Information Sciences 280 (2014) 322–337

Contents lists available at ScienceDirect

Information Sciences

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

A mixed integer programming formulation and solution for traffic analysis zone delineation considering zone amount decision

Linqing Wang [a], Jiafu Tang [a,b,*], Xiang Fei [c], Min Gong [c]

[a] State Key Lab of Synthetical Automation of Process Industry, Dept. of System Engineering, Northeastern University (NEU), Shenyang, PR China
[b] College of Management Science and Engineering, Dongbei University of Finance and Economics (DUFE), Dalian, PR China
[c] Smart Mobility & Intelligent Transportation, IBM China Research Lab, Beijing, PR China

ARTICLE INFO

Article history:
Received 27 October 2013
Received in revised form 18 April 2014
Accepted 22 April 2014
Available online 15 May 2014

Keywords:
Traffic analysis zone delineation
Amount of zone
Mixed integer programming
Domain reduction method
Aggregation hierarchical clustering

ABSTRACT

Due to the spatial aggregation effect, the amount of zones (AoZ) has a direct influence on the optimality of traffic analysis zone delineation (TAZD) problems. Unlike existing studies, this paper addresses a TAZD problem in which both the AoZ and zone partitions are optimized. For the TAZD problem, we propose a mixed integer programming (MIP) model with the optimization objective of minimizing the sum of geographical errors within each zone. To solve the TAZD model, the following works have been done. First, the TAZD model is strengthened by our proposed IR (influence region) construction algorithm. Second, based on the strengthened model, we analyse the lower bound of the optimal AoZ from the theoretical perspective, and then the DRMI (domain reduction method I) is proposed to specify a tighter lower bound. Also, to reduce the amount of decision variables, we put forward the DRMII (domain reduction method II) which enumerates all feasible AoZ values. Third, to settle a non-solution situation, we propose the AHCBHA (aggregation hierarchical clustering based heuristic algorithm) in order to rebuild the solution space. Fourth, we then design two methods to solve the TAZD model. The former method combines the AHCBHA with the DRMI and the latter method combines the AHCBHA with the DRMII. Finally, the performance of the proposed TAZD and the developed methods is explored using a numerical example and a real-world case. The obtained results constitute good solutions.

1. Introduction

Zoning system can greatly influence any observed spatial relationships or patterns, which are the effects known as the modifiable areal unit problem (MAUP) and the scale aggregation problem [19–23]. Thus, the zoning system design has become an important research field in studies that employ geographical areas as the analysis units [1,8,9,15,17,20,24,29,30]. In transportation science, the design of such zoning system is known as the traffic analysis zone delineation (TAZD) problem. TAZD has long been recognized as an essential component for transport policy evaluations and for real-time traffic operation and management [23].

* Corresponding author at: College of Management Science and Engineering, Dongbei University of Finance and Economics (DUFE), Shahekou, Dalian 116026, PR China. Tel.: +86 24 83673810.

E-mail addresses: tangjiafu@dufe.edu.cn, jftang@mail.neu.edu.cn (J. Tang).

http://dx.doi.org/10.1016/j.ins.2014.04.040
0020-0255/

Aiming to divide a large set of spatial objects into a number of spatially contiguous regions, the TAZD problem involves a certain optimization objective, homogeneity constraints, spatial contiguity constraints and any other constraints that the planner is free to manipulate. The optimization objective could be the minimization of the distances or travel cost from each point to its zone center, the minimization of intra-zonal trips, the maximization of internal homogeneity and the minimization of internal geographical errors [3,5,14,19,31]. The definition of the homogeneity depends mainly on the requirements of the planners or the TAZD context, such as the homogeneity of socio-economic, the homogeneity of accessibility, the homogeneity of traffic demand and the homogeneity of traffic habit [3,24,26]. The contiguity represents the quality of a single zone being connected (see comprehensive reviews in [28]). The contiguity constraint is difficulty to formulate in a mathematical way. Some recent researches have studied exact contiguity modeling for TAZD problems [10,19,28].

Regarding the approaches used to solve TAZD problems, there are primarily six delineation methods that are encountered in the literature: clustering [13], heuristic search [19], meta-heuristic search [4,7,15,22,27], graphical and numerical simulation [2,12], graph theory [3], and mathematical programming [9,28]. A clustering or heuristic/meta-heuristic method usually explores solution space directly on the basis of certain algorithm framework, considering less about analytical properties of optimal solutions. Thus, it is hard to value the optimality of solutions from the mathematical perspective. An areal interpolation [2] and GPU-based algorithm [12] performs in a simulation way. It uses model to generate useful data and simulates the spatial propagation and distribution patterns of data. Then, zones are designed according to the borders between different patterns. The zoning solutions obtained by simulation methods are qualified only when information was abundant and the correctness of simulation model can be assured. A graph theory method focuses mainly on how to build the abstract structure of the search space in TAZD problems using relative concepts, such as treating a zone as a spanning tree or subgraph, but the method takes little care about how to solve the problem. Mathematical programming methods were not recommended due to the computational complexity in earlier days when it was first proposed in the area of TAZD, even if the problem itself has long been noted to be analogous to several classical operational research problems, such as the: Set Partitioning Problem and Set Covering Problem. Nevertheless it is believed in this paper that using mathematical programming methods to solve TAZD problems makes sense in both theory and practice. On the one hand, some types of problems that are similar to TAZD problems have been solved using mathematical programming [7,21] hitherto, and the increasing sophistication of mathematical programming and the availability of more powerful computers has allowed such problems to be handled ever more efficiently with respect to formalized single-objective optimization or multi-objective optimization. On the other hand, mathematical programming model is needed to determine an optimal solution, and it is a basic criterion to choose between a heuristic algorithm and an exact one.

With respect to the amount of zones (AoZ), it is apparent that if AoZ cannot be fixed in advance, the resulting TAZD problem becomes much more difficult to be solved. AoZ has a direct influence on the optimality of a TAZD problem because of the spatial aggregation effect. It is often treated as a constant or with a pre-defined range in most TAZD studies. However, the best optimal configuration of the AoZ should inherently be decided according to a concrete context of TAZD. There are primarily two types of researches study on the topic of determining the optimal AoZ. One type of these studies treats the problem as a part of the MAUP and scale aggregation problem [6,11,25], which allows the planners to explore the aggregation element of the MAUP by considering the sensitivity of results to different zone designs at each scale. Focusing on the error while paying less attention on the mathematical model, this type of research involves: (1) identifying and investigating errors of the solutions [6] and (2) developing methods to eliminate or reduce the errors [11,25]. The other type of these studies involves hierarchical grouping to determine the optimal AoZ, i.e., form several zone partition solutions with different AoZ configurations for one TAZD case and then let the planners or algorithms make the final decision according to the requirements of the optimization objective [19]. Martínez [19] has reported that the TAZ number had a significant influence on the global indicators of zone system and adds an outer zone-number-loop on his proposed algorithm to give different zone solutions with several chosen AoZ configurations. Being different from the above research, this paper raises the AoZ decision problem during the process of the zone partition decision by considering the TAZD scenario without a pre-defined AoZ.

In this paper, we focus on the mixed integer programming (MIP) formulation and solution for the TAZD problem considering the AoZ decision. Within the framework of the classical *k*-median model, we propose a mixed integer programming model for the TAZD problem in which the AoZ is not pre-defined and is treated as one of the decision variables. The model minimizes the sum of geographical errors within each zone while guaranteeing the accessibility homogeneity and the area homogeneity. After implementing an influence region construction algorithm (IRCA), the initial model, which considers the basic geographical unit (BGU) as a building block, is subsequently strengthened. Both the analytical and numerical results show that the IRCA would indeed reduce the size of the TAZD problem by changing the decision elements from the BGU scale to the IR scale. Then, the analytical nature of the optimal AoZ is derived within the framework of the strengthened model. Based on the analytical nature of the optimal AoZ, two domain reduction methods (DRM) are given. One narrows the range of the AoZ variable by specifying the maximum lower bound. The other provides a way to enumerate all feasible AoZ values according to homogeneity constraints of the proposed model. To deal with a non-solution situation of the proposed model, a heuristic algorithm based on the aggregation hierarchical clustering (AHCBHA) is then developed. Combining this AHCBHA with each DRM, we design two methods to solve the proposed model. We demonstrate the performance of our proposed methods by a numerical example and a real-world case (TAZD in Suzhou, China). The results indicate that the proposed method is greatly effective to solve the large-size TAZD problem without a pre-defined AoZ, and the methods could provide not only the optimal zone partition but also the optimal AoZ.