Analysing OpenStreetMap data for topological errors

Sukhjit Singh Sehra*

I.K.G. Punjab Technical University,
Jalandhar, Punjab, India
Email: sukhjitsehra@gmail.com
*Corresponding author

Jaiteg Singh

School of Computer Sciences,
Chitkara University,
Punjab, India
Email: jaitegkhaira@yahoo.co.in

Hardeep Singh Rai

Department of Civil Engineering,
Guru Nanak Dev Engineering College,
Ludhiana, Punjab, India
Email: hardeep.rai@gmail.com

Abstract: OpenStreetMap is producing huge spatial data contributed by users of different backgrounds and varying level of mapping experiences. Due to this generated map data may be topologically incorrect, which explicitly expresses the spatial relationship between features. To make the map data navigable, it is important that data is free from topological errors. The current work has been conducted to detect topological errors in OpenStreetMap data. OpenStreetMap data of Punjab (India) has been taken as test data for finding topological errors. For cleaning the topological errors, map data has been processed using different algorithms of open source geographic information systems, and it has been concluded that OpenStreetMap data is not free from topological errors and need a thorough preprocessing before being used for navigation purposes.

Biographical notes: Sukhjit Singh Sehra is currently pursuing his PhD at I.K.G. Punjab Technical University, Jalandhar, Punjab and working as an Assistant Professor at Guru Nanak Dev Engineering College, Ludhiana, India. He is an OpenStreetMap contributor. His areas of research are assessment of spatial databases, spatial data mining and high performance computing.

Jaiteg Singh is working with Chitkara University as an Associate Professor in the School of Computer Sciences. His area of research includes data mining, data quality and ETL testing in datawarehouses.

Hardeep Singh Rai is a Professor at Guru Nanak Dev Engineering College, Ludhiana, India. His areas of research are OpenStreetmap, BRL-CAD, BIM./Structural Design, GIS, CAD, BIM and FOSS.

1   Introduction

O'Reilly (2005) discussed the concept of Web 2.0 or 'read-write' web. Web 2.0 encouraged greater collaboration among internet users and other users, content providers and enterprises (Hudson-Smith et al., 2009). This movement provided new methods of sharing and computing data (Hudson-Smith et al., 2009; Rana and Joliveau, 2009; Turner, 2006; Walsh, 2008; Goodchild, 2009; Haklay et al., 2008) by crowdsourcing movement similar to Wikipedia (Howe, 2006). In regard to the geographical data the crowd-sourced movement is known as Volunteered Geographic Information (VGI) and others call it collaborative mapping (Fischer, 2008), so it is a special case of this web phenomenon and has been applied in many popular websites such as Wikimapia, OpenStreetMap, Google Map, Flickr (Kounadi, 2009). The CrowdSourced OpenStreetMap produces huge spatial data, with the help of numerous users of varying level of mapping experiences. Due to varying level of mapping experience of the users, various errors are introduced and one of the most common errors is topological error, also called logical inconsistency. In recent years some studies on quality have been undertaken (Neis et al., 2011), the results of the studies on OpenStreetMap data collection quality concluded overall heterogeneous quality. Heterogeneous quality means that the quality and completeness of the map data varies highly from country to country. The results showed that in urban areas of Europe, i.e., UK, Germany, Austria, and Switzerland, the OpenStreetMap data prove to be at par with the completeness to commercial or governmental data providers. But rural areas showed lower data concentration in the OpenStreetMap data. An exception is also concluded in case of the USA, where rural areas are better covered and less completeness in urban areas (Neis et al., 2011).

The current work focuses on topological relationships of OpenStreetMap data of Punjab (India), because topological relationships are concerned with determining the faithfulness of the data set. These relations typically involve spatial data inconsistencies such as incorrect line intersections, polygons not properly closed, duplicate lines or boundaries, or gaps in lines. These types of errors must be corrected to avoid incomplete features and to ensure the data integrity, as geographic information system (GIS) analyst working on transportation and navigation should use the topologically correct map data.

The main reasons for analysing map data of Punjab (India) are, firstly, societal, as this project is a community project i.e. of users and by the users and provides open and free map data. Secondly, most of the Indian map data, i.e., basic roads network and city name, is donated by Automotive Navigation Data (AND) (OpenStreetMap.org, 2013). It is possible that the topological errors may have come from AND. Lastly, in the recent years huge amount is invested by the Indian Government on building a stronger road network infrastructure. Governmental mapping agency's way of collecting the information is slow and is also not available for public visualisation. The crowdsourced OpenStreetMap platform can work as a framework for fast changes with quick response, like in case of disaster like Haiti City (Sehra et al., 2013).

This paper has been divided into different sections. The next section discusses briefly the OpenStreetMap statistics about Indian users. Section 3 discusses the importance of map topology and various topological errors. Section 4 elaborates on the methodology used for identification and correction of errors. Section 5 discusses the results and in the last are the conclusion and future directions of research work.

2 OpenStreetMap

OpenStreetMap, started in 2004, is an editable and freely available database of mapping information which can be used for many different purposes with very few restrictions (Neis and Zipf, 2012).

OpenStreetMap has three main components (Haklay and Weber, 2008), these are node, way and relation. Relation is the most important element of OpenStreetMap data structure, consisting of other members, i.e., node and ways. Relation is a multi-purpose data structure that describes the logical relationship between elements. Any member of relation can have an optional role to describe the part it plays in a relation. All types of data elements (nodes, ways and relations) can have tags. Tags describe the meaning of the particular element to which they are attached.

2.1 OpenStreetMap users

OpenStreetMap is getting popular and many enthusiastic users are joining this crowdsourced movement. This revolutionary movement has encouraged nearly 1,742,729 registered users, 2,470,855,510 uploaded nodes, 246,912,924 ways and 2,738,864 relations (OpenStreetMap.org., 2014a). For India, the total number of registered users are 2,877, 7,561,749 uploaded nodes, 433,747ways and 6,094 relations (OpenStreetMap.org., 2014b). As shown in Table 1, v1 means objects created and never modified (version 1) in this region and last edit means objects last edited i.e. version is greater than 1, in this region.

Table 1   Number of nodes, ways and relations contributed by Indian OpenStreetMap users

	Nodes	*Ways*	*Relation*
v1	6,780,201	295,329	4,879
last edit	781,548	138,418	1,215
Total (v1 + last edit)	7,561,749	433,747	6,094

> *Source:*   OpenStreetMap.org. (2014b)

2.2   *Map data contribution methods*

The contribution of newdata to the OpenStreetMap data can be accomplished using different approaches by users of varying backgrounds of mapping experiences and using different devices. The most common approach is recording map data using a Global Positioning System (GPS) receiver and edit the recorded information using one of the various freely available editors such as JOSM (Steiniger and Hunter, 2012). The users provide additional information about the collected data by adding attributes and store the final results in the OpenStreetMap database. Users do not require any specialised GPS receiver for mapping, as it has been tested that smartphones can be considered as device for mapping (Sehra et al., 2013; Golicher, 2013). The accuracy of GPS has been checked and found satisfactory, but the accuracy depends upon the quality of GPS chip. In addition if the accuracy of any smartphone is not good as compared to the professional GPS receiver, the mapper can map to good accuracy as Microsoft Bing supports the project (Neis and Zipf, 2012) by providing various aerial images as background layer. It allows the OpenStreetMap users to digitise data such as streets from the traces very easily, but for completeness of the attributes local knowledge is still required. The only problem using satellite imagery is that it can be outdated. Other web based tools are Polatch 2 and iD editor, but only registered users can upload the changes to OpenStreetMap. Figure 1 shows different methods of uploading and processing the OpenStreetMap data.

Figure 1   Methods of uploading data to OpenStreetMap server
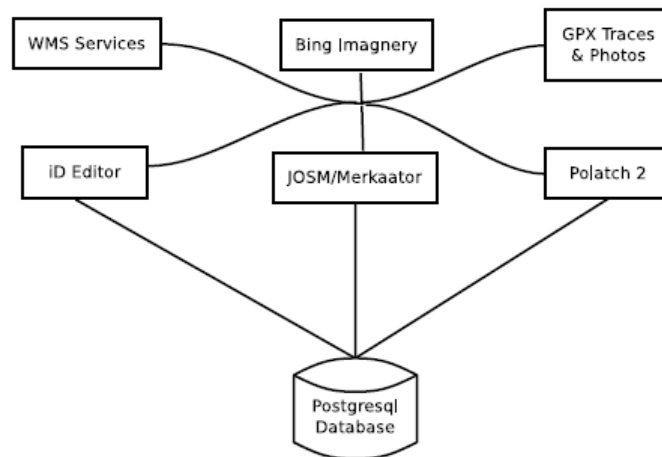
2.3   OpenStreetMap data retrieval and test map data

OpenStreetMap data of Punjab, India has been taken as study data for checking the topological errors, because the future direction of this research would be to provide the navigable data of Punjab. The full dataset is available on website Geofabrik and downloads.cloudmade.com. The format of downloaded data is in the form XML formatted .osm files.

3   Topological errors

A map is a representation of what is around us. All map features must be updated completely by maintaining the level of accuracy. The accuracy is the most important aspect for maps (Chrisman, 1991). The working group of ISO/TC211 (Wolfgang and Kian, 2003) is responsible for standardisation of map data quality parameters. The most important standards related to quality are:

- ISO 19113: Quality principles.

- ISO 19114: Quality evaluation procedures.

ISO 19113 establishes the principles for describing the quality of map data (Gong et al., 2004). According to this standard, the data quality elements identified are completeness, logical or topological consistency (Servigne et al., 2000), positional accuracy, attribute accuracy and lineage (Girres and Touya, 2010). So for grading any map data, all these elements must be carefully assessed (Neis et al., 2011). But most of these elements in case of OpenStreetMap data depend on the accuracy of GPS device used by the user during mapping and his mapping experience (Van Exel, 2012). So within the scope of this paper, logical or topological inconsistencies are considered in the OpenStreetMap data. This component is most important element for a map data to be used for navigational needs (Maraş et al., 2010).

Topological data structure is explained through graph theory and uses diagrams or graphs to describe the arrangement of geometric objects (Chen and Schneider, 2011). It is used to represent the logical relations such as neighbourhood, coincidence, inclusion, intersection in addition to metric relationship such as geometrically identifiable coordinate, length, area (Maraş et al., 2010; Laurini and Milleret-Raffort, 1994). Topological relationships between points, lines and polygons that represent the features of a geographic region provide mechanism for navigating between features (Neis et al., 2011; Maraş et al., 2010). This is also advantageous as it reduces data storage for polygons as boundaries between adjacent polygons are stored only once and enables advanced spatial analysis such as adjacency, connectivity, and containment (Clementini et al., 1993).

Inconsistencies in map data exist due to violation of predefined topology rules. The common topological errors are classified by the type of features, i.e., point, linestring (line), polygon (Egenhofer and Herring, 1992). These topological errors are detected by enforcing various rules on the map data (Chang, 2008). The possible topological errors feature-wise, except point errors are given as:
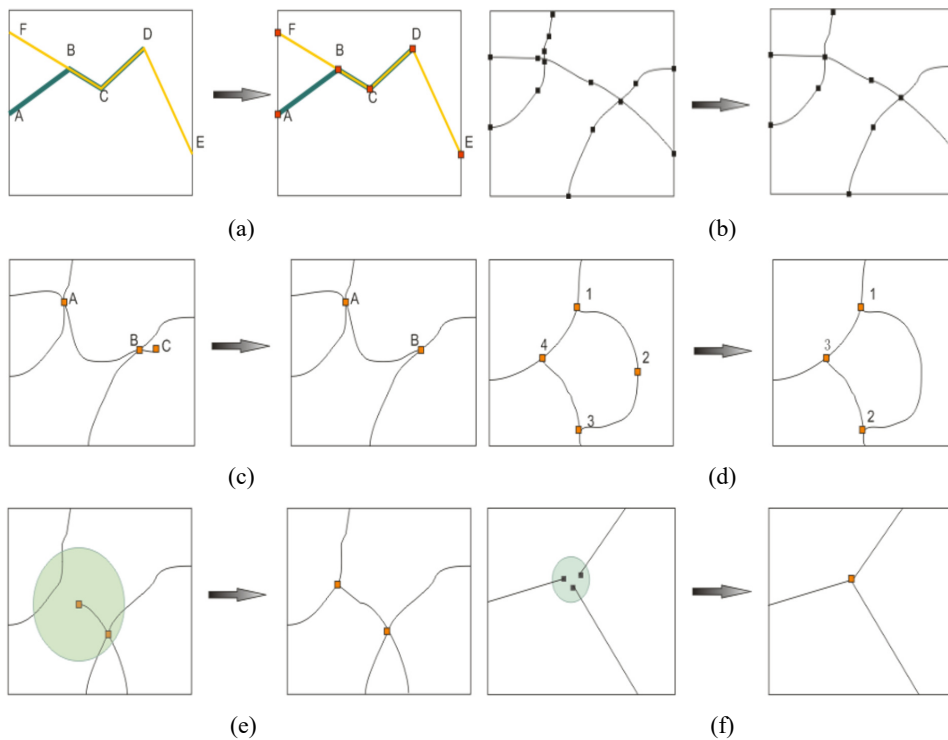
Polygon errors

- Unclosed gaps
- Gaps between polygons
- Overlapping polygons
- Self-intersection

Linestring errors

- Overshoots
- Undershoots
- Micro segments
- Pseudo nodes
- Adjacent nodes

The linestring errors and the removal of these errors are shown in Figure 2. The mapped nodes are not just points between line segments that show the directional changes in the line, in addition they have specific topological meaning. Pseudo nodes occur where a line connects itself as shown in Figure 2(d) or where two lines intersect along a parallel path rather than crossing each other (Theobald, 2001).

Figure 2   Various topological errors, (a) duplicate line removal (b) microsegments removal (c) overshoot removal (d) pseudo nodes removal (e) undershoot removal (f) merge adjacent nodes (see online version for colours)

(a)   (b)

(c)   (d)

(e)   (f)

Source: Chang (2008)

Another type of error, called dangling node or arc is the outcome of three possible mistakes: failure to make polygon closed, failure to connect the node to the object it was supposed to be connected (called an undershoot), or going beyond the object it was supposed to be connected (called an overshoot). So, lines which are supposed to meet at a node but failing to do so results in undershoot or overshoot (Siejka et al., 2013). These

nodes introduced due to error in digitisation or data conversion. Such errors reduce the data integrity and quality and have bad influence on data analysis (Maraş et al., 2010; Chen and Walter, 2009; Schmitz et al., 2008; Hope and Kealy, 2008).

Since OpenStreetMap is using topological data structure. It may contain errors (Neis et al., 2011), which occur due to quality of source map data and data collection techniques by users of varying level of mapping experience (Ubeda and Egenhofer, 1997). So, there is necessity of methods and tools to identify such errors and hence remove them. The topology rules must be enforced such as lines must not overlap (Hodgson et al., 1989) and intersection of features must be a node, e.g., the line must end as node at intersection (Cowen, 1988).

4　Methodology used

To analyse test map data for topological errors, topology rules (Servigne et al., 2000) are enforced on map data, depending on the spatial relationships that are considered as most important (Chen and Walter, 2009). The algorithms (Michaud, 2014b) used for the detection of the topological errors are executed in OpenJump (Steiniger and Hunter, 2012). OpenJump utilises Java Topology Suite (JTS) (Davis, 2007) which is a Java API that implements robust geometric algorithms and provides a complete model for specifying 2-D linear geometry.

4.1　Linestring errors

Topology extension based on JTS capabilities (Michaud, 2014b) is used to detect and remove various linestring errors from OpenStreetMap data.

4.1.1　Errors of micro-segments

The prerequisite for detection and correction of topological errors is identification and removal of micro-segments (small line segments). The presence of micro-segments less than distance threshold inside single geometry often disrupts the GIS analysis process (Michaud, 2014b). The micro-segments are closely mapped nodes at about same spot, thus creating small line segments. The micro-segment removal algorithm uses the distance parameter (threshold) to decide whether two features closed enough or not. The threshold distance taken during reported work is +/–1 metre.

As shown in Figure 3(a), three closely mapped points are detected and in Figure 3(b) those small segments (2–3, 3–4) have been removed, with minimal deformation. The algorithm will never remove the micro-segments if they are located on the geometry boundary and if both segments are strictly inside the linestring. But, it removes the segment in case interior angle is the closest to flat angle (180°). Figure 4 shows the Punjab (India) road network data and detected micro-segments in it.

Figure 3   Identification and cleaning of micro-segments, (a) micro-segments identification
            (b) cleaning of micro-segments (see online version for colours)

(a)   (b)
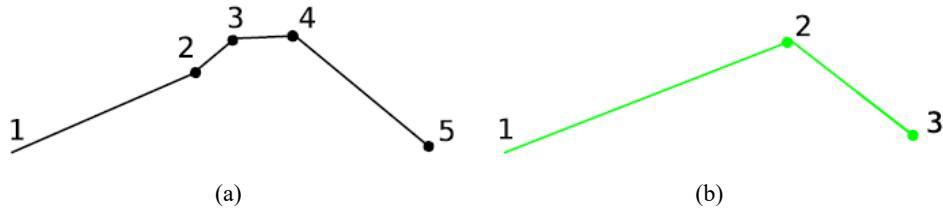
Figure 4   Outcome of micro-segment detection in Punjab map data, (a) road data of Punjab
            (b) identified micro-segments for removal (see online version for colours)

(a)   (b)
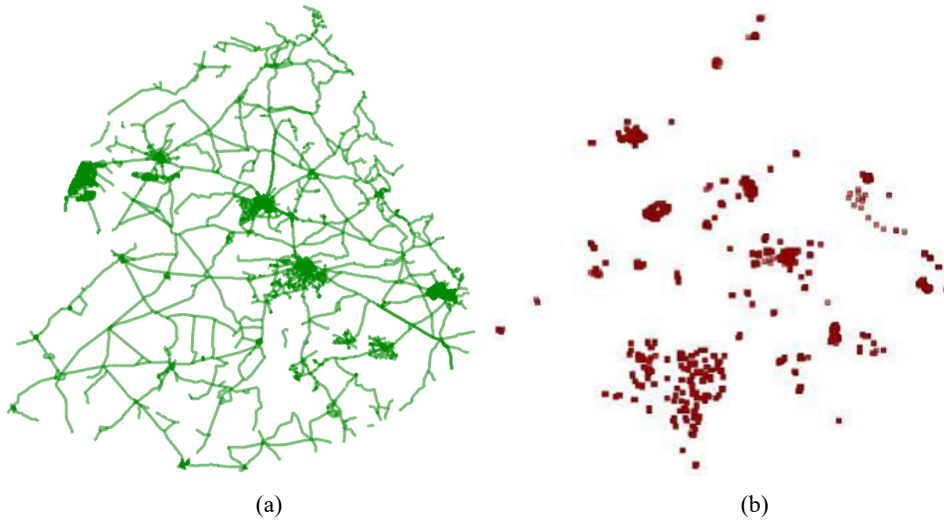
The problem arises if the segment removed by the algorithm is also used by another geometry. This removal may break further topology consistency. The solution used for this problem is that all shared nodes between multiple geometries are identified, thereafter all such segment nodes are marked (Michaud, 2014b, 2014a) and then executed the algorithm only for unmarked nodes.

The output of micro-segment removal algorithm is stored as layer in shapefile format and given as input to network topology cleaning algorithm.
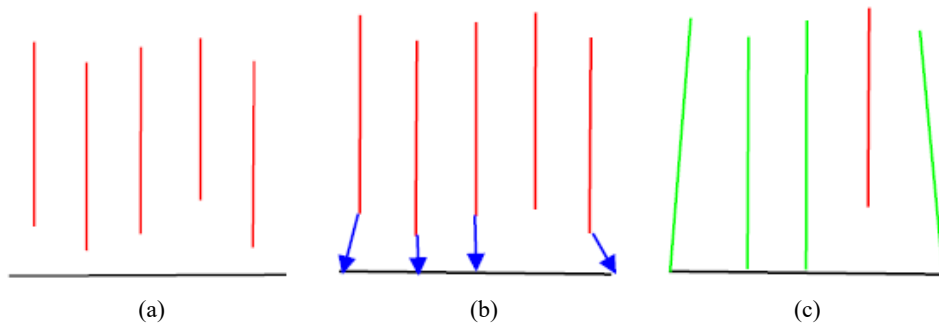
4.1.2 Dangling errors

The network topology cleaning algorithm processes dataset representing a road network, obtained after micro-segment removal and detects dangling errors like node mismatches, undershoot and overshoots (Michaud, 2014b) as shown in Figure 2. It searches nodes in network which are close enough to another linestring, but are not snapped on this linestring. The main parameters considered for snapping is distance threshold between a node and feature of the layer (Steiniger and Hunter, 2012).

The algorithm has been executed with distance threshold (tolerance) value of 3 metres. As described in Figure 5, first given feature is at less than 3 metres from a reference vertex, whereas second and third features are located at 1 metre from a reference linestring, and is also 2 metres from another one and at greater than 3 metres from any vertex. The fourth feature is at distance greater than 3 metres from the reference feature. The last feature is at less than 3 metres from a reference vertex.

Figure 5   Detection and correction of dangling errors, (a) map features (b) detected features
(c) after snapping (see online version for colours)

(a)   (b)   (c)

Figure 6   Dangling errors identified in small region of Punjab (see online version for colours)
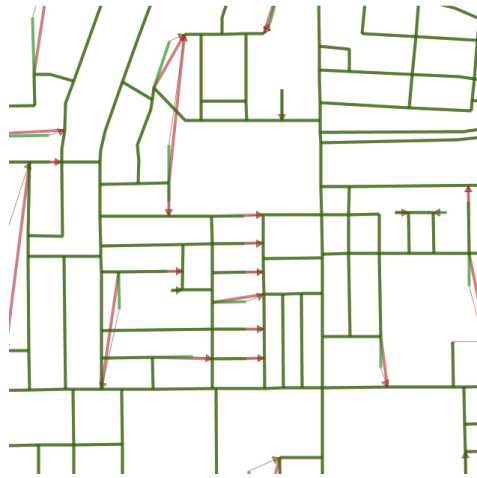

Once nodes within tolerance value have been detected, then algorithm will snap them to their right position. Figure 6 illustrates that first and last features have been properly snapped to reference, second and third features are also snapped, but a vertex needed to be inserted into the reference feature. The fourth feature is not snapped as it is at greater distance than threshold. It is demonstrated in Figure 6 that using 3 metres tolerance value, large number of nodes are detected for snapping in OpenStreetMap data of Punjab and have been snapped to the nearby feature (vertex).

4.2   Polygon errors

Polygon errors are also considered as topological errors. One very common error is self-intersection. Micro-segment removal algorithm is able to fix invalid polygons with small self-intersection as shown in Figure 7. Self-intersection is an error, but this does not pose any problem for map-render. But all tools may not have the same error tolerance. So, it is always better, to have a clean error free data structure. Nowadays, the spatial databases come with built-in geometry validation functions to correct self-intersecting polygons at the time of uploading to the database.

Figure 7   Detection of self-intersection (see online version for colours)

(a)   (b)
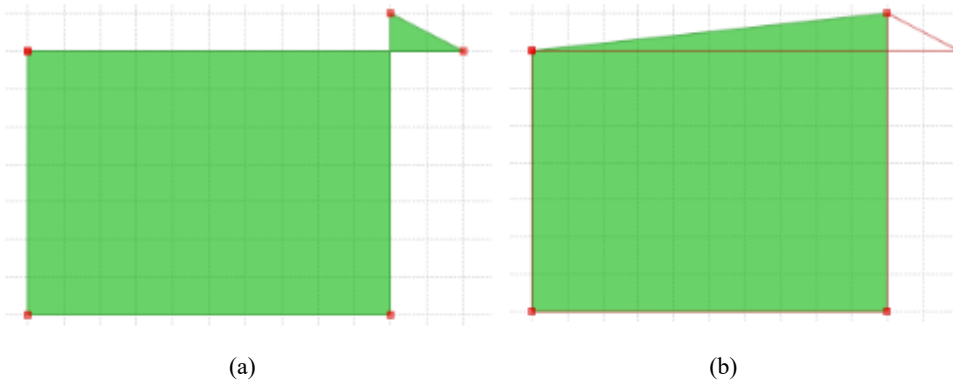
GIS map data may contain neighbouring polygons which do not share common boundary, i.e., may have spaces and holes. These may be allowed to represent the real features such as streets between different blocks. There are two types of polygon errors found in GIS data (VividSolutions.com, 2003):

1   *Gaps:* Gap is an area where two polygons are separated by too small amount along some or all the boundary as shown in Figure 8(a). There may be cases when two polygons legally contain a defined relative distance. So, spaces which have adjacent line segments separated by distance greater than tolerance are considered as gaps.

2   Overlaps: As shown in Figure 8(b), overlap is the area where two polygons overlap and is an error.

Figure 8   Detection of gaps and overlap errors, (a) gap in polygons (b) overlaps in polygon (see online version for colours)

(a)   (b)
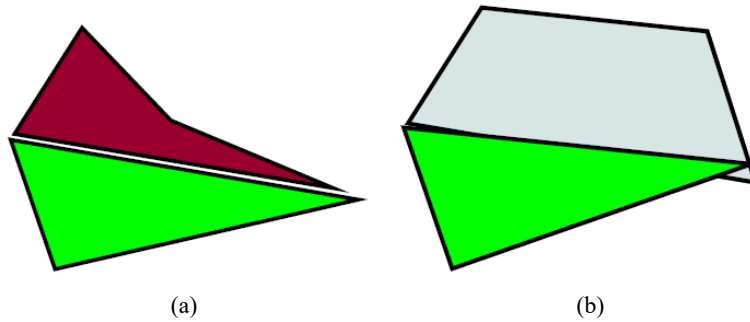
The above errors are undetectable by the naked eye. The gap detection algorithm matches each segment of polygons with segments of neighbour polygons. Segments are considered matched based on following heuristic rules:

- segments are not topologically equal (AB != CD or AB != DC)

- minimum distance between segments is less than the defined tolerance (1.0 m)

- the angle between both segments is less than the defined tolerance (22.5 degree)

- the orthogonal projection of each segment on the matched one is non null.

Figure 9 Results of coverage gaps algorithm, (a) coverage overlaps detected (b) gap segments detected (see online version for colours)

(a)   (b)
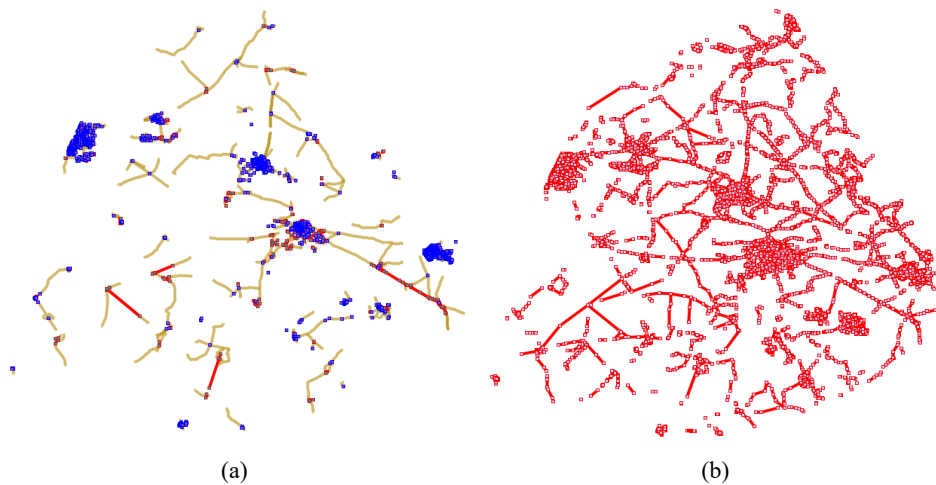
The gaps and overlaps are identified using heuristic rules and nearly 95% errors are automatically corrected (Steiniger and Hunter, 2012). The rest of 5% can be fixed manually in OpenJUMP. Figure 9 shows the gaps and overlaps in Punjab (India) map data. Still, a few errors were left unresolved that might be due to smaller distance tolerance (1 m) used. The solution is to increase the distance tolerance, but that will include the valid geometries, which may be unacceptable.

5   Results

The data has been tested by enforcing various heuristic rules. The test data comprises of 117,756 features out of which, 8911 segments are detected as redundant micro-segments. These micro-segments are removed for making the analysis by GIS analyst faster. A total of 8,492 logical errors such as undershoot, overshoot and mismatches have been identified under the constraint of tolerance value of 3 metres and are corrected by snapping them to vertex with tolerance distance. Few other rules are applied on the map data and identified total of 44,036 errors in which 39,643 are dangling errors, one duplicate node and rest are pseudo nodes. Further 20,105 gaps segments and 6,493

overlap of features are detected using tolerance distance of 1.0 metre and angle 22.5 degree.

6   Conclusions

The correct topology is a vital feature for map data to be considered for navigation as well as commercial acceptability. The emphasis of the paper is to identify the logical inconsistencies, i.e. topological errors in the Punjab OpenStreetMap data. So, the data has been processed by applying the heuristic rules on data. It has been concluded that large number of topological errors exists in the Punjab map data. The detection of these inconsistencies tells us about the amount of effort required to correct the existing data available in the public domain. Few corrections have been made to the data by enforcing rules and snapping the segments to the nearest possible vertex. But, the accuracy of such action is still questionable. Any claim for improvement can be made only after the comparison of the data with a reference dataset.

The future work will focus on removing topological errors from the OpenStreetMap data of Punjab and devising a plug-in to automatically correct the errors. Comparison of proprietary map data and governmental map data with OpenStreetMap data for topological errors will also be carried out. Further, initiatives will be taken to spread awareness about OpenStreetMap data as very few Indians are contributing and aware about it.

References

Chang, K.T. (2008) *Introduction to GIS*, McGraw-Hill [online] http://www.knu.edu.tw/lrm/English/faculty/link/Kang-Tsung%20Chang.htm (accessed 15 August 2014).

Chen, H. and Walter, V. (2009) 'Quality inspection and quality improvement of large spatial datasets', *Proceedings of the GSDI 11World Conference: Spatial Data Infrastructure Convergence: Building SDI Bridges to Address Global Challenges* [online] http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/374.pdf (accessed 17 August 2013).

Chen, T. and Schneider, M. (2011) 'The neighborhood configuration model: a framework to distinguish topological relationships between complex volumes', *Advances in Conceptual Modeling: Recent Developments and New Directions*, pp.251–260, Springer [online] http://link.springer.com/chapter/10.1007/ 978-3-642-24574-9_32 (accessed 12 August 2014).

Chrisman, N.R. (1991) 'The error component in spatial data', *Geographical Information Systems*, Vol. 1, pp.165–174 [online] http://gis.depaul.edu/ shwang/teaching/geog458/Cr1991.pdf (accessed 25 October 2014).

Clementini, E., Di Felice, P. and van Oosterom, P. (1993) 'A small set of formal topological relationships suitable for end-user interaction', *Advances in Spatial Databases*, pp.277–295, Springer [online] http://link.springer. com/chapter/10.1007/3-540-56869-7_16 (accessed 12 August 2014).

Cowen, D.J. (1988) 'GIS versus CAD versus DBMS: what are the differences?', *Photogramm. Eng. Remote Sens.*, Vol. 54, No. 11, pp.1551–1555 [online] http://www.rc.unesp.br/igce/aplicada/DIDATICOS/PAULINA/DBMS.pdf (accessed 17 August 2014).

Davis, M. (2007) 'Secrets of the JTS topology suite', *Free and Open Source Software for Geospatial* [online] http://tsusiatsoftware.net/jts/files/jts_secrets_foss4g2007.pdf (accessed 15 August 2014).

Egenhofer, M.J. and Herring, J. (1992) *Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases*, Department of Surveying Engineering, University of Maine, Orono, ME [online] http://didattica.univaq.it/moodle/file.php/1508/ 9intReport_Egenhofer_1992.pdf (accessed 28 October 2014).

Fischer, F. (2008) 'Collaborative mapping – how wikinomics is manifest in the geoinformation economy', *Geoinformtics*, Vol. 11, No. 2, pp.28–31 [online] http://www.gisaci.upol.cz/filesftp/Geoinformatics-02-2008.pdf (accessed 3 April 2013).

Girres, J.F. and Touya, G. (2010) 'Quality assessment of the French OpenStreetMap dataset', *Transaction in GIS*, Vol. 14, pp.435–459 [online] http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2010.01203.x/abstract (accessed 3 April 2013).

Golicher, D. (2013) 'Accuracy of an android cell phone GPS in the UK' [online] http://duncanjg.wordpress.com/2011/05/08/ accuracy-of-an-android-cell-phone-gps-in-the-uk/ (accessed 3 April 2013).

Gong, J., Shi, L., Du, D. and de By, R.A. (2004) 'Technologies and standards on spatial data sharing', *Proceedings of 20th ISPRS: Geo-Imagery Bridging Continents*, Vol. 34, pp.118–128 [online] http://cartesia.org/geodoc/ isprs2004/comm4/papers/327.pdf (accessed 24 October 2014).

Goodchild, M. (2009) 'NeoGeography and the nature of geographic expertise', *Journal of Location Based Services*, Vol. 3, No. 2, pp.82–96 [online] http://www.tandfonline.com/doi/abs/10.1080/17489720902950374 (accessed 3 April 2013).

Haklay, M. and Weber, P. (2008) 'OpenStreetMap: user-generated street maps', *Pervasive Computing*, Vol. 7, No. 4, pp.12–18, IEEE [online] http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4653466 (accessed 12 August 2014).

Haklay, M., Singleton, A. and Parker, C. (2008) 'Web mapping 2.0: the neogeography of the GeoWeb', *Geography Compass*, Vol. 2, No. 6, pp.2011–2039 [online] http://onlinelibrary.wiley.com/doi/10.1111/j. 1749-8198.2008.00167.x/full (accessed 8 April 2014).

Hodgson, M.E., Barrett, A.L. and Plews, R.W. (1989) 'Cartographic data capture using CAD', *Auto-carto: Proceedings of the International Conference on Computer-Assisted Cartography, American Congress on Surveying and Mapping*, Vol. 9, p.357 [online] http://mapcontext.com/autocarto/proceedings/auto-carto-9/pdf/ cartographic-data-capture-using-cad.pdf (accessed 17 August 2014).

Hope, S and Kealy, A. (2008) 'Using topological relationships to inform a data integration process', *Transactions in GIS*, Vol. 12 [online] http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2008.01098.x/abstract (accessed 25 October 2014).

Howe, J. (2006) 'The rise of crowdsourcing', *Wired Magazine*, Vol. 14, No. 6, pp.1–4, [online] http://sistemas-humano-computacionais.wikidot.com/local--files/ capitulo:redes-sociais/Howe_The_Rise_of_ Crowdsourcing.pdf (accessed 3 April 2013).

Hudson-Smith, A., Crooks, A., Gibin, M., Milton, R. and Batty, M. (2009) 'NeoGeography and Web 2.0: concepts, tools and applications', *Journal of Location Based Services*, Vol. 3, No. 2, pp.118–145 [online] http://www. tandfonline.com/doi/abs/10.1080/17489720902950366 (accessed 3 April 2013).

Kounadi, O. (2009) *Assessing the Quality of OpenStreetMap Data*, MSc dissertation, University College of London, Department of Civil, Environmental and Geomatic Engineering [online] ftp://ftp.cits.nrcan.gc.ca/pub/ cartonat/Reference/VGI/Rania_OSM_dissertation.pdf (accessed 14 November 2014).

Laurini, R. and Milleret-Raffort, F. (1994) 'Topological reorganization of inconsistent geographical databases: a step towards their certification', *Computers & Graphics*, Vol. 18, No. 6, pp.803– 813, ISSN 0097-8493 [online] doi:http://dx.doi.org/10.1016/0097-8493(94)90006-X, http://www.sciencedirect.com/science/article/pii/009784939490006X (accessed 12 August 2014).

Maraş, H., Maraş, H., Bahadir, A., Maraş, E.E. and Yildiz, F. (2010) 'Topological error correction of GIS vector data', *International Journal of Physical Sciences*, Vol. 5, No. 5, pp.476–483 [online] http://www.academicjournals.org/journal/IJPS/article-abstract/ 58118B225792 (accessed 27 November 2014).

Michaud, M. (2014a) 'Remove micro-segment (topology extension)' [online] https://mail.google.com/mail/u/1/#apps/michael+michaud+%3Cm.michael.michaud%40orang e.fr%3E/149480141e3b35a0 (accessed 7 November 2014).

Michaud, M. (2014b) *Personal Website: Topology Library* [online] http://geo.michaelm.free.fr (accessed 27 November 2013).

Neis, P. and Zipf, A. (2012) 'Analyzing the contributor activity of a volunteered geographic information project – the case of OpenStreetMap', *ISPRS International Journal of Geo-Information*, Vol. 1, No. 2, pp.146–165 [online] http://www.mdpi.com/2220-9964/1/2/146/pdf (accessed 27 January 2014).

Neis, P., Zielstra, D. and Zipf, A. (2011) 'The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011', *Future Internet*, Vol. 4, No. 1, pp.1–21 [online] http://www.mdpi.com/1999-5903/4/1/1 (accessed 3 April 2013).

O'Reilly, T. (2005) 'What is Web 2.0: design patterns and business models for the next generation of software', *O' Reilly Media*, Cambridge, USA [online] http://oreilly.com/web2/archive/what-is-web-20.html (accessed 11 April 2013).

OpenStreetMap.org (2013) *Automotive Navigation Data* [online] http://wiki.openstreetmap.org/wiki/AND_Data (accessed 13 August 2014).

OpenStreetMap.org. (2014a) *OpenStreetMap Statistics* [online] http://www.openstreetmap.org/stats/data_stats.html (accessed 13 August 2014).

OpenStreetMap.org. (2014b) *OpenStreetMap ODbL Acceptance and User Ranks for the Region of India* [online] http://odbl.de/india.html (accessed 13 August 2014).

Rana, S. and Joliveau, T. (2009) 'NeoGeography: an extension of mainstream geography for everyone made by everyone?', *Journal of Location Based Services*, Vol. 3, No. 2, pp.75–81 [online] http://www.tandfonline.com/doi/abs/10.1080/17489720903146824 (accessed 1 May 2014).

Schmitz, S., Zipf, A. and Neis, P. (2008) 'New applications based on collaborative geodata – the case of routing', *XXVIII INCA International Congress on Collaborative Mapping and Space Technology*, Gandhinagar, India [online] http://koenigstuhl.geog.uni-heidelberg.de/publications/bonn/conference/ cmap2008.cartography-bonn.subm.pdf (accessed 1 May 2014).

Sehra, S.S., Singh, J. and Rai, H.S. (2013) 'Assessment of OpenStreetMap data – a review', *International Journal of Computer Applications*, Vol. 76, No. 16, pp.17–20, ISSN 0975 – 8887, doi: 10.5120/13331-0888 [online] http://research.ijcaonline.org/volume76/number16/pxc3890888.pdf (accessed 21 October 2014).

Servigne, S., Ubeda, T., Puricelli, A. and Laurini, R. (2000) 'A methodology for spatial consistency improvement of geographic databases', *GeoInformatica*, Vol. 4, No. 1, pp.7–34 [online] http://link.springer.com/article/10.1023/A:1009824308542 (accessed 25 October 2014).

Siejka, M., Slusarski, M. and Zygmunt, M. (2013) 'Correction of topological errors in geospatial databases', *International Journal of Physical Sciences*, Vol. 8, No. 12, pp.498–507 [online] http://www.academicjournals.org/article/ article1380709682_Siejka%20et%20al.pdf (accessed 12 August 2014).

Steiniger, S. and Hunter, A.J.S. (2012) 'OpenJUMP HoRAE – a free GIS and toolbox for home-range analysis', *Wildlife Society Bulletin*, Vol. 36, No. 3, pp.600–608 [online] ISSN 1938-5463, doi: 10.1002/wsb.168 [online] http://dx.doi.org/10.1002/wsb.168 (accessed 12 August 2014).

Theobald, D.M. (2001) 'Topology revisited: representing spatial relations', *International Journal of Geographical Information Science*, Vol. 15, No. 8, pp.689–705, doi: 10.1080/13658810110074519 [online] http://dx.doi.org/10.1080/13658810110074519 (accessed 25 October 2014).

Turner, A.J. (2006) 'Introduction to neogeography', *O'Reilly Media*, MA, USA [online] http://shop.oreilly.com/product/9780596529956.do (accessed 14 November 2014).

Ubeda, T. and Egenhofer, M.J. (1997) 'Topological error correcting in GIS', in Scholl, M. and Voisard, A. (Eds.): *Advances in Spatial Databases, Lecture Notes in Computer Science*, Vol. 1262, pp.281–297, ISBN 978-3-540-63238-2, Springer, Berlin, Heidelberg [online] http://dx.doi.org/10.1007/3-540-63238-7_35 (accessed 30 September 2014).

Van Exel, M. (2012) 'Detecting highway trouble in OpenStreetMap' [online] http://oegeo.wordpress.com/2012/04/07/detecting-highway-trouble-in-openstreetmap/ (accessed 17 October 2014).

VividSolutions.com (2003) *Java Conflation Suite – Technical Report* [online] http: //www.vividsolutions.com (accessed 27 October 2014).

Walsh, J. (2008) 'The beginning and end of neogeography', *GEOconnexion International Magazine*, No. 7 [online] http://www.highbeam.com/doc/1G1-178486013.html (accessed 14 November 2014).

Wolfgang, K. and Kian, F. (2003) *ISO Standards for Geographic Information*, Springer Science & Business Media Edition, ISBN 978-3-643-05763-2, Springer [online] http://books.google.com (accessed 14 November 2014).