Analysing OpenStreetMap Data for Topological Errors

Sukhjit Singh Sehra

Assistant Professor,
Department of Computer Sci. & Engg,
Guru Nanak Dev Engineering College,
Ludhiana, Punjab, India.
E-mail: sukhjitsehra@gmail.com

Jaiteg Singh

Associate Professor,
School of Computer Science,
Chitkara University,
Punjab, India.
E-mail: jaitegkhaira@yahoo.com

Hardeep Singh Rai

Professor,
Dean, Testing & Consultancy,
Guru Nanak Dev Engineering College,
Ludhiana, Punjab, India.
E-mail: 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.

Keywords: CrowdSourcing; OpenStreetMap; Geographic Information Systems.

1 Introduction

Tim O'Reilly [1] 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 [2]. This movement provided new methods of sharing and computing data [2, 3, 4, 5, 6, 7] by crowdsourcing movement similar to Wikipedia [8]. In regard to the geographical data the crowd-sourced movement is known as VGI (Volunteered Geographic Information) and others call it collaborative mapping [9], 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 [10]. 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 error is topological error, also called logical inconsistency. In recent years some studies on quality have been undertaken [11], 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. United Kingdom, 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 USA, where rural areas are better covered and less completeness in urban areas [11].

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 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) [12]. 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 [13].

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

2 OpenStreetMap

OpenStreetMap started in 2004, is a editable and freely available database of mapping information which can be used for many different purposes with very few restrictions [14].

OpenStreetMap has three main components [15], which are Node, Way and Relation. Relation is 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 day by day and many enthusiastic users are joining this crowdsourced movement. This revolutionary movement has encouraged nearly 1,742,729 as registered users, 247,0855,510 uploaded nodes, 246,912,924 ways and 2,738,864 relations [16]. For India the total number of registered users are 2,877, 7,561,749 uploaded nodes, 433,747 ways and 6,094 relations [17]. 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 & Relations Contributed by Indian OpenStreetMap Users [17]

	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

2.2 Map Data Contribution Methods

The contribution of new data 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 [18]. 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 specialized GPS receiver for mapping, as it has been tested that smartphones can be considered as device for mapping [13, 19]. 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 [14] 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. The above fig. 1 shows different methods of uploading and processing the OpenStreetMap data.

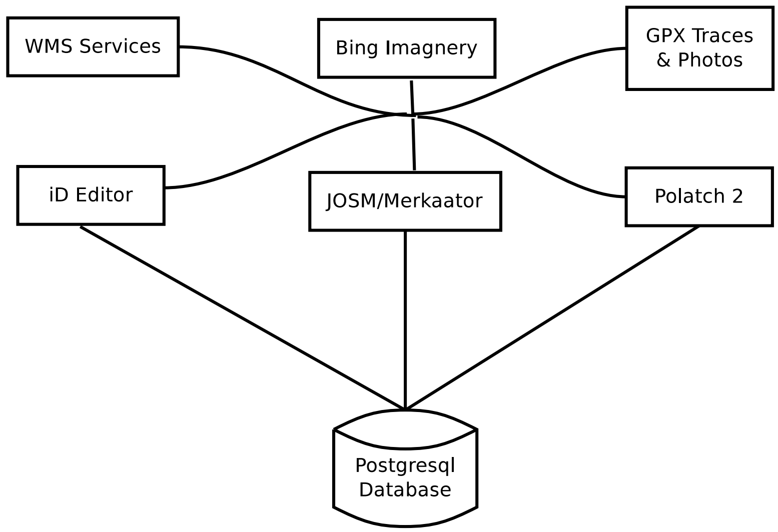


Figure 1: Methods of Uploading Data to OpenStreetMap Server

2.3 OpenStreetMap Data Retrieval & 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

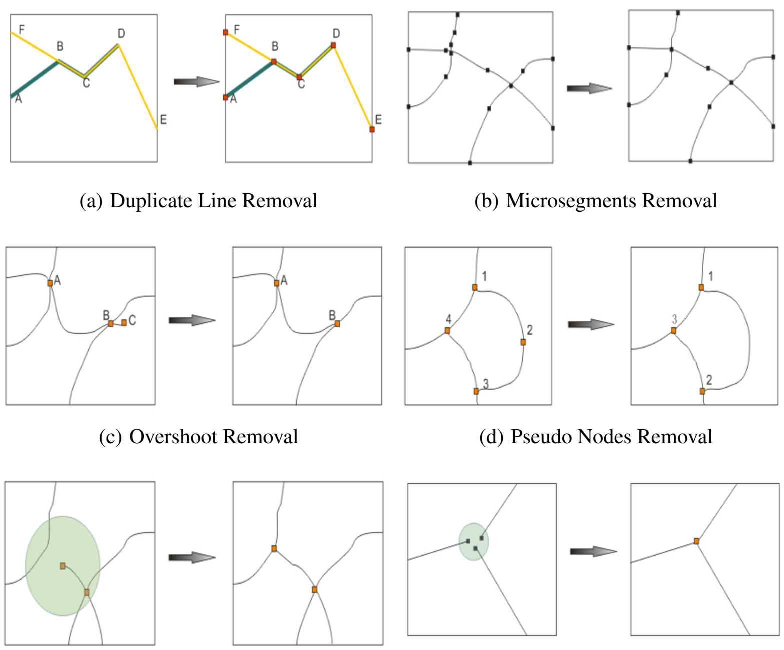
Map is representation of what is around us. All map features must be updated completely by maintaining the level of accuracy. The accuracy is most important aspect for map [20]. The working group of ISO/TC211 [21] is responsible for standardisation of map data quality parameters. Most important standards 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 [22]. According to this standard, the data quality elements identified are completeness, logical or topological consistency [23], positional accuracy, attribute accuracy and lineage [24]. So for grading any map data, all these elements must be carefully assessed [11]. 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 [25]. 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 [26].

Topological data structure is explained through graph theory and uses diagrams or graphs to describe the arrangement of geometric objects [27]. It is used to represent the logical relations such as neighborhood, coincidence, inclusion, intersection in addition to metric relationship such as geometrically identifiable coordinate, length, area [26, 28]. Topological relationships between points, lines and polygons that represent the features of a geographic region provide mechanism for navigating between features [11, 26]. 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 [29].

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 [30]. These topological errors are detected by enforcing various rules on the map data [31]. The possible topological errors feature-wise, except point errors are given as: -



(a) Duplicate Line Removal (b) Microsegments Removal

(c) Overshoot Removal (d) Pseudo Nodes Removal

(e) Undershoot Removal (f) Merge Adjacent Nodes

Figure 2: Various Topological Errors [31].

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 fig. 2. The mapped nodes are not just points between line segments that shows the directional changes in the line, in addition they have specific topological meaning. Pseudo nodes occur where a line connects itself as shown in fig. 2(d) or where two lines intersect along a parallel path rather than crossing each other [32].

Another type of error, called dangling node or arc is 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 overshoots [33]. 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 [26, 34, 35, 36].

Since OpenStreetMap is using topological data structure. it may contain errors [11], which occur due to quality of source map data and data collection techniques by users of varying level of mapping experience [37]. 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 [38] and intersection of features must be a node e.g. the line must end as node at intersection [39].

4 Methodolgy Used

To analyse test map data for topological errors, topology rules [23] are enforced on map data, depending on the spatial relationships that are considered as most important [34]. The algorithms [40] used for the detection of the topological errors are executed in OpenJump [18]. OpenJump utilises Java Topology Suite (JTS) [41] which is a Java API that implements robust geometric algorithms and provides a complete model for specifying 2-D linear geometry.

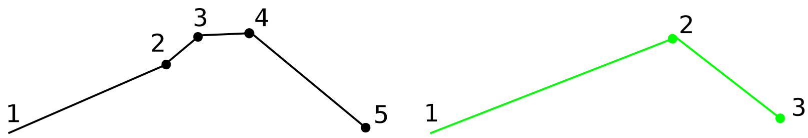
4.1 Linestring Errors

Topology extension based on JTS capabilities [40] is used to detect and removes various linestring errors from OpenStreetMap data.

4.1.1 Errors of Micro-segments

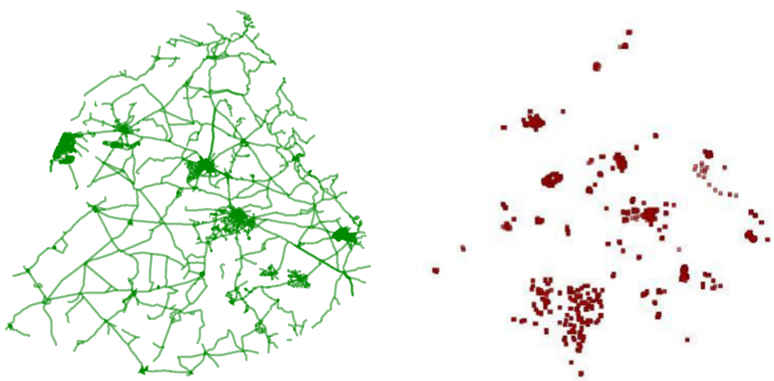
The prerequisite for detection and correction of topological errors is identification and removal of micro-segments (small line segments). The presence of micro-segments less than distance threshold inside single geometry often disrupt the GIS analysis process [40]. The micro-segments are closely mapped 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 meter.

As shown in fig. 3(a) that three closely mapped points are detected and in fig. 3(b) those small segments (2-3, 3-4) have been removed, with minimal deformation. 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°). The fig. 4 shows the Punjab (India) road network data and detected micro-segments in it.



(a) Micro-segments Identification (b) Cleaning of Micro-segments

Figure 3: Identification and Cleaning of Micro-segments



(a) Road data of Punjab (b) Identified Microsegments for removal

Figure 4: Outcome of micro-segment Detection in Punjab Map Data

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 [40, 42] 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 detect dangling errors like node mismatches, undershoot and overshoots [40] as shown in fig. 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 [18].

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

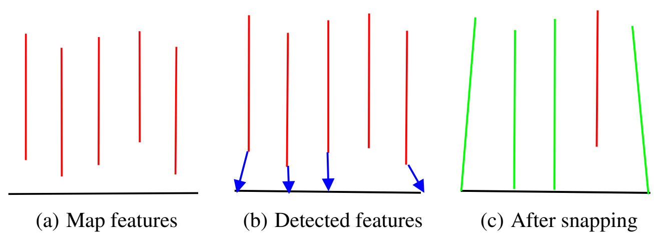


Figure 5: Detection & Correction of Dangling Errors

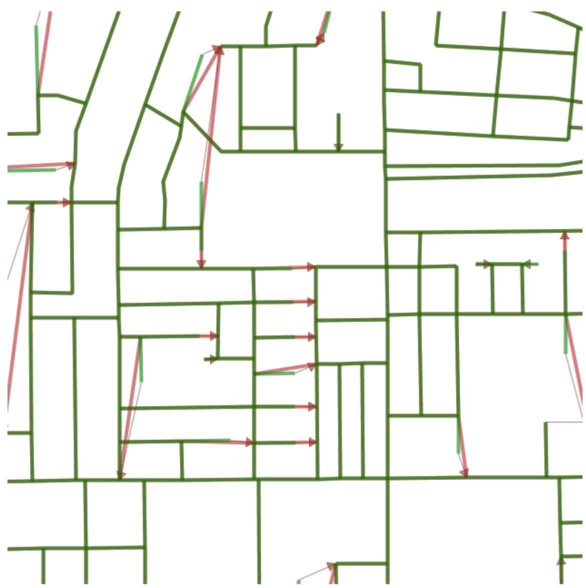


Figure 6: Dangling Errors identified in small region of Punjab

Once nodes within tolerance value has 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 fig. 6 that using 3 meter 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 of very common error is self-intersection, micro-segment removal algorithm is able to fix invalid polygons with small self-intersection as shown in fig. 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

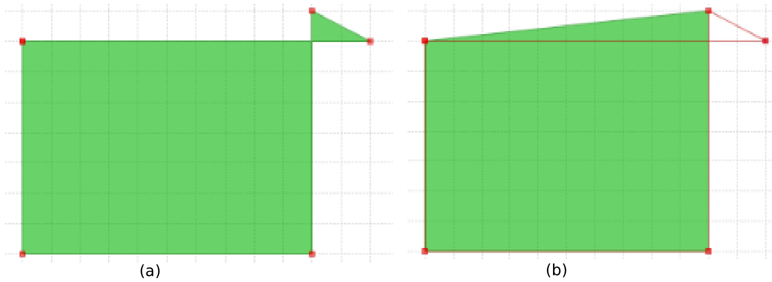


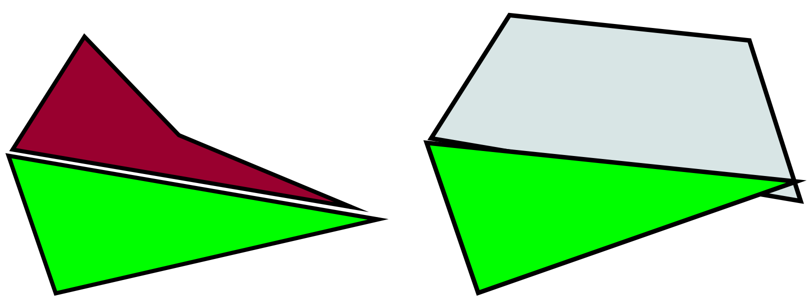
Figure 7: Detection of Self-intersection

built-in geometry validation functions to correct self-intersecting polygons at the time of uploading to the database.

GIS map data may contain neighboring 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 [43] :-

1 *Gaps*: - Gap is area where two polygons are separated by too small amount along some or all the boundary as shown in fig. 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 fig. 8(b), overlap is the area where two polygons overlap and is an error.



(a) Gap in Polygons (b) Overlaps in polygon

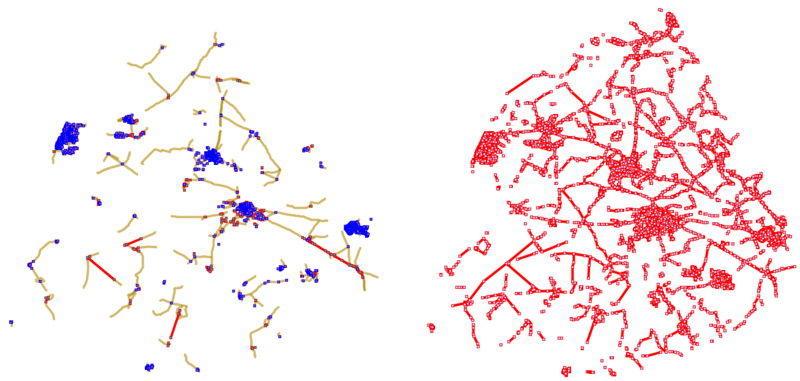
Figure 8: Detection of Gaps and Overlap Errors

The above errors are undetectable by the naked eye. The gap detection algorithm matches each segment of polygons with segments of neighbor 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 (1.0m).

- 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.



(a) Coverage Overlaps Detected (b) Gap Segments Detected

Figure 9: Results of Coverage Gaps Algorithm

The gaps and overlaps are identified using heuristic rules and nearly 95% errors are automatically corrected [18]. Rest of 5% can be fixed manually in OpenJUMP. The fig. 9 shows the gaps and overlaps in Punjab (India) map data. Still few errors were left unresolved that might be due to smaller distance tolerance (1m) used. The solution is to increase the distance tolerance, but that will include the valid geometries, which may 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 meters 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 meter and angle 22.5 degree.

6 Conclusion

The correct topology is vital feature for map data to be considered for navigation as well as commercial acceptability. The emphasis of 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 exist 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 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 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

[1] Tim O'Reilly. What is web 2.0: Design patterns and business models for the next generation of software. In *O' Reilly Media:*, Cambridge, USA, 2005. URL `http://oreilly.com/web2/archive/what-is-web-20.html`.

[2] Andrew Hudson-Smith, Andrew Crooks, Maurizio Gibin, Richard Milton, and Michael Batty. NeoGeography and web 2.0: concepts, tools and applications. *Journal of Location Based Services*, 3(2):118–145, 2009. URL `http://www.tandfonline.com/doi/abs/10.1080/17489720902950366`.

[3] Sanjay Rana and Thierry Joliveau. NeoGeography: an extension of mainstream geography for everyone made by everyone? *Journal of Location Based Services*, 3(2):75–81, 2009. URL `http://www.tandfonline.com/doi/abs/10.1080/17489720903146824`.

[4] Andrew J. Turner. Introduction to neogeography. *O'Reilly Media, MA, USA*, 2006. URL `http://shop.oreilly.com/product/9780596529956.do`.

[5] Jo Walsh. The beginning and end of neogeography. *GEOconnexion International Magzine*, (7), 2008. URL `http://www.highbeam.com/doc/1G1-178486013.html`.

[6] Michael Goodchild. NeoGeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2):82–96, 2009. URL `http://www.tandfonline.com/doi/abs/10.1080/17489720902950374`.

[7] Muki Haklay, Alex Singleton, and Chris Parker. Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass*, 2(6):2011–2039, 2008. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1749-8198.2008.00167.x/full`.

[8] Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006. URL `http://sistemas-humano-computacionais.wikidot.com/local--files/capitulo:redes-sociais/Howe_The_Rise_of_Crowdsourcing.pdf`.

[9] Florian Fischer. Collaborative mapping–How wikinomics is manifest in the geo-information economy. *Geoinformtics*, 11(2):28–31, 2008. URL `http://www.gisaci.upol.cz/filesftp/Geoinformatics-02-2008.pdf`.

[10] Ourania Kounadi. *Assessing the quality of OpenStreetMap data*. MSc dissertation, University College of London, Department of Civil, Environmental And Geomatic Engineering. URL ftp://ftp.cits.nrcan.gc.ca/pub/cartonat/Reference/VGI/Rania_OSM_dissertation.pdf.

[11] Pascal Neis, Dennis Zielstra, and Alexander Zipf. The street network evolution of crowdsourced maps: OpenStreetMap in germany 2007–2011. *Future Internet*, 4(1): 1–21, 2011. URL http://www.mdpi.com/1999-5903/4/1/1.

[12] OpenStreetMap.org. Automotive navigation data, 2013. URL http://wiki.openstreetmap.org/wiki/AND_Data.

[13] Sukhjit Singh Sehra, Jaiteg Singh, and Hardeep Singh Rai. Assessment of openstreetmap data - a review. *International Journal of Computer Applications*, 76 (16):17–20, 2013. ISSN 0975 - 8887. doi: 10.5120/13331-0888. URL http://research.ijcaonline.org/volume76/number16/pxc3890888.pdf.

[14] Pascal Neis and Alexander Zipf. Analyzing the contributor activity of a volunteered geographic information project–The case of OpenStreetMap. *ISPRS International Journal of Geo-Information*, 1(2):146–165, 2012. URL http://www.mdpi.com/2220-9964/1/2/146/pdf.

[15] Mordechai Haklay and Patrick Weber. Openstreetmap: User-generated street maps. *Pervasive Computing, IEEE*, 7(4):12–18, 2008. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4653466.

[16] OpenStreetMap.org. OpenStreetMap statistics, 2014. URL http://www.openstreetmap.org/stats/data_stats.html.

[17] OpenStreetMap.org. OpenStreetMap ODbL acceptance and user ranks for the region of india, 2014. URL http://odbl.de/india.html.

[18] Stefan Steiniger and Andrew J. S. Hunter. OpenJUMP HoRAE–a free GIS and toolbox for home-range analysis. *Wildlife Society Bulletin*, 36(3):600–608, 2012. ISSN 1938-5463. doi: 10.1002/wsb.168. URL http://dx.doi.org/10.1002/wsb.168.

[19] Duncan Golicher. Accuracy of an android cell phone GPS in the UK, 2013. URL http://duncanjg.wordpress.com/2011/05/08/accuracy-of-an-android-cell-phone-gps-in-the-uk/.

[20] N. R. Chrisman. The error component in spatial data. *Geographical information systems*, 1:165–174, 1991. URL http://gis.depaul.edu/shwang/teaching/geog458/Cr1991.pdf.

[21] Kresse Wolfgang and Fadaie Kian. *ISO Standards for geographic information*. Springer, springer science & business media edition, 2003. ISBN 978-3-643-05763-2. URL books.google.com.

[22] Jianya Gong, Lite Shi, Daosheng Du, and Rolf A. de By. Technologies and standards on spatial data sharing. In *Proceedings of 20th ISPRS: Geo-imagery bridging continents*, volume 34, pages 118–128, 2004. URL http://cartesia.org/geodoc/isprs2004/comm4/papers/327.pdf.

[23] Sylvie Servigne, Thierry Ubeda, Alain Puricelli, and Robert Laurini. A methodology for spatial consistency improvement of geographic databases. *GeoInformatica*, 4(1): 7–34, 2000. URL http://link.springer.com/article/10.1023/A:1009824308542.

[24] J.F Girres and G Touya. Quality assessment of the french OpenStreetMap dataset. *Transaction in GIS*, 14:435–459, 2010. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2010.01203.x/abstract.

[25] Martin Van, Exel. Detecting highway trouble in OpenStreetMap, 2012. URL http://oegeo.wordpress.com/2012/04/07/detecting-highway-trouble-in-openstreetmap/.

[26] Süleyman Sırrı Maraş, Hakan Hadi Maraş, Aktuğ Bahadir, Erdem Emin Maraş, and Ferruh Yildiz. Topological error correction of GIS vector data. *International Journal of Physical Sciences*, 5(5):476–483, 2010. URL http://www.academicjournals.org/journal/IJPS/article-abstract/58118B225792.

[27] Tao Chen and Markus Schneider. The neighborhood configuration model: a framework to distinguish topological relationships between complex volumes. In *Advances in Conceptual Modeling. Recent Developments and New Directions*, pages 251–260. Springer, 2011. URL http://link.springer.com/chapter/10.1007/978-3-642-24574-9_32.

[28] Robert Laurini and Francoise Milleret-Raffort. Topological reorganization of inconsistent geographical databases: A step towards their certification. *Computers & Graphics*, 18(6):803 – 813, 1994. ISSN 0097-8493. doi: http://dx.doi.org/10.1016/0097-8493(94)90006-X. URL http://www.sciencedirect.com/science/article/pii/009784939490006X.

[29] Eliseo Clementini, Paolino Di Felice, and Peter van Oosterom. A small set of formal topological relationships suitable for end-user interaction. In *Advances in Spatial Databases*, pages 277–295. Springer, 1993. URL http://link.springer.com/chapter/10.1007/3-540-56869-7_16.

[30] Max J. Egenhofer and John Herring. *Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases*. Department of Surveying Engineering, University of Maine, Orono, ME, 1992. URL http://didattica.univaq.it/moodle/file.php/1508/9intReport_Egenhofer_1992.pdf.

[31] Kang Tsung Chang. *Introduction to GIS*. McGraw-Hill, 2008. URL http://www.knu.edu.tw/lrm/English/faculty/link/Kang-Tsung%20Chang.htm.

[32] David M. Theobald. Topology revisited: representing spatial relations. *International Journal of Geographical Information Science*, 15(8):689–705, 2001. doi: 10.1080/13658810110074519. URL http://dx.doi.org/10.1080/13658810110074519.

[33] M. Siejka, M. Slusarski, and M. Zygmunt. Correction of topological errors in geospatial databases. *International Journal of Physical Sciences*, 8(12): 498–507, 2013. URL http://www.academicjournals.org/article/article1380709682_Siejka%20et%20al.pdf.

[34] H Chen and V Walter. 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*, 2009. URL http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/374.pdf.

[35] Sebastian Schmitz, Alexander Zipf, and Pascal Neis. New applications based on collaborative geodata–the case of routing. In *XXVIII INCA international congress on collaborative mapping and space technology, Gandhinagar, India*, 2008. URL http://koenigstuhl.geog.uni-heidelberg.de/publications/bonn/conference/cmap2008.cartography-bonn.subm.pdf.

[36] Hope S and Kealy A. Using topological relationships to inform a data integration process. *Transactions in GIS*, 12, 2008. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2008.01098.x/abstract.

[37] Thierry Ubeda and MaxJ. Egenhofer. Topological error correcting in GIS. In Michel Scholl and Agnès Voisard, editors, *Advances in Spatial Databases*, volume 1262 of *Lecture Notes in Computer Science*, pages 281–297. Springer Berlin Heidelberg, 1997. ISBN 978-3-540-63238-2. URL http://dx.doi.org/10.1007/3-540-63238-7_35.

[38] Michael E. Hodgson, Ann L. Barrett, and Reese W. Plews. Cartographic data capture using CAD. In *Auto-carto: Proceedings of the International Conference on Computer-Assisted Cartography*, volume 9, page 357. American Congress on Surveying and Mapping, 1989. URL http://mapcontext.com/autocarto/proceedings/auto-carto-9/pdf/cartographic-data-capture-using-cad.pdf.

[39] David J. Cowen. GIS versus CAD versus DBMS: What are the differences?. *Photogramm. Eng. Remote Sens.*, 54(11):1551–1555, 1988. URL http://www.rc.unesp.br/igce/aplicada/DIDATICOS/PAULINA/DBMS.pdf.

[40] Michaël Michaud. Personal website : Topology library, 2014. URL http://geo.michaelm.free.fr.

[41] Martin Davis. Secrets of the JTS topology suite. *Free and Open Source Software for Geospatial*, 2007. URL http://tsusiatsoftware.net/jts/files/jts_secrets_foss4g2007.pdf.

[42] Michaël Michaud. Remove micro-segment (topology extension), 2014. URL https://mail.google.com/mail/u/1/#apps/michael+michaud+%3Cm.michael.michaud%40orange.fr%3E/149480141e3b35a0.

[43] VividSolutions.com. Java conflation suite - technical report, 2003. URL http://www.vividsolutions.com.