Robust multisource geographic entity matching through the maximization of geometric and semantic similarities

Yao Cheng (China, PR)

Key words: Cartography; Geoinformation/GI

SUMMARY

Geographic entity matching is an important method for performing multisource spatial data fusion and information association and sharing. Corresponding matching methods have been designed in the existing studies for different types of entity data characteristics, such as lines and areas. However, these approaches are often limited in terms of their generalizability for matching heterogeneous data derived from multiple sources and their accuracy in complex pattern matching tasks. To resolve these problems, a robust multisource geographic entity matching method implemented by maximizing geometric and semantic similarities is proposed. First, the entire target entity is segmented based on its shape features, and the partitioned feature segments are extracted as matching primitives. Second, the feature segments are grouped into patterns, encompassing three major categories and fourteen subcategories. Following this, pattern matching is performed based on spatial similarity metrics such as the maximum projection distance. Finally, the spatial matches are detected and refined through semantic similarity calculations. The proposed method is tested on two datasets derived from Southeast China and Northwest China. The experimental results demonstrate that our method can be effectively applied to both area and line entity matching tasks with strong generalizability and application capabilities and significantly improved matching accuracy. Specifically, nine feature segment matching patterns for matching area entities and six patterns for line entities are utilized, and the precision and recall values are nearly 90%.
