

Computing Semantic Similarity between ECG-Information Concepts Based on an Entropy-Weighted Concept Lattice

Kai Wang* and Shu Yang*

Abstract

Similarity searching is a basic issue in information processing because of the large size of formal contexts and their complicated derivation operators. Recently, some researchers have focused on knowledge reduction methods by using granular computing. In this process, suitable information granules are vital to characterizing the quantities of attributes and objects. To address this problem, a novel approach to obtain an entropy-weighted concept lattice with inclusion degree and similarity distance (ECLisd) has been proposed. The approach aims to compute the combined weights by merging the inclusion degree and entropy degree between two concepts. In addition, another method is utilized to measure the hierarchical distance by considering the different degrees of importance of each attribute. Finally, the rationality of the ECLisd is validated via a comparative analysis.

Keywords

Concept Lattice Theory, ECG Concept, Entropy, Inclusion-Degree, Semantic Computing

1. Introduction

Searching similar patterns between various concepts is a major issue in medical information science and systems [1-4]. In this regard, some researchers have tried to discover useful patterns by measuring semantic similarity in Chinese medicine [5], biological knowledge discovery [6], and health record systems [7]. All of these methods are unable to represent and analyze the inner relationships between various semantic levels [8]. To address these relationships, concept lattice theory (CLT) was proposed to process domain knowledge with generalization and instantiation relationships among the attributes. In other words, the formal context might create a large number of concepts, generating considerable information. In that case, CLT is combined with other models [9,10]. Each of the available approaches provide a way to assign conceptual weights to the attributes, without the ability to differentiate attributes based on their values in different applications, such as assigning different weights according to various similarities between concepts. Therefore, a novel approach to compute the combined weight of the attributes is introduced in the domain context to obtain similarity distances among formal concepts.

To characterize the properties of similar symptoms to known diseases in medical sets, at present, formal concept analysis (FCA)-based similarity methods are proposed to process the concepts. This motivates

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received September 13, 2018 ; first revision December 11, 2018; second revision March 8, 2019; accepted April 14, 2019.

Corresponding Author: Shu Yang (shuyang_prof@sina.com)

* School of Health Management, Bengbu Medical College, Bengbu, China (wangkai0552@126.com, shuyang_prof@sina.com)

the proposed paper to concentrate on similarity measurements among different entities. Considering the numerous research methods, we divide the main semantic calculation methods into four categories [11]. The distance-based approach is the most common model for calculating semantic distances among semantic nodes, where the main parameter is the path and its connection. The feature-based approach introduces the feature of attributes as a main variable [12]. The information content (IC)-based model exploits normalized entities via information classification knowledge, according to Shannon information theory [13]. The integrated approach fully harnesses the advantages of the previous measures. Gao et al. [14] considered the cooccurrence frequency of concepts to optimize the measurement of conceptual similarity distance. Ghafourian et al. [15] partitioned a given ontology by mixing the feature-based approach and structural network approach according to its semantic structure to achieve reusability. Comparisons of these approaches are shown in Table 1. In addition, the approaches mentioned above have two drawbacks. First, these approaches rely on relatively clear classification information or a stable ontology structure. Specifically, the concepts in the semantic graph are represented in a certain way without considering the nonbinary relationships between the concepts, which fails to assign weights of non-taxonomic relations between the concepts. Second, the feature-based approach does not flexibly differentiate concepts that belong to different classification relationships, and it is difficult to assign corresponding semantic weights based on different application contexts.

Table 1. Comparisons among different approaches for similarity measurement

	Distance-based	Feature-based	IC-based	FCA-based
Graph	√	√	√	√
Lattice	√	×	√	√
Information pre-classification	×	×	×	√
Non-binary relationship	×	×	×	√
Flexibility	×	√	×	√

The objective of the current paper is to find more relevant concept sets in the electrocardiogram (ECG) domain ontology. To achieve this goal, the ECLisd is proposed to measure semantic similarity on the basis of an entropy-weighted concept lattice with inclusion degree and similarity distance, which introduces the ECG ontologies as a semantic representation to extract essential domain properties. Actually, considering specific application backgrounds, it is much easier to obtain the semantic relationship among entities in the domain ontology. Specifically, the ECLisd first computes the combined weight of the attributes of each concept by integrating its inclusion-degree importance and entropy-degree importance in the domain context; then, the weighted concept lattice with a hierarchical structure is constructed based on the formal concept lattice. Finally, the semantic similarity between weighted formal concepts is calculated.

The current paper focuses on the following steps to achieve the above goals:

- (1) The specific classification of the combined weights of attributes is extracted and represented as a formal context;
- (2) The inclusion degree and entropy degree are combined to obtain the importance degree of weighted attributes;
- (3) The hierarchical relationship for weighted entities can be computed to analyze the semantic distance with an illustration.

The rest of this article is arranged as follows: a weighted concept lattice with combined weights is built on the basis of information entropy in Section 2. Section 3 measures the hierarchical distances of each entity. Section 4 calculates the similarity distances of formal concepts based on the ECLisd and discusses the results concisely. Finally, in Section 5, the contributions of our approach are summarized, and the direction of future work is outlined.

2. Weighted Concept Lattice

2.1 Extraction of the Combined Weight of Attributes

In this section, we mainly discuss the specific domain knowledge to make the demonstration clearer. Ontology is viewed as a useful method for representing knowledge in many studies [16]. In ECG information science, various scholars have managed to establish and extract essential attributes of ECG-related concepts from prominent ECG standards. The necessary attributes of ontologies based on categories are extracted from GB/T14396 (disease classification and coding standards) and definitions of these categories from T27733-2011 (medical material reference standard). The ECG ontologies representing ECG-related concepts can effectively express the implicit semantic relationship between the ontology concepts and their attributes [17]. ECG-related concepts with attributes are listed in Table 2. The attributes with marks in Table 2 are listed in Table 3.

Table 2. ECG-related concepts with attributes table

Concept	Time	Amplitude	Lead	Waveform	Segment	Reference beat	Complex
AMI	a	c	e	g	j	l	o
CAD	a	c	e	g	k	n	p
IE	a	d	f	g	j	l	o
AF	a	c	f	h	j	l	p
RHD	a	d	f	h	j	l	p
CHD	b	d	f	i	j	m	∅

Letters “a” to “p” represent the values, ∅ is the object without the corresponding data.

AMI=acute myocardial infarction, CAD=coronary heart disease, IE=infective endocarditis, AF=atrial fibrillation; RHD=rheumatic heart disease, CHD=congenital heart disease.

Table 3. Attributes of Table 2 and their marks

Identifier	Attribute	Identifier	Attribute
a	Time/peak separation	i	Waveform/P-loss
b	Time/band width	j	Segment/ST-depression
c	Amplitude/single	k	Segment/RR-inordinance
d	Amplitude/double	l	Reference beat/QRS
e	Lead/limb	m	Reference beat/ST
f	Lead/chest	n	Reference beat/QSP
g	Waveform/T-inversion	o	Complex/widening
h	Waveform/Q-duration	p	Complex/duration

Assigning and determining conceptual weights is important for obtaining precise results. In the decision table, the properties with different weights are referred to as a feature, which are represented as an attribute with values. In our model, we compute the weight of an attribute by combining two factors with independent influences. We first introduce the inclusion-degree importance as the first factor to quantify the importance of different attributes. Then, the entropy-degree importance is proposed, denoted as the overall information of weighted concepts. With this in mind, it is easy to determine how an attribute affects the semantic distance of the concept by integrating two different factors.

2.1.1 The inclusion-degree-importance of attributes

To obtain attributes of different importance through a decision table, the degree of attribute importance is introduced in rough set theory. Some necessary knowledge of rough set theory is briefly described as follows.

Definition 1 [18]. Let $S = (U, A, V, f)$ be an information system, where $A=C\cup D$ is a nonempty set of attributes, in which C is a set of conditional attributes and D is a set of decision properties; $f : U \times A \rightarrow V$ is a mapping function that designates a nonempty value to each attribute of its object. S can be denoted as a decision table only if $C \cap D = \emptyset$.

For instance, if Table 1 contains the decision properties that satisfy the conditions in Definition 1, it can be defined as a decision table, where U contains *AMI*, *CAD*, *IE*, *AF*, *RHD* and *CHD* and A includes *time*, *amplitude*, *lead*, *waveform*, *segment*, *reference beat*, and *complex*.

Definition 2 [18]. Let $S = (U, C\cup D, V, f)$ be a decision table, where C is a set of conditional attributes and D is a set of decision properties; then, $\gamma_C(D)$ denotes the degree of dependence of D with respect to C :

$$\gamma_C(D) = \frac{|POS_C(D)|}{|U|} \quad (1)$$

where $POS_C(D)$ represents the positive region of partition U / D and $|\cdot|$ represents the number of the set.

For example, in Table 1, $U / \{time\} = \{\{AMI, CAD, IE, AF, RHD\}, \{CHD\}\}$, and $U / \{amplitude\} = \{\{AMI, CAD, AF\}, \{IE, RHD, CHD\}\}$. In addition, let $B \subseteq C$ and $B = \{time\}$ and $D = \{reference beat\}$ be a decision property; thus, $POS_B(D) = \{CAD, CHD\}$, $|POS_B(D)| = 2$, $|U| = 6$ and $\gamma_B(D) = 1/3$.

Definition 3 [18]. For a given decision table $S = (U, C\cup D, V, f)$, if $B \subseteq C$, $sig_\gamma(B, C, D)$ denotes the importance of U / D with respect to C .

When $B=\{a\}$, the importance of attribute a with respect to decision attribute D is:

$$sig_\gamma(a, C, D) = \gamma_C(D) - \gamma_{C-\{a\}}(D) \quad (2)$$

For instance, in Table 1, $|POS_C(D)| = 6$, $|U| = 6$ and $\gamma_C(D) = 1$. $\gamma_{C-\{time\}}(D) = |POS_{C-\{time\}}(D)| / |U| = 6 / 6 = 1$; thus, $sig_\gamma(\{time\}, C, D) = 0$.

In our work, the decision table is constructed by introducing related domain specifications and distinguishing the decision properties in the object attribute table. For example, the decision properties

can be obtained by extracting essential and electrophysiological properties and relations, based on which a decision table (Table 1) is built by tagging condition attributes and decision properties. Basically, domain specifications with related knowledge, such as the prominent ECG standards (AHA/MIT-BIH [19], SCP-ECG [20], and HL7 aECG [21]), can represent the decision properties, which reflect the general understanding of a specific context to a great extent. Consequently, the inclusion-degree importance of the attributes, which is calculated from the decision properties, reflects the degree of influence of this conditional property of the semantic concept relatively and precisely.

When there are redundant attributes in the decision table, the importance of the conditional attribute will decrease. Therefore, we introduce the inclusion degree and the importance of the inclusion degree to diminish the attributes that are unnecessary for the semantic representation.

Definition 4 [22]. Let $S = (U, A, V, f)$ be a decision table, in which $A_1, A_2 \subseteq A$, and U / A_1 and U / A_2 are partitions of U , defined as $\{X_1, X_2 \dots X_n\}$ and $\{Y_1, Y_2, \dots, Y_m\}$, respectively; $CON(A_1 / A_2)$ is denoted as the inclusion degree of U / A_2 to U / A_1 :

$$CON(A_1 / A_2) = \sum_{0 \leq i \leq n, 0 \leq j \leq m} con(X_i / Y_j) \tag{3}$$

We assume that the given set, X and Y , satisfies a function as follows: when $X \not\subseteq Y$, $con(X / Y) = 0$, and if $X \subseteq Y$, $con(X / Y) = 1$. In the above condition, the range of the function satisfies the following: $0 \leq CON(A_1 / A_2) \leq n$.

For instance, in Table 1, $U / \{time\} = \{\{AMI, CAD, IE, AF, RHD\}, \{CHD\}\}$ and $U / \{reference\ beat\} = \{\{AMI, IE, AF, RHD\}, \{CAD\}, \{CHD\}\}$; thus, $CON(time / reference\ beat) = 1$ and $CON(reference\ beat / time) = 2$.

Definition 5 [22]. Let $S = (U, C \cup D, V, f)$ be a decision table, where a is a property of the conditional properties C ; $SIG^R_\gamma(a, C, D)$ is denoted as the importance of the inclusion degree:

$$SIG^R_\gamma(a, C, D) = (|U|^2 sig^R_\gamma(a, C, D) + |U| m_a - n_a + 1) / (|U|^2 + |U| + 1) \tag{4}$$

where

$$m_a = (|POS_a(D)| + CON(a/D) / |U|) / (|U| + 1),$$

$$n_a = (|POS_D(a)| + CON(D/a) / |U|) / (|U| + 1),$$

m_a and n_a represent the influence of the positive region and the inclusion to the property importance, respectively. Notably, $0 \leq m_a, n_a \leq 1$.

For instance, in Table 1, $m_{time} = (1+1/6) / (6+1) = 0.1667$ and $n_{time} = (6+2/6) / (6+1) = 0.905$; thus, $SIG^R_\gamma(time, C, D) = (0+6*0.1667-0.905+1) / (36+6+1) = 0.0255$.

2.1.2 The entropy-degree-importance of attributes

The partial order relation is needed to construct a concept lattice, especially when we employ semantic integration between ECG data standards. This relation can be achieved through attributes of reasonable

importance. Thus, we propose another representation of combined weights to quantify the inclusion degree importance of conditional properties more precisely, which is called entropy-degree importance on the basis of Shannon entropy theory [23].

Definition 6 [23]. Let $K_m = (O, P, I, R)$ be a formal context, where $O = \{o_1, o_2, \dots, o_n\}$ and $P = \{p_1, p_2, \dots, p_k\}$; then, the probability of the i -th object (o_i) processing the corresponding j -th attribute (p_j) is computed by $P(p_j/o_i)$. The average information entropy weight of attribute p_j can be represented by $E(p_j)$ as follows:

$$E(p_j) = -\sum_{i=1}^m P(p_j / o_i) * \log_2(P(p_j / o_i)) \quad (5)$$

where m represents the total number of attributes in the given formal context K_m . For example, in Table 1, $E(a) = -\sum_{j=1}^{10} (5/6) * \log_2(5/6) = 0.219$.

Definition 7. Let $K_m = (O, P, I, R)$ be a formal context with the conditional properties, where $P = \{p_1, p_2, \dots, p_k\}$ denotes a set of conditional properties. Then, $SIG(p_j, C, D)$ denotes the inclusion degree importance of a conditional property p_j :

$$SIG(p_j, C, D) = SIG_{\gamma}^R(N(p_j, C, D)) / |N(p_j, C, D)| \quad (6)$$

where $N(p_j, C, D)$ represents a function to obtain the value of the conditional property $p_j \in C$ in decision table S . $|N(p_j, C, D)|$ is the number of elements of $N(p_j, C, D)$.

For instance, considering the decision table (Table 1), $(a, C, D) = (b, C, D) = time$; thus, $|N(a, C, D)| = 2$. Therefore, $SIG(a, C, D) = SIG_{\gamma}^R(N(time, C, D)) / |N(time, C, D)| = 0.0255/2 = 0.0128$.

To evaluate the attribute importance of formal concepts, we define a set of the combined weights of the entropy-degree importance (Eq. (5)) and the inclusion-degree importance of a conditional property (Eq. (6)).

Definition 8. Let $K_m = (O, P, I, R)$ be a formal context, where $P = \{p_1, p_2, \dots, p_k\}$ and $i(p_j)$ ($1 \leq j \leq k$) indicates the importance of an attribute p_j in P . $w_c \in W_c$ is defined as the combined weight in K_m . Then, a weighted formal context is defined as $K'_m = (O, P, I, R, W_c)$, which is generated from K_m , where the combined weight of p_j can be denoted as follows:

$$w_{c(p_j)} = SIG(p_j, C, D) * E(p_j) / \sum_{j=1}^k SIG(p_j, C, D) * E(p_j) \quad (7)$$

where $E(p_j)$ and $SIG(p_j, C, D)$ can be calculated by using Eq. (5) and Eq. (6), respectively. For example, the combined weights of attribute a are calculated as follows:

$$w_{c(a)} = SIG(a, C, D) * E(a) / \sum_{j=1}^k SIG(a, C, D) * E(a) = 0.0236.$$

2.2. Construction of the Weighted Concept Lattice

2.2.1 A formal context based on the decision table

Based on the analysis above, we first introduce a formal context consisting of an object set and an

attribute set according to FCA. Within the formal context, whole concepts can be constructed into a concept lattice using a partial order relation.

Definition 9. A formal context is described by a triple $K_m = (O, P, I, R)$, where O and P are two nonempty sets of objects and attributes, respectively, W represents the attribute values, and $R \subseteq O \times P \times I$ represents a subset of the Cartesian product of O and P ($I \subseteq O \times P$). R is a ternary relation between the three sets for which $(o, p, i) \in R$. When $o \in O$ and $p \in P$, oRp means that object o has attribute p and that attribute p belongs to object o .

For example, Table 4 is a formal context, where $O = \{m_1, m_2, m_3, m_4, m_5, m_6\}$ and $P = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p\}$. Table 4 illustrates partial objects of ECG concepts and their essential attributes by FCA in which each row and column of the formal context represent the disease related to the ECG diagnosis (extents) and their attributes (intents), according to the clinical effect of different elements on the diagnosis results. If the relation between an object and an attribute exists, then the relation is denoted as ‘*’. On the basis of Definition 9, the value of the decision properties from the decision table can be transformed into attributes in the formal context. Therefore, ‘*’ marked in Table 4 can also be viewed as an attribute that holds for a specific concept.

Table 4. A formal context in binary converted from the decision table

Mark	Object	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
m ₁	AMI	*		*		*		*			*		*			*	
m ₂	CAD	*		*		*		*				*			*		*
m ₃	IE	*			*		*	*			*		*			*	
m ₄	AF	*		*			*		*		*		*				*
m ₅	RHD	*			*		*		*		*		*				*
m ₆	CHD		*		*		*			*	*	*		*			

Each letter from “a” to “o” represents attributes in Table 2. * represents criterion satisfied.

2.2.2 A weighted concept lattice

Based on Definition 9, we define the weighted formal context by assigning the combined weight to each attribute of the formal context. Then, the weighted concept lattice from the weighted formal context is constructed.

Definition 10. Given a weighted formal context $K'_m = (O, P, I, R, W_c)$ and a complete concept lattice $(L(O, P, I, R), \leq)$, $C = (A, B)$ is a formal concept based on $(L(O, P, I, R), \leq)$. The triple $C_w = (A, B, w_c(B))$ can be defined as a weighted formal concept with combined weights generated from $(L(O, P, I, R), \leq)$, in which $w_c(B)$ represents the weights of attribute B .

$$w_c(B) = \sum_{i=1}^n w_{c(b_i)} \tag{8}$$

For instance, let $C = (\{m_2, m_4\}, \{a, c, p\})$ be a formal concept, where the sum of the combined weight of the attribute $\{a, c, p\}$ is $w_c(a, c, p) = 0.1649$. Thus, $C_w = (\{m_2, m_4\}, \{a, c, p\}, 0.1649)$ is a weighted formal concept.

3. Semantic Similarity Measurement

After constructing a weighted concept lattice according to a formal context, we calculate the weight of each weighted concept via Eq. (8). In this section, the semantic similarity between weighted concepts is measured to analyze the semantic distance in the weighted concept lattice. Moreover, partial order relations among the weighted concepts are represented by weighted concepts with the same attributes in the concept lattice. Before calculating the semantic similarity between the weighted concepts, we first introduce the hierarchical distance between formal concepts to calculate the semantic similarity according to the general special relation between the concepts.

3.1 Hierarchical Distance between Formal Concepts

Definition 11. Let $(L(O, P, I, R), \leq)$ be a complete concept lattice. Given that $L_{\text{sup}} \subseteq L$ is a subset of L , $\text{sup}(L_{\text{sup}})$ is denoted as a supremum of the subset L_{sup} . In particular, $\text{sup}(a, b)$ is defined as the supremum of concepts a and b only if $L_{\text{sup}} = \{a, b\}$.

For instance, in the complete concept lattice (L, \leq) (Fig. 1), $\text{sup}(C3, C4) = \text{sup}(C4, C3) = C1$.

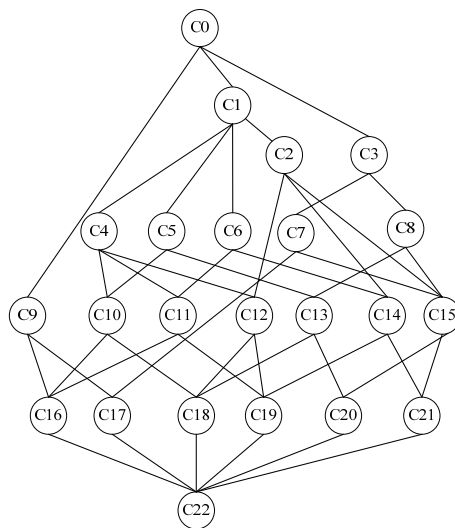


Fig. 1. A complete concept lattice.

Definition 12. Let $(L(O, P, I, R), \leq)$ be a complete concept lattice, where a and c are defined as two formal concepts belonging to L . Given that a is the subconcept of c , $\text{dis}(a, c)$ is denoted as the shortest distance from a to c in the concept lattice, which is the least number of connecting edges from vertices a to c . In particular, we define $\text{dis}(a, a) = 0$.

For example, in Fig. 1, one can calculate $\text{dis}(C11, C2) = 2$ and $\text{dis}(C6, C2) = 1$, as $C11 < C2$ and $C6 < C2$ in the concept lattice. Because the shortest distance in Definition 12 is merely defined from the lower vertex (subconcept) to its higher vertex (superconcept) or itself, $\text{dis}(C2, C11)$ is not defined.

In the classic transformation model, the distance between concepts is normally an indicator that reflects their semantic similarities [24]. Meanwhile, the vertices representing concepts are usually measured by a specific type of distance, which is also used to reflect the similarity between these two concepts.

Similarly, the shortest distance defined in Definition 12 possesses a function to calculate the semantic similarity according to the general special relation between the concepts. In the formal concept lattice, the intent set of the subconcept basically inherits the intent of its superconcept. As a result, the subconcepts always share the same subset of characteristics of the superconcept. For example, in Fig. 1, concepts $C6$ and $C7$ have the same superconcept $C0$; however, $dis(C6, C0) = 2$, while $dis(C7, C0) = 1$. Therefore, concept $C7$ has a smaller degree of difference from $C0$ than concept $C6$.

A general concept with a smaller depth is usually more abstract, which tends to cover less information content [25]. Conversely, if the underlying concept is more specific, it always inherits more information, and the probability of the upper information being shared between concepts is much higher. Therefore, the semantic similarity between underlying concepts is generally greater than that between high-level concepts. Based on the analysis above, we propose a comparative hierarchical distance in Definition 13 to distinguish the hierarchical differences of formal concepts.

Definition 13. Let $(L(O, P, I, R), \leq)$ be a complete concept lattice, where a and c are defined as two formal concepts belonging to L . When $a \leq c$, $chd(a, c)$ is denoted as the comparative hierarchical distance from a to c in the concept lattice, which is defined as follows:

$$chd(a, c) = dis(a, c) / (dis(a, c) + dis(c, sup(L_{sup}))) \tag{9}$$

where $sup(L_{sup})$ is the supremum of the subset L_{sup} . In particular, we define $chd(a, a) = 0$.

For instance, in Fig. 1, one can calculate $chd(C11, C1) = 2 / (2+1) = 2/3$, while $chd(C2, C1) = 1 / (1+1) = 1/2$, as $C11 < C1$ and $C2 < C1$ in the concept lattice. The comparative hierarchical distance in Definition 13 depends on the shortest distance parameter in Definition 12, which is only defined from the subconcept to its superconcept or itself. Therefore, $chd(C1, C11)$ is not defined. The purpose of the comparative hierarchical distance is to distinguish the differences between concepts when they have equal shortest distances. For example, $dis(C6, C1) = 1$, and $dis(C11, C6) = 1$. However, it is obvious that the information shared between $C6$ and $C2$ and $C11$ and $C6$ is basically different. Thus, by taking into consideration the hierarchical distance, $chd(C6, C1) = 1/2$ and $chd(C11, C6) = 1/3$, which indicates that the concept with a higher hierarchy often leads to a larger impact on the degree of difference.

3.2 Similarity Distance Model

Tversky measured the degree of similarity between concepts by using the shared feature sets of entities [26]. The computational model is as follows.

$$Sim(x, y) = \frac{f(X \cap Y)}{f(X \cap Y) + \alpha \cdot f(X - Y) + \beta \cdot f(Y - X)} \tag{10}$$

where $Sim(x, y)$ denotes the similarity between the concepts m and n ; M and N are the feature sets of x and y , respectively; f is a metric function of the feature sets; $(X - Y)$ represents a set of attributes that belongs to concept M but not to concept N ; and similarly, $(Y - X)$ represents a set of attributes that belongs to concept Y but not to concept N .

In this section, we introduce the semantic ratio similarity between formal concepts in the entropy-weighted concept lattice, which enables to address the measurement of an asymmetric property of concepts to some extent.

Definition 14. Let $(L(O, P, I, R), \leq)$ be a complete concept lattice, where a and c are defined as two formal concepts belonging to L . Given that $a = (A_1, B_1)$ and $b = (A_2, B_2)$ are two formal concepts, $sPath(a, b)$ is denoted as the weighted entropy distance from a to b in the concept lattice, which represents the weighted relative shortest path from the concepts from a to b . In particular, we define $sPath(a, a) = 0$.

$$sPath(a, b) = (w_{c(a)} - w_{c(b)}) * chd(a, b) \quad (11)$$

where $w_{c(a)}$ ($w_{c(b)}$) is the combined weight of the importance and entropy of the attributes that belong to its formal concept a (b); $chd(a, b)$ is the comparative hierarchical distance from a to b .

For example, in Fig. 1, based on Eq. (11), one can calculate $sPath(C11, C6) = (w_{c(C11)} - w_{c(C6)}) * 1/2 = 0.048$ and $sPath(C6, C1) = (w_{c(C6)} - w_{c(C1)}) * 1/2 = 0.0202$, which indicates that the concept with a higher weighted entropy distance often leads to a larger impact on the path difference.

Accordingly, we propose asymmetric semantic similarity to evaluate the combined weight of each attribute.

Definition 15. Let $K'_m = (O, P, I, R, W_c)$ be a weighted concept lattice, where a and b are defined as two formal concepts and r represents root concepts. Given $c, r \in L$ and $c = \sup(a, b)$, $sim(a, b)$ is denoted as the asymmetric semantic similarity between a and b , which is defined as follows:

$$sim(a, b) = \frac{sPath(a, b) + 2 * sPath(c, r)}{sPath(a, b) + sPath(a, c) + sPath(b, c) + 2 * sPath(c, r)} \quad (12)$$

For instance, let the lattice shown in Fig. 1 be a weighted concept lattice, with the weighted concepts $C11 = (\{m_1, m_2\}; \{d, l, n\}; 0.16)$ and $C6 = (\{m_1, m_3\}; \{l, n\}; 0.064)$. It is assumed that the combined weight of the importance and entropy of the attributes is as follows: $sPath(C11, C6) = 0.048$, $sPath(C6, C0) = 0.427$ and $sPath(C6, C6) = 0$. Therefore, $sim(C11, C6) = 0.735$. In addition, the range of $sim(a, b)$ is from zero to one ($sim(a, b) \in (0, 1]$).

In Eq. (12), the entropy-weighted attributes and the comparative hierarchical distance between formal concepts are considered. The semantic similarity between two concepts has a symmetric property, which only depends on the semantic relationship of the same attributes that they share. Therefore, $sim(a, a) = 1$, and, in particular, $sim(a, b) = 0.5$ indicates that the similarity and dissimilarity distances between concepts a and b are equal.

4. Results and Discussion

4.1 Results

In Section 4, we first construct the attributes table (Table 1) on the basis of the essential properties extracted from the specifications for disease classification (GB/T14396). Afterwards, the decision property is labeled to obtain a decision table under the guidance of the definitions of these concepts from GB/T27733-2011. Eventually, a formal context (Table 3) in binary is converted from the decision table.

Using Eq. (4), the inclusion-degree importance of each attribute in the decision table is calculated, which is shown in Table 5. Then, we compute the entropy-degree importance of the attributes via Eq. (5)

and Eq. (6) to substitute the value of each $SIG^R_\gamma(property, C, D)$ in Table 4. The results of the combined weights of each attribute are calculated by using Eq. (7) in Table 6.

Table 5. The inclusion-degree importance of each property in the decision table

Inclusion degree	Time	Amplitude	Lead	Waveform	Segment	Reference	Complex
$SIG^R_\gamma(property, C, D)$	0.0255	0.0221	0.0221	0.0287	0.0249	0.0975	0.0674

Table 6. Computed weight value of each attribute of the formal context

Attribute	P(x)	E(x)	SIG(x,C,D)	w _c (x)
a	0.833	0.219	0.0128	0.0236
b	0.167	0.431	0.0128	0.0464
c	0.500	0.500	0.0111	0.0467
d	0.500	0.500	0.0111	0.0467
e	0.333	0.528	0.0111	0.0493
f	0.667	0.390	0.0111	0.0364
g	0.500	0.500	0.0096	0.0404
h	0.333	0.528	0.0096	0.0426
i	0.167	0.431	0.0096	0.0348
j	0.667	0.390	0.0125	0.0410
k	0.333	0.528	0.0125	0.0555
l	0.667	0.39	0.0325	0.1066
m	0.167	0.431	0.0325	0.1178
n	0.167	0.431	0.0325	0.1178
o	0.333	0.528	0.0225	0.0999
p	0.500	0.500	0.0225	0.0946

After computing all the combined weights of the formal concepts based on Eq. (8), the weighted formal concepts are demonstrated as follows.

- C0 = ($\{m_1, m_2, m_3, m_4, m_5, m_6\}; \emptyset; 0$);
- C1 = ($\{m_1, m_2, m_3, m_4, m_5\}; \{a\}; 0.0236$);
- C2 = ($\{m_1, m_3, m_4, m_5\}; \{a, l, j\}; 0.0364$);
- C3 = ($\{m_3, m_4, m_5, m_6\}; \{f\}; 0.1712$);
- C4 = ($\{m_1, m_2, m_4\}; \{a, c\}; 0.0703$);
- C5 = ($\{m_2, m_4, m_5\}; \{a, p\}; 0.1182$);
- C6 = ($\{m_1, m_2, m_3\}; \{a, g\}; 0.0640$);
- C7 = ($\{m_3, m_5, m_6\}; \{f, d\}; 0.0831$);
- C8 = ($\{m_3, m_4, m_5\}; \{a, f, j, l\}; 0.2076$);
- C9 = ($\{m_2, m_6\}; \{k\}; 0.0555$);
- C10 = ($\{m_2, m_4\}; \{a, c, p\}; 0.1649$);
- C11 = ($\{m_1, m_2\}; \{a, c, e, g\}; 0.1600$);
- C12 = ($\{m_1, m_4\}; \{a, l, j, c\}; 0.2179$);
- C13 = ($\{m_4, m_5\}; \{a, f, h, j, l, p\}; 0.3448$);
- C14 = ($\{m_1, m_3\}; \{a, g, j, l, o\}; 0.3115$);
- C15 = ($\{m_3, m_5\}; \{a, d, f, j, l\}; 0.2543$);
- C16 = ($\{m_2\}; \{a, c, e, g, k, n, p\}; 0.4279$);

- C17 = ($\{m_6\}$; $\{b, d, f, i, k, m\}$; 0.3376);
- C18 = ($\{m_4\}$; $\{a, c, f, h, j, l, p\}$; 0.3915);
- C19 = ($\{m_1\}$; $\{a, c, e, g, j, l, o\}$; 0.4075);
- C20 = ($\{m_5\}$; $\{a, d, f, h, j, l, p\}$; 0.3915);
- C21 = ($\{m_3\}$; $\{a, d, f, g, j, l, o\}$; 0.3946);
- C22 = (\emptyset ; $\{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p\}$; 1).

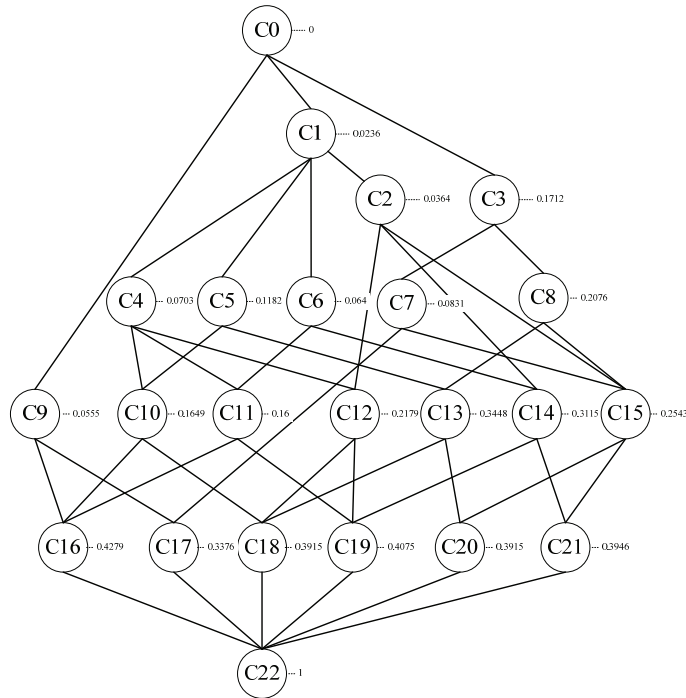


Fig. 2. A weighted concept lattice.

We build a weighted concept lattice with the whole weighted formal concepts (Fig. 2) by using a partial order relation. It is easy to observe that the formal concepts C19, C16, C21, C18, C20, and C17 contain only one original ECG-related concept in their extents, representing the concepts of m_{1-6} in Table 4. Therefore, the semantic similarities of the concepts represent the weighted formal concepts in the weighted concept lattice.

To evaluate the effectiveness of the above method, we first calculate the degree of similarity between concepts via the shared feature sets according to Eq. (11) as a comparison, and the parameters α and β are equal to 0.5. For instance, in Table 3, $f(m_1) = 7$, and $Sim(m_1, m_2) = 4/(4+0.5*3+0.5*3) = 0.57$. With this in mind, we compute the semantic similarities between each pair of formal concepts in Table 4 (m_{1-6}) and those among various concepts of different levels in the weighted concept lattice (C1 and other concepts), which are listed in Tables 7 and 8, respectively.

After obtaining the values of the formal concepts, we then compute the semantic similarities of the same weighted formal concepts based on Definitions 14 and 15. The semantic similarities of concepts m_{1-6} and the semantic similarities of the concepts (C1 to other concepts) are demonstrated in Tables 9 and 10, respectively.

Table 7. Semantic similarity of concepts in Table 3 (feature-based approach)

Sim(row,col)	m ₁ (C19)	m ₂ (C16)	m ₃ (C21)	m ₄ (C18)	m ₅ (C20)	m ₆ (C17)
m ₁ (C19)	1	-	-	-	-	-
m ₂ (C16)	0.57	1	-	-	-	-
m ₃ (C21)	0.83	0.29	1	-	-	-
m ₄ (C18)	0.57	0.43	0.57	1	-	-
m ₅ (C20)	0.43	0.29	0.71	0.86	1	-
m ₆ (C17)	0.14	0.14	0.43	0.29	0.43	1

The value in each cell represents the semantic similarity of the concept in its row to that of its column; “-” represents null.

Table 8. Semantic similarity of the concept C1 and other concepts (feature-based approach)

Sim(C1,col)	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	0.5	0	0.67	0.67	0.67	0	0.4	0	0.5	0.4
Sim(C1,col)	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21
	0.4	0.29	0.33	0.33	0.25	0	0.25	0.25	0.25	0.25

The value in each cell represents the semantic similarity of the concept C1 to that of its column.

Table 9. Semantic similarity of concepts in Table 3 (ECLisd approach)

Sim(row,col)	m ₁ (C19)	m ₂ (C16)	m ₃ (C21)	m ₄ (C18)	m ₅ (C20)	m ₆ (C17)
m ₁ (C19)	1	-	-	-	-	-
m ₂ (C16)	0.713	1	-	-	-	-
m ₃ (C21)	0.908	0.296	1	-	-	-
m ₄ (C18)	0.692	0.658	0.719	1	-	-
m ₅ (C20)	0.663	0.415	0.691	0.926	1	-
m ₆ (C17)	0.345	0.366	0.672	0.337	0.812	1

The value in each cell represents the semantic similarity of the concept in its row to that of its column; “-” represents null.

Table 10. Semantic similarity of the concept C1 and other concepts (ECLisd approach)

Sim(C1,col)	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	0.824	0.503	0.668	0.555	0.684	0.399	0.461	0.350	0.556	0.443
Sim(C1,col)	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21
	0.542	0.526	0.530	0.532	0.518	0.499	0.479	0.520	0.521	0.520

The value in each cell represents the semantic similarity of the concept C1 to that of its column.

4.2 Comparative Analysis

In Table 7 and Table 9, the results of the semantic similarities between the original ECG-related concepts are listed based on the feature-based approach and ECLisd approach, respectively. When comparing the results, we can obtain the line chart in Fig. 3 and conclude that the distribution between the results is relatively similar. In addition, the correlation coefficient between the results is 0.941, which means that the two results are highly correlated. Meanwhile, we analyze some differences between these

two results to prove the validity of the ECLisd approach.

In Fig. 3, we can see that the proposed approach mainly has two advantages. First, the ECLisd approach can reveal small differences in the semantic similarity. According to the categories from GB/T14396, the results illustrate that the concepts that belong to the same supercategory have relatively larger semantic similarities to one another. For example, the objects *AMI* (m_1) and *IE* (m_3) in the lattice both have the same supremum ($\text{sup}(C19, C21) = C14$), which are also in the same supercategory in the specification. There are comparatively large semantic similarities ($\text{Sim}(C19, C21) = 0.908$). Similarly, the pair of *AF* (m_4) and *RHD* (m_5) also has a large semantic similarity ($\text{Sim}(C18, C20) = 0.926$), where both have the same supremum ($\text{sup}(C18, C20) = C13$). However, these semantic similarities cannot be revealed by the feature-based approach. For instance, $\text{Sim}(C19, C18) = \text{Sim}(C21, C18) = 0.57$, while in the ECLisd approach, $\text{Sim}(C19, C18) = 0.692$ and $\text{Sim}(C21, C18) = 0.719$. Furthermore, the ECLisd approach can also reveal the importance of certain attributes. For example, we can see that the similarity distance between the objects *AMI* (m_1) – *CHD* (m_6) and *CAD* (m_2) – *CHD* (m_6) is different. The reason for this difference is that the shared attributes of m_1 and m_6 are *Segment* (j), while m_2 and m_6 do not have the same attributes. In the ECLisd approach, we can obtain the same conclusion ($\text{Sim}(m_1, m_6) = 0.345$ and $\text{Sim}(m_2, m_6) = 0.366$). However, the result from the feature-based approach cannot be obtained ($\text{Sim}(m_1, m_6) = \text{Sim}(m_2, m_6) = 0.14$).

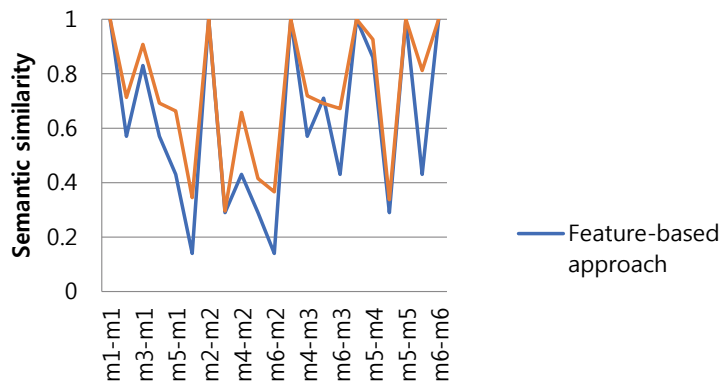


Fig. 3. Line chart of semantic similarities from Tables 6 and 8.

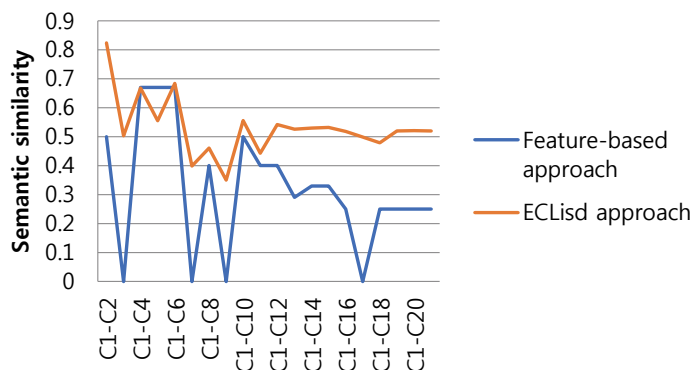


Fig. 4. Line chart of semantic similarities from Tables 7 and 9.

Moreover, to compare these two methods in another aspect, we represent the semantic relationship between the concept nodes in the vertical dimension of the weighted concept lattice. We choose the concept node C1 and its associated nodes to compute the semantic similarities to reveal more implied information in Fig. 4. We can also see that the ECLisd approach is better than the feature-based approach in two other ways. On one hand, one can reveal implicit domain knowledge of the classification. In Fig. 4, according to the feature-based approach, between C1 and other concepts that belong to the same category at different levels, the similarities between the concepts gradually decrease as the distances between the vertices (concept nodes) increase. For instance, $Sim(C1, C6) = 0.67$, $Sim(C1, C11) = 0.4$ and $Sim(C1, C16) = 0.25$. However, this result is not always correct, especially when the attributes between two concepts have various weights. For example, it is reasonable to see that the semantic similarity between C1 and C16 ($Sim(C1, C16) = 0.518$) is greater than that between concepts C1 and C11 ($Sim(C1, C11) = 0.443$), considering the weighted attributes that have relatively high combined weights (Table 4) in the lattice. On the other hand, one might also be able to consider semantic weights and feature distances comprehensively to correct certain misunderstandings. For instance, in the feature-based approach, $Sim(C1, C3) = Sim(C1, C7) = Sim(C1, C9) = Sim(C1, C17) = 0$, which is obviously incorrect, because there is always an intrinsic association among universes in the same domain. In regard to the proposed method, semantic similarities among concepts are totally different. In this scenario, the ECLisd approach is a valid complement to the existing methods.

5. Conclusions

This research attempts to construct an entropy-weighted concept lattice of ECG ontologies. In addition, the combined weights of attribute features are measured by introducing a decision table and entropy theory, according to the semantic similarity model and the specifications for disease classification (GB/T14396). Specifically, the proposed method (ECLisd) shows the semantic characteristics more intuitively and accurately by distinguishing the semantic weights of conceptual features and assigning various weights to the concepts.

Regarding the proposed method, there are three main innovative points. The first, the method merges the importance of inclusion-degree and entropy-degree of the concepts in semantic similarity measurement. Second, the method converts the impact of an ECG-based relevant concept on the semantic assignment into the combined weights of the attributes. Third, the method includes the hierarchical depth of weighted concepts as an important parameter of the similarity calculation.

Although the ECLisd can be applied to extract properties of concepts from the domain ontology, it is not easy to obtain essential properties from the relevant text or specification. Therefore, our future study will aim to develop an approach to extract properties from text analysis that can represent not only the concept classes with the entity objects but also their classification relationships. In addition, as ECG information is widely applied in different fields, semantic similarity should be adapted to measure domain concepts in various application areas. Furthermore, the algorithm that is used to construct the weighted concept lattice needs much more time and space resources if the context contains large concepts. Thus, our future work will also focus on the adaptation and extension of calculations of the semantic relatedness.

Acknowledgement

This work is supported by the Key Natural Science Projects (No. KJ2017A223, SK2018A1072, and KJ2019A0371).

References

- [1] S. Fan, X. An, and W. Lu, "The study on method for topic semantic similarity based on medical literature," *Library and Information Service*, vol. 61, no. 8, pp. 96-105, 2017.
- [2] Z. Wei, G. Song, H. Sun, G. Qiao, Q. Qi, M. He, and A. Song, "Kinematic modeling and trotting gait planning for the quadruped robot with an active spine," *Journal of Southeast University(Natural Science Edition)*, vol. 49, no. 6, pp. 1019-1025, 2019
- [3] H. Deng, X. Zhu, Q. Li, and Q. Peng, "Sentence similarity calculation based on syntactic structure and modifier," *Computer Engineering*, vol. 43, no. 9, pp. 240-249, 2017.
- [4] W. Shi, J. Zhou, and Y. Lan, "A novel fuzzy clustering algorithm based on similarity of attribute space," *Journal of Electronics & Information Technology*, vol. 41, no. 11, pp. 2722-2728.
- [5] Z. Wu, T. Yu, H. Chen, X. Jiang, Y. Feng, Y. Mao, H. Wang, J. Tang, and C. Zhou, "Information retrieval and knowledge discovery on the semantic web of traditional Chinese medicine," in *Proceedings of the 17th International Conference on World Wide Web*, Beijing, China, 2008, pp. 1085-1086.
- [6] M. She, Z. Zhang, and J. Hu, "Domain ontolog-based personalized recommendation research in the health care system," *Electronic Design Engineering*, vol. 18, no. 11, pp. 23-26, 2010.
- [7] C. Pechsiri, S. Phainoun, and R. Piriyakul, "Extraction of ObjectProperty-UsageMethod relation from web documents," *Journal of Information Processing Systems*, vol. 13, no. 5, pp. 1103-1125, 2017.
- [8] W. Lou, H. Wang, and Y. Ju, "An ontology fusion method based on binary similarity calculation," *Journal of the China Society for Scientific and Technical Information*, vol. 38, no. 6, pp. 622-631, 2019.
- [9] A. Burusco and R. Fuentes-Gonzalez, "Construction of the L-fuzzy concept lattice," *Fuzzy Sets and Systems*, vol. 97, no. 1, pp. 109-114, 1998.
- [10] M. Wolski, "Formal concept analysis and rough set theory from the perspective of finite topological approximations," in *Transactions on Rough Sets III*. Heidelberg: Springer, 2005, pp. 230-243.
- [11] D. H. Wei, T. Kang, H. A. Pincus, and C. Weng, "Construction of disease similarity networks using concept embedding and ontology," *Studies in Health Technology and Informatics*, vol. 264, pp. 442-446, 2019.
- [12] D. Sanchez, M. Batet, D. Isern, and A. Valls, "Ontology-based semantic similarity: a new feature-based approach," *Expert Systems with Applications*, vol. 39, no. 9, pp. 7718-7728, 2012.
- [13] J. Dai, W. Wang, Q. Xu, and H. Tian, "Uncertainty measurement for interval-valued decision systems based on extended conditional entropy," *Knowledge-Based Systems*, vol. 27, pp. 443-450, 2012.
- [14] J. B. Gao, B. W. Zhang, and X. H. Chen, "A WordNet-based semantic similarity measurement combining edge-counting and information content theory," *Engineering Applications of Artificial Intelligence*, vol. 39, pp. 80-88, 2015.
- [15] S. Ghafourian, A. Rezaeian, and M. Naghibzadeh, "Graph-based partitioning of ontology with semantic similarity," in *Proceedings of International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, Iran, 2013, pp. 80-85.
- [16] M. Donnelly, *Formal Ontology in Information Systems: Proceedings of the 7th International Conference (FOIS2012)*. Amsterdam: IOS Press, 2012.
- [17] Z. W. Ras and A. Dardzinska, "Ontology-based distributed autonomous knowledge systems," *Information Systems*, vol. 29, no. 1, pp. 47-58, 2004.

- [18] W. Wang, X. Kang, and Y. Wu, "Knowledge acquisition model of concept lattice in an incomplete formal context," *CAAI Transactions on Intelligent Systems*, vol. 14, no. 5, pp. 1048-1055, 2019.
- [19] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, 2001.
- [20] J. D. Trigo, F. Chiarugi, A. Alesanco, M. Martinez-Espronedada, L. Serrano, C. E. Chronaki, J. Escayola, I. Martinez, and J. Garcia, "Interoperability in digital electrocardiography: harmonization of ISO/IEEE x73-PHD and SCP-ECG," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 6, pp. 1303-1317, 2010.
- [21] N. Stockbridge and B. D. Brown, "Annotated ECG waveform data at FDA1," *Journal of Electrocardiology*, vol. 37, pp. 63-64, 2004.
- [22] J. Ji, G. X. Wu, and W. LI, "Significance of attribute based on inclusion degree," *Journal of Jiangxi Normal University (Natural Sciences Edition)*, vol. 33, no. 6, pp. 656-660, 2009.
- [23] B. Jiang, J. Ma, and X. Wang, "Inclusion degree and support degree of fuzzy sets," *Journal of Fuzzy Mathematics*, vol. 13, no. 3, pp. 747-755, 2005.
- [24] A. Q. Andrade, W. Blonde, J. Hastings, and S. Schulz, "Process attributes in bio-ontologies," *BMC Bioinformatics*, vol. 13, article no. 217, 2012.
- [25] H. Shi, "Semantic similarity computation of an information retrieval," *Journal of Chizhou College*, vol. 30, no. 3, pp. 26-29, 2016.
- [26] H. H. Huang and Y. H. Kuo, "Cross-lingual document representation and semantic similarity measure: a fuzzy set and rough set based approach," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 6, pp. 1098-1111, 2010.



Kai Wang <https://orcid.org/0000-0002-9098-3017>

He graduated from Anhui Agricultural University with a master's degree in 2011. He is currently a lecturer in Bengbu Medical College, whose current research focuses on information processing and information retrieval.



Shu Yang <https://orcid.org/0000-0002-3461-5594>

He graduated from Hefei University of Technology with a master's degree in 2012. He is currently a professor in Bengbu Medical College, whose current research focuses on ontology learning and information integration.