

C3IT-2012

## A Decision Support System for Industry based upon Multivariate Spatial Outlier Detection Techniques

Arvind Sejwal<sup>a</sup>, Surender Jangra<sup>b</sup>, Yashwant Singh Sangwan<sup>c</sup>

<sup>a</sup> Head, Department of CSE, ACE Ambala-133101, INDIA

<sup>b,c</sup> Associate Professor, Department of IT, HCTM, Kaithal-136027, INDIA

---

### Abstract

Decision support systems are computer based programs which assists decision makers in effective and efficient decision making. Outliers are the data sets which are highly irrelevant with respect to main data sets. Spatial outliers are the spatial objects with distinct features from their surrounding neighbours. In this paper we have proposed a decision support system for strategic decision makers to establish an industry. We have shown how spatial outlier detection techniques may be used to aid decision makers. Each site is evaluated based upon multiple decision variables. Here, we use two different outlier detection algorithms to detect a site which is most unsuitable to establish an industry. We also perform an experiment on a synthetic dataset of fifteen different sites to locate ambiguous sites. We found that decision making process for site selection may be improved by using spatial outlier detection techniques.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of C3IT

*Keywords:* Decision support system; spatial database; site index, clustering; outlier detection; database systems.

---

### 1. Introduction

Industrial location is an increasingly important factor facing both national and international firms. To establish any industry a manager has to analyse large amount of information regarding available list of sites. Some of the critical factors of industrial location are transportation, labour, raw material, markets, utilities, electricity, climate, tax structure, technology, government policies & etc. The sites which are not suitable are completely discarded from the decision making process and such sites are called as spatial outlier sites. Spatial outlier is an object whose non-spatial attribute value is significantly different from the values of its spatial neighbours. Spatial outlier detection algorithms may be applied to find such ambiguous sites.

## 2. Related work & background

An approach for discovering outliers using distance metrics was first presented by Knorr *et al.*. They define a point to be a *distance outlier* if at least a user-defined fraction of the points in the data set are further away than some user-defined minimum distance from that point. In their experiments, they primarily focus on data sets containing only continuous attributes. Related to distance-based methods are methods that cluster data and find outliers as part of the process of clustering [1]. Points that do not cluster well are labelled as outliers. Recently, density-based approaches to outlier detection have been proposed. In this approach, a local outlier factor (LOF) is computed for each point. The LOF of a point is based on the ratios of the local density of the area around the point and the local densities of its neighbours. A Spatial-Temporal approach has been used in medical science to differentiate epidemic risk patterns [2]. Outlier detection approach has also been used for PCB Testing in industry [3]. An application of such techniques has discussed in [4] for outlier detection in high dimensional data. Another application was proposed by [5] for temporal outlier detection in vehicle traffic data. [6] Discuss application of spatial outlier detection for malicious user detection in a cognitive radio cooperative sensing systems. Another application of such techniques has been discussed in [8] where hierarchical clustered outlier detection methods have been used in laser scanner point clouds. Such applications may also be used in crowds. The [9] used one such approach where anomaly detection in crowded scenes is illustrated. Some improved techniques have been discovered for special kind of data. Spatial outlier detection may also be used in security based computing [10]. One such approach is cluster-based congestion outlier detection method on trajectory data [11]. Recently [12] used spatial outlier detection techniques to locate less developed regions in the state of haryana.

## 3. Problem formulations

Suppose we have a dataset of  $n$  sites, denoted as  $D = \{S_1, S_2, S_3, \dots, S_n\}$ . Each site has a set of  $m$  attributes,  $A = \{a_1, a_2, \dots, a_m\}$ , Where  $a_i \in A$  represents a decision parameter of the site. Problem is to find a site  $S_a \in D$  which is most ambiguous to establish an industry.  $S_a$ , obviously is a site which does not fit into minimum criteria set by the decision maker.



Fig. 1 An example of spatially distributed sites with non spatial features

Supposing that minimum requirement for choosing a site is denoted by a consolidated numerical threshold, denoted as  $\Phi$ , and then  $S_a$  be a site whose measures is below  $\Phi$ . Such a site is called as spatial outlier site. In Fig. 1, a set of seven sites are shown. Sites 2, 4, 6 and 7 are normal one, while site 1, 3 and 5 are inconsistent sites where establishing an industry, is extremely difficult or risky. Problem in our case is comprised of locating sites with red dots. These are the sites where establishing an industry is very difficult or not feasible.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات