Conflation of national Mapping and Crowd-Sourced Data – a Comparison of two different Approaches

Zihan Liu, Volker Walter, Dieter Fritsch

Institute for Photogrammetry, Stuttgart University, Geschwister-Scholl-Str. 24D, D-70174 Stuttgart, Germany

Abstract. Along with the increasing power of Geographical Information Systems (GIS) there is an increasing demand for spatial data. National and private institutions collect spatial data in different data models and scales in order to meet this demand. Additionally, huge amounts of spatial data are collected in Web 2.0 mapping portals. The result is a multiple representation of the same topographic objects of the landscape. The aim of this project is to investigate the integration of such datasets. We use datasets from a National Mapping Agency (NMA) and from a Web 2.0 mapping portal. The integration will be done with conflation techniques. Conflation can be summarized as the process of integrating geographical datasets, combining multisource data, improving data quality, and updating spatial information. In this study, two programs (ifp conflation program and Radius Studio) are used to conflate ATKIS and OSM datasets in two test areas. The relative accuracy of the source datasets and the conflated datasets are evaluated by overlap analysis. A basic buffer overlap and an increment buffer overlap are calculated to examine the difference of datasets.

Keywords: Conflation, ATKIS, OSM matching, and overlap analysis

1. Introduction

There are mainly two types of inconsistencies between two spatial datasets: (1) heterogeneous feature fitting error and (2) homogeneous feature representation error (Volz, 2006). The focus of this paper is on the second inconsistence. Conflation consists of feature alignment, matching, and deconfliction. Matching - including geometrical, topological, and semantic matching - is the core of conflation (Wiemann and Bernard, 2010). The principle of matching is to identify the correspondence entities from source layers.

This work focuses on the conflation of two vector datasets. Firstly, the two datasets are compared based on quantitative and qualitative evaluations. Buffers were created with increment distances to analyse the percentage of overlap changes. Then the datasets were conflated by software with distinct functionality. The comparison of the two conflation methods is evaluated, regarding to the program capability and the conflation quality.

2. Data

Two test areas with a size 2×2 km^2 were chosen in Baden-Württemberg, Germany. Stuttgart is the capital of the state and one of the test areas is located in the downtown of the city with relative complex street elements. The other test area is Calw in the southwest of Stuttgart with less complex street distribution.

3. Software

In the ifp conflation program developed by Hainan Chen (Chen and Walter, 2009), the datasets are first matched manually and then conflated automatically. Radius Studio is a fully automatic rule-based spatial data evaluation and processing software. The reference and target datasets are imported into the system. Then predefined rules and actions are executed to conflate the data. Figure 1 shows the two different conflation approaches in a flow chart.

4. Results

The data accuracy is assessed by measuring the relative accuracy of ATKIS and OSM data. Buffer analysis based on increment radius reveals the positional accuracy: 10m buffers are created for both datasets and the percentage of overlap is calculated. This procedure is repeated with different buffer sizes. To assess the completeness, the number of elements and the total length of all elements is calculated. Finally, we compared manually the datasets and identified three different classes of inconsistencies: completeness difference, endpoints difference and collection accuracy difference

The ifp conflation program matches the data manually and combines the data either in simple or complicated scenarios. The Radius Studio software is based on fully automatic geometric and attribute matching. The automatic matching is more efficient than manual matching, but the accuracy of automatic matching is lower. The ifp conflation program combines data

with geometry correction and the connectivity of streets are reserved. Therefore, the ifp conflation program can be used for applications that require very high quality datasets. For application which require high updating rate, Radius Studio is ideal.

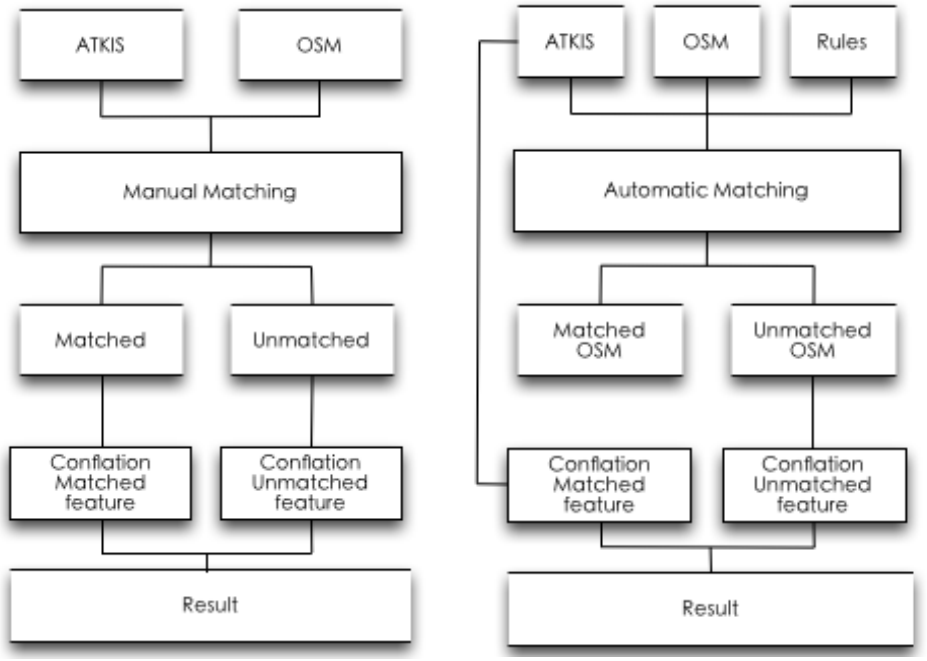

Figure 1. Flowchart of ifp conflation program and Radius Studio

References

Adv, (2010), Authorities Topographic Cartographic Information System. Accessed date: 4, Dec, 2012.

Chen, H., Walter, V. (2009). Quality Inspection and Quality Improvement of Large Spatial Datasets. In Proceedings of the GSDI 11 World Conference: Spatial Data Infrastructure Convergence: Building SDI Bridges to Address Global Challenges.

Volz, S. (2006). An iterative approach for matching multiple representations of street data. In ISPRS Workshop, Multiple representation and interoperability of spatial data :101-110.

Wiemann, S., Bernard, L. (2010). Conflation Services within Spatial Data Infrastructures. In 13th AGILE International Conference on Geographic Information Science.