designing the experiments that will allow us to combine more useful information in other fields and use some other approaches to enhance the ability of this algorithm.

REFERENCES

- [1] M. Akiyama, "Risk management and measuring productivity with POAS-Point of Act System-A medical information system as ERP (enterprise resource planning) for hospital management," *Methods of information in medicine*, vol. 46, no. 6, pp. 686-693, 2007.
- [2] A. B. McCoy, L. R. Waitman, J. B. Lewis, J. A. Wright, D. P. Choma, R. A. Miller, and J. F. Peterson, "A framework for evaluating the appropriateness of clinical decision support alerts and responses," *Journal of the American Medical Informatics Association*, 2011.
- [3] M. Poch, J. Comas, I. Rodriguez-Roda, M. Sanchez-Marre, and U. Cortés, "Designing and building real environmental decision support systems," *Environmental Modelling & Software*, vol. 19, no. 9, pp. 857-873, 2004.
- [4] C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing Cloud Computing and Android OS," in *Proc. Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pp. 1037-1040, IEEE, Aug, 2010.
- [5] I. L. Wu, J. Y. Li, and C. Y. Fu, "The adoption of mobile healthcare by hospital's professionals: an integrative perspective," *Decision Support Systems*, vol. 51, no. 3, pp. 587-596, 2011.
- [6] K. R. Saverno, L. E. Hines, T. L. Warholak, A. J. Grizzle, L. Babits, C. Clark, and D. C. Malone, "Ability of pharmacy clinical decision-support software to alert users about clinically important drug-drug interactions," *Journal of the American Medical Informatics Association*, vol. 18, no. 1, pp. 32-37, 2011.
- [7] A. Moxey, J. Robertson, D. Newby, I. Hains, M. Williamson, and S. A. Pearson, "Computerized clinical decision support for prescribing: provision does not guarantee uptake," *Journal of the American Medical Informatics Association*, vol. 17, no. 1, pp. 25-33, 2010.
- [8] G. J. Kuperman, A. Bobb, A. J. Payne, T. K. Gandhi, G. Burns, and D. W. Bates, "Medication-related clinical decision support in computerized provider order entry systems: a review," *Journal of the*

- American Medical Informatics Association, vol. 14, no. 1, pp. 29-40, 2007.
- [9] M. A. Musen, M. Blackford, and R. A. Greenes, "Clinical decision-support systems," *Biomedical informatics*, Springer, London, 2014, pp. 643-674.
- [10] K. Khan, R. Kunz, J. Kleijnen, and G. Antes, *Systematic Reviews to Support Evidence-Based Medicine*, CRC Press, 2011.
- [11] C. Safran, M. Bloomrosen, W. E. Hammond, S. Labkoff, S. Markel-Fox, P. C. Tang, and D. E. Detmer, "Toward a national framework for the secondary use of health data: An American Medical Informatics Association white paper," *Journal of the American Medical Informatics Association*, vol. 14, no. 1, pp. 1-9, 2007.
- [12] P. L. Elkin, M. Liebow, B. A. Bauer, S. Chaliki, D. Wahner-Roedler, J. Bundrick, and G. Barnett, "The introduction of a diagnostic decision support system (DXplain™) into the workflow of a teaching hospital service can decrease the cost of service for diagnostically challenging Diagnostic Related Groups (DRGs)," *International Journal of Medical Informatics*, vol. 79, no. 11, pp. 772-777, 2010.
- [13] G. O. Barnett, J. J. Cimino, J. A. Hupp, and E. P. Hoffer, "DXplain: An evolving diagnostic decision-support system," *Jama*, vol. 258, no. 1, pp. 67-74, 1987.
- [14] D. Heckerman, "A tractable inference algorithm for diagnosing multiple diseases," 2013.
- [15] M. A. Shwe, B. Middleton, D. E. Heckerman, M. Henrion, E. J. Horvitz, H. P. Lehmann, and G. F. Cooper, "Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base," *Methods of information in Medicine*, vol. 30, no. 4, 241-255, 1991.
- [16] D. Klimov and Y. Shahar, "iALARM: An intelligent alert language for activation, response, and monitoring of medical alerts," *Process Support and Knowledge Representation in Health Care*, pp. 128-142, Springer International Publishing, 2013.
- [17] N. Elhadad, M. Y. Kan, J. L. Klavans, and K. R. McKeown, "Customization in a unified framework for summarizing medical literature," *Artificial Intelligence in Medicine*, vol. 33, no. 2, pp. 179-198, 2005.
- [18] M. W. Jaspers, M. Smeulders, H. Vermeulen, and L. W. Peute, "Effects of clinical decision-support systems on practitioner performance and patient outcomes: A synthesis of high-quality systematic review findings," *Journal of the American Medical Informatics Association*, vol. 18, no. 3, 327-334, 2011.
- [19] H. M. Seidling, S. Hansalkar, D. Leger, M. D. Paterno, S. Shaykevich, W. E. Haefeli, and W. Bates, "Factors influencing alert acceptance: A novel approach for predicting the success of clinical decision support," *Journal of the American Medical Informatics Association*, vol. 18, no. 4, pp. 479-484, 2011.
- [20] A. Wright, D. F. Sittig, J. S. Ash, D. W. Bates, J. Feblowitz, G. Fraser, and B. Middleton, "Governance for clinical decision support: case studies and recommended practices from leading institutions," *Journal of the American Medical Informatics Association*, 2011.
- [21] A. Wright, D. F. Sittig, J. S. Ash, J. Feblowitz, S. Meltzer, C. McMullen, and B. Middleton, "Development and evaluation of a comprehensive clinical decision support taxonomy: Comparison of front-end tools in commercial and internally developed electronic health record systems," *Journal of the American Medical Informatics Association*, vol. 18, no. 3, pp. 232-242, 2011.
- [22] L. J. Hoeksema, A. Bazzy-Asaad, E. A. Lomotan, D. E. Edmonds, G. Ramirez-Garnica, R. N. Shiffman, and L. I. Horwitz, "Accuracy of a computerized clinical decision -support system for asthma assessment and management," *Journal of the American Medical Informatics Association*, vol. 18, no. 3, pp. 243-250, 2011.
- [23] A. S. Raja, I. K. Ip, L. M. Prevedello, A. D. Sodickson, C. Farkas, R. D. Zane, and R. Khorasani, "Effect of computerized clinical decision support on the use and yield of CT pulmonary angiography in the emergency department," *Radiology*, vol. 262, no. 2, pp. 468-474, 2012.
- [24] M. J. Romano and R. S. Stafford, "Electronic health records and clinical decision support systems: Impact on national ambulatory care quality," *Archives of Internal Medicine*, vol. 171, no. 10, pp. 897-903, 2011.
- [25] S. Y. Ou, "An entailment-based question answering method in a restricted domain," *Journal of the China Society for Scientific and Technical Information*, vol. 30, no. 5, pp. 540-547, 2011.
- [26] D. Demner-Fushman, *Complex Question Answering Based on a Semantic Domain Model of Clinical Medicine*, 2006.
- [27] X. Huang, J. Lin, and D. Demner-Fushman, "Evaluation of PICO as a knowledge representation for clinical questions," *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, 2006.
- [28] F. Li, S. J. Han, and D. Zhang, "The construction of sea cucumber disease diagnosis inference engine," *Computer Applications and Software*, vol. 29, no. 12, pp. 211-213, 2012.
- [29] Z. X. Huang, C. Zhong, and X. R. Li, "Simulation study of respiratory disease diagnosis based on BP neural network," *Journal of Hefei University of Technology (Natural Science)*, vol. 35, no. 3, pp. 347-349, 2012.
- [30] S. Q. Li, X. Xu, and M. J. Xu, "The measures of books' recommending quality and personalized book recommendation service based on bipartite network of readers and books' lending relationship," *Journal of Library Science in China*, no. 3, pp. 83-95, 2013.
- [31] W. Q. Chen, J. A. Lu, and J. Liang, "Research in disease-gene network based on bipartite network projection," *Complex Systems and Complexity Science*, vol. 6, no. 1, pp. 13-19, 2009.



Ying Sun received her B.S. in information science from Peking University in 1996. In 1999, she received her MLS degree from the Information Management Department at Peking University. She entered the Ph.D. program in School of Communication, Information and Library Studies, Rutgers University, in 1999. She received her Ph.D from Rutgers, the State University of New Jersey in October 2005. She also obtained a master degree in computer science (also at Rutgers) while completing her Ph.D studies in 2003. Dr. Ying Sun is an associate professor of information studies at the University at Buffalo, the State University of New York. Her research interests are in the areas of information retrieval (searching for information), text mining (detecting significant patterns or trends) and formal information system evaluation.