Engineering Applications of Artificial Intelligence 133 (2024) 108198

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Research paper

Attribute granules-based object entropy for outlier detection in nominal data

Chang Liu [a], Dezhong Peng [a,c], Hongmei Chen [b], Zhong Yuan [a,*]

[a] College of Computer Science, Sichuan University, Chengdu 610065, China
[b] School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China
[c] Sichuan Newstrong UHD Video Technology Co., Ltd., Chengdu 610095, China

ARTICLE INFO

Keywords:
Concept lattice theory
Formal context
Information entropy
Outlier detection
Nominal data

ABSTRACT

Concept lattice theory, which is one of the key mathematical models of granular computing, is capable of successfully dealing with uncertain information in nominal data. It has been applied to machine learning tasks such as data reduction, classification, and association rule mining. For the problem of outlier detection in nominal data, this paper presents a concept lattice theory-based approach for detecting outliers in nominal data. First, subcontexts and concept lattices based on subsets of objects are discussed. Then, information entropy is introduced into the formal context, and an object entropy based on attribute granules is proposed. Finally, a nominal data-oriented outlier detection method is explored based on the proposed object entropy. The experimental results show that the proposed detection method can effectively detect outliers in nominal data. Besides, the results of the hypothesis testing indicate that the proposed method is statistically significantly different from the other methods. The code is publicly available online at https://github.com/from-china-to/OEOD.

1. Introduction

Wille created Formal Concept Analysis (FCA) in 1982 (Wille, 1982), which has since become a useful method for data analysis and processing. To the best of our knowledge, Concept Lattice (CL) is one of the most concerning elements in FCA. Currently, research on CL theory and techniques is mostly focused on the extension, construction, and application of the model.

As the research progressed, scholars realized that FCA is mainly oriented to Boolean data. However, the data in reality is often more complex, which limits the wide application of the method. For this reason, some scholars have successively proposed various extended concept operators, such as L-fuzzy CL (Juandeaburre and Fuentes-González, 1994), rough CL (Yao, 2004), three-way CLs (Qi et al., 2014), and so on. Due to the enormous expansion of concept lattice building time and scale, how to enhance concept lattice construction efficiency has become a primary research objective of FCA theory. At present, the mainstream methods of concept lattice construction mainly include batch (Nourine and Raynaud, 1999), incremental (Merwe et al., 2004), and parallel (Zou et al., 2022) construction methods.

Concept lattice theory is becoming more and more mature, which has been widely used in information retrieval (Huang et al., 2018), knowledge discovery (Hao et al., 2021), recommender systems (Cordero et al., 2020), and data mining (Zou et al., 2018). Among them, concept reduction and rule extraction are important research directions in concept lattice theory. CL reduction is to avoid the redundancy of object, attribute, or concept knowledge on the premise of keeping specific information unchanged. At present, there are mainly three types of existing concept lattice reduction: object reduction (Li et al., 2014; Trnecka and Trneckova, 2018), attribute reduction (Wu et al., 2008; Mi et al., 2010; Shao and Li, 2017) and concept reduction (Stumme et al., 2002; Wei et al., 2020). The research on rule extraction in formal concept analysis mainly focuses on how to discover implication rules (Hu et al., 2000; Li et al., 2013). For example, Wang et al. used the strategy of incrementally constructing CLs to realize dynamic extraction of rules (Wang et al., 1999). Xie and Liu investigated the connection between CL and association discovery and gave an association rule extraction algorithm for the corresponding lattice structure (Xie and Liu, 2000). Li et al. proposed a new algorithm for mining non-redundant decision rules from decision formalization environments (Li et al., 2013).

Outlier detection (OD) is a critical study area in data mining. In most knowledge discovery research, outliers are usually discarded as noise. However, finding outliers may be more valuable in many applied studies such as industrial production (Wang and Mao, 2019), secured routing (Thangaramya et al., 2020), medical diagnosis (Alaverdyan et al., 2020), and load forecasting (Yue et al., 2019). Currently, OD

* Corresponding author.
E-mail addresses: liuchangai@stu.scu.edu.cn (C. Liu), pengdz@scu.edu.cn (D. Peng), hmchen@swjtu.edu.cn (H. Chen), yuanzhong@scu.edu.cn (Z. Yuan).

https://doi.org/10.1016/j.engappai.2024.108198
Received 26 June 2023; Received in revised form 19 November 2023; Accepted 3 March 2024

methods are loosely classified into four groups depending on the assumptions they make regarding anomalies and other objects, i.e., (1) statistics-based (Edgeworth, 1887), (2) distance-based (Knox and Ng, 1998; Knorr et al., 2000), (3) density-based (Breunig et al., 2000), and (4) cluster-based (He et al., 2003).

Most of the above detection models are constructed according to the Euclidean metric. However, it may not be reasonable to use Euclidean metrics to characterize the similarities and differences of nominal attributes. As a result, there may be a better strategy for detecting outliers in nominal data. To this end, scholars began to study rough computing methods for OD (Jiang et al., 2009; Chen et al., 2010; Jiang and Chen, 2015; Jiang et al., 2016; Yuan et al., 2021, 2022). CL theory, which is one of the key mathematical models of Granular Computing (GrC), is capable of successfully dealing with uncertain information in nominal data. The CL has become an effective tool for data analysis and knowledge discovery due to its simplicity, conciseness, and completeness of knowledge expression. However, the use of CL for OD is less common (Zhang et al., 2009; Hu et al., 2023) and requires further research.

As seen in the literature, most traditional detection models use Euclidean metrics such as DIS, KNN, and LOF. The use of Euclidean metrics is not appropriate to characterize similarities and differences in nominal data. In addition, some existing OD methods involve parameters, such as GrC, OMkNN, and FGAS (Liu et al., 2023) algorithms, which are prone to problems such as parameter dependence and difficulty in parameter setting. Finally, real-world rules often contain uncertainty, and conceptual lattice theory can deal with uncertain information in nominal data.

Based on the above discussion, this article proposed a method of detecting anomalies for nominal data utilizing the concept of lattice theory. First, subcontexts and concept lattices based on object subsets are discussed. Then, the Information Entropy (IE) is introduced into the formal context, and the object entropy based on attribute granules is proposed. Based on the proposed object entropy, its relative uncertainty measures are studied and their relationships are discussed. Furthermore, a nominal data-oriented OD method is explored based on object entropy and its relative uncertainty measures. In this method, object mutual information is used to define correlation, conditional correlation, and aggregation degree of objects, and then the object anomaly evaluation index of minimum correlation-maximum aggregation degree is constructed. Finally, an Object-oriented Entropy-based OD (OEOD) algorithm is given. Experimental evaluation demonstrates OEOD's applicability and effectiveness for OD in nominal data.

This article is organized as follows. Section 2 reviews relevant knowledge. The third section discusses object-oriented formal subcontexts and their associated uncertainty measures. In the fourth section, we propose an OD based on object entropy. The fifth section shows the experimental comparison results. Finally, the sixth section gives the conclusion.

2. Related works

With the rapid development of GrC, more and more scholars have studied the application of GrC in OD. In this section, we will discuss the traditional OD methods and discuss the research related to Rough Set (RS) and CL as tools in the theory of GrC in the OD methods.

Traditional OD methods can generally be divided into four categories: statistics-based, distance-based, density-based, and cluster-based. The statistics-based methods (Edgeworth, 1887), are usually based on the assumption that outliers are extreme values or deviate significantly from the expected statistical properties of the data. These methods are generally sensitive to data distribution, although they are relatively easy to understand and implement. Knorr et al. (2000) investigated the concept of distance-based outliers for datasets with more than two attributes. Ramaswamy et al. (2000) proposed a distance-based representation of outliers based on the distance between a point and its kth nearest neighbor. However, these methods typically require the calculation of pairwise distances for all objects and are therefore computationally inefficient. Breunig et al. (2000) proposed a local OD algorithm based on an object's local outlier factor. The extent to which an outlier is assigned to each object depends on how isolated the object is relative to its surroundings. These methods can effectively handle datasets with varying data densities, but these typically require extensive hyperparameter tuning, which greatly affects performance and efficiency. For local data, He et al. (2003) proposed a cluster-based OD method. These methods can often account for and be robust to data distributions. However, they can be sensitive to the initial clustering conditions. The traditional four algorithms usually construct detection models based on Euclidean distances, and thus their associated detection models are only applicable to numerical data.

RS theory is an effective mathematical tool for dealing with imprecise, uncertain, and incomplete nominal data. It has been widely used in the fields of attribute reduction, feature selection (Wang et al., 2021; Jiang et al., 2015), and OD. It has been shown to be effective in detecting outliers, such as GrC-based (Chen et al., 2008), IE-based (Jiang et al., 2010), Rough Sequence-based OD (RSOD) (Jiang et al., 2011), and approximate accuracy-based (Jiang et al., 2019). For example, Chen et al. (2008) introduced an OD technique based on GrC which uses information granules to measure the extent of outliers. Jiang et al. (2005) first used RSs to solve the OD problem, the boundary-based OD method by using the RS framework. Jiang et al. (2009) applied RS theory to OD by proposing sequence-based and distance-based OD. Jiang et al. (2010) employed IE to measure the level of data uncertainty in their outlier detection model. Using the concept of RS, Yang and Zhu (2011) proposed an outlier approximation method based on an OD and analysis system, which can mine the outliers in the key genus subsets by defining the similarity of peripheral partitions. Albanese et al. (2012) extended OD to spatiotemporal data by using the RS approximations to define outliers. Jiang and Chen (2015) put forward an extension OD algorithm by rough approximation accuracy. Uncertain information in nominal data can be successfully handled by CL theory. Zhang et al. (2009) proposed a concept lattice-based outlier mining algorithm for low-dimensional subspaces, which treats the intent of each concept lattice node as a subspace. Hu et al. (2023) proposed a granular concept-based OD model. These methods address to some extent the problem of OD in nominal data. However, these methods do not take into account the advantage that information theory can effectively deal with uncertain information. Therefore, this paper introduces information theory into the CL theory to propose the attribute granules-based object entropy and applies it to OD.

3. Preliminaries

Some basic concepts related to the FCA are reviewed first (Wille, 1982; Ganter and Wille, 1999; Wu et al., 2008) to enhance the readability of this article.

3.1. Formal context and its CL

A CL is usually generated by a formal context. A formal context can be represented by a data table, which can be defined formally as follows.

Definition 1. A formal context is a triple $F = (OB, AT, I)$, where $OB = \{o_1, o_2, \dots, o_n\}$ is a nonempty finite set of objects; $AT = \{a_1, a_2, \dots, a_m\}$ is a nonempty finite set of attributes; and I is a binary relation of $OB \times AT \rightarrow \{0, 1\}$, where for any $o \in OB$ and $a \in AT$, $I(o, a) = 1$ and $I(o, a) = 0$ imply that the object o has and does not have the attribute a, respectively.

In this article, it is assumed that the formal context being discussed is regular (Wu et al., 2008). To effectively analyze the intrinsic information structure contained in the formal context, Wille introduced the following pair of concept forming operators.

Definition 2. Given a formal context $F = (OB, AT, I)$, for any $O \subseteq OB$ and $A \subseteq AT$, their corresponding concept formatting operators are defined respectively as

$$O^* = \{a \in AT | \forall o \in O, I(o,a) = 1\}; \tag{1}$$

$$A^* = \{o \in OB | \forall a \in A, I(o,a) = 1\}, \tag{2}$$

where O^* is the maximal set of attributes common to all objects in O, and A^* denotes the maximal set of objects with all attributes in A.

The following formal concepts are defined based on the above two operators.

Definition 3. Let $O \subseteq OB$ and $A \subseteq AT$, a binary pair (O, A) is called a formal concept if $O^* = A$ and $A^* = O$, where O and A are called the extension and the intension of (O, A), respectively.

Proposition 1. *If $O, O_1, O_2 \subseteq OB$ and $A, A_1, A_2 \subseteq AT$, then there are some conclusions that hold true as follows.*

(1) $O_1 \subseteq O_2 \Rightarrow O_1^* \supseteq O_2^*, A_1 \subseteq A_2 \Rightarrow A_1^* \supseteq A_2^*$;
(2) $O \subseteq O^{**}, A \subseteq A^{**}$;
(3) $O^{***} = O^*, A^{***} = A^*$;
(4) $(O_1 \cup O_2)^* = O_1^* \cap O_2^*, (A_1 \cup A_2)^* = A_1^* \cap A_2^*$;
(5) $O \subseteq A^* \Leftrightarrow A \subseteq O^*$.

The formal concepts of F make up a complete lattice known as the CL. This lattice is denoted as $L(OB, AT, I)$. Further, the sets consisting of all extensions and all intensions are denoted as $L_{OB}(OB, AT, I) = \{O|(O,A) \in L(OB, AT, I)\}$ and $L_{AT}(OB, AT, I) = \{A|(O,A) \in L(OB, AT, I)\}$, respectively.

For any $(O_1, A_1), (O_2, A_2) \in L(OB, AT, I)$, their corresponding partial order relation is defined as $O_1 \subseteq O_2 (A_1 \supseteq A_2) \Leftrightarrow (O_1, A_1) \leq (O_2, A_2)$. Besides, the infimum and the supremum between two concepts are respectively given by $(O_1, A_1) \wedge (O_2, A_2) = (O_1 \cap O_2, (A_1 \cup A_2)^{**})$ and $(O_1, A_1) \vee (O_2, A_2) = ((O_1 \cup O_2)^{**}, A_1 \cap A_2)$.

(O^{**}, O^*) is a formal concept for any $O \subseteq OB$. Consequently, O^* is known as the object intension of O in (OB, AT, I). The smallest extension containing O is denoted as O^{**}. For any given $A \subseteq AT$, the corresponding formal concept (A^*, A^{**}) has an extension of some concept A^*, which we refer to as the attribute extension of the A in (OB, AT, I). Additionally, A^{**} is the smallest intension that contains A. It is easy to prove that for any $o \in OB$ and $a \in AT$, the formal concepts (o^{**}, o^*) and (a^*, a^{**}) exist referred to as the object concept and attribute concept, respectively (To simplify notation, we denote $\{o\}^*$ as o^* and $\{a\}^*$ as a^*).

3.2. Formal subcontext and its CL

Definition 4. Let $C \subseteq AT$, $F_C = (OB, C, I_C)$ is called a formal subcontext of $F = (OB, AT, I)$, where $I_C = I \cap (OB \times C)$. Similarly, for any $O \subseteq OB$ and $A \subseteq C$, two concept formatting operators on the formal subcontext $F_C = (OB, C, I_C)$ are defined as follows, respectively.

$$O^{*C} = \{a \in C | \forall o \in O, I_C(o,a) = 1\}; \tag{3}$$

$$A^{*C} = \{o \in OB | \forall a \in A, I_C(o,a) = 1\}. \tag{4}$$

It is easy to obtain $O^{*C} = O^* \cap C$ and $O^{*AT} = O^*$. Similarly, by Definition 3, we can obtain all concepts in the formal subcontext. However, the above formal subcontexts are discussed mainly based on attribute subsets. There are few discussions of formal subcontexts based on object subsets.

4. Object-oriented formal subcontext and its uncertainty measures

In this section, we first discuss the relationship between the object-oriented formal subcontext and the formal context and then discuss the object-oriented IE and its corresponding uncertainty measures.

4.1. Object-oriented formal subcontext

Definition 5. Let $V \subseteq OB$, $F_V = (V, AT, I_V)$ is called an object-oriented formal subcontext of $F = (OB, AT, I)$, where $I_V = I \cap (V \times AT)$. Similarly, for any $O \subseteq V$ and $A \subseteq AT$, two concept formatting operators on the formal subcontext $F_V = (V, AT, I_V)$ are defined as follows, respectively.

$$O^{*V} = \{a \in AT | \forall o \in O, I_V(o,a) = 1\}; \tag{5}$$

$$A^{*V} = \{o \in V | \forall a \in A, I_V(o,a) = 1\}. \tag{6}$$

It can easily be observed that $A^{*V} = A^* \cap V$ and $A^{*OB} = A^*$. Next, the relationships between the concept formatting operators of the formal context and the ones of object-oriented formal subcontext are discussed.

Proposition 2. *Let $O \subseteq V \subseteq OB$. For any $A \subseteq AT$ and $a \in A$, there are the following conclusions.*

(1) $O^{*V} = O^{*OB}$;
(2) $A^{*O} \subseteq A^{*V}$;
(3) $a^{*O} \subseteq a^{*V}$;
(4) $A^{*V*V} \subseteq A^{*O*O}$;
(5) $a^{*V*V} \subseteq a^{*O*O}$.

Proof. By Eqs. (1) and (5), it is obvious that $O^{*V} = O^{*OB}$, so Item (1) holds. Since $A^{*O} = A^{*OB} \cap O$ and $A^{*V} = A^{*OB} \cap V$, so from $O \subseteq V$, there is $A^{*O} \subseteq A^{*V}$. Therefore, Item (2) holds. We can readily verify that Item (3) holds for Item (2). Since $O \subseteq V \subseteq OB$, from Item (1), there is $A^{*O} \subseteq A^{*V}$. Then from Item (1) in Proposition 1, there is $A^{*V*OB} \subseteq A^{*O*OB}$. Further, by Item (1), there is $A^{*V*OB} = A^{*V*V}$ and $A^{*O*OB} = A^{*O*O}$. Therefore, Item (4) holds. We can readily verify that Item (5) holds for Item (4). □

From Proposition 2, we can see that the attribute extension of an attribute set in a formal context contains that of the same attribute set in its formal subcontext. The smallest intension of an attribute set in a context is contained in that of the same attribute set in any of its subcontexts.

4.2. Attribute granules-based uncertainty measures

Definition 6. Let $O \subseteq OB$, the attribute granules-based object entropy with respect to F_O is defined as

$$OE(O) = OE(F_O) = -\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|a^{*O*O}|}{|AT|}. \tag{7}$$

It is easy to get $0 < OE(O) \leq \log_2 |AT|$. Since the formal context is regular, so for any $a \in AT$, there is $\{a\} \subseteq a^{*O*O} \subset AT$. Thus, there is $-\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|AT|}{|AT|} < -\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|a^{*O*O}|}{|AT|} \leq -\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|\{a\}|}{|AT|}$, i.e., $0 < -\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|a^{*O*O}|}{|AT|} \leq \log_2 |AT|$. Therefore, there is $0 < OE(O) \leq \log_2 |AT|$.

Proposition 3. *If $O_1 \subseteq O_2 \subseteq OB$, then $OE(O_1) \leq OE(O_2)$.*

Proof. Let $O_1 \subseteq O_2 \subseteq OB$, by Proposition 2, for any $a \in AT$, there is $a^{*O_1*O_1} \supseteq a^{*O_2*O_2}$. So there is $|a^{*O_1*O_1}| \geq |a^{*O_2*O_2}|$, i.e., $-\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|a^{*O_1*O_1}|}{|AT|} \leq -\frac{1}{|AT|} \sum_{a \in AT} \log_2 \frac{|a^{*O_2*O_2}|}{|AT|}$. Therefore, there is $OE(O_1) \leq OE(O_2)$. □

Proposition 3 reflects the monotonic variation of object entropy with the increasing number of objects, which indicates its capacity to measure uncertainty in formal contexts.

The above object entropy in the definition is generated for a formal subcontext. Therefore, the object joint entropy is defined for multiple formal subcontexts.

Definition 7. Let $O, P \subseteq OB$, the attribute granules-based object joint entropy between O and P are defined as

$$OE(O,P) = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|AT|}. \tag{8}$$

Corollary 1. *Let $O_1, O_2, \ldots, O_k \subseteq OB$, the object joint entropy among $O_1, O_2, \ldots, O_k$ is*

$$OE(O_1,O_2,\ldots,O_k) = OE(F_{\cup_{s=1}^{k} O_s}) = -\frac{1}{|OB|}\sum_{a\in AT}\frac{|\cap_{s=1}^{k} a^{*O_s*O_s}|}{|AT|}. \tag{9}$$

After O is known, the uncertainty of P can be defined by conditional entropy.

Definition 8. The attribute granules-based object conditional entropy of O on P is defined as

$$OE(P|O) = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|a^{*O*O}|}. \tag{10}$$

Proposition 4. $OE(P|O) = OE(P,O) - OE(O)$.

Proof. According to the above definition, we have $OE(P,O) - OE(O) = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|AT|} - \left(-\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}|}{|AT|}\right) = -\frac{1}{|AT|}\sum_{a\in AT}\left(\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|AT|} - \log_2\frac{|a^{*O*O}|}{|AT|}\right)$
$= -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|a^{*O*O}|} = OE(P|O)$. □

Definition 9. The attribute granules-based object mutual information between O and P is defined as

$$OMI(P;O) = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}|\cdot|a^{*P*P}|}{|AT|\cdot|a^{*O*O}\cap a^{*P*P}|}. \tag{11}$$

Proposition 5. *Let $O, P \subseteq OB$, we have*

(1) $OMI(P;O) = OMI(O;P)$;
(2) $OMI(P;O) = OE(P) - OE(P|O) = OE(O) - OE(O|P)$;
(3) $OMI(P;O) = OE(P) + OE(O) - OE(P,O)$.

Proof. By Eq. (11), it is easy to $OMI(P;O) = OMI(O;P)$, so List (1) holds.

$OMI(P;O) = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}|\cdot|a^{*P*P}|}{|AT|\cdot|a^{*O*O}\cap a^{*P*P}|} = -\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*P*P}|}{|AT|} - \left(-\frac{1}{|AT|}\sum_{a\in AT}\log_2\frac{|a^{*O*O}\cap a^{*P*P}|}{|a^{*O*O}|}\right) = OE(P) - OE(P|O)$. Similarly, $OMI(P;O) = OE(O) - OE(O|P)$, so (2) holds.

From Proposition 4, there is $OMI(P;O) = OE(P) + OE(O) - OE(P,O)$. Therefore, (3) holds. □

In Proposition 5, $OMI(P;O)$ is equal to $OE(P)$ minus $OE(P|O)$, which reflects the amount of information held jointly in P and O. Thus, $OMI(P;O)$ represents the degree of the correlation between P and O. The detection method proposed in this paper utilizes object mutual information as an indicator to describe some anomalous property.

5. Object-oriented Entropy-based OD

In this section, we will propose an OEOD based on the above uncertainty metric.

5.1. Relevant definitions

The goal of OD is to find objects whose behavior is significantly different from other objects. The set of these objects is called an outlier set. Objects in the outlier set should be far away from other objects in the dataset. That is, outliers should have minimal correlation with other objects. The object-oriented mutual information defined above can characterize the correlation between objects subset. Therefore, those objects that hold the least mutual information with other objects should be identified as outliers by the OD algorithm. To express the correlation between objects, the correlation of an object is first defined based on object-oriented mutual information.

Definition 10. For any $o_i \in OB$, the correlation of o_i is defined as

$$Corr(o_i) = \frac{1}{|OB|}\sum_{j=1}^{|OB|} OMI(o_i;o_j). \tag{12}$$

In the above definition, $Corr(o_i)$ is defined as the average of mutual information with all other objects, which can characterize the abnormal degree of an object. The greater the $Corr(o_i)$ of an object o_i, the less likely it is to become an outlier. Therefore, the OD method in this paper starts with an empty object set OS and adopts a step-by-step strategy to select one object at a time. In the first step, the object x_{i_1} with the smallest correlation is selected, which satisfies the following conditions:

$$o_{i_1} = \min_{o_i\in OB}\{Corr(o_i)\}. \tag{13}$$

Since the object o_{i_1} has the smallest correlation, it is most likely to be an outlier. That is, if there is only one outlier, o_{i_1} is the optimal choice. At this point, let $OS_1 = \{o_{i_1}\}$.

Let OS_u denote the set of objects that are not currently selected, and OS_{r-1} denote the set of $(r-1)$ objects that have been selected, how to select the rth object? In the proposed method, the selection of the rth object adopts the following strategy: Relative to the objects in $OS_u - \{o_{i_r}\}$, o_{i_r} should be the smallest in the entire object set, and it should be maximally similar to the selected objects in OS_{r-1}. To this end, the similarity between objects is further defined.

Definition 11. The conditional correlation of a selected object o_{i_k} with respect to the candidate object o is defined as

$$Corr(o_{i_k}|o) = \frac{OE(o_{i_k}|o)}{OE(o_{i_k})}Corr(o_{i_k}). \tag{14}$$

Further, the difference between correlation and conditional correlation is defined as the degree of aggregation between candidate objects and selected objects.

Definition 12. The degree of aggregation between a candidate object o and the selected object o_{i_k} is defined as

$$Aggr(o,o_{i_k}) = Corr(o) - Corr(o_{i_k}|o). \tag{15}$$

The greater the degree of aggregation of a candidate object with the selected objects, the more likely it is to become an outlier. Therefore, when selecting the rth object, the correlation degree of the candidate object and its aggregation degree with the selected objects can be comprehensively considered, and the object anomaly evaluation index of the minimum Correlation-Maximum Aggregation (mCMA) can be obtained as follows.

$$o_{i_r} = \min_{o\in OS_u}\{w(o)(Corr(o) - \frac{1}{r-1}\sum_{k=1}^{r-1} Aggr(o,o_{i_k}))\}, \tag{16}$$

where $w(o)$ is a weight function such that $w(o) = \sqrt{\frac{|o^{**}|}{|OB|}}$. Then the rth object can be selected as o_{i_r}. When selecting subsequent objects, the similar strategy can be used to select one by one, and finally, an ordered object set $OS = \{o_{i_1}, o_{i_2}, \ldots, o_{i_n}\}$ is obtained.

Generally, in practical applications, it is more flexible and convenient to output outlier scores for objects. To this end, the object's outlier score is further defined as follows.

Definition 13. For any $o_{i_k} \in OS$, the outlier score of o_{i_k} is defined as

$$\mathrm{Score}(o_{i_k}) = \frac{1}{k}. \tag{17}$$

Further, OD may be defined based on the above outlier scores.

Definition 14. Given a threshold $\mu > 0$. For any $o \in OB$, if $Score(o) > \mu$, then o is called an OEOD.

5.2. Corresponding algorithm

In this subsection, based on the content of the previous subsection, we design an OD algorithm based on object-oriented entropy.

Algorithm 1: OEOD algorithm

Input: A formal context $F = (OB, AT, I)$
Output: Score

1 $OS \leftarrow \emptyset, OS_u \leftarrow OB$;
2 **for** $i \leftarrow 1$ *to* $|OB|$ **do**
3   **for** $i \leftarrow 1$ *to* $|OB|$ **do**
4   Compute $OMI(o_i, o_j)$;
5   **end**
6 **end**
7 **for** $i \leftarrow 1$ *to* $|OB|$ **do**
8   Compute $Corr(o_i)$;
9 **end**
10 Select o_{i_1} so that $Corr(\{o_{i_1}\})$ has a minimal value ;
11 $OS \leftarrow \langle o_{i_1} \rangle, OS_u \leftarrow OS_u - \{o_{i_1}\}$;
12 **while** $|OS_u| \neq 0$ **do**
13   **for** $l \leftarrow 1$ *to* $|OS_u|$ **do**
14   **for** $s \leftarrow 1$ *to* $|OS|$ **do**
15   Calculate $Aggr(o_l, o_{i_s})$;
16   **end**
17   **end**
18   Select o_{i_r} so that $w(o)(Corr(o_{i_r}) - \frac{1}{|OS|}\sum_{s=1}^{|OS|} Aggr(o_{i_r}, o_{i_s}))$ has a minimal value ;
19   $OS \leftarrow \mathrm{strcat}(OS, o_{i_r}), OS_u \leftarrow OS_u - \{o_{i_r}\}$ // o_{i_r} is selected to add to the end of OS ;
20 **end**
21 **for** *each* $o_{i_k} \in OS$ **do**
22   $Score(o_{i_k}) = \frac{1}{k}$;
23 **end**
24 **return** $Score$.

Algorithm 1 starts with an empty object set. First, the correlation of each object is calculated, and the object with the smallest correlation is selected to be added to the object set. Then, the minimum correlation-maximum aggregation evaluation index is used to select subsequent objects until the set of unselected objects is empty. Further, an outlier score for each object is calculated based on the obtained ordered set of outliers. Finally, the algorithm outputs the result. The framework of the OEOD model is shown in Fig. 1.

In Algorithm 1, the number of loops in Steps 2–5 is $|OB| \times |OB|$, the number of loops in Step 4 is $|AT|$, the number of loops in Steps 7–9 is $|OB|$, and the number of loops in Steps 12–20 is $|OS_u| \times |OS|$. Thus, the total number of cycles of Algorithm 1 is $|OB| \times |OB| \times |AT| + |OB| + |OS_u| \times |OS|$. Therefore, its time complexity is $O(|AT||OB|^2)$.

Since algorithm OEOD is designed based on classical formal contexts, it can only be applied to data sets described by nominal attributes. For data sets containing numerical values, this paper converts the original data sets into nominal data sets through discretization methods. Furthermore, it is easy to convert the nominal data set into the corresponding formal context. Finally, the algorithm OEOD is used to detect outliers in the formal context.

Table 1
A data table and its formal context.

OB	a	b	c	a_1	a_2	b_1	b_2	c_1	c_2
o_1	b	B	f	0	1	1	0	0	1
o_2	a	C	e	1	0	0	1	1	0
o_3	a	C	e	1	0	0	1	1	0
o_4	b	C	e	0	1	0	1	1	0
o_5	a	C	f	1	0	0	1	0	1

5.3. A corresponding example

In this subsection, we give a corresponding example to demonstrate the algorithm proposed above.

Example 1. Raw data is represented on the left side of Table 1, where $OB = \{o_1, o_2, \ldots, o_5\}$, $AT = \{a, b, c\}$. By transforming the original data, its corresponding formal context is shown on the right side of Table 1.

According to Definition 10, the correlation of each object is calculated as follows.

$Corr(o_1) \approx 0.0423$, $Corr(o_2) \approx 0.1770$, $Corr(o_3) \approx 0.1770$, $Corr(o_4) \approx 0.0806$, and $Corr(o_5) \approx 0.0806$.

Next, the object with the smallest correlation is selected and added to OS, i.e., o_1 is added to OS, and $OS = \langle o_1 \rangle, OS_u = \{o_2, o_3, o_4, o_5\}$.

By Definition 11, the conditional correlation of o_1 relative to each object in OS_u is calculated as follows.

$Corr(o_1|o_2) \approx 0.0505$, $Corr(o_1|o_3) \approx 0.0505$, $Corr(o_1|o_4) \approx 0.0423$, $Corr(o_1|o_5) \approx 0.0423$.

According to Definition 12, the degree of aggregation between a candidate object $o \in OS_u$ and the selected object o_1 is calculated as follows.

$Aggr(o_2, o_1) = 0$, $Aggr(o_3, o_1) = 0$, $Aggr(o_4, o_1) \approx -0.0082$, $Aggr(o_5, o_1) \approx -0.0082$.

From this, we can get

$w(o_2) \times (Corr(o_2) - Aggr(o_2, o_1)) = 0.1119$, $w(o_3) \times (Corr(o_3) - Aggr(o_3, o_1)) = 0.1119$, $w(o_4) \times (Corr(o_4) - Aggr(o_4, o_1)) = 0.0398$,
$w(o_5) \times (Corr(o_5) - Aggr(o_5, o_1)) = 0.0398$.

Next, o_4 is selected to add to OS, there is $OS = \langle o_1, o_4 \rangle, OS_u = \{o_2, o_3, o_5\}$.

The similar strategy is used to select objects one by one and an ordered object set $OS = \langle o_1, o_4, o_5, o_2, o_3 \rangle$ is obtained.

Finally, according to Definition 13, the outlier score for each subject is calculated as follows.

$Score(o_1) = 1$, $Score(o_2) = 0.25$, $Score(o_3) = 0.20$, $Score(o_4) = 0.50$, $Score(o_5) \approx 0.33$.

6. Experiments

We conduct an extensive experimental comparison with the existing algorithms. First, some preparations for the experiments are presented. Then, the experimental results are compared and analyzed. Finally, hypothesis testing is performed to verify the statistical validity of the proposed algorithm.

6.1. Experimental preparation

To verify the effectiveness of the proposed method, we compared the following seven methods:

(1) Overlap Metric-based DIStance (OMDIS) algorithm (Jiang et al., 2009): The method first gives a modified definition of overlap metric in RS theory for the traditional overlap metric and then constructs an OD model based on the modified overlap metric.
(2) Overlap Metric-based k-Nearest Neighbor (OMkNN) (Jiang et al., 2009): The overlap metric is used to perform k-nearest neighbor OD.

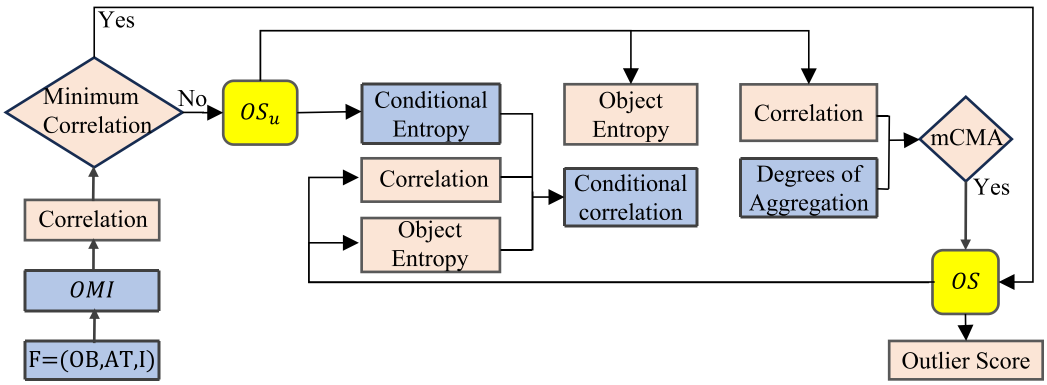

Fig. 1. The framework of OEOD model.

Table 2
Dataset characteristics.

Datasets	Abbr.	Attributes	Objects	Outliers
CreditA_plus_42_variant1	Cred	15	425	42
Diabetes_tested_positive_26	Diab	8	526	26
Ecoli	Ecoli	7	336	9
Glass	Glass	9	214	9
Heart270_16	Heart	13	166	16
Horse_1_12	Horse	27	256	12
Letter	Letter	32	1600	100
Mushroom_p_221_variant1	Mush	22	4429	221
Pima_TRUE_55	Pima	9	555	55
Vowels	Vowels	12	1456	50
Wdbc_M_39	Wdbc	31	396	39
Wine	Wine	13	129	10

Table 3
Experimental evaluation of AUC.

Dataset	OMDIS	OMkNN	GrC	IE	ODGrCR	RSOD	GCOD	OEOD
Cred	0.950	0.925	0.605	0.941	0.949	0.880	**0.971**	0.947
Diab	0.950	0.945	0.419	0.950	**0.954**	0.641	**0.954**	0.953
Ecoli	0.813	0.794	0.764	0.760	0.816	0.520	0.814	**0.827**
Glass	0.759	0.739	0.712	**0.834**	0.739	0.275	0.734	0.793
Heart	0.988	0.965	0.988	0.984	0.983	0.974	0.988	**0.990**
Horse	0.958	0.890	**0.967**	0.962	0.947	0.915	0.965	**0.967**
Letter	0.590	**0.824**	0.561	0.586	0.596	0.611	0.592	0.610
Mush	0.955	0.548	0.951	0.979	**0.985**	0.970	0.979	0.948
Pima	0.931	0.906	0.691	0.917	0.935	0.570	0.937	**0.938**
Vowels	0.679	**0.895**	0.679	0.754	0.705	0.655	0.706	0.814
Wdbc	0.997	0.793	**0.999**	0.997	0.997	0.994	0.997	0.998
Wine	0.859	0.387	0.871	0.854	0.855	0.818	0.869	**0.884**
Average	0.869	0.801	0.767	0.877	0.872	0.735	0.876	**0.889**
1st order	0	2	2	1	2	0	2	**5**
2nd order	1	0	1	0	1	1	4	**4**

(3) GrC-based outliers (Chen et al., 2008): In the GrC model, a new definition of outliers is proposed, i.e., outliers based on GrC
(4) IE-based outliers (Jiang et al., 2010): The method utilizes the IE model in RSs and proposed a new definition of outliers, i.e., outliers based on IE.
(5) RSOD (Jiang et al., 2011): The method uses the concepts of knowledge entropy and attribute importance in RS theory to construct three types of sequences, and detects outliers by analyzing the changes of elements in the sequence.
(6) OD based on GrC and rough set theory (ODGrCR) (Jiang and Chen, 2015): An OD method is proposed from the perspective of GrC and RS Theory.
(7) Granular Concept-based OD (GCOD) (Hu et al., 2023): A granular concept-based OD model. It introduces formal concept analysis to address the problem of OD.

On 12 datasets, we experimentally compared the OEOD algorithm with 7 algorithms. In order to get the optimal detection performance, different algorithms need to set appropriate parameters. In the experiments, OMkNN mainly involved the parameter k; thus, we vary the value of k from 1 to 60 with step size 1 to determine the optimal setting. For the GrC algorithm, the experimental parameters were set to the default values as described in the original paper. The other comparison algorithms OMDIS, IE, ODGrCR, RSOD, and GCOD are parameter-free.

Since the above comparative methods only work on nominal data, we discretize the numerical attributes using a discretization technique of equal width, where the number of intervals is 3. In addition, 12 datasets were downloaded from public web pages[1,2] for comparative experiments, and Table 2 presents essential information regarding the data used in this study.

[1] https://odds.cs.stonybrook.edu
[2] https://github.comBElloney/Outlier-detection

The Receiver Operating Curve (ROC) and Area Under Curve (AUC) indexes are used in the tests to evaluate the algorithm's performance extensively (Liu et al., 2023; Yuan et al., 2023b,a). OD algorithms generate an Anomaly Score (AS) for each sample and use a threshold to identify outliers. The ROC curve plots False Positive Rate (FPR) on the x-axis and True Positive Rate (TPR) on the y-axis, and the closer it is to the upper left corner, the better the detection performance. AUC is used to evaluate the performance of the comparison algorithm, and its value ranges from 0 to 1, with a higher AUC indicating better performance. AUC is a parameter-free metric that represents the average of the probabilities of outlier-normal point pairs.

We provide the F1-score, which is the harmonic mean of precision and recall. The F1-score is dependent upon the threshold, and we report the largest F1-score across all thresholds. An F1-score of 1 means there exists a threshold that gives both precision and recall equal to 1, i.e., a perfect separation of inliers and outliers.

6.2. Experimental result

We analyzed the performance of the OD algorithms using ROC curves. Fig. 2 shows the ROC curves of the eight algorithms, with the black curves corresponding to the algorithms studied here. All comparison algorithms displayed monotonically increasing ROC curves, as shown in Fig. 2. A detection algorithm's ROC curve closer to the upper left corner of the figure indicates better performance, as discussed earlier.

Results presented in Fig. 2 indicate that the proposed OEOD algorithm outperforms other OD algorithms on several datasets, including Pima, Wdbc, etc, where its ROC curve is closest to the upper left corner, i.e., the optimal point. Moreover, OEOD's ROC curve also demonstrates

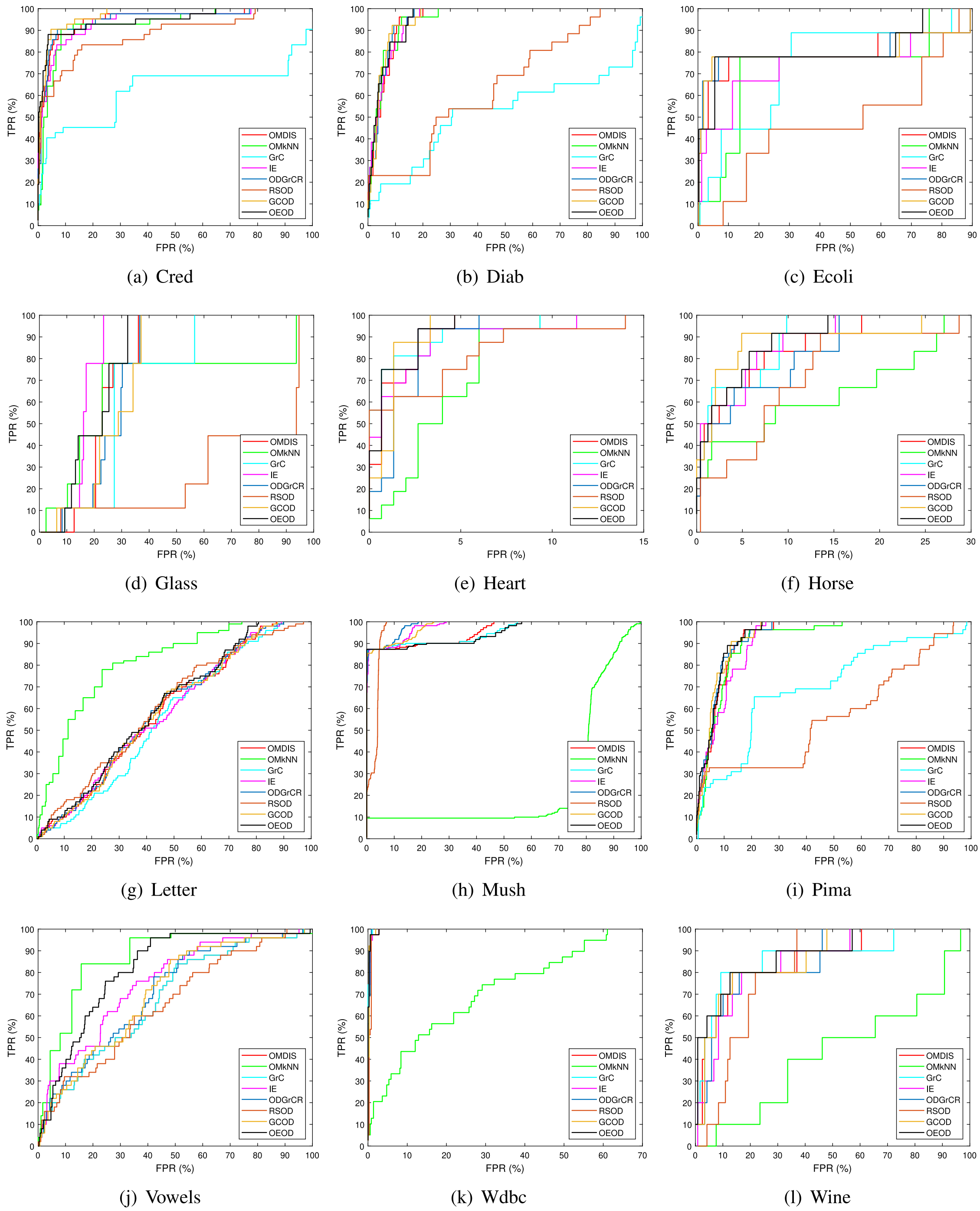

Fig. 2. Experimental comparison results on ROC.

its relative proximity to the optimal point on other datasets such as Cred, Diab, Glass, Horse, and Vowels. These observations suggest that OEOD exhibits superior performance compared to other algorithms in most cases.

However, some datasets exhibit similar performance in the ROC curves of the two algorithms, such as Ecoli, Heart, Wine, and so on. It is tough to determine which algorithm is completely superior in this circumstance. As a consequence, we present below the AUC comparison findings of 8 detection methods.

Table 3 presents the experimental results, with the best results highlighted in bold, and the second highest results highlighted in underline. OEOD achieves better results in most cases. For example, OEOD achieved the highest AUC values on 5 datasets, while the OMDIS, OMkNN, GrC, IE, ODGrCR, RSOD, and GCOD only achieved 0, 2, 2, 1, 2, 0, and 2 optimal AUC values, respectively. OEOD achieved the second highest AUC value on 4 datasets while OMDIS, OMkNN, GrC, IE, ODGrCR, RSOD, and GCOD only achieved the second highest AUC values of 1, 0, 1, 0, 1, 1, and 4 respectively. In addition, the proposed algorithm also achieves the optimal value for the mean value. These findings indicate that OEOD exhibits superior performance both overall and statistically.

Table 4 presents the experimental results of F1-score, with the best results highlighted in bold, and the second highest results highlighted in underline. OEOD achieves better results in most cases. For example, OEOD achieved the highest F1-score values on 4 datasets, while the OMDIS, OMkNN, GrC, IE, ODGrCR, RSOD, and GCOD only achieved

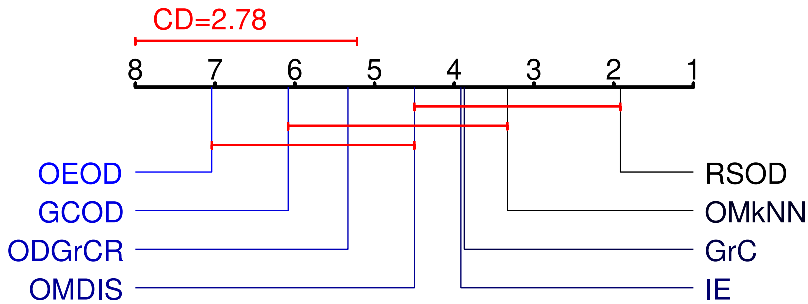

Fig. 3. Nemenyi's test figure on AUC.

Table 4
Experimental evaluation of F1-score.

Dataset	OMDIS	OMkNN	GrC	IE	ODGrCR	RSOD	GCOD	OEOD
Cred	0.731	0.655	0.479	0.683	0.747	0.667	0.783	**0.796**
Diab	0.491	0.509	0.133	0.493	0.554	0.364	**0.563**	0.548
Ecoli	0.500	0.215	0.211	0.375	**0.600**	0.094	0.571	0.571
Glass	0.197	0.219	0.192	**0.275**	0.194	0.085	0.191	0.214
Heart	0.857	0.750	0.839	0.833	0.857	0.720	**0.875**	0.857
Horse	0.600	0.455	0.667	0.632	0.545	0.393	**0.692**	0.609
Letter	0.151	**0.337**	0.143	0.140	0.153	0.153	0.154	0.155
Mush	0.921	0.174	0.904	0.863	0.881	0.649	0.914	**0.932**
Pima	0.578	0.548	0.359	0.533	0.616	0.391	0.622	**0.623**
Vowels	0.155	**0.328**	0.155	0.241	0.180	0.176	0.167	0.200
Wdbc	0.962	0.391	0.962	0.949	**0.963**	0.950	0.962	0.962
Wine	0.526	0.144	0.552	0.462	0.483	0.364	0.526	**0.625**
Average	0.556	0.394	0.466	0.540	0.564	0.417	0.585	**0.591**
1st order	0	2	0	1	2	0	3	**4**
2nd order	3	1	3	1	2	0	4	**4**

0, 2, 0, 1, 2, 0, and 3 optimal F1-score values, respectively. OEOD achieved the second highest F1-score value on 4 datasets while OMDIS, OMkNN, GrC, IE, ODGrCR, RSOD, and GCOD only achieved the second highest F1-score value of 3, 1, 3, 1, 2, 0, and 4 respectively. In addition, the proposed also achieves the optimal value for the mean value. These findings indicate that OEOD exhibits superior performance both overall and statistically.

6.3. Statistical analysis

According to the strategy in Yuan et al. (2023b), we first use Friedman test to estimate whether there is a significant difference between all algorithms. Then, Nemenyi test is used to distinguish them. In the Nemenyi test figure, the average ordinal values of 8 different OD algorithms determine their corresponding positions on the horizontal axis. If a horizontal line segment connects a group of algorithms, it indicates that there is no significant difference between this group of algorithms.

From Table 3, we can see that there are 8 algorithms and 12 datasets in the experiment, from which we can obtain F distribution with degrees of freedom of 7 and 77. According to Friedman test, the value of $\tau_F = 8.5988$ is greater than the critical value of 1.7963 when significance level $\alpha = 0.1$. Therefore, the null hypothesis "all algorithms perform equally" does not hold. It illustrates a significant difference in performance between these OD algorithms, at which point a post-hoc test is needed to further differentiate them.

In this work, the critical distance is calculated as $CD_\alpha = 2.7800$ where significance level $\alpha = 0.1$. Finally, Nemenyi test figure on AUC is shown in Fig. 3. As can be seen from Fig. 3, OEOD is statistically significantly different from most other algorithms. For example, according to the results in Fig. 3, it can be analyzed that OEOD has no horizontal line segment connection with IE, GrC, OMkNN, or RSOD, indicating that OEOD is statistically significantly different from these algorithms. However, there is no consistent evidence for statistical differences among OEOD, GCOD, ODGrCR, and OMDIS.

7. Conclusions

In this paper, we propose a new OD method based on attribute granules. Through the definition of concepts such as object entropy in FCA, we give a definition of outliers for nominal data and proposed OEOD. Experiments on 12 real data sets show that OEOD outperforms existing algorithms in most cases. Since OD using FCA is rare, this paper extends the application of FCA in data mining and other fields and opens up a new application space for FCA theory. Since the OD method proposed in this paper is based on the classical FCA model, this model cannot be directly used to process numerical data. Therefore, in our future work, we plan to perform OD under the fuzzy concept analysis model to address this limitation.

CRediT authorship contribution statement

Chang Liu: Conceptualization, Methodology, Software, Investigation, Writing – original draft. **Dezhong Peng:** Supervision, Resources, Project administration, Funding acquisition. **Hongmei Chen:** Resources, Funding acquisition, Project administration. **Zhong Yuan:** Supervision, Structuralization, Validation, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have shared the data link in the article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62306196, 62372315, and 62376230), Sichuan Science and Technology Program (2023YFQ0020, 2022YFQ0014, 2022YFH0021, and 2023ZYD0143), and the Fundamental Research Funds for the Central Universities, China (YJ202245).

References

Alaverdyan, Z., Jung, J., Bouet, R., Lartizien, C., 2020. Regularized siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: application to epilepsy lesion screening. Med. Image Anal. 60, 101618.

Albanese, A., Pal, S.K., Petrosino, A., 2012. Rough sets, kernel set, and spatiotemporal outlier detection. IEEE Trans. Knowl Eng. 26 (1), 194–207.

Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J., 2000. LOF: identifying density-based local outliers. Acm Sigmod Record 29 (2), 93–104.

Chen, Y.M., Miao, D.Q., Wang, R.Z., 2008. Outlier detection based on granular computing. In: International Conference on Rough Sets and Current Trends in Computing. Springer, pp. 283–292.

Chen, Y.M., Miao, D.Q., Zhang, H.Y., 2010. Neighborhood outlier detection. Expert Syst. Appl. 37 (12), 8745–8749.

Cordero, P., Enciso, M., López, D., Mora, A., 2020. A conversational recommender system for diagnosis using fuzzy rules. Expert Syst. Appl. 154, 113449.

Edgeworth, F.Y., 1887. Xli. on discordant observations. Lond. Edinb. Dublin Philos. Mag. J. Sci. 23 (143), 364–375.

Ganter, B., Wille, R., 1999. Formal Concept Analysis: Mathematical Foundations. Springer, Berlin.

Hao, F., Yang, Y.X., Min, G., Loia, V., 2021. Incremental construction of three-way concept lattice for knowledge discovery in social networks. Inform. Sci. 578, 257–280.

He, Z.Y., Xu, X.F., Deng, S.C., 2003. Discovering cluster-based local outliers. Pattern Recognit. Lett. 24 (9–10), 1641–1650.

Hu, K.Y., Lu, Y.C., Shi, C.Y., 2000. An integrated mining approach for classification and association rule based on concept lattice. J. Softw. 11 (11), 1478–1484.

Hu, Q., Yuan, Z., Qin, K.Y., Zhang, J., 2023. A novel outlier detection approach based on formal concept analysis. Knowl.-Based Syst. 268, 110486.

Huang, M., Lin, J.J., Peng, Y., Xie, X., 2018. Design a batched information retrieval system based on a concept-lattice-like structure. Knowl.-Based Syst. 150, 74–84.

Jiang, F., Chen, Y.M., 2015. Outlier detection based on granular computing and rough set theory. Appl. Intell. 42 (2), 303–322.

Jiang, F., Du, J.W., Ge, Y., Sui, Y.F., Cao, C.G., 2011. Sequence outlier detection based on rough set theory. ACTA Electron. Sinica 39 (2), 345–350.

Jiang, F., Liu, G.Z., Du, J.W., Sui, Y.F., 2016. Initialization of K-modes clustering using outlier detection techniques. Inform. Sci. 332, 167–183.

Jiang, F., Sui, Y.F., Cao, C.G., 2005. Outlier detection using rough set theory. In: International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing. Springer, pp. 79–87.

Jiang, F., Sui, Y.F., Cao, C.G., 2009. Some issues about outlier detection in rough set theory. Expert Syst. Appl. 36 (3), 4680–4687.

Jiang, F., Sui, Y.F., Cao, C.G., 2010. An information entropy-based approach to outlier detection in rough sets. Expert Syst. Appl. 37 (9), 6338–6344.

Jiang, F., Sui, Y.F., Zhou, L., 2015. A relative decision entropy-based feature selection approach. Pattern Recognit. 48 (7), 2151–2163.

Jiang, F., Zhao, H., Du, J., Xue, Y., Peng, Y., 2019. Outlier detection based on approximation accuracy entropy. Int. J. Mach. Learn. Cybern. 10, 2483–2499.

Juandeaburre, A.B., Fuentes-González, R., 1994. The study of the L-fuzzy concept lattice.. Math. Soft Comput. 1 (3), 209–218.

Knorr, E.M., Ng, R.T., Tucakov, V., 2000. Distance-based outliers: algorithms and applications. VLDB J. 8 (3), 237–253.

Knox, E.M., Ng, R.T., 1998. Algorithms for mining distancebased outliers in large datasets. In: Proceedings of the International Conference on Very Large Data Bases. Morgan Kaufmann, Burlington, pp. 392–403.

Li, J.H., Mei, C.L., Kumar, C.A., Zhang, X., 2013. On rule acquisition in decision formal contexts. Int. J. Mach. Learn. Cybern. 4 (6), 721–731.

Li, J.H., Mei, C.L., Wang, J.H., Zhang, X., 2014. Rule-preserved object compression in formal decision contexts using concept lattices. Knowl. Based Syst. 71, 435–445.

Liu, C., Yuan, Z., Chen, B.Y., Chen, H.M., Peng, D.Z., 2023. Fuzzy granular anomaly detection using Markov random walk. Inform. Sci. 646, 119400.

Merwe, D.V.D., Obiedkov, S., Kourie, D., 2004. Addintent: A new incremental algorithm for constructing concept lattices. In: International Conference on Formal Concept Analysis. Springer, pp. 372–385.

Mi, J.S., Leung, Y., Wu, W.Z., 2010. Approaches to attribute reduction in concept lattices induced by axialities. Knowl.-Based Syst. 23 (6), 504–511.

Nourine, L., Raynaud, O., 1999. A fast algorithm for building lattices. Inf. Process. Lett. 71 (5–6), 199–204.

Qi, J.J., Wei, L., Yao, Y.Y., 2014. Three-way formal concept analysis. In: International Conference on Rough Sets and Knowledge Technology. Springer, pp. 732–741.

Ramaswamy, S., Rastogi, R., Shim, K., 2000. Efficient algorithms for mining outliers from large data sets. ACM SIGMOD Record 29 (2), 427–438.

Shao, M.W., Li, K.W., 2017. Attribute reduction in generalized one-sided formal contexts. Inform. Sci. 378, 317–327.

Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., Lakhal, L., 2002. Computing iceberg concept lattices with titanic. Data Knowl. Eng. 42 (2), 189–222.

Thangaramya, K., Kulothungan, K., Indira Gandhi, S., Selvi, M., Santhosh Kumar, S., Arputharaj, K., 2020. Intelligent fuzzy rule-based approach with outlier detection for secured routing in WSN. Soft Comput. 24 (21), 16483–16497.

Trnecka, M., Trneckova, M., 2018. Data reduction for boolean matrix factorization algorithms based on formal concept analysis. Knowl.-Based Syst. 158, 75–80.

Wang, Z.H., Hu, K.Y., Hu, X.G., Liu, Z.T., Zhang, D.C., 1999. General and incremental algorithms of rule extraction based on concept lattice. Chinese J. Comput. 22 (1), 66–70.

Wang, B., Mao, Z.Z., 2019. Outlier detection based on Gaussian process with application to industrial processes. Appl. Soft Comput. 76, 505–516.

Wang, C.Z., Qian, Y.H., Ding, W.P., Fan, X.D., 2021. Feature selection with fuzzy-rough minimum classification error criterion. IEEE Trans. Fuzzy Syst. 30 (8), 2930–2942.

Wei, L., Cao, L., Qi, J.J., Zhang, W.X., 2020. Concept reduction and concept characteristics in formal concept analysis. Sci. China (Inf. Sci.) 50, 1817–1833.

Wille, R., 1982. Restructuring lattice theory: an approach based on hierarchies of concepts. In: International Conference on Formal Concept Analysis. Springer, pp. 445–470.

Wu, W.Z., Leung, Y., Mi, J.S., 2008. Granular computing and knowledge reduction in formal contexts. IEEE Trans. Knowl. Data Eng. 21 (10), 1461–1474.

Xie, Z.P., Liu, Z.T., 2000. Concept lattice and association rule discovery. J. Comput. Res. Devel. 37 (12), 1415–1421.

Yang, P., Zhu, Q.S., 2011. Finding key attribute subset in dataset for outlier detection. Knowl.-Based Syst. 24 (2), 269–274.

Yao, Y.Y., 2004. Concept lattices in rough set theory. In: IEEE Annual Meeting of the Fuzzy Information, 2004. Processing NAFIPS'04, Vol. 2. IEEE, pp. 796–801.

Yuan, Z., Chen, H.M., Li, T.R., Liu, J., Wang, S., 2021. Fuzzy information entropy-based adaptive approach for hybrid feature outlier detection. Fuzzy Sets and Systems 421, 1–28.

Yuan, Z., Chen, H.M., Li, T.R., Sang, B.B., Wang, S., 2022. Outlier detection based on fuzzy rough granules in mixed attribute data. IEEE Trans. Cybern. 52 (8), 8399–8412.

Yuan, Z., Chen, B.Y., Liu, J., Chen, H.M., Peng, D.Z., Li, P.L., 2023a. Anomaly detection based on weighted fuzzy-rough density. Appl. Soft Comput. 134, 109995.

Yuan, Z., Chen, H.M., Luo, C., Peng, D.Z., 2023b. MFGAD: Multi-fuzzy granules anomaly detection. Inf. Fusion 95, 17–25.

Yue, M., Hong, T., Wang, J.H., 2019. Descriptive analytics-based anomaly detection for cybersecure load forecasting. IEEE Trans. Smart Grid 10 (6), 5964–5974.

Zhang, J.F., Jiang, Y.Y., Chang, K.H., Zhang, S.L., Cai, J.H., Hu, L.H., 2009. A concept lattice based outlier mining method in low-dimensional subspaces. Pattern Recognit. Lett. 30 (15), 1434–1439.

Zou, C.F., Deng, H.F., Wan, J.F., Wang, Z.R., Deng, P., 2018. Mining and updating association rules based on fuzzy concept lattice. Future Gener. Comput. Syst. 82, 698–706.

Zou, L.G., He, T.T., Dai, J.H., 2022. A new parallel algorithm for computing formal concepts based on two parallel stages. Inform. Sci. 586, 514–524.