Journal of information and
communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 15(2): 85-90, Jun. 2017

Regular paper

Developing an Intelligent Health Pre-diagnosis System for Korean Traditional Medicine Public User

Kwang Baek Kim[*]**,** ***Member, KIICE***

Division of Computer Software Engineering, Silla University, Busan 46958, Korea

Abstract

Expert systems for health diagnosis are only for medical experts who have deep knowledge in the field but we need a self-checking pre-diagnosis system for preventive public health monitoring. Korea Traditional Medicine is popular in use among Korean public but there exist few available health information systems on the internet. A computerized self-checking diagnosis system is proposed to reduce the social cost by monitoring health status with simple symptom checking procedures especially for Korea Traditional Medicine users. Based on the national reports for disease/symptoms of Korea Traditional Medicine, we build a reliable database and devise an intelligent inference engine using fuzzy c-means clustering. The implemented system gives five most probable diseases a user might have with respect to symptoms given by the user. Inference results are verified by Korea Traditional Medicine doctors as sufficiently accurate and easy to use.

Index Terms: Disease/symptoms database, FCM, Fuzzy inference, Korean traditional medicine, Self-health diagnosis

I. INTRODUCTION

Computer-assisted medical diagnosis system has long history of research and requires analyzing bulky test cases, gathering field experts' opinions, and uncertainty management algorithms. Usually such diagnostic systems are designed for medical doctors giving more accurate decision-making and therefore the diagnosis/treatment consultation is limited to a certain set of diseases or body parts [1, 2] with deep knowledge.

However, from the view of public health management, more important issue is to find the possible disease as early as possible by the general public based on vague but noticeable symptoms like chronic abdominal pain for example. That is, people may need pre-diagnosis type healthcare software accessible online with some degree of professional knowledge/diagnostic functions with comprehensible vocabulary for general public. The system may not need to be as accurate as the medical doctor but its role is to draw attention to the public to listen to their body for preventive public healthcare.

Another concern of ours in this research is that for East Asian countries like Korea and China, traditional complementary and alternative medicine (CAM) is widely used. A report was published with usage rates ranging from 29% to 53% among various patient populations. CAM also accounts for a large share of healthcare costs [3]. Korean traditional medicine (KTM, frequently referred as Hanbang) was originated in ancient and prehistoric times and can be traced back as far as 3000 B.C. and the treatment includes herbal medicine, acupuncture, moxibustion, and aromatherapy. The legendary textbook of Korean traditional medicine, Dongeui Bogam [4], written in 17th century, was registered to UNESCO Memory of the world in 2009 [5].

Received 29 May 2017, **Revised** 31 May 2017, **Accepted** 07 June 2017
***Corresponding Author** Kwang Baek Kim (E-mail: gbkim@silla.ac.kr, Tel: +82-51-999-5052)
Division of Computer Software Engineering, Silla University, 140, Baegyang-daero 700beon-gil, Sasang-gu, Busan 46958, Korea.

Open Access http://doi.org/10.6109/jicce.2017.15.2.85 print ISSN: 2234-8255 online ISSN: 2234-8883

Even for Korean American elderly society, health service utilization pattern includes or both Western and traditional clinics [6]. Thus, there have been great deals of effort to have cooperated treatment with western medicine [7] in the last decade. Interestingly, similar patterns of combining traditional medicine and western medicine are found in China [8] and North Korea [9].

However, diagnosis based on KTM is not easy to understand by the public as its inference mechanism is largely metaphorical or abstract. One of the main difficulties in building informative pre-diagnosis or self-diagnosis system for KTM is to build a reliable symptom-disease database and disease classification system available for western medicine [10, 11].

Recently the government agent [12] published the Korean Standard Causes of Death Disease Classification Index (KCD) for such a Korean specialized database build-up. Thus, we can construct a reliable database based on various reports submitted to the government about “Diseases burdensome to Korean Patients” in 2005 and medical contents “Engel Pharm” with 60 diseases [5] and continuously evolved since then.

In this paper, we propose a self-health pre-diagnosis system based on KTM and related symptom/disease database we have built upon [5]. Since any medical diagnosis by computer software should deal with highly uncertain decision making, there have been various mechanisms attempted such as fuzzy neural network [13] and a hybrid system combining case based reasoning and fuzzy decision tree [14]. The uncertainty management system we adopt is fuzzy c-means clustering (FCM) [15, 16] that has been further developed and optimized to adjust this application.

Since the proposed system is not an expert system for medical practitioners but a pre-diagnosis system for general public who has limited knowledge in medicine, we provide a simple and convenient user interface to guide the user. A user simply chooses the uncomfortable body parts and major symptoms he/she feels. With that answer, the system computes the similarity for diseases collected in the database (60 in total) and shows top five most probable diseases to the requested user. That inference is based on FCM.

II. METHODS

A. Standardizing Symptom/Disease Database

Various reports and clinical statistics were collected by the Korean government [12, 17] in 2005 and our database is constructed based on those reports. We collected 60 major diseases and 161 representative symptoms on 22 different body parts related with those diseases frequently found among Korean people. Collected information is verified by KTM doctors for reliability.

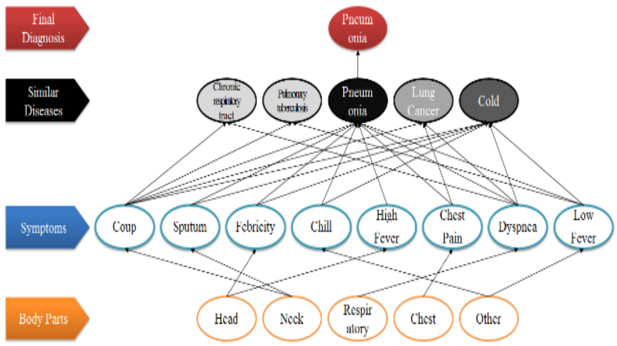


Fig. 1. Disease-symptoms-body parts associations.

For those collected 60 major diseases, representative symptoms are collected by textbook analysis and other medical contents collecting agencies and the associations between then is shown as Fig. 1 that is verified by medical doctors. The database structure can be found from our previous studies [5, 16].

When a user chooses a body part that one feels uncomfortable, the user is guided to select one of the representative symptom he/she feels. Then the system enumerates all symptoms of diseases that include user’s choice as one of the symptoms. Then the user further chooses other symptoms that occur to him/her. Such collected information is given as an input vector to FCM algorithm to infer most similar disease with respect to symptoms the user gives through that query process. The darkness of the uppermost disease part in Fig. 1 demonstrates the similarity measured by FCM as an illustration of out inference system.

B. Proposed FCM Algorithm

In FCM algorithm, every point has a degree of belonging to clusters. In our system, we check the similarity of input vector that consists of a set of symptoms the user feels with disease clusters that have their own symptom sets constructed by standardized database. The distance is in general controlled by fuzzy inference rules provided. However, in this medical domain, there might exist vague clustering results because of outlier distribution of patterns. That may result an odd-one-out disease extraction while others are mutually related.

Thus, in this paper, we apply fuzzy theory to the symmetry measure of clusters and practically perform re-clustering the result from the original FCM to improve the clustering accuracy. The symmetry measure can be represented as formula (1).

$$Symmetric(x_i, c) = \max_{j=\forall pattern(i \neq j)} \left((1-\alpha)(1-\frac{\deg(x_i, x_j, c)}{180} - (\alpha \cdot ratio_d(x_i, x_j, c)) \right), \quad (1)$$

where $\deg(x_i, x_j, c)$ denote the angles of point x_i and x_j from point ratio(x) is obtained by formula (2) and α is the weighting factor.

$$ratio_d(x) = \begin{cases} \dfrac{d(x_j, c)}{d(x_i, c)} & if \quad d_i > d_j \\ \dfrac{d(x_i, c)}{d(x_i, c)} & if \quad d_i < d_j \end{cases}. \quad (2)$$

Let $u(x)$ be the Symmetric number obtained by formula (1) and the center of the cluster is computed by formula (3).

$$u^{(p)} = u(x)x_k / u(x). \quad (3)$$

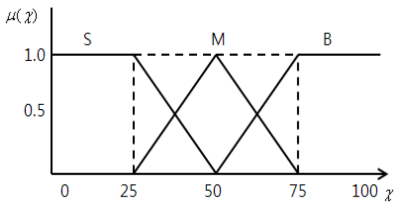
The re-clustering of original FCM clustering is performed based on fuzzy inferences.

We provide membership functions for the number of agreement symptoms between input and the cluster and total number of symptoms the user provided to the system. Then the defuzzified value is used to control the distance as shown in formula (4).

$$D_k = \sum_{t=1}^{c} \left(\left| x_k - v_i \right|^2 \right) \times (1.0 - O_t), \quad (4)$$

where x_k is the user input vector and v is the center of cluster. The value of O_t is the defuzzified value obtained by designed fuzzy inference rules combining the membership degree of symptom agreement and degree of number of symptoms provided.

The membership function for the agreement of user input and disease symptoms is defined as below where the intervals are represented as [Small, Medium, Large].



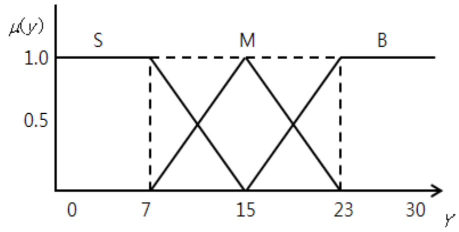
The membership degree of S (small agreement interval), M (medium agreement interval), and L (large agreement interval) are computed by formula (5), (6), and (7), in respectively.

$$\begin{cases} if\ (X \geq 50) & then\ u(X) = 0 \\ else\ if\ (X \geq 50) & then\ u(X) = -1 \cdot (X - 50)/(50 - 25) \\ else & then\ u(X) = 1 \end{cases}, \quad (5)$$

$$\begin{cases} if\ (X \geq 75 or X \leq 25) & then\ u(X) = 0 \\ else\ if\ (X < 25) & then\ u(X) = (X - 25)/(50 - 25) \\ else\ if\ (X \geq 50) & then\ u(X) = -1 \cdot (X - 75)/(75 - 50) \end{cases}, \quad (6)$$

$$\begin{cases} if\ (X \leq 50) & then\ u(X) = 0 \\ else\ if\ (X < 75) & then\ u(X) = (X - 50)/(75 - 50) \\ else & then\ u(X) = 1 \end{cases}. \quad (7)$$

The number of symptoms provided by the user is also a factor in decision making. Again, the input intervals are separated by Small, Medium, and Large as shown below and the computation formulas are represented from formula (8) to (10) in respectively.



where Y denote the number of symptoms the user provided.

$$\begin{cases} if\ (Y \geq 15) & then\ u(Y) = 0 \\ else\ if\ (Y \geq 7) & then\ u(Y) = -1 \cdot (Y - 15)/(15 - 7) \\ else & then\ u(Y) = 1 \end{cases}, \quad (8)$$

$$\begin{cases} if\ (Y \geq 23 or Y \leq 25) & then\ u(Y) = 0 \\ else\ if\ (Y < 7) & then\ u(Y) = (Y - 23)/(23 - 7) \\ else\ if\ (Y \geq 23) & then\ u(Y) = -1 \cdot (Y - 23)/(23 - 15) \end{cases}, \quad (9)$$

$$\begin{cases} if\ (Y \leq 25.5) & then\ u(Y) = 0 \\ else\ if\ (Y < 27) & then\ u(Y) = (Y - 25.5)/(27 - 25.5) \\ else & then\ u(Y) = 1 \end{cases}. \quad (10)$$

The value of O_t is then computed by membership function defined by disease membership degree with respect to user chosen symptoms using agreement membership degree $u(x)$ and the number of symptoms provided degree $u(y)$.

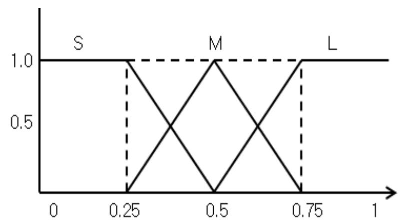


Table 1 shows the inference rules and Fig. 2 demonstrates the FCM learning procedure.

Table 1. Fuzzy inference rules

R1	If X is S and Y is S Then W is S
R2	If X is S and Y is M Then W is S
R3	If X is S and Y is L Then W is M
R4	If X is M and Y is S Then W is S
R5	If X is M and Y is M Then W is M
R6	If X is M and Y is L Then W is L
R7	If X is L and Y is S Then W is M
R8	If X is L and Y is M Then W is L
R9	If X is L and Y is L Then W is L

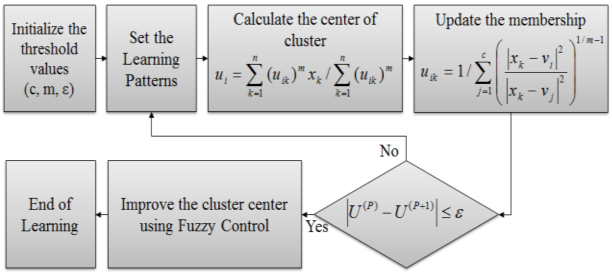


Fig. 2. Learning procedure of enhanced FCM.

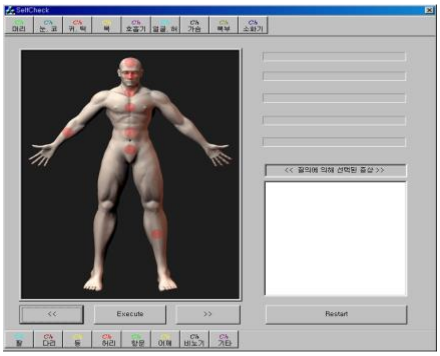
III. EXPERIMENT AND RESULTS

The implementation environment is as following: IBM compatible PC with Intel Pentium IV 2 GHz CPU and 1 G RAM is used with JDK 1.6 and Oracle 10g are used in implementation and the system is available for on-line environment using JSP. For this specific experiment, we use 60 diseases associated with 167 symptoms and 22 different body parts. Fig. 3 demonstrates the snapshot of the implemented proposed system waiting for the uncomfortable body part selection.

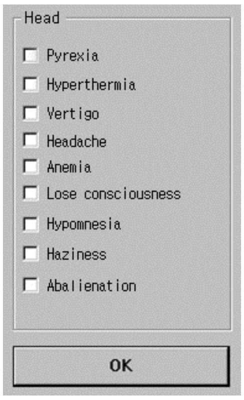
When a user chooses the body part from the screen shown in Fig. 3(a), the system asks if the user feels one of the representative symptoms on that body part as shown in Fig. 3(b) that is an example of selecting head part uncomfortable.

After collecting user's symptoms through such queries, FCM extracts five most probable diseases the user might have with confidence rate (fuzzy membership degree as shown in Fig. 4.

The proposed system infers probable diseases through improved FCM algorithm. The original FCM, if applied directly, has disadvantages in this domain when cluster patterns are much differently distributed. Thus, we apply fuzzy theory to measure the symmetry in clustering to avoid that problem. Fig. 5 compares the original FCM and our improved fuzzy controlled FCM for the same user input—sputum, cough, chill, and respiratory.



(a)



(b)

Fig. 3. User interface of the proposed software. (a) Starting screen and body parts, (b) example symptom query for head part.

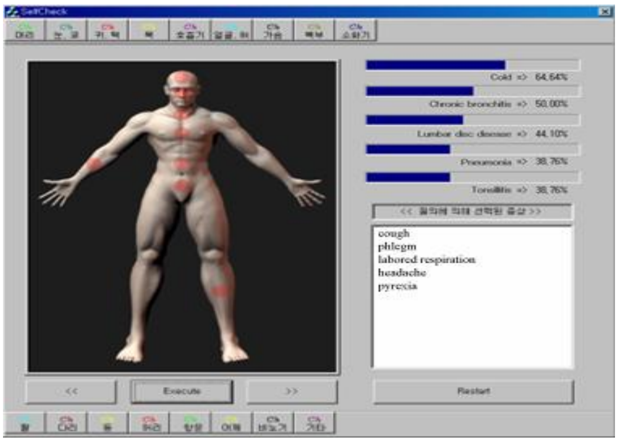
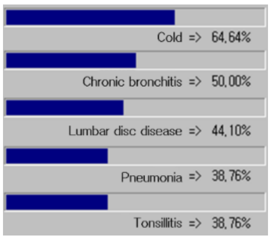
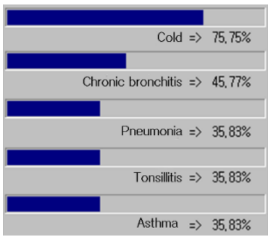


Fig. 4. Probable diseases extracted.

In Fig. 5(a), the original FCM extracts disc as the third probable disease that does not seen to relate with other four results—cold, chronic bronchitis, pneumonitis, and quinsy. However, our improved FCM extracts asthma as the fifth probable one while omitting disc in diagnosis. Furthermore, the fuzzy membership degrees of five chosen diseases are in general higher than the original FCM result.



(a)



(b)

Fig. 5. Comparing the original FCM (a) and proposed FCM (b) in Inference.

Table 2. Initialization of the proposed FCM

	Patterns	Input nodes	m	# of clusters
Proposed	50	125	1000	60

Table 2 summarizes the parameter setting of our experiment where m denotes the weight factor.

IV. CONCLUSION

Hanbang or Korea traditional medicine has a long history and popularly used among Koreans for more than thousand years. Like many reports about ethnically specific medicines, KTM is widely used concurrently with western medicine and it can be now standardized and computerized. In this paper, we developed a self-pre-diagnosis model of KTM for general public who does not have deep knowledge about KTM nor computer technology.

The proposed health pre-diagnosis system consists of two major parts: disease-symptom database and classification/ learning algorithm. We construct the database based on KCD with KTM based on government reports and gathered information is verified by the KTM doctors for their validity.

Simple user interface is provided and the user picks his/her uncomfortable body part and gives symptoms with respect to that body part. Then the system answers back with five most probable diseases with confidence rate.

The main inference engine that explores the database and finds five most probable diseases the user might have is based on FCM algorithm. The FCM algorithm is though changed to adopt fuzzy inference rules to control the distance. This improvement is for extracting only mutually related diseases and excluding false positive diseases that have different patterns from other similar diseases in clustering step.

This system is designed to stimulate general public to check one's health without deep knowledge. It may not need to be as accurate as the medical doctor but its role is to draw attention to the public to listen to their body for preventive public healthcare.

Currently this system is available with online access and in the developing phase for the mobile platform.

REFERENCES

[1] A. G. Karegowda, A. S. Manjunath, and M. A. Jayaram, "Application of genetic algorithm optimized neural network connection weights for medical diagnosis of Pima Indians diabetes," *International Journal on Soft Computing*, vol. 2, no. 2, pp. 15-23, 2011.

[2] M. Fathi-Torbaghan and D. Meyer, "MEDUSA: a fuzzy expert system for medical diagnosis of acute abdominal pain," *Methods of Information in Medicine*, vol. 33, no. 5, pp. 522-529, 1994.

[3] C. D. Hong, "Complementary and alternative medicine in Korea: current status and future prospects," *The Journal of Alternative & Complementary Medicine*, vol. 7, no. 1, pp. 33-40, 2001.

[4] S. K. Bong, "A study on the preservation and utilization of Dongeuibogam," *Journal of Korea Institute of Oriental Medicine*, vol. 15, no. 1, pp. 31-42, 2009.

[5] K. B. Kim and J. W. Kim, "Self health diagnosis system with Korean traditional medicine using fuzzy ART and fuzzy inference rules," in *Intelligent Information and Database Systems, Lecture Notes in Computer Science*, vol. 7198, pp. 326-335, 2012.

[6] M. Kim, H. R. Han, K. B. Kim, and D. N. Duong, "The use of traditional and Western medicine among Korean American Elderly," *Journal of Community Health*, vol. 27, no. 2, pp. 109-120, 2006.

[7] D. Y. Chung, D. K. Baek, S. I. Hwang, S. H. Shin, D. W. Kim, and M. A. Han, "One case of systemic lupus erythematosus treated by integrated therapy of western medicine with oriental differential diagnosis of symptoms and signs," *Journal of Korean Traditional Internal Medicine*, vol. 23, no. 2, pp. 306-312, 2002.

[8] H. Xu and K. J. Chen, "Integrating traditional medicine with biomedicine towards a patient-centered healthcare system," *Chinese Journal of Integrative Medicine*, vol. 17, no. 2, pp. 83-84, 2011.

[9] B. Lim, J. Park, and C. Han, "Attempts to utilize and integrate traditional medicine in North Korea," *The Journal of Alternative & Complementary Medicine*, vol. 15, no. 3, pp. 217-223, 2009.

[10] D. L. Berry, L. J. Trigg, W. B. Lober, B. T. Karras, M. L. Galligan, M. Austin-Seymour, and S. Martin, "Computerized symptom and quality-of-life assessment for patients with cancer. Part I. Development and pilot testing," *Oncology Nursing Forum*, vol. 31, no. 5, pp. E75-E83, 2004.

[11] K. H. Mullen, D. L. Berry, and B. K. Zierler, "Computerized symptom and quality-of-life assessment for patients with cancer. Part II. Acceptability and usability," *Oncology Nursing Forum*, vol. 31, no. 5, pp. E84-E89, 2004.

[12] Korea Ministry of Health and Welfare [Internet], Available: http://www.mohw.go.kr.

[13] S. Moein, S. A. Monadjemi, and P. A. Moallem, "A novel fuzzy-neural based medical diagnosis system," in *Proceedings of World Academy of Science, Engineering and Technology*, Tokyo, Japan, 2009.

[14] C. Y. Fan, P. C. Chang, J. J. Li, and J. C, Hsieh, "A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification," *Applied Soft Computing*, vol. 11, no. 1, pp. 632-644, 2011.

[15] K. B. Kim, S. Kim, and G. H. Kim, "Nucleus classification and recognition of uterine cervical pap-smears using FCM clustering algorithm," in *Adaptive and Natural Computing Algorithms*, Heidelberg: Springer, pp. 290-299, 2007.

[16] B. H. Chung, T. R. Chun, H. C. Kim, and K. B. Kim, "Self health diagnosis system using enhanced FCM algorithm," in *Proceedings of KIICE Conference*, pp. 143-149, 2006.

[17] Korea National Statistical Office [Internet], Available: http://www.kostat.go.kr.

Kwang Baek Kim
received his M.S. and Ph.D. degrees from the Department of Computer Science, Pusan National University, Busan, Korea, in 1993 and 1999, respectively. From 1997 to the present, he is a professor at the Division of Computer Software Engineering, Silla University, Korea. He is currently an associate editor for Journal of Intelligence and Information Systems and The Open Computer Science Journal (USA). His research interests include fuzzy Logic and applications, bioinformatics, and image processing.