Towards Identifying Multicriteria Outliers: An Outranking Relation-Based Approach

Baroudi Rouba, University of Abdelhamid Ibn Badis, Mostaganem, Algeria & LITIO Laboratory, University of Oran 1
Ahmed Benbella, Oran, Algeria
Safia Nait-Bahloul, LITIO Laboratory, University of Oran 1 Ahmed Benbella, Oran, Algeria

ABSTRACT

This article tackles the problem of outlier detection in the multicriteria decision aid (MCDA) field. The authors propose an outlier detection method based on binary outranking relations and Local Outlier Factor (LOF) algorithm. The outlier is detected by applying LOF algorithm on the distribution of the outranking relations generated by a multicriteria outranking method. The proposed approach is illustrated on an artificial example and evaluated on a real life financial problem, the country risk problem.

KEYWORDS

LOF Algorithm, Multicriteria Decision, Outlier Detection, Outranking Relation

1. INTRODUCTION

Outlier detection is one of the most important techniques in knowledge discovery in databases (KDD), which aims to detect objects that are substantially different from the rest of the data.

Many algorithms have been employed in Data Mining, Machine Learning, and Statistics to find outliers. These algorithms can be classified as either distance-based (Knorr & Ng, 1997; Knorr & Ng, 1998; Knorr & Ng, 1999; Knorr, Ng & Tucakov, 2000), density-based (Breunig, Kriegel, Ng & Sander, 2000), clustering (Jain, Murty & Flynn, 1999), depth-based (Johnson, Kwok & Ng, 1998) or distribution-based (Branett & Lewis, 1994). These algorithms have been applied in several domains such as network intrusion (Ding et al, 2012), (Idé & Kashima, 2004), credit card fraud (Richard & Hand, 2001), email spam (Castillo et al, 2007), customer activity monitoring (Fawcett & Provost, 1999), and many other domains. Nevertheless, a little attention has been paid to detect outliers in multicriteria decision aid (MCDA) field.

A large number of multicriteria methods have been developed (Brans & Vincke, 1985), (Figueira, Mousseau & Roy, 2005), (Dyer, 2005) and applied in several domains (Joerin, 1997), (Belacel, 2000), (Zopounidis, 1999) but the study of the robustness of these methods has not been sufficiently tackled (Hites et al, 2006). Most of these methods are based on subjective parameters which represent decision makers’ preferences (like preference, indifference, weight and veto thresholds). Changes in the values of one of these parameters could affect the results of these methods. In addition, if the values of these parameters are not estimated carefully, the results will be less sensitive to outliers. For instance, adding only one arbitrarily high (or low) action to a given data-set is sufficient to completely modify...
the result (ranking, choice or sorting) of the method. As consequence, the detection of outliers can be exploited to improve the estimation of such parameters. Furthermore, in multicriteria classification and sorting methods, the presence of outliers is not benefit for improving classification results. These objects affect other objects to be classified correctly. Thus, detecting multicriteria outliers can be exploited as a pretreatment step before executing multicriteria classification methods.

For these reasons, it is important to study the presence of multicriteria outliers. This new research direction has not been enough treated in the literature. The only work that dealt with this problem has been developed by (De Smet, Hubinont & Rosenfeld, 2017). The authors proposed a distance-based model to detect multicriteria outliers. The authors extended the multicriteria distance proposed in (De Smet & Guzman, 2004) to different samplings of the set of objects. Outliers were detected by the identification of bi-modal distributions of the distance values.

The present paper proposes a new idea for detecting outliers in MCDA field. The paper introduces the concept of relation-based outlier. This concept permits to detect outliers using the preference relations that characterize the multicriteria outranking methods. More explicitly, in our approach, each object is identified by its relations with the rest of the objects in the data-set. The outliers will be detected by applying the local outlier factor (LOF) algorithm (Breunig, Kriegel, Ng & Sander, 2000) on the distributions of the outranking relations.

The rest of the paper is organized as follows. The main concepts related to preference relations in MCDA field are presented at first. Then, a brief description of LOF algorithm is given. Next, the algorithm of the proposed approach is detailed followed by an experimental study. Finally, we conclude by some general remarks and future works.

2. OUTRANKING RELATION-BASED APPROACH FOR IDENTIFYING MULTICRITERIA OUTLIER

2.1. Preference Relation in MCDA Problems

In what follows, we will focus on multicriteria outranking methods. In a MCDA problem, the objects submitted to a decision-making process are called actions. An action can be a potential decision, a candidate, a project, a feasible solution, an item, etc. The set of actions will be noted \( A = \{a_1, a_2, \ldots, a_n\} \). These actions are compared by defining a preference structure.

The comparison between two actions \( a_i, a_j \) of \( A \) may be expressed by a preference (P), incomparability (I) or incomparability (R) statement (Vincke, 1992). These relations are built from the comparison of every couple of actions with the respect of a set \( G = \{g_1, g_2, \ldots, g_m\} \) of \( m \) criteria using a given multicriteria outranking method (PROMETHEE I (Brans & Vincke, 1985) or ELECTRE I (Figueira, Mousseau & Roy, 2005) for example).

These preference relations have to satisfy the following conditions to be accepted as reflecting the feeling of a decision maker:

\[
\forall a_i, a_j \in A, \begin{cases}
    a_iPa_j \Rightarrow a_j Pa_i \\
a_iIa_j \Rightarrow a_j Ia_i \\
a_iRa_j \Rightarrow a_j Ra_i
\end{cases}
\]

(1)
Harnessing New Product Development Processes through Strategic Thinking Initiatives
[www.igi-global.com/article/harnessing-new-product-development-processes-through-strategic-thinking-initiatives/136284?camid=4v1a](www.igi-global.com/article/harnessing-new-product-development-processes-through-strategic-thinking-initiatives/136284?camid=4v1a)

DMSS Implementation Research: A Conceptual Analysis of the Contributions and Limitations of the Factor-Based and Stage-Based Streams
[www.igi-global.com/chapter/dmss-implementation-research/8077?camid=4v1a](www.igi-global.com/chapter/dmss-implementation-research/8077?camid=4v1a)

Online Spatial HIV/AIDS Surveillance and Monitoring System for Nigeria
[www.igi-global.com/chapter/online-spatial-hiv-aids-surveillance-and-monitoring-system-for-nigeria/138639?camid=4v1a](www.igi-global.com/chapter/online-spatial-hiv-aids-surveillance-and-monitoring-system-for-nigeria/138639?camid=4v1a)