Artificial Intelligence in Medicine 77 (2017) 31–47

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

A case-based reasoning system based on weighted heterogeneous value distance metric for breast cancer diagnosis

Dongxiao Gu[a], Changyong Liang[a,*], Huimin Zhao[b]

[a] School of Management, Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui, 230009, China
[b] Sheldon B. Lubar School of Business, University of Wisconsin-Milwaukee, 3202 North Maryland Avenue, Milwaukee, WI, 53201, USA

ARTICLE INFO

Article history:
Received 5 May 2016
Received in revised form 20 January 2017
Accepted 5 February 2017

Keywords:
Medical decision support system
Case-based reasoning
Case matching
Weighted heterogeneous value distance metric
Breast cancer

ABSTRACT

Objective: We present the implementation and application of a case-based reasoning (CBR) system for breast cancer related diagnoses. By retrieving similar cases in a breast cancer decision support system, oncologists can obtain powerful information or knowledge, complementing their own experiential knowledge, in their medical decision making.
Methods: We observed two problems in applying standard CBR to this context: the abundance of different types of attributes and the difficulty in eliciting appropriate attribute weights from human experts. We therefore used a distance measure named weighted heterogeneous value distance metric, which can better deal with both continuous and discrete attributes simultaneously than the standard Euclidean distance, and a genetic algorithm for learning the attribute weights involved in this distance measure automatically. We evaluated our CBR system in two case studies, related to benign/malignant tumor prediction and secondary cancer prediction, respectively.
Result: Weighted heterogeneous value distance metric with genetic algorithm for weight learning outperformed several alternative attribute matching methods and several classification methods by at least 3.4%, reaching 0.938, 0.883, 0.933, and 0.984 in the first case study, and 0.927, 0.842, 0.939, and 0.989 in the second case study, in terms of accuracy, sensitivity × specificity, F measure, and area under the receiver operating characteristic curve, respectively.
Conclusion: The evaluation result indicates the potential of CBR in the breast cancer diagnosis domain.

1. Introduction

Today, breast cancer is one of the most common cancers and one of the top causes of cancer deaths, especially in women [1–4]. Development of techniques for earlier diagnosis and better treatment has been helping to reduce breast cancer death rates [5]. In recent years, the survival rate of sufferers without cancer cell transferred is getting rapidly increased and most patients can now survive many years after early diagnosis and timely treatment [6].

The decision making accuracy of an oncologist depends largely on her knowledge, experience, and the quality and quantity of the information she can acquire. Accordingly, tools providing support with knowledge for oncologists are extremely helpful in the process of their decision making. To provide diagnosis assistance to physicians, various assistive technologies have been applied to the process of diagnosis, treatment, and prognosis [7–9]. However, technologies still cannot replace physicians. The experience and knowledge the oncologists have determine to a great extent the diagnosis quality and subsequent success of treatment.

Essentially, the medical decision making process itself is highly dependent on historical experiential knowledge. This is exactly where case-based reasoning (CBR) holds its technical superiority over other machine learning classification techniques for medical diagnosis. CBR provides the physician not only a prediction but also similar historical cases, which are perhaps more valuable than the prediction itself for the physician in making the diagnosis decision, as the consequences of a wrong diagnosis can be dire and the physician cannot simply delegate her responsibility of decision making to the prediction of a system. CBR is capable of retrieving experience and knowledge from historical cases. This closely resembles the real thinking process in the human brain. Hence, in essence, the idea of CBR is in accordance with the diagnosis decision making process of physicians. Moreover, using CBR, knowledge-based systems (usually a case-based explanation system and a problem-solving system included) for medical diagnosis decision support can be developed. In the past twenty years, CBR has been consid-

* Corresponding author.
E-mail addresses: gudongxiao@hfut.edu.cn (D. Gu), cyliang@hfut.edu.cn (C. Liang), hzhao@uwm.edu (H. Zhao).

http://dx.doi.org/10.1016/j.artmed.2017.02.003
0933-3657/

ered a useful decision support technology for medical diagnosis, prognosis, and prescription [10].

In this paper, we present the implementation and application of a CBR system for breast cancer related diagnoses. We designed and implemented the system, named BTCBRsys (for breast tumor CBR system), in collaboration with two hospitals and applied it in two studies related to benign/malignant tumor prediction and secondary cancer prediction, respectively. We observed two problems in applying standard CBR to this context: the abundance of different types of attributes and the difficulty in eliciting appropriate attribute weights from human experts. We strived to address these problems in our system.

While there have been several studies on applying CBR to breast cancer [11–15], many of the previous related studies are preliminary investigations, and are limited to presenting an application framework only [16]. There are also some case matching methods proposed in previous research associated with CBR. Although these methods have shown their extraordinary performance in a large number of CBR applications, their advantages are largely restricted to case retrieval problems with only continuous and symbolic attributes [17]. For example, the widely used Euclidian distance is effective in measuring the distance among continuous attributes but not among discrete attributes, which prevail in the context of breast cancer diagnosis. For discrete attributes, although Euclidian distance still works by generally regarding them as symbolic attributes, the results are not ideal. Taking for example the computing of the local similarity of a discrete attribute, such as the smoothness of tumor border with the possible values "highly smooth", "moderately smooth", and "not smooth at all", by Euclidian distance, matching discrete attribute values contribute maximally to the similarity while nonmatching attribute values do not contribute at all. That is, the difference between "highly smooth" and "not smooth at all" is not differentiated from that between "highly smooth" and "moderately smooth" or that between "moderately smooth" and "not smooth at all". The information in the order of the possible values is totally ignored. Obviously, that is unreasonable and unreliable. Hence, this method normally does a poor job with discrete attributes.

Another problem with practical application of CBR in breast cancer diagnosis is the difficulty in eliciting weights reflecting the relative importance of attributes in measuring the distance between cases. Attribute weighting has become an indispensable constituent part of case retrieval in CBR. Generally, the attribute weights have a significant influence on the results. Experts' evaluation is still a common way for weight determination in practical CBR applications. In such complex applications as breast cancer diagnosis, it is extremely difficult for human experts to precisely quantify the relative importance of the attributes. Also, their perceived attribute importance may not match in scale to weights that will be used in the distance measure of the CBR system. Such weight elicitation process is subjective, and the weights elicited from different expert groups may differ greatly.

To address these problems, we used a distance measure named weighted heterogeneous value distance metric (WHVDM) [18], which can better deal with both continuous and discrete attributes simultaneously than the standard Euclidean distance, and a genetic algorithm for learning the attribute weights involved in this distance measure automatically [18–21], avoiding the need for experts' evaluation. Based on two real-world breast cancer data sets, we have evaluated our system in comparison with a standard CBR and several standard classification methods. Results show that the WHVDM distance measure and genetic algorithm for attribute weight learning led to better prediction accuracy, compared to the standard Euclidean distance and weight elicitation from experts using the Delphi method. In addition to the advantage of providing the physician not only a prediction but also similar cases, our CBR system also outperformed the classification methods in terms of prediction accuracy.

Our work makes a contribution to the application domain, i.e., breast cancer diagnosis. Most of the previous related studies have used a mammographer's BI-RADS(TM)[1] features [22,23], which are from a breast imaging reporting and data system, as well as age, to predict breast biopsy. Few studies have investigated the use of clinical symptoms and signs, such as location, density, and regulation of tumor, occurrence of transferring of nodes, and smoothness of tumor border, in a CBR system for breast cancer diagnosis. In practice, clinical symptoms and signs play an important role in the diagnosis of physicians [24,25]. In our application, besides mammographic data, we also collected data about clinical symptoms and signs. Our evaluation results shed light on the utility of such data in the application of CBR to breast cancer diagnosis.

The rest of the paper is organized as follows. The next section briefly reviews the background and literature related to the application of case-based reasoning in the area of health care. Section 3 introduces the framework of BTCBRsys. Section 4 describes the system design of BTCBRsys. Section 5 presents the methodology for case matching in BTCBRsys in detail. Section 6 presents the application and evaluation of BTCBRsys in two studies and discusses the results. The last section concludes the paper and suggests directions for future research.

2. Background

In reality, not all doctors are experienced, and young and unskilled oncologists exist in almost every hospital. So the question is how inexperienced oncologists can improve their diagnosis quality by utilizing the experiential knowledge from senior doctors or successful health care cases, and how historical experience and knowledge can be fully utilized. Although it is impossible for experienced doctors to directly copy their experiential knowledge to young oncologists, it is still feasible for inexperienced doctors to get experience for reference from senior ones through an intelligent knowledge system, which integrates valuable experience and knowledge of senior oncologists and can be used to support decision making. Such systems may be designed to provide decision support in various scenarios, such as breast cancer early detection, diagnosis, recurrence prognosis, and treatment. Experienced oncologists can also benefit from such systems and make better decisions as they are better informed.

CBR resolves new problems based on similar past experience within the same field. It completes the reasoning by the use of experiential knowledge, represents knowledge using problem cases, and combines problem solving and studying. CBR has the ability to retrieve experience and knowledge from historical cases, and its reasoning process is essentially very close to the real thinking process in the human brain. Hence, it is considered an effective artificial intelligence technique for medical diagnosis, prognosis, and prescription [10]. If the numerous breast cancer cases, along with the rich health care data, such as chemical examinations report, pathological report, and medical imaging data, are organized, breast cancer case base is built, and relevant knowledge reasoning system based on CBR is implemented, it can be helpful in various stages of the care of patients, such as early detection, diagnosis, recurrence prognosis, and even nursing.

Health care is a major application area of CBR [26]. Successful applications have been achieved in this field [27,28]. Specific research topics include medical decision tasks, medical diagnosis, hospital admission authorization management, ward supervision,

[1] The Breast Imaging Reporting and Data System developed by the American College of Radiology to standardize mammographic reporting.

and mixed paradigm design of medical records. We conducted a bibliometric analysis (reported in Appendix A) to get an overall understanding of the recent literature on CBR in cancer-related medicine. Examining the co-occurrence frequency and centrality of hot subjects based on the bibliometric analysis result, we identify some main recent directions of research on CBR in cancer as follows.

(1) Key CBR techniques, such as similarity algorithms [29,30]. In the past few years, CBR has been used to support the diagnosis or prognosis of many kinds of cancers, such as breast cancer, colon cancer, and gastrointestinal cancer. Much of the literature has focused on investigating how to improve the performance of CBR through various ways, such as combining with neural networks [31], logistic regression, and rule-based reasoning (RBR) [32,33]. Saraiva et al. built a gastrointestinal cancer diagnosis support model by combining CBR and RBR and the accuracy got improved by 23% [34]. Ghany et al. used CBR combining rough set to aid breast cancer diagnosis and showed that their proposed method had an advantage over k-nearest neighbors (kNN) and fuzzy neural network [35].
(2) Development and application of CBR systems. In the past few years, much effort has been devoted to the design, implementation, and application of CBR in medicine. El-Sappagh et al. proposed a fuzzy ontology-based CBR framework and implemented it on the diabetes monitoring and diagnosis problem [36]. Petrovic et al. [37–39] developed a CBR system for radiotherapy treatment planning of brain cancer in collaboration with medical physicists at the Nottingham University Hospitals NHS Trust, City Hospital Campus, UK. This system can generate the parameters of a treatment plan by capturing the intuitive knowledge of medical physicists and automatically capture and record the expertise and experience of oncologists during the radiotherapy treatment planning process. Mishra et al. developed a self-adaptive CBR system, which can capture the experience of doctors during dose planning in prostate cancer radiotherapy [40]. Bulu et al. designed an ontology-based CBR for breast cancer and validated its feasibility and accuracy [41]. Chuang investigated CBR for the early detection of liver cancer and developed a system combining CBR with data mining methods, such as neural networks, logistic regression, and discriminant analysis [32]. van den Branden et al. integrated CBR with an electronic patient record system and developed a decision support system called ExcelicareCBR [42]. Moreover, several general frameworks for implementing concrete CBR systems are now available. Some examples are jColibri2, eXiT* CBR, and Ahn and Kim's CBR framework. jColibri2 is a toolkit for the reuse of previous problem solutions, result of several years of framework development and evolution. Recio-García et al. developed a distributed CBR tool (eXiT* CBR. v2) for medical prognosis [43]. Ahn and Kim proposed a CBR framework for breast cytology diagnosis to deal with simultaneous optimization of feature weights, case selection, and the number of the most similar neighbors by combining genetic algorithms [44]. Lieber and Bresson proposed a CBR framework for aiding breast cancer treatment decisions [13].
(3) Feature selection and weight determination. Weight learning is an important issue within the machine learning community. CBR is no exception. Several approaches, such as evolutionary algorithms, entropy method, and Delphi method, have been taken to address the weight determination problem [26]. For example, Gu et al. proposed a gray CBR with a weight determination method based on information entropy and conducted comparative study with Delphi-based weight determination method in breast cancer diagnosis [45]. Yan et al. proposed a membrane computing based approach to optimize the attribute weights of CBR [46]. The CBR system developed by Petrovic et al. for radiation therapy treatment planning can also generate feature weights of cases according to treatment plan [39].

3. BTCBRsys framework

CBR originated by Roger Schank from Yale University can be traced back to 1980s [47]. CBR is generally a process using historical similar problems to solve new problems. When a doctor needs to make decision for a new patient, generally based on the current symptoms and characteristics of the patient, she will recall historical similar cases and similar treatment methods from other patients. This is actually a CBR process. CBR is not only a computer-based reasoning method, but also a common behavior existing in daily life [48].

Our BTCBRsys can be used to conduct knowledge discovery based on historical breast tumor case bases. It can provide oncologists with valuable knowledge during their diagnosis. The CBR process can be described with four "Re" steps in a reasoning cycle [49–51], as illustrated in Fig. 1. There are the following nine basic steps in CBR for breast cancer diagnosis.

1. Organize and build an initial breast cancer diagnosis case base based on historical electronic healthcare records. An expert panel may be organized to provide advice for the construction of the case base.
2. Present a new case with its values for characteristic attributes.
3. Conduct case matching (the *Retrieve* step in Fig. 1, described in more detail in section 5) and obtain the most similar cases.
4. If a selected case satisfies the requirements of diagnosis decision support, go to step 5; otherwise, go to step 6. When the case base is initially constructed with a moderate number of cases, the chance that the retrieved cases are not satisfactory may be high. As the system cumulates more and more cases, the chance should gradually decrease. Since breast cancer is one of the most common cancers, the breast cancer case base is expected to grow rapidly.
5. Carry on case reuse (the *Reuse* step in Fig. 1) by performing an adaptation of the retrieved case and obtain a satisfactory solution based on related knowledge, as well as experts experience.
6. Take the corresponding alternative of the satisfactory case as a solution to the new case. This historical case is used as a reference during doctors diagnosis decision making.
7. Based on the satisfactory solution and the real situation, conduct further revisions (the *Revise* step in Fig. 1) for the case until the alternative fully satisfies the requirements of the new problem.
8. Evaluate the quality of the new case. Determine whether the case is qualified to be entered into the case base. If yes, go to step 9. Otherwise, the case is not retained in the case base (note that all the relevant data are still stored in the EHR (electronic health record) system, outside of BTCBRsys).
9. Retain the case in the case base (the *Retain* step in Fig. 1).

According to NCCN[2] Clinical Practice Guidelines in Oncology and CACA[3] Breast Cancer Diagnosis and Treatment: Guidelines and specifications (2015), the diagnosis and treatment process for a breast cancer patient typically includes early detection, early diagnosis, clinical diagnosis, imaging diagnosis, pathological diagnosis, treatment, rehabilitation, and prognosis. This is a complex long, dynamic, specialized practical process in which doctors need different information in each stage for better decision making. The different decision stages of the diagnosis, prognosis, and treatment

[2] NCCN: National Comprehensive Cancer Network.
[3] CACA: Chinese Anti-Cancer Association.

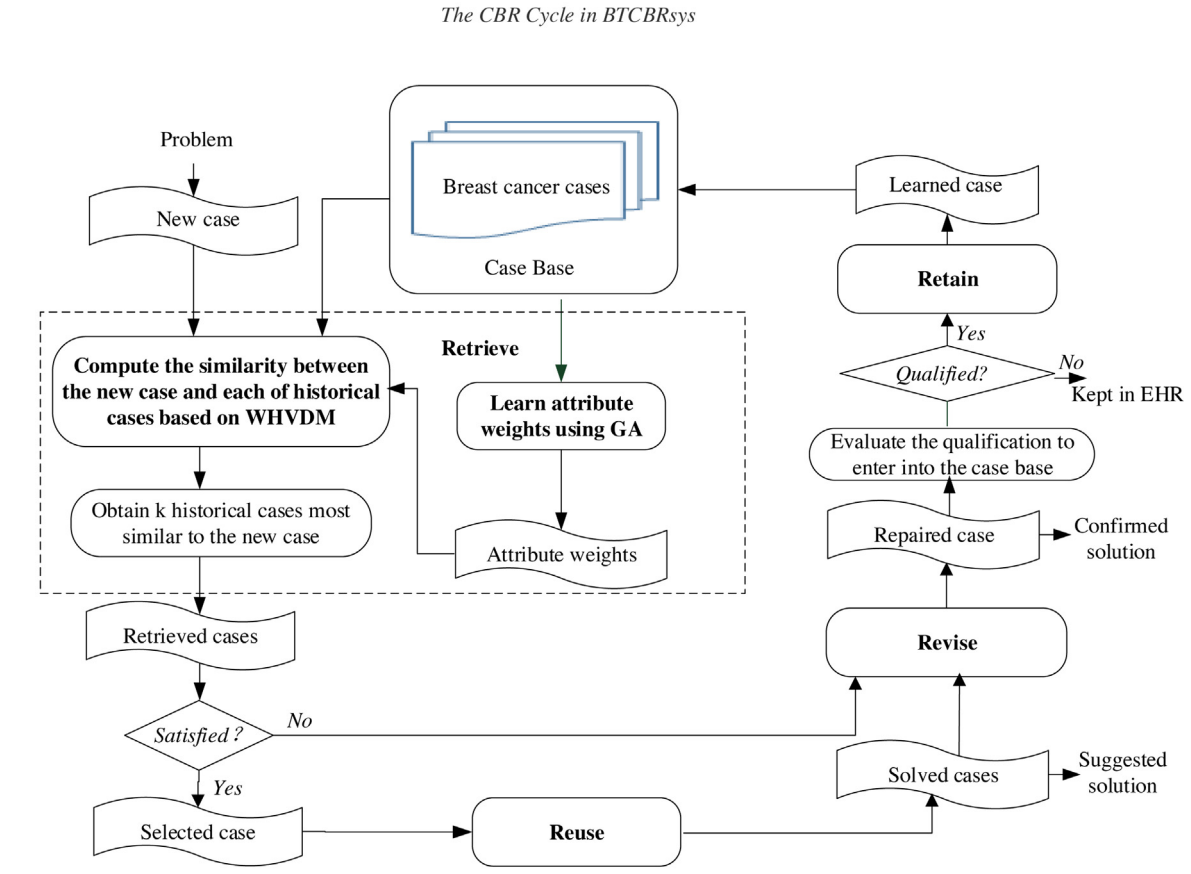

Fig. 1. CBR for breast cancer diagnosis.

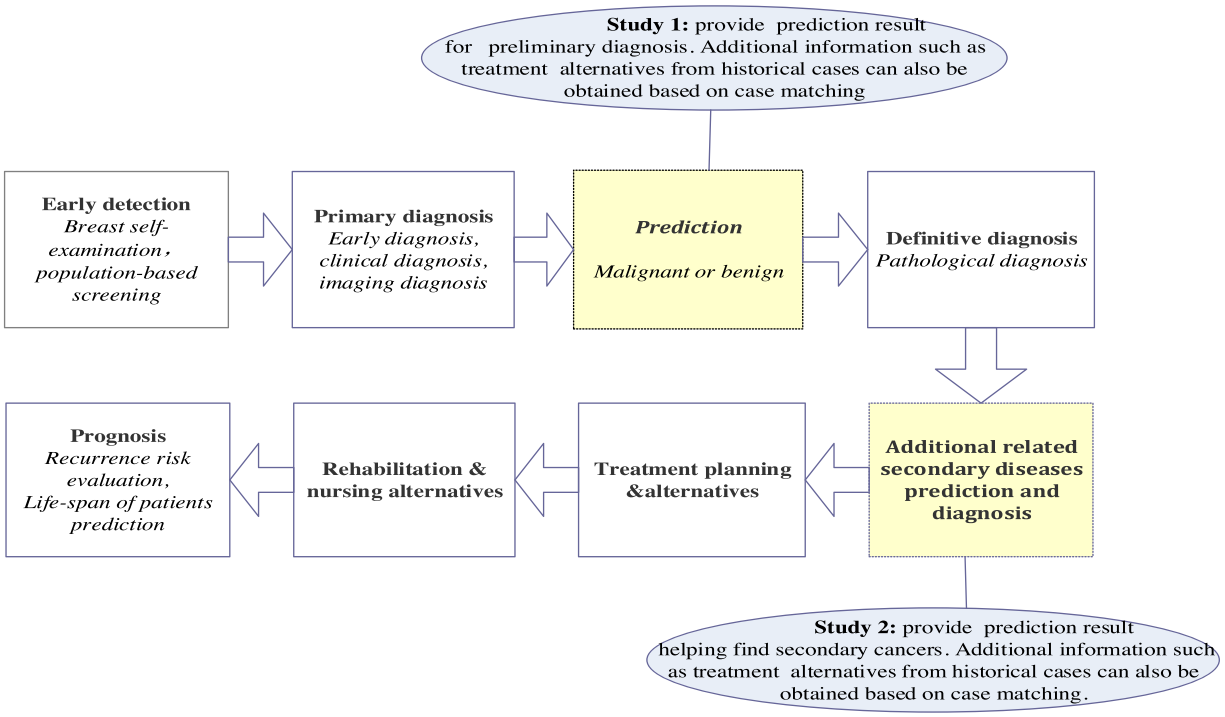

Fig. 2. Different stages of breast cancer diagnosis and treatment.

of breast cancer are outlined in Fig. 2. We will report two case studies later in section 6. In the first study, BTCBRsys was used to provide predictive results (malignant or benign) before a pathological diagnosis, the final diagnosis. In the second study, BTCBRsys was used to help find other related secondary cancers before treatment.

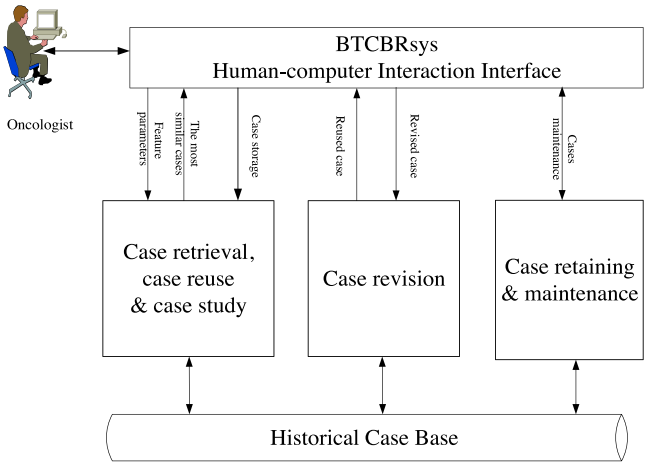

Fig. 3. BTCBRsys core components and interaction mechanism.

BTCBRsys can provide information or knowledge support in three different stages of diagnosis decision making. The first stage starts before the beginning of the diagnosis decision making. An oncologist is able to obtain some knowledge, such as characteristic attributes, and get prepared for the subsequent diagnosis work by studying in the BTCBRsys. The second stage is during the eventual decision making to reach a conclusion. An oncologist seeks historical knowledge through case matching and draws a predictive conclusion. The last stage is after the conclusion. The oncologist can study the successful experience from the most similar cases, such as alternatives for treatment, suggestions for health promotion, and suggestions for nursing. Among the three stages, the support of the second stage is the focus in our current study.

4. BTCBRsys system design

Our system design follows such principals as practicability (conforming to practical breast cancer diagnosis decision making process), usefulness, and ease of use. The core components and interaction mechanism are shown in Fig. 3.

Historical cases are stored in the case base. The case retrieval and reuse subsystem performs case matching and case reuse. By default, the most similar case for a new case will be reused. The related content of this case will be copied to the new case. To provide more references for experts, the system does not prescribe a threshold on the similarity or a limit on the number of cases the doctor can see. The system presents the historical cases in ascending order of similarity page by page. The doctor can set the number of cases listed on each page (default is 10) and can always look at more pages if needed. The case revision subsystem carries out revision of new cases. The case retaining and maintenance subsystem evaluates new cases for possible retaining and maintains the case base.

The communication between users and all the components of BTCBRsys is bidirectional. The interaction involves information acquisition, confirmation, and feedback. The users or team of decision making (oncologists) not only should ensure the correctness of inputted information associated with case feasures, but also play an important role in interaction with BTCBRsys. Hence, they should be familiar with the reasoning process of BTCBRsys and know how to guide it to conduct knowledge reasoning according to the actual situation. The tasks of users include: (1) Select the features of the current diagnosis case; (2) Conduct case matching. If succeeded, reuse the most similar case, carry on some necessary revision to make an adaption to the actual situation of the current problem; (3) Otherwise, analyze possible reasons, adjust the parameters of case retrieval, and conduct case matching again; (4) Conduct further improvement on the alternative of the target case and obtain the final solution after the revision is done by the system.

There is an evaluation process to determine if a new case should be stored into the case base after the final solution is obtained. Not all the new cases are accepted into the case base. Only those considered having high quality, value, and necessity by oncologists will be retained in the case base. This will make the case base gradually expand, improving CBR quality over time. Based on these idea, we designed the main functions of BTCBRsys (shown in Fig. 4).

5. The methodology for case matching in BTCBRsys

In this section, we describe two major components in case matching (highlighted in Fig. 1) in detail: the distance measure WHVDM, which can deal with both continuous and discrete attributes appropriately, and a genetic algorithm for learning the weights of attributes in this new distance measure.

5.1. CBR problem specification

A problem case is described by a pair (x, y), where $x = (x_1, x_2, ..., x_n)$ is a vector of independent variables (also called *characteristic attributes* or *features*) and $y \in Y$ is a discrete dependent variable (also called *class*). A case base stores a set of historical solved cases where the class values are already known. Given a new unsolved target case where the class value is not yet known, the objective of CBR is to retrieve a set of cases from the case base that are deemed most similar to the new case and to support the decision maker in making a prediction on the class value. The relative importance of the attributes in measuring the distance between a new target case and a stored case is reflected in a vector of weights $w = (w_1, w_2, ..., w_n)$, where $0 \le w_i \le 1$ $(i = 1, 2, ..., n)$ and $\sum_{i=1}^{n} w_i = 1$.

An attribute may be continuous or discrete. Strictly speaking, any value stored in a computer is discrete at some level [52]. Some really discrete attributes that are measured on a linear scale, such as the age of a patient, are often treated as continuous attributes, because they can have many possible values such that each value may be observed very infrequently [52]. We follow this convention and treat such attributes as continuous attributes. We use the term discrete attribute to refer to a nominal (or symbolic) attribute or ordinal attribute that has only a small number of possible values and is not measured on a linear scale, such as the smoothness of tumor border with the possible values "highly smooth", "moderately smooth", and "not smooth at all".

5.2. Distance measure

At the core of CBR is the measure of the distance between two cases. The Euclidean distance has been the most widely used distance measure in various CBR systems [53]. Given a new target case t and a stored case r, the Euclidean distance between the two cases is defined as:

$$EU(t, r) = \sqrt{\sum_{i=1}^{n} d_i^2(t, r)} \quad \text{, where}$$
$$d_i(t, r) = diff(x_{t,i}, x_{r,i}),$$
$$diff(x_{t,i}, x_{r,i}) = x_{t,i} - x_{r,i}. \tag{1}$$

To avoid the effect of measurement scales on the distance measure, the attributes may be normalized or standardized [52]. It has also become a common practice in CBR to weight the attributes

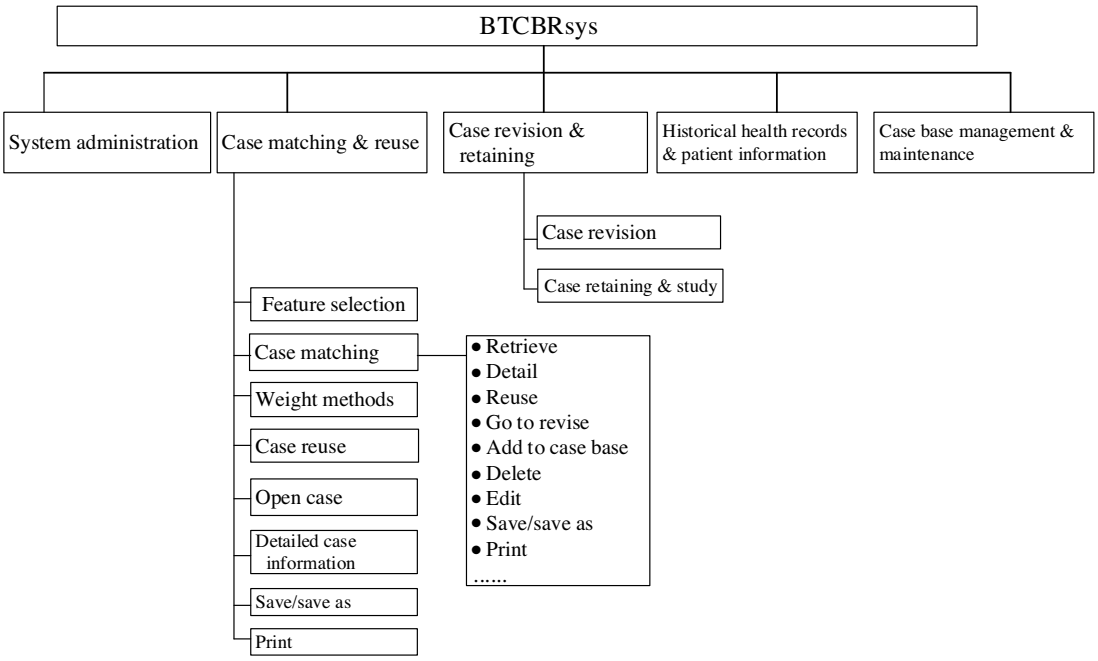

Fig. 4. The main functional structure of BTCBRsys.

according to their relative importance. The weighted Euclidean distance between a target case t and a stored case r is defined as:

$$WEU(t,r)=\sqrt{\sum_{i=1}^{n} w_i d_i^2(t,r)}. \tag{2}$$

The Euclidean distance is defined for continuous attributes and needs to be extended when both continuous and discrete attributes are present. A simple yet widely used solution is the Heterogeneous Euclidean-Overlap Metric (HEOM), which uses the overlap metric on discrete attributes and Euclidean distance on continuous attributes [52].

$$HEOM(t,r)=\sqrt{\sum_{i=1}^{n} d_i^2(t,r)}, \text{ where}$$
$$d_i(t,r)=\begin{cases} overlapp\ (x_{t,i},x_{r,i}), \text{ if } x_i \text{ is discrete} \\ diff\ (x_{t,i},x_{r,i}), \text{ if } x_i \text{ is continuous} \end{cases}, \tag{3}$$
$$overlapp(x_{t,i},x_{r,i})=\begin{cases} 0, \text{ if } x_{t,i}=x_{r,i} \\ 1, \text{ otherwise} \end{cases}$$
and $diff(x_{t,i},x_{r,i})$ **is the same as that in** (1).

HEOM is overly simplistic in dealing with discrete attributes [52]. Nonmatching discrete attribute values contribute maximally to the distance while matching attribute values do not contribute at all [54]. To address this issue, Stanfill and Waltz [55] proposed the value difference metric (VDM). The VDM between a target case t and a stored case r on a discrete attribute x_i is defined as:

$$vdm_i(t,r)=\sum_{a\in Y}(Pr(y=a|x_i=x_{t,i})-Pr(y=a|x_i=x_{r,i}))^2$$
$$\cdot\sqrt{\sum_{a\in Y} Pr(y=a|x_i=x_{t,i})^2}. \tag{4}$$

VDM reflects the distance between the conditional probability distributions of the class given the attribute values of the two cases, weighted by a term reflecting the degree of certainty of the class given the attribute value of the target case. Two cases are deemed closer to each other with regard to an attribute, if the two attribute values lead to closer conditional probability distributions on the class. The conditional probabilities involved in the definition can be estimated based on the stored cases in the case base.

VDM is defined for discrete attributes and has been extended to deal with continuous attributes. For example, Wilson and Martinez [52] introduced several new metrics based on VDM: heterogeneous VDM (HVDM), which is similar to the similarity metric of Gower [56], interpolated VDM (IVDM), and windowed VDM (WDVM), which is a similar but more sophisticated version of IVDM. IVDM and WDVM extend VDM to deal with continuous attributes by discretizing them to estimate conditional probabilities [52]. Such discretization inevitably leads to loss of information contained in the continuous attributes. Also, the result varies with the extent of discretization to continuous attributes.

The Euclidean distance is appropriate for continuous attributes but inappropriate for discrete attributes, whereas VDM is appropriate for discrete attributes but inappropriate for continuous attributes [52]. HVDM synthesizes both by using Euclidean distance on continuous attributes and a simplified version of VDM (without the weighting scheme) on discrete attributes. It is similar to HEOM, except that it uses VDM rather than the simple overlap metric on discrete attributes.

$$HVDM(t,r)=\sqrt{\sum_{i=1}^{n} d_i^2(t,r)}, \text{ where}$$
$$d_i^2(t,r)=\begin{cases} vdm'_i(t,r), \text{ if } x_i \text{ is discrete} \\ diff^2(x_{t,i},x_{r,i}), \text{ if } x_i \text{ is continuous} \end{cases}, \tag{5}$$
$vdm'_i(t,r)$ is the $vdm_i(t,r)$ in (4) without the weighting term $\sqrt{\sum_{a\in Y} Pr(y=a|x_i=x_{t,i})^2}$, and
$diff(x_{t,i},x_{r,i})$ *is the same as that in* (1).

A remaining difficulty with HVDM is the appropriate scaling of the two different types of measurements, for continuous and discrete attributes, respectively. Wilson and Martinez [52] proposed three alternative scaling functions, which are all arbitrary and do not take into account the relative importance of different attributes.

Observing that weighting of attributes has proven to be a useful mechanism and has become a common practice in CBR, we use a weighted HVDM (WHVDM) [18] in BTCBRsys. When appropriately chosen, weights on attributes can serve two purposes simultaneously, reflecting the relative importance of difference attributes and scaling the two different types (continuous and discrete) of measurements. The WHVDM between a new target case t and a stored case r is defined as:

$$WHVDM(t,r) = \sqrt{\sum_{i=1}^{n} w_i d_i^2(t,r)} \ , \ where$$
$$d_i^2(t,r) = \begin{cases} vdm_i(t,r), \text{ if } x_i \text{ is discrete} \\ diff^2(x_{t,i}, x_{r,i}), \text{ if } x_i \text{ is continuous} \end{cases}, \quad (6)$$

$vdm_i(t, r)$ is the original VDM with the weighting term [9] in (4), and $diff(x_{t,i}, x_{r,i})$ is the same as that in the Euclidean distance (1).

Note that the attribute weighted HVDM used in [18] is very similar, except that it uses $vdm'_i(t, r)$ rather than $vdm_i(t, r)$.

5.3. Genetic algorithm for weight determination

The weights of attributes in the distance measure play an important role in CBR and significantly influence the case retrieval results. Weight elicitation from domain experts is still a common way for weight determination in real-world CBR applications, with Delphi being the commonly used approach [57]. However, as is in most knowledge engineering practices, knowledge acquisition from experts often becomes a bottleneck. In complex applications such as breast cancer diagnosis, it is extremely difficult for experts (oncologists) to draw appropriate weights on the attributes that will eventually lead to highly accurate predictions through CBR. Weights obtained from different expert groups may differ greatly.

With WHVDM, weight elicitation from domain experts becomes even more difficult, since the attribute weights not only reflect relative importance of the attributes as before, but also serve the purpose of scaling the two different types (continuous and discrete) of measurements. We therefore adopt a learning-based approach and use a genetic algorithm (GA) to learn the attribute weights based on sample cases. Derived from the principles of natural selection and evolutionary theory, GAs have been theoretically and empirically proven to be robust search techniques. They are inspired by the principle of "survival of the fittest", where the fittest individuals are selected to produce offspring for the next generation [58], and in essence are "search algorithms based on the mechanics of natural selection and natural genetics" [59]. GA has been successfully used for weight learning in lazy learning methods, CBR, or ensemble CBR (e.g., [18–21]).

A standard GA can be used to learn the weight vector in WHVDM, once the chromosome encoding and fitness metric for an individual are determined. The chromosome of each individual in GA encodes the attribute weight vector to be determined. The fitness of each individual is the objective to be maximized, i.e., the prediction accuracy.

Given a reference case set R and a testing case set T (the stored case base can be split into these two case sets), we predict the class value of each testing case with that of the most similar reference case and count the number of correctly predicted testing cases. Let $s(t)$ denote the most similar reference case for testing case t, i.e., $s(t) = \arg\min_{r \in R} WHVDM(t, r)$. The number of correctly predicted testing cases is $\sum_{t \in T} I(y_t = y_{s(t)})$, where I is the indicator function, $I(e) = \begin{cases} 1, \text{if } e \\ 0, \text{otherwise} \end{cases}$.

Table 1
The attributes used in the first study.

Attribute	Type	Description
Age	Continuous	The age of the patient.
Location	Discrete	The location of the patient's mass.
Node	Continuous	The number of metastatic lymph nodes.
Density	Discrete	The density of the patient's mass.
Clarity	Discrete	The clarity of the patient's mass margin.
Area	Continuous	The area of the patient's mass.
Regulation	Discrete	The regulation of the patient's mass border.
Surface_Smoothness	Discrete	The smoothness of the patient's mass surface.
Nipple	Discrete	Whether a woman with breast tumor has nipple discharge.
Family_History	Discrete	Whether the patient has a family history of breast cancer.

6. Application and evaluation

We implemented BTCBRsys in collaboration with two hospitals in East China and evaluated it in two case studies, in comparison with a standard CBR and several standard machine learning classification techniques. The system is owned by the hospitals and is currently not publicly available.

6.1. Data

The objective of the first study is for oncologists to reduce unnecessary breast biopsy. Currently, the only definitive method for diagnosing breast cancer is pathological diagnosis with a breast biopsy. Biopsy involves the removal of tissue or cells for microscopic examination. The specimen can be obtained from a symptomatic area or from an area identified with breast imaging. However, approximately 70–85% of breast biopsies are performed on benign lesions [60]. Breast biopsy is after all one of the most invasive examinations. Numerous women performed with breast biopsy are unnecessarily subjected to the discomfort, pain, high anxiety, potential complications, cost, and altered cosmetic appearance [61]. Reducing the number of benign biopsies is tremendously valuable for patients.

We spent over seven months on collecting a data set from one of the large-scale hospitals in East China. The data set contains ten attributes obtained from clinically preliminary and regular examinations: age, location, node, density, clarity, area, regulation, border smoothness, nipple, and family history (see Table 1 for descriptions of the attributes). Most attributes characterize symptoms and signs from oncologists' clinical examinations. The class to be predicted describes the ground truth about the breast cancer severity: benign (negative) or malignant (positive). There are totally 450 cases. We reserved 150 cases, with 82 (54.67%) benign cases and 68 (45.33%) malignant ones, for use in the GA training only. We then organized a case base using the remaining 300 cases, with 164 (54.67%) benign cases and 136 (45.33%) malignant ones. By retrieval from the case base using the CBR system, an oncologist may acquire valuable knowledge and powerful support for her decision making on whether a patient needs further examinations and whether a biopsy for pathological analysis is necessary.

Table 2
The attributes used in the second study.

Attribute	Type	Description
Age	Continuous	The age of the patient.
Tumor_Size	Continuous	The size of the mass/lesion. The larger the tumor size, the more aggressive the disease and the more aggressive the treatment.
No_Nodes	Continuous	The number of lymph nodes that contain cancer cells.
Histological_Grade	Discrete	Measures the aggressivity of the tumor based on the degree of histologic deviation from normal.
ER_Percent	Continuous	Used to determine if the cancer is sensitive to hormone therapy.
PR_Percent	Continuous	Measures the intensity of positivity and indicates the probability of response to endocrine therapy.
Cancer_Type	Discrete	Provides the diagnosis and indicates the prognosis and treatment.
PR_Status	Discrete	Measures the amount of receptors to hormone and consequently the sensitivity of the tumor cells to progesterone, which makes the tumor grow.
Menopause_Status	Discrete	This comes from medical history. The values are pre and post menopause. It is related to hormone status of the patient. It helps to assess risk and to decide on therapy.

The second study concerns the prediction of whether a breast cancer patient also has secondary cancers. We spent two years on collecting data from two hospitals: a large hospital in Midwestern United States and one of the 3-A (top grade) hospitals in China. All the data are associated with confirmed breast cancer cases. The data set includes nine characteristic attributes: age, tumor, number of nodes, histological grade, ER percent, PR percent, cancer type, PR status, and menopause status (see Table 2 for descriptions of the attributes). The class to be predicted provides a more detailed description of the diagnosis reflecting the progressive nature of the disease: without other secondary cancers (negative) or with other secondary cancers (positive). There are totally 301 cases. We reserved 100 cases, with 45 (45%) negative cases and 55 (55%) positive ones, for use in the GA training only. We then organized a case base based on the remaining 201 cases, with 86 (42.78%) negative cases and 115 (57.22%) positive ones. An oncologist may use the CBR system to acquire useful information to support decision on whether further examinations are necessary to a breast cancer patient. For example, for a patient with Invasive Lobular Carcinoma (ILC), her doctor can predict whether she has other secondary cancers and needs further examinations, based on the knowledge discovered from similar cases. This helps to reduce the number of unnecessary further examinations without missing other secondary cancers.

The case bases in both studies include not only the information listed in Tables 1 and 2, but also additional supplemental or subsequent information, such as the proof of diagnosis, the planned alternatives for treatments, the reports (notes) of operation, doctors' advice, suggestions for nursing, and suggestions for health promotion. Although we did not use the other data in our tests for prediction performance, they are really useful for providing doctors more detailed information for decision making. This is an advantage of CBR over other classification methods, because it provides not only a prediction (benign or malignant; with or without secondary cancer) but also more detailed information on similar cases.

BTCBRsys provides physicians similar historical cases for their reference in their decision making but does not make any decision on their behalf no matter how certain it is with its prediction. The consequences of a wrong diagnosis can be dire (missing a true breast cancer case may kill the patient) and the physician can never delegate her responsibility of decision making to a system. Generally, when doctors obtain the most similar cases by case matching, they will conduct a careful study on these cases. They will examine all the related information instead of only the classification result, such as the differences in features, doctors' diagnosis conclusions reports, the treatment plans or alternatives the doctors used in the matching cases. Sometimes, the doctors will also have a consultation to discuss the case matching results if necessary. Based on all the related information, the doctors will make a preliminary decision of diagnosis and prediction on the new breast cancer patient. In addition, since the matching cases carry additional information associated with treatment plans or alternatives, survival duration, recurrence, rehabilitation planning, and nursing alternatives, the doctors generally also use these cases as reference when they make decision on the alternatives for treatment, rehabilitation, nursing, or prognosis (such as survival prediction and recurrence prediction).

6.2. Experiments

The experiments focused on evaluating the prediction performance of BTCBRsys, in comparison with a standard CBR and a recent gray CBR with information weighting scheme [45]. We therefore compared WHVDM with attribute weights learned using GA (denoted WHVDM-GA) with the following baselines: Euclidean distance without attribute weighting (i.e., all attributes are given equal weight) (denoted Euclidean), weighted Euclidean distance with attribute weights elicited from experts using a Delphi method (denoted WE-Expert), gray CBR with a weighting scheme based on information entropy (denoted GCBR-IE) [45], WHVDM with attribute weights elicited from experts using a Delphi method (denoted WHVDM-Expert), and WHVDM with attribute weights based on information entropy (denoted WHVDM-IE). All continuous attributes were standardized (zero mean and unit standard deviation) beforehand. kNN was used as the retrieval method.

For the sake of robustness, we gauged prediction performance using three measures: accuracy (i.e., the proportion of predictions that are correct), sensitivity × specificity, and F-measure. Sensitivity × specificity is the product of a pair of measures, sensitivity (also called the true positive rate) and specificity (also called the true negative rate), which are widely used in the field of medical diagnosis. F-measure is the harmonic mean of a pair of measures, precision (also called the positive predictive value) and recall (same as sensitivity), which are widely used in the field of information retrieval [62].

Let TP, TN, FP, and FN denote the number of true positives (*hits*), the number of true negatives, the number of false positives (*false alarms*), and the number of false negatives (*misses*), respectively. The performance measures are defined as follows.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$recall = Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$F-\text{measure} = \frac{2 * Precision * Recall}{Precision + Recall}$$

We used repeated random subsampling validation with five iterations to evaluate the methods. In each of the two studies, we randomly split the case set into a reference set and a test set. We randomly selected two-thirds of the cases as reference cases and tested the prediction performance on the other one-third of cases. In the first study, 200 cases (109 benign cases and 91 malignant ones) were randomly selected into the reference set, and the other 100 cases (55 benign cases and 45 malignant ones) were left in the test set. In the second study, 135 cases (77 cases with secondary cancers and 58 without) were randomly selected into the reference set, and the other 66 cases (38 cases with secondary cancers and 28 without) were left in the test set. We repeated this performance estimation procedure five times, with independent random splits across the five iterations. All performance results presented later are averages over the five iterations.

6.3. Weight elicitation using a Delphi method

Delphi is commonly used in CBR applications to elicit attribute weights from domain experts, especially in domains where the availability of experts is scarce while the problem is nontrivial. In the domain of breast cancer diagnosis, quantifying the relative importance of various attributes is extremely difficult, even for experienced oncologists, while access to expert oncologists is extremely scant. Methods that require intensive direct interactions among experts are infeasible. Delphi does not require direct interactions among experts, but can still reach some level of consensus among the experts on a complex problem through an iterative procedure, allowing the experts to repeatedly revise their opinions in light of the collective opinion of their peers.

The selection of panel members is crucial in Delphi projects. In our studies, the following requirements were enforced on a person to be qualified as a panel member: (1) extensive professional background and knowledge on breast cancer, (2) substantial working experience in breast cancer early detection, diagnosis, prognosis, and treatment, and (3) well-regarded authority in breast cancer clinical diagnosis and treatment decision making.

In the first study, we formed a panel of eight expert oncologists to participate in a Delphi project. Six of the experts were from the hospital where our study was conducted, and the other two were from another major hospital in East China. All the panel members are distinguished experts in the field of breast cancer diagnosis and treatment. They are all chief physicians or professors from 3A (top grade in China) hospitals with over 20 years of clinical diagnosis and treatment experiences on breast cancer.

We strictly followed the standard process of Delphi. We asked the experts to rate the relative importance of the ten attributes (Table 1) in making a decision on whether breast biopsy is needed for a patient. After each round of expert evaluation, we computed the average weights across the experts, and assessed the degree of consensus among the experts using Kendall's coefficient of concordance test (Kendall's W), a nonparametric measure of correlation. Kendall's W is a normalization of the statistic of the Friedman test and can be used to assess the degree of agreement among raters. It ranges from 0 (no agreement at all) to 1 (complete agreement). Intermediate values of W indicate a greater or lesser degree of unanimity among the various responses [63,64]. If needed, we executed another round, sending the average weights to the experts, and asking the experts to further deliberate over their ratings in light of the averages and revise their ratings as they feel appropriate. After three rounds, a satisfactory degree of consensus was achieved (Kendall's $W = 0.79$, $p < 0.001$), and the Delphi project was concluded. The average weights from the last round were subsequently used in our experiments.

In the second study, we conducted a Delphi project similarly to elicit weights of the nine attributes (Table 2) in predicting whether a breast cancer patient may have other secondary cancers too. The panel consisted of two famous experts in the area of oncology from a medical college and six senior oncologists. Strictly following the standard process of Delphi and experiencing three rounds of opinion giving and revising, we reached a satisfactory degree of consensus among the eight experts (Kendall's $W = 0.72$, $p < 0.001$).

The attribute weights elicited using the Delphi method for the two studies are as follows.

For the first study:

(0.0521, 0.0521, 0.1391, 0.1130, 0.1184, 0.0587, 0.1163, 0.1228, 0.1369, 0.0902).

For the second study:

(0.041, 0.055, 0.0823, 0.1329, 0.1451, 0.1648, 0.1371, 0.1123, 0.1275).

6.4. Attribute weight learning using GA

We used a standard GA with a typical parameter setting (similar to that in [18]). Each individual represents a possible attribute weight vector. A real-valued representation of individuals was used. The fitness of each individual is the prediction accuracy, as described earlier (in 5.3). A 5-fold cross-validation was used in evaluating the prediction accuracy. The population size was 20. The initial population was randomly generated. Roulette-wheel selection was used in parent selection. Uniform crossover with discrete recombination was used in offspring creation. The crossover rate was 80%. The multiple crossover points in the uniform crossover were selected randomly and independently for each gene in the chromosome. The mutation rate was 5%. Mutation was done in a uniform manner in which a random value from the range [0,1] was drawn and set as the new value in the current position. Elitism was used (the fittest individual in each generation was retained) [18]. The evolution was terminated after 2000 generations or if there was no improvement for 20 successive generations. We repeated this entire procedure five times. The average of the attribute weight vectors of the best individuals, as evaluated using the 5-fold cross-validation, over the five runs was used as the final weight vector.

The final attribute weights learned by GA for the two studies are as follows.

For the first study:

(0.0003, 0.0001, 0.1637, 0.2105, 0.0000, 0.0209, 0.0600, 0.2907, 0.2148, 0.0389).

For the second study:

(0.0001, 0.0001, 0.0000, 0.1263, 0.0942, 0.4576, 0.0001, 0.1526, 0.1690).

The weights elicited from experts and those learned by GA are highly correlated (0.70 in the first study and 0.67 in the second study), showing that they agreed to some extent on which attributes are more important. However, the human experts were much more conservative in expressing their opinions and there was much lower variation in their weights (the standard deviation of weights was 0.03 (expert) vs. 0.11 (GA) in the first study and 0.04 (expert) vs. 0.15 (GA) in the second study).

6.5. Results

Table 3 summarizes the results from the first study. The results are consistent across the three performance measures. In terms of every performance measure, BTCBRsys (i.e., WHVDM-GA) outperformed the five baselines. Attribute weighting using expert provided weights improved the prediction performance of the Euclidean distance. Information entropy weighting (GCBR-IE) [45]

Table 3
Results from the first study.

Setting	Accuracy	Sensitivity × specificity	F-measure
Euclidean	0.800 (0.022)***	0.622 (0.050)***	0.770 (0.036)***
WE-Expert	0.800 (0.024)***	0.634 (0.058)***	0.778 (0.048)***
GCBR-IE	0.848 (0.022)***	0.713 (0.042)***	0.828 (0.029)***
WHVDM-Expert	0.884 (0.049)*	0.810 (0.036)**	0.890 (0.022)**
WHVDM-IE	0.902 (0.011)***	0.811 (0.022)***	0.887 (0.018)***
WHVDM-GA	0.938 (0.011)	0.883 (0.022)	0.933 (0.012)

Standard deviation is enclosed in parentheses.
* indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.1 level, based on *t* test.
** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.05 level, based on *t* test.
*** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.01 level, based on *t* test.

Table 4
Results from the second study.

Setting	Accuracy	Sensitivity × specificity	F-measure
Euclidean	0.827 (0.042)***	0.684 (0.084)**	0.846 (0.032)***
WE-Expert	0.809 (0.025)***	0.642 (0.045)***	0.826 (0.023)***
GCBR-IE	0.887 (0.021)**	0.782 (0.039)*	0.903 (0.017)**
WHVDM-Expert	0.875 (0.029)**	0.767 (0.059)*	0.890 (0.024)**
WHVDM-IE	0.884 (0.031)*	0.780 (0.056)	0.899 (0.028)*
WHVDM-GA	0.927 (0.013)	0.842 (0.022)	0.939 (0.011)

Standard deviation is enclosed in parentheses.
* indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.1 level, based on *t* test.
** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.05 level, based on *t* test.
*** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.01 level, based on *t* test.

further improved the prediction performance. The WHVDM distance outperformed the Euclidean distance, with expert provided attribute weights, as well as GCBR-IE. Finally, the GA for attribute weight learning further improved performance. Both WHVDM and GA contributed to performance improvement.

Table 4 summarizes the results from the second study. In terms of every performance measure, BTCBRsys (i.e., WHVDM-GA) outperformed the five baselines. In this study, adding attribute weighting using expert provided weights to the Euclidean distance actually lowered prediction performance, highlighting the difficulty in appropriately weighting the attributes in such complex applications even for expert oncologists. The WHVDM distance outperformed the Euclidean distance, with expert provided attribute weights, but the performance was still worse than GCBR-IE. Finally, the GA for attribute weight learning further improved performance and led to better performance than GCBR-IE. Again, both WHVDM and GA, contributed to performance improvement.

The receiver operating characteristic (ROC) curves from the two studies are shown in Figs. 5 and 6 respectively. The area under the ROC curve (AUC) of Euclidean, WE-Expert, GCBR-IE, WHVDM-Expert, WHVDM-IE and WHVDM-GA in the first study is 0.666, 0.704, 0.774, 0.888, 0.904, and 0.984, respectively, and in the second study is 0.707, 0.774, 0.851, 0.868, 0.922 and 0.989, respectively. This also shows that BTCBRsys (i.e., WHVDM-GA) outperformed the five baselines.

6.6. Comparison with commonly used classification methods

Although the main advantage of CBR over other classification methods is not necessarily in prediction performance but in its ability to return similar historical cases for reference in the decision making of a physician, we also compared CBR (WHVDM-GA) with a few classification methods in terms of prediction performance, for the sake of completeness. Specifically, we compared our CBR with RBF neural network, CART decision tree, logistic regression, and Naïve Bayes in the Weka machine learning toolkit [65]. These are some widely used classification methods and are representatives of different types of methods (there are certainly many others that could have been chosen for comparison). A brief description regarding these methods is available in Appendix B. Readers interested in more details about these methods and the actual implementations in Weka are referred to [69] and the associated toolkit.

Table 5
Results of comparison with classification methods in the first study.

Method	Accuracy	Sensitivity × specificity	F-measure
CBR(WHVDM-GA)	0.938 (0.011)	0.883 (0.022)	0.933 (0.012)
RBF Network	0.906 (0.005)***	0.821 (0.011)***	0.898 (0.010)***
CART	0.772 (0.011)***	0.633 (0.077)***	0.736 (0.079)***
Logistic Regression	0.850 (0.008)***	0.706 (0.017)***	0.825 (0.012)***
Naive Bayes	0.852 (0.029)***	0.704(0.058)***	0.821 (0.038)***

Standard deviation is enclosed in parentheses.
*** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the 0.01 level based on *t* test.

Table 6
Results of comparison with classification methods in the second study.

Method	Accuracy	Sensitivity × specificity	F-measure
CBR(WHVDM-GA)	0.927 (0.013)	0.842 (0.022)	0.939 (0.011)
RBF Network	0.779 (0.026)***	0.606 (0.033)***	0.801 (0.029)***
CART	0.838 (0.037)***	0.713 (0.059)***	0.847 (0.039)***
Logistic Regression	0.808 (0.037)***	0.649 (0.059)***	0.831 (0.035)***
Naive Bayes	0.808 (0.018)***	0.656 (0.030)***	0.827 (0.017)***

Standard deviation is enclosed in parentheses.
*** indicates that the performance advantage of WHVDM-GA over the method is statistically significant at the .01 level based on *t* test.

We used the same three performance measures: accuracy, F-measure, and sensitivity × specificity. The method used for performance estimation is consistent with that for CBR methods. Tables 5 and 6 summarize the comparison results. The ROC curves of the different methods for the two studies are shown in Figs. 7 and 8. The AUC of Naïve Bayes, Logistic Regression, CART, RBF Network, and BTCBRsys (WHVDM-GA) for the first study are 0.739, 0.802, 0.760, 0.903, and 0.984, respectively, and in the second study are 0.762, 0.814, 0.887, 0.806, and 0.989, respectively. In both studies, BTCBRsys (WHVDM-GA) outperformed all four classification methods in terms of every performance measure.

6.7. User evaluation of BTCBRsys application

After the implementation of BTCBRsys in the two collaborating hospitals, some oncologists have begun to use the system on trial. To evaluate the effectiveness of its application in the hospitals, we conducted a survey of the users. The items we used in the questionnaire include the usefulness to improve medical quality, ease of use, satisfaction of use, usefulness to reduce medical cost, usefulness to save medical resource, usefulness to improve oncologists' specialization level, usefulness to improve patients' satisfaction, and intention of continued use. We distributed 15 questionnaires and collected 12 valid responses (the results are shown in Table 7). Most of the participants approved the usefulness to improve medical quality, as well as the role of BTCBRsys in saving medical resource and improving oncologists' specialization level. All of them had the intention to continue their use of the system. Future work should focus on improving the ease of use through further optimization, design, and visualization.

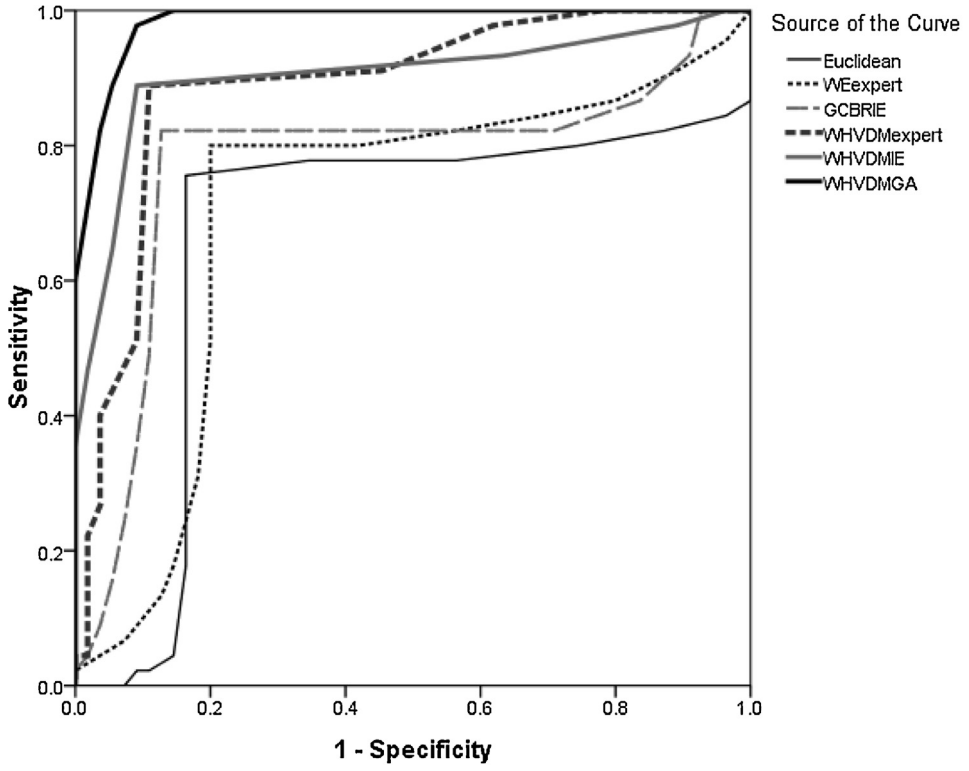

Fig. 5. ROC curves in the first study.

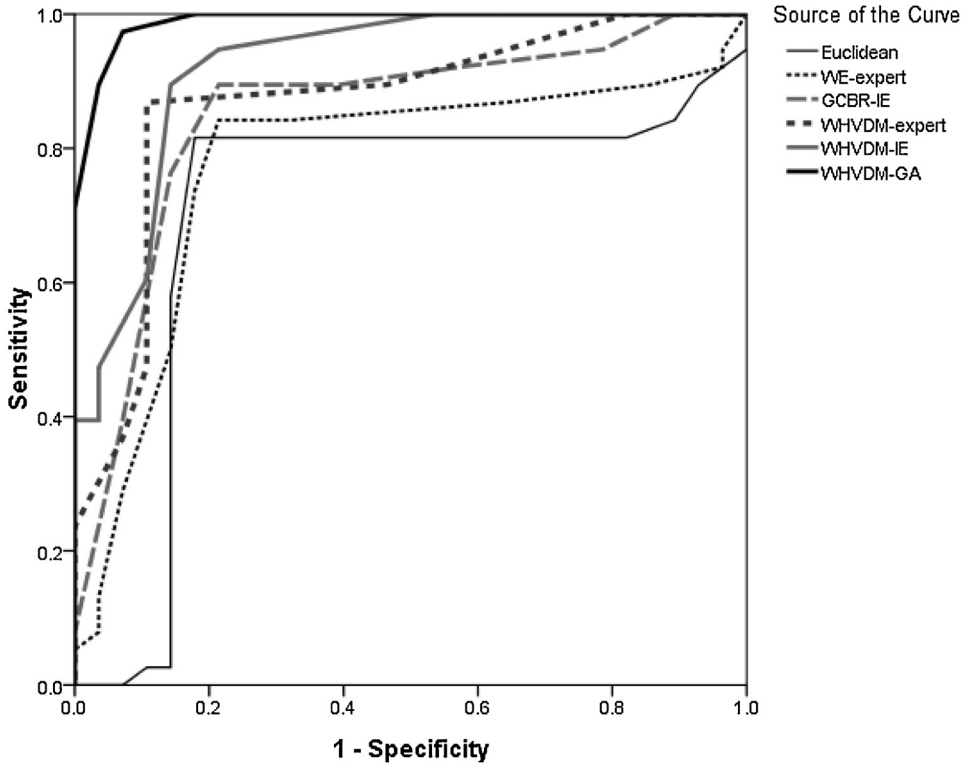

Fig. 6. ROC curves in the second study.

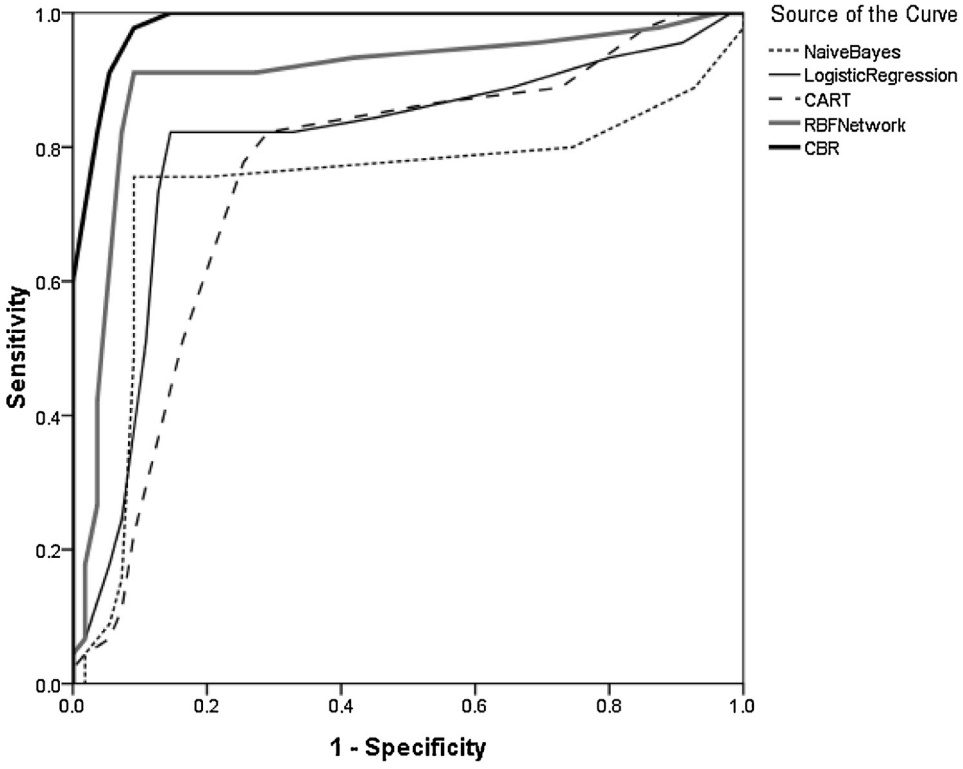

Fig. 7. ROC curves in the first study.

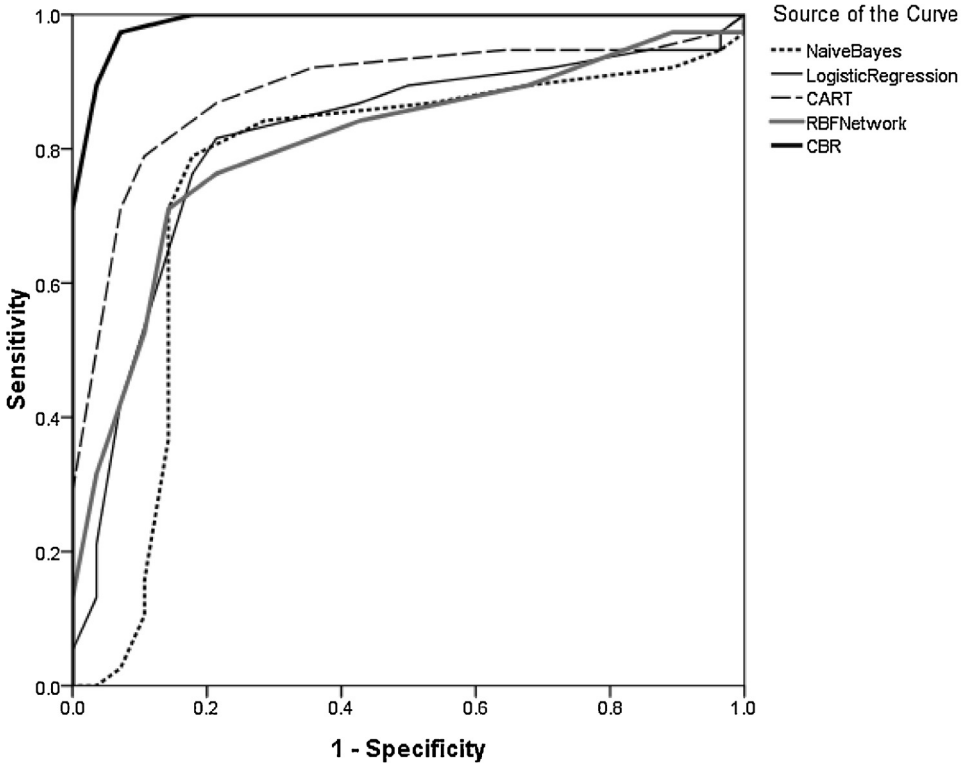

Fig. 8. ROC curves in the second study.

7. Conclusion and future research

CBR has proven to be especially applicable to problem solving and decision support in medical diagnosis. CBR research, as well as its practical applications in medicine, is currently experiencing rapid growth [66]. Case retrieval is a key component of CBR. Studying more high-performance retrieval algorithms for better knowledge mining is one of the highlighted trends of CBR development.

Table 7
The user evaluation result.

Item of Evaluation	Percentage of participants answering "Yes"
Usefulness to improve medical quality	91.7
Ease of use	75.0
Satisfaction of use	83.3
Usefulness to reduce medical cost	83.3
Usefulness to save medical resource	91.7
Usefulness to improve oncologists' specialization level	91.7
Usefulness to improve patients' satisfaction	66.7
Intention of continued use	100

In this paper, we have presented a CBR system for breast cancer diagnosis. Our goal is to provide physicians an effective way to support their clinical decision making during their diagnosis or treatment process, especially in early detection of breast cancer. Motivated by our observations of practical difficulties during CBR applications, we used a distance measure named WHVDM for dealing with both continuous and discrete attributes and a genetic algorithm for learning attribute weights involved in this distance measure. Experiments in two studies, concerning the diagnoses of breast cancer and other secondary cancers, respectively, using data from historical electronic health care records, showed the utility of the distance measure and the genetic algorithm for attribute weight learning. For such common, life-threatening diseases, even a small improvement in prediction performance has significant impacts.

As the reasoning ability of CBR is dependent on the size of the case base, it can be expected that the case matching performance will get further improved with the accumulation of cases in continued applications. Our system has been taken on trial and is favorably considered and acclaimed by physicians. It is hopeful that, with further enhancements, such as friendlier human-computer interface, more cases, and visual representations, the system will be put into clinical use.

CBR appears to have an advantage over other knowledge-based methods in ease of use [67]. Reasoning from historical cases is natural for clinical staff members. The way of reasoning in CBR resembles that in the human mind. For physicians, CBR is more natural and easier to understand than other classification methods, such as logistic regression, neural network, and decision tree. This makes it easier for clinical use by doctors.

CBR also provides more valuable information support for physicians' decision making [65,68]. Physicians can obtain complete historical cases similar to a new case. According to our practical investigation during our field studies, the cases contain plenty of information, not only the attributes used in case matching, but also data on diagnosis processes and main consultations, treatment alternatives and processes, treatment effects, post-treatment symptoms and signs, suggested nursing strategy, etc. All the information is extremely valuable to the further diagnosis, treatment, and even nursing for the physicians and nurses. In contrast to other methods that only offer a prediction, CBR offers more valuable information and knowledge, beyond the attributes used to make the prediction, to the oncologists for their reference in their decision making.

Our work opens up several avenues for further research. First, more large-scale field experiments may be conducted to further validate the utility of BTCBRsys, considering that we could collect only a few hundreds of cases in our experiments. Second, future research may further integrate natural language processing and other intelligent technologies into CBR to make use of clinical texts, which are prevalent in patients' health care records, in case matching. Third, although the CBR system was developed in the context of breast cancer diagnosis, it is general in principal and can be adapted and validated in other contexts. Fourth, while we have compared BTCBRsys with a few other retrieve methods (Euclidean, WE-Expert, GCBR-IE, WHVDM-Expert, and WHVDM-IE), it would be interesting to compare it with more different CBR configurations (such as different case reuse and retrieve methods). For example, probabilistic approaches for case reuse, such as Pous et al.'s method [14], may be evaluated in future research. Fifth, the two collaborating hospitals decided to keep case maintenance completely under the control of authoritative experts through a manual process. When an oncologist sees a new case that has good quality and should be retained in the case base based on her judgement, she submits a recommendation with justifications. The director of the department of clinical oncology will review the recommendation and decide whether to approve or reject it. The director may consult other experts before making the decision if needed. Future research may further investigate the effectiveness of automated or semi-automated case maintenance. Sixth, future research may use ontologies to deal with the evaluation of case similarity. In artificial intelligence and in the medical domain, there is a trend of using ontologies to better evaluate the difference between discrete attributes [36]. Ontologies can be used for case representation, storage, and evaluation. Using breast cancer ontologies will enhance the interoperability and integration with medical systems. Moreover, the use of fuzzy or imprecise knowledge further advances the semantic effectiveness of medical CBR systems. Future research effort may be devoted to the development of an ontology-based semantic CBR system. Seventh, usually the CBR approach allows to learn about good behavior, but from mistakes as well. It will be very interesting to investigate how to discover knowledge from adverse historical cases (i.e., mistakes).

Conflicts of interest

None.

Acknowledgements

The authors would like to thank Dr. Carmelo Gaudioso for his help with data collection. This research is partially supported by the China National Natural Science Foundation under grants Nos. 71331002, 71301040, 71573071, 71601061, as well as NSFC Creative Research Group Program under grant No. 71521001.

Appendix A.

A Bibliometric Analysis of Case-Based Reasoning in Cancer-related Medicine (2011- 2016)

For an overall understanding of the recent literature on CBR in cancer-related medicine, we conducted a bibliometric analysis based on 531 journal articles published in the past few years (from 2011 to 2016) and indexed in the *Web of Science* database (SCI-E, CPCI-S).[4] According to Bradford's law, most key studies are published in core journals. We therefore collected data on articles published in core journals from the Web of Science for our literature review.

Using CiteSpace, we performed a co-word network analysis of keywords. We extracted the 50 most frequent keywords in each year and constructed a co-word network based on them. We trimmed the network using minimum spanning tree. The results are shown in Fig. A1. The colorful banners on the top represent

[4] The search strategy is ((case-based reason) or (CBR) or (case retrieval)) and ((cancer) or (tumor) or (tumour) or (neoplasm) or (carcinoma)).

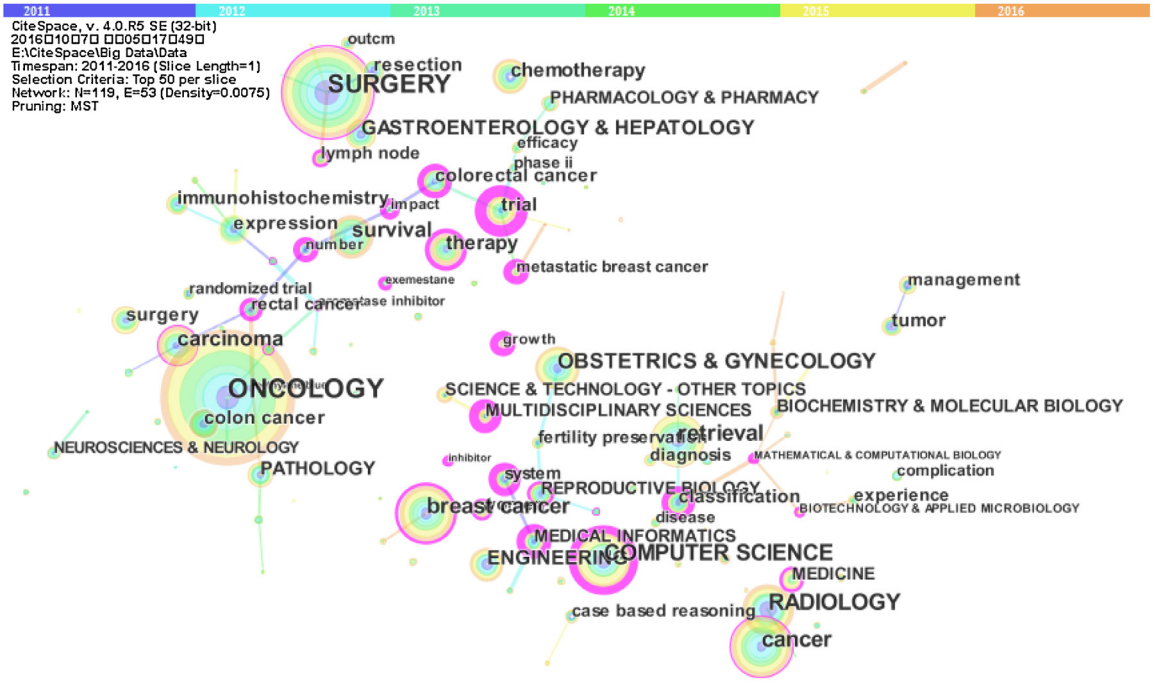

Fig. A1. Co-word analysis result.

different years (2011–2016). A ring consisting of multiple color layers represents a keyword (called node). The size of a ring corresponds to the frequency of the keyword; a bigger ring indicates a more frequent keyword. The thickness of a color layer in a ring represents the frequency of the keyword in one year. The line connecting two nodes denotes the co-occurrence relation between the two keywords. The thickness of the line corresponds to the co-occurrence frequency; the thicker the line; the more frequently the two keywords co-occurred in the same articles. The color of the line indicates the year in which the two keywords co-occurred in the same article for the first time. Some of the top hot keywords are case retrieval, oncology, surgery, computer science, radiology, survival, therapy, trial, and breast cancer. Based on betweenness centrality, keywords in important positions in the co-word network include trial, breast cancer, system, classification, and therapy.

Appendix B.

Classification Methods Used in the Experiments

Besides CBR, a variety of other classification methods have also been used for classification decision making in medicine, especially in cancer diagnosis, treatment, and prognosis [1–3]. Some of the methods are artificial neural network (ANN) [4,5], decision tree [6], and logistic regression [5]. For example, Zhang et al. [7] presented an integrative approach to building neural network models for the diagnosis of breast cancer. The ANN classifier they proposed can produce a single-valued diagnostic index to be used to discriminate malignant from benign pelvic masses. Tan et al. [8] proposed a complementary learning fuzzy neural network, which exploits the lateral inhibition between positive and negative samples for ovarian cancer diagnosis. Chen et al. [9] proposed evolving hierarchical RBF neural networks for breast cancer detection. Cruz-Ramirez et al. [6] used decision trees and Bayesian networks to discover interobserver variability in the cytodiagnosis of breast cancer.

We used the following classification methods in our experiment: RBF neural network, CART decision tree, logistic regression, and Naïve Bayes. Naïve Bayes estimates the conditional probability of the class given the attribute values based on training data. The estimation is simplified by assuming that the attributes are conditionally independent given the class value. Logistic regression is a no-parametric linear model, where the logit (i.e., log odds) is regressed as a linear function of the attributes. The regression coefficients are usually estimated using maximum likelihood estimation. Decision tree methods take a "divide and conquer" approach and build tree-like sequential models. Each intermediate node tests an attribute to select the next subtree to traverse, and a leaf leads to a prediction on the class. The selection of an attribute for an intermediate node is based on a measure of the dependence between the class and an attribute (e.g., information gain and gain ratio). A radial basis function (RBF) neural network [10] is an artificial neural network in which radial basis functions are used as activation functions. It typically has an input layer, a hidden layer with a non-linear RBF activation function, and a linear output layer. RBF networks are typically trained in two steps. First, the center vectors of the RBF functions in the hidden layer are determined in an unsupervised manner. Next, the linear model from the hidden layer to the output layer is fit with respect to the objective function (e.g., least square for regression and error rate for classification).

These methods deal with continuous and discrete attributes in different ways. For example, decision trees and Naïve Bayes handle discrete attributes naturally but typically discretize continuous attributes. On the other hand, logistic regression and RBF neural network handle continuous attributes naturally but convert each discrete attribute into several binary dummy variables.

We used the default parameter values for the classification methods in Weka. For example, for RBF Network, there is one hidden layer of basis functions, the basis functions are clusters generated by the k-means clustering algorithm, there are two nodes (i.e., clusters) on the hidden layer, the minimum standard deviation for the clusters is 0.1, and logistic regression is used to fit the linear model from the hidden layer to the output layer. For more details on the implementation and configuration of the classifica-

tion methods in Weka, please refer to the documentation of the toolkit (http://www.cs.waikato.ac.nz/ml/weka/).

REFERENCES

[1] Salchenberger LM, Cinar EM, Lash NA. Neural networks: a new tool for predicting thrift failures. Decision Sciences 1992; 23(4): 899–916.

[2] West D, Mangiameli P, Rampal R, West V. Ensemble strategies for a medical diagnostic decision support system: A breast cancer diagnosis application. European Journal of Operational Research 2005; 162(2): 532–551.

[3] Lee S. Using data envelopment analysis and decision trees for efficiency analysis and recommendation of B2C controls. Decision Support Systems 2010; 49(4): 486–497.

[4] Shin KS, Han I. Case-based reasoning supported by genetic algorithms for corporate bond rating. Expert Systems with applications, 1999; 16(2): 85–95.

[5] Chang CL, Hsu MY. The study that applies artificial intelligence and logistic regression for assistance in differential diagnostic of pancreatic cancer. Expert Systems with Applications, 2009; 36(7): 10663–10672.

[6] Cruz-Ramirez N, Acosta-Mesa HG, Carrillo-Calvet H, Barrientos-Martínez RE. Discovering interobserver variability in the cytodiagnosis of breast cancer using decision trees and Bayesian networks. Applied Soft Computing 2009; 9(4): 1331–1342.

[7] Zhang Z, Zhang H, Bast Jr RC. An Application of Artificial Neural Networks in Ovarian Cancer Early Detection. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00)-Volume 4, p.4107, July 24–27, 2000. URL. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=860758 [Accessed 17.01.17].

[8] Venkatesh V, Speier C, Morris MG. User acceptance enablers in individual decision making about technology: Toward an integrated model. Decision Sciences, 2002; 33(2): 297–316.

[9] Chen YH, Wang Y, Yang B. Evolving Hierarchical RBF Neural Networks for Breast Cancer Detection. In: Proceedings of the International Conference on Neural Information Processing (ICONIP) 2006; pp. 137–144. URL. http://link.springer.com/chapter/10.1007/11893295_16 (Accessed 17.01.17).

[10] Bugmann G. Normalized Gaussian radial basis function networks. Neurocomputing, 1998; 20(1): 97–110.

Appendix C

Experiment with the Breast Cancer Wisconsin Dataset

We have experimented with the Breast Cancer Wisconsin dataset in the UC Irvine machine learning repository (http://archive.ics.uci.edu/ml/). The results are summarized in Table C1 and Fig. C1.

Table C1
Results from the Experiment with the Breast Cancer Wisconsin Dataset.

Setting	Accuracy	Sensitivity × specificity	F-measure
Euclidean	0.928	0.843	0.895
WE-expert	0.940	0.871	0.913
GCBR-IE	0.942	0.876	0.916
WHVDM-expert	0.968	0.934	0.954
WHVDM-IE	0.967	0.932	0.952
WHVDM-GA	0.971	0.937	0.958

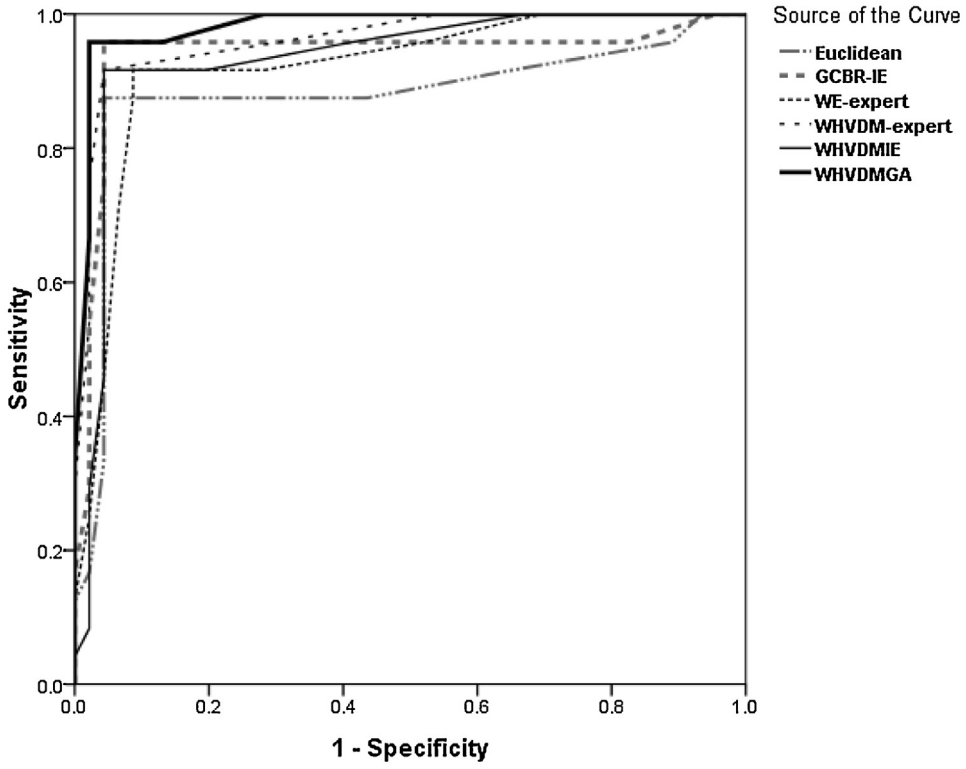

Fig. C1. ROC Curves from the Experiment with the Breast Cancer Wisconsin Dataset.

References

[1] American Cancer Society. Cancer Facts & Figures. Atlanta: American Cancer Society, Inc.; 2015 (Accessed 17 January 2017) http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf.

[2] American Cancer Society. Breast Cancer Facts & Figs 2015–2016. Atlanta: American Cancer Society, Inc.; 2015 (Accessed 17 January 2017) http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-046381.pdf.

[3] Fan L, Strasser-Weippl K, Li JJ, Finkelstein DM, Yu KD, Chen WQ, et al. Breast cancer in China. Lancet Oncol 2014;15(7):e279–89.

[4] Salchenberger LM, Cinar EM, Lash NA. Neural networks: a new tool for predicting thrift failures. Decis Sci 1992;23(4):899–916.

[5] Yarnold J. Early and locally advanced breast cancer: diagnosis and treatment National Institute for Health and Clinical Excellence guideline 2009. Clin Oncol 2009;21(3):159–60.

[6] Houssami N, Ciatto S, Martinelli F, Bonardi R, Duffy SW. Early detection of second breast cancers improves prognosis in breast cancer survivors. Ann Oncol 2009;20(9):1505–10.

[7] Ayer T, Chhatwal J, Alagoz O, Kahn Jr CE, Woods RW, Burnside ES. Comparison of logistic regression and artificial neural network models in breast cancer risk estimation. Radiographics 2010;30(1):13–22.

[8] Hung M, Shanker M, Hu M. Estimating breast cancer risks using neural networks. J Oper Res Soc 2002;53(2):222–31.

[9] Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. Oper Res 1994;43(4):570–7.

[10] West D, Mangiameli P, Rampal R, West V. Ensemble strategies for a medical diagnostic decision support system: a breast cancer diagnosis application. Eur J Oper Res 2005;162(2):532–51.

[11] Lotfy Abdrabou EAM, Salem ABM. A breast cancer classifier based on a combination of case-based reasoning and ontology approach. In: Proceedings of the 2010 International Multiconference on Computer Science and Information Technology (IMCSIT). 2010. p. 3–10. IEEE. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5680045 (Accessed 17 January 2017).

[12] Bilska-Wolak AO, Floyd Jr CE. Investigating different similarity measures for a case-based reasoning classifier to predict breast cancer. In: Proc. SPIE Medical Imaging 2001 4322: 1862–1866. 2017 (Accessed 17 January 2017) http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=906894.

[13] Lieber J, Bresson B. Case-based reasoning for breast cancer treatment decision helping. In: Blanzieri E, Portinale L, editors. Advances in Case-Based Reasoning- Proceedings of the fifth European Workshop on Case-Based Reasoning (EWCBR-2000). 2000. p. 173–85. LNAI, 1898, Springer.

[14] Pous C, Gay P, Pla A, Brunet J, Sanz J, Cajal TR, et al. Modeling reuse on case-based reasoning with application to breast cancer diagnosis. In: Dochev D, Pistore M, Traverso P, editors. International Conference on Artificial Intelligence: Methodology Systems and Applications AIMSA 2008. 2008. p. 322–32. LNAI.5253, Lecture Notes in Computer Science (Accessed 17 January 2017) http://link.springer.com/chapter/10.1007%2F978-3-540-85776-1_27.

[15] Salamó M, Golobardes E. Analysing rough sets weighting methods for case-based reasoning systems. Intel Artif: Rev Iberoam Intel Artif 2002;6(15):3.

[16] Floyd Jr CE, Lo JY, Tourassi GD. Case-based reasoning computer algorithm that uses mammographic findings for breast biopsy decisions. Am J Roentgenol 2000;175(5):1347–52.

[17] Cunningham P. A taxonomy of similarity mechanisms for case-based reasoning. IEEE Trans Knowl Data Eng 2009;21(11):1532–43.

[18] Varpa K, Iltanen K, Juhola M. Genetic algorithm based approach in attribute weighting for a medical data set. Journal of Computational Medicine 2014:1–11. Article ID 526801.

[19] Shin KS, Han I. Case-based reasoning supported by genetic algorithms for corporate bond rating. Expert Syst Appl 1999;16(2):85–95.

[20] Pla A, LóPez B, Gay P, Pous C. eXiT* CBR. v2: Distributed case-based reasoning tool for medical prognosis. Decis Support Syst 2013;54(3):1499–510.

[21] Beddoe GR, Petrovic S. Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering. Eur J Oper Res 2006;175(2):649–71.

[22] Choi JY, Kim DH, Plataniotis KN, Ro YM. Classifier ensemble generation and selection with multiple feature representations for classification applications in computer-aided detection and diagnosis on mammography. Expert Syst Appl 2016;46:106–21.

[23] Song L, Hsu W, Xu J, van der Schaar M. Using contextual learning to improve diagnostic accuracy: application in breast cancer screening. IEEE J Biomed Health Inform 2016;20(3):902–14.

[24] Din NU, Ukoumunne OC, Rubin G, Hamilton W, Carter B, Stapley S, et al. Age and gender variations in cancer diagnostic intervals in 15 cancers: analysis of data from the UK Clinical Practice Research Datalink. PLoS One 2015;10(5):e0127717.

[25] Walter FM, Rubin G, Bankhead C, Morris HC, Hall N, Mills K, et al. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study. Br J Cancer 2015;112:S6–13.

[26] Gu DX, Liang CY, Chen WE, Lu WX, Zuo CR. Using information entropy as weight determination approach in breast cancer case retrieval. Adv Mater Res 2011;268–270:1619–22.

[27] Montani S. Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support. Appl Intel 2008;28(3):275–85.

[28] Nilsson M, Funk P, Olsson EMG, von Schéele B, Xiong N. Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system. Artif Intell Med 2006;36(2):159–76.

[29] Ayeldeen H, Shaker O, Hegazy O, Hassanien AE. Distance similarity as a CBR technique for early detection of breast cancer: an Egyptian case study. In: Information Systems Design and Intelligent Applications. India: Springer; 2015. p. 449–56 (Accessed 17 January 2017) http://link.springer.com/chapter/10.1007/978-81-322-2247-7_46.

[30] Yazid H, Kalti K, Benamara NE. A new similarity measure based on Bayesian network signature correspondence for brain tumors cases retrieval. Int J Comput Intell Syst 2014;7(6):1123–36.

[31] Huang ML, Hung YH, Lee WM, Li RK, Wang TH. Usage of case-based reasoning: neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. J Med Syst 2012;36(2):407–14.

[32] Chuang CL. Case-based reasoning support for liver disease diagnosis. Artif Intell Med 2011;53(1):15–23.

[33] Sharaf-El-Deen DA, Moawad IF, Khalifa ME. A new hybrid case-based reasoning approach for medical diagnosis systems. J Med Syst 2014;38(2):1–11.

[34] Saraiva R, Perkusich M, Silva L, Almeida H, Siebra C, Perkusich A. Early diagnosis of gastrointestinal cancer by using case-based and rule-based reasoning. Expert Syst Appl 2016;61:192–202.

[35] Ghany KKA, Ayeldeen H, Zawbaa HM, Shaker O. A rough set-based reasoner for medical diagnosis. IEEE International Conference on Green Computing and Internet of Things 2015:429–34 (Accessed 17.01.17) http://ieeexplore.ieee.org/document/7380502/.

[36] El-Sappagh S, Elmogy M, Riad A. A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis. Artif Intell Med 2015;65(3):179–208.

[37] Petrovic S, Khussainova G, Jagannathan R. Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. Artif Intell Med 2016;68:17–28.

[38] Khussainova G. Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. In: Proceedings of the ICCBR conference 2014 Doctoral Consortium. ICCBR. 2014 (Accessed 17.01.17) https://idea.library.drexel.edu/islandora/object/idea%3A5974/datastream/OBJ/download/Proceedings_of_the_ICCBR_2014_Doctoral_Consortium.pdf#page=29.

[39] Petrovic S, Mishra N, Sundar S. A novel case based reasoning approach to radiotherapy planning. Expert Syst Appl 2011;38(9):10759–69.

[40] Mishra N, Petrovic S, Sundar S. A self-adaptive case-based reasoning system for dose planning in prostate cancer radiotherapy. Med Phys 2011;38(12):6528–38.

[41] Bulu H, Alpkocak A, Balci P. Ontology-based mammography annotation and case-based retrieval of breast masses. Expert Syst Appl 2012;39(12):11194–202.

[42] van den Branden M, Wiratunga N, Burton D, Craw S. Integrating case-based reasoning with an electronic patient record system. Artif Intell Med 2011;51(2):117–232.

[43] Recio-García JA, González-Calero PA, Díaz-Agudo B. jcolibri2: A framework for building Case-based reasoning systems. Sci Comput Program 2014;79:126–45.

[44] Ahn H, Kim KJ. Global optimization of case-based reasoning for breast cytology diagnosis. Expert Syst with Appl 2009;36(1):724–34.

[45] Gu DX, Liang CY, Kim KS, Yang CH, Cheng WJ, Wang J. Which is more reliable: expert experience or information itself? weight scheme of complex cases for health management decision making. International J Inf Technol Dec Mak 2015;14(3):597–620.

[46] Yan AJ, Shao HS, Guo Z. Weight optimization for case-based reasoning using membrane computing. Inf Sci 2014;287:109–20.

[47] Riesbeck CK, Schank RC. Inside case-based reasoning. Hillsdale, NJ, US: Lawrence Erlbaum Associates; 1989.

[48] Hüllermeier E, Minor M. Case-Based reasoning research and development. In: Proceedings of 23rd International Conference, ICCBR 2015. 2015.

[49] Aamodt A, Plaza E. Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun 1994;7(1):39–59.

[50] Gupta UG. How case-based reasoning solves new problems. Interfaces 1994;24(6):110–9.

[51] Kolodner JL, Penberthy TL. A case-based approach to creativity in problem solving. Proceedings of the Twelfth Annual Conference of the Cognitive Science Society. Cambridge, Massachusetts, vol. 978. 1990.

[52] Wilson DR, Martinez TR. Improved heterogeneous distance functions. J Artif Intel Res 1997;6:1–34.

[53] Gu DX, Liang CY, Li XG, Yang SL, Zhang Z. Intelligent technique for knowledge reuse of dental medical records based on case-based reasoning. J Med Syst 2010;34(2):213–22.

[54] McCane B, Albert M. Distance functions for categorical and mixed variables. Pattern Recognit Lett 2008;29(7):986–93.

[55] Stanfill C, Waltz D. Toward memory-Based reasoning. Commun ACM 1986;29(12):1213–28.

[56] Gower JC. A general coefficient of similarity and some of its properties. Biometrics 1971;27:857–74.

[57] Lee S. Using data envelopment analysis and decision trees for efficiency analysis and recommendation of B2C controls. Decis Support Syst 2010;49(4):486–97.
[58] Fidelis MV, Lopes HS, Freitas AA. Discovering comprehensible classification rules with a genetic algorithm. Proceedings of the 2000 Congress on Evolutionary Computation 2000;vol. 1:805–10.
[59] Goldberg DE. Genetic Algorithms in Search, Optimization and Machine Learning. Reading, MA: Addison-Wesley; 1989.
[60] Bilska-Wolak AO, Floyd Jr CE. Development and evaluation of a case-based reasoning classifier for prediction of breast biopsy outcome with BI-RADS™ lexicon. Med Phys 2002;29(9):2090–100.
[61] Dixon JM, John TG. Morbidity after breast biopsy for benign disease in a screened population. Lancet 1992;339(8785):128.
[62] Mellios C, Six P, Lai AN. Dynamic speculation and hedging in commodity futures markets with a stochastic convenience yield. Eur J Oper Res 2015;250(2):493–504.
[63] Sexton RS, Dorsey RE. Reliable classification using neural networks: a genetic algorithm and backpropagation comparison. Decis Support Syst 2000;30(1):11–22.
[64] Zhang Z, Zhang H, Bast Jr RC. An application of artificial neural networks in ovarian cancer early detection. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN'00) 2000;vol. 4:107–12 (Accessed 17 January 17) http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=860758.
[65] Venkatesh V, Speier C, Morris MG. User acceptance enablers in individual decision making about technology: toward an integrated model. Decis Sci 2002;33(2):297–316.
[66] Bichindaritz I, Marling C. Case-based reasoning in the health sciences: what's next? Artif Intell Med 2006;36(2):127–35.
[67] Witten IH, Frank E. Data mining: practical machine learning tools and techniques. 2nd ed. San Francisco, CA: Morgan Kaufmann; 2005.
[68] Venkatesh V, Morris MG. Why don't men ever stop to ask for directions? gender: social influence, and their role in technology acceptance and usage behavior. MIS Q 2000;24(1):115–39.