Real-time Linear Simplification under Space Constraints

Titus Tienaah
Geodesy and Geomatics Engineering,
University of New Brunswick, Fredericton, New Brunswick.
ttienaah@unb.ca

Abstract

Linear features on a map are commonly represented as polylines (e.g. road centreline, stream, trajectory, and boundary of polygonal shapes). Various research efforts in cartography and temporal GIS have focused on algorithms to reduce linear features to a subset of useful interest or critical polygonal chains. These techniques of filtering, generalization, reduction or compression are useful for rendering dense datasets, data sampling, external storage, limited memory, and transmission over a limited bandwidth. Despite advancements, efficient automatic simplification of linear features under space constraints is intractable.

Most research efforts consider simplification of linear features in isolation, this lead to topological inconsistencies. Other heuristics in maintaining topological consistency are near or above quadratic complexity. Furthermore, there is limited research in simultaneous handling of geometric, metric, and direction relations in contextual linear simplification. This research aims to solve these challenges through efficient heuristic implementation with geometric, direction and distance constraints. The main contribution of this research is to provide accessible tools for consistent reduction of linear features to their essential characteristics in real-time for pattern mining, online transmission, and variable scale representation.

Background and Relevance

There is an increasing need to manage, process, and analyse mass collection of data with significant volume, velocity, and variety - termed as “Big Data” (Brown et al. 2011; Manyika et al. 2011; Krishnan 2013). To reduce data to a subset of useful features or features of interest, massive redundant data requires some form of filtering, generalization, reduction, or compression. The process of reducing cartographic detail is often termed as map or cartographic generalization. It involves a reduction of complexity in a map, emphasizing the essential while suppressing redundancy, maintaining logical and unambiguous relations between map objects, and preserving aesthetic quality (Weibel 1997).

With the emergence of location based services, Global Positioning Systems (GPS), wireless communication networks, radio frequency identification, and mobile devices, the mass online and offline collection of geographic data such as position, trajectory, velocity, orientation, proximity, and activity has become commonplace (Madden 2012). Mobile devices have location sensors with varying levels of precision and temporal resolution (Coleman et al. 2009). Data collected may be redundant in describing shape or other semantic patterns of movement. Redundancies also exist in linear features collected at a large scale. It is beneficial to collect data once at the highest possible resolution, and then derive a consistent representation without having to collect new data at the coarser scale (Nickerson 1988). For example, creating data at a small scale of

1:50,000 from data collected at 1:10,000 is similar to creating a map from another map as input. Simplification if often used in cartographic generalization.

In cartography, simplification is only one generalization operator. Linear simplification algorithms start with a polyline L made up of two end points and an arbitrary set of vertices V. With a given criteria, L is simplified into a polyline L' by reducing the number of vertices V to V', while keeping the end points fixed. V' is a subset of V, and no further vertex locations are introduced nor displaced (Douglas and Peucker 1973, Weibel 1997). From literature review, most simplification algorithms are performed in isolation or out of context. This leads to internal geometric inconsistency and external relational conflicts (Muller 1990; Saalfeld 1999; Stefanakis 2012).

The process of simplification often leads to geometric, direction, and distance relation conflict. Figure 1 shows some relational conflicts because of unconstrained linear simplification. In Figure 1(a), the original polyline intersects the neighbouring polygon; the simplified polyline has a disjoint relation. The reverse geometric relation in Figure 1(a) occurs in Figure 1(b). In Figure 1(c) a direction conflict is introduced by simplification, the original polyline has a westward relation with the neighbouring polygon but the simplified polyline has an eastward relation. Minimum distance constraint and self-intersection conflicts are illustrated in Figures 1(d) and 1(e) respectively. Conflict resolution is a core problem of cartographic generalization and has been proven non-deterministic polynomial hard (Estkowski 1998; Saalfeld 2000; Estkowski and Mitchell 2001). There is limited research in simultaneous handling of geometric, metric, and direction relations in contextual linear simplification. Generalization has many operators. This research focuses on the simplification operator- one of the most important generalization operators (Weibel 1997). The research scope focuses on linear simplification because lines are composed of points and can compose into closed polygonal structures. Furthermore, heuristics developed for linear features can be extended to trajectory simplification for moving objects.

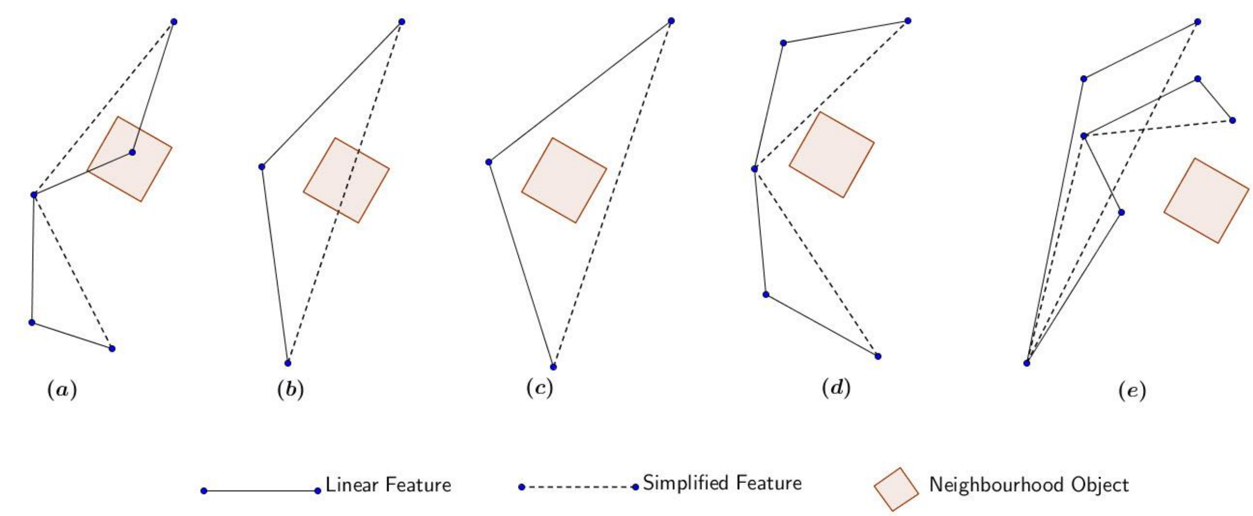

Figure 1: Relational conflicts in unconstrained linear simplification

Methods and Data

The proposed implementation of this research involves a critical review of published linear simplification techniques in research areas such as cartography, geographic information systems, computational geometry, computer graphics, computer vision, spatial data structures, mathematics (Euclidean and fractal geometry), geography, and other spatio temporal research areas (gaming and moving object databases). This review will help identify and classify the essential characteristic elements of linear features and simplification algorithms in to these classes: independent point algorithms, local algorithms, constrained extended local algorithms, unconstrained extended local algorithms, and global algorithms.

By creating a classification of characteristic elements of linear features and linear simplification algorithms, a *posteriori* empirical analysis of each algorithm will be performed and evaluated based on geometric mathematical measures, areal, and vector indices (McMaster 1987). Data to be used in testing offline and simulated on-line performance will be prepared from MarineTraffic (http://www.marinetraffic.com/ais/) and OpenStreetMap (http://www.openstreetmap.org/). Development of loosely coupled algorithms will facilitate linear simplification in cartography and temporal GIS. In addition, binary search trees, PMR quadtree and loose octree indexing will be implemented to provide efficient heuristic development (Samet, 2006; Samet et al. 2013). The implementation will also address conflicts such as geometric, direction and distance relations.

To achieve real-time efficiency (without perceptible delay), a programming environment that supports non-blocking input/output (I/O) operations is critical to this implementation. From literature review, a cross platform library libuv was identified with support for asynchronous I/O. Network platform such as Node.js (nodejs.org) and the Julia programming language (julialang.org) use libuv as the core asynchronous module. A non-blocking implementation will ensure requests are CPU bound and will facilitate high concurrency. A high-performance computing language such as Julia will handle the processing workload of the evented server implemented in JavaScript. JavaScript will be used on both the client and server using the Node.js built on Google's V8 JavaScript engine.

Results

An Initial prototype of this proposed method shows some promising results. The prototype was implemented as a binary search tree using Douglas-Peucker algorithm (Douglas and Peucker 1973). The tree structure is constructed at zero minimum offset distance to keep all vertices in the polyline. In addition, the distance offset at each vertex is stored at each node to enable variable simplification by performing a binary search. For N number of vertices and T vertices in a simplified line, this spatial structure provides a search (simplification) complexity of $O(T.\log N + C)$ for a balanced tree and $O(T.N + C)$ for degenerate cases. C is the cost of resolving relational conflicts for T vertices.

Figure 2 shows an illustration of Douglas-Peucker simplification using a binary tree constructed at zero distance offset of a polyline A, B, C,…N. The simplification requires a minimum distance constraint of 0.5 units, direction, and geometric relations. Using the binary tree representation, a normal Douglas-Peucker simplification at 5 units offset is represented by the red dash-dot line. Douglas-Peucker simplification (red dash-dot line) is done in isolation of the space constraint objects O, S, and line PQR. Circled regions show direction constraint violation at O, geometric and distance constraint violation at polyline PQR. A space-constrained simplification consistent with geometric, direction, and minimum distance offset is illustrated with the green dash line.

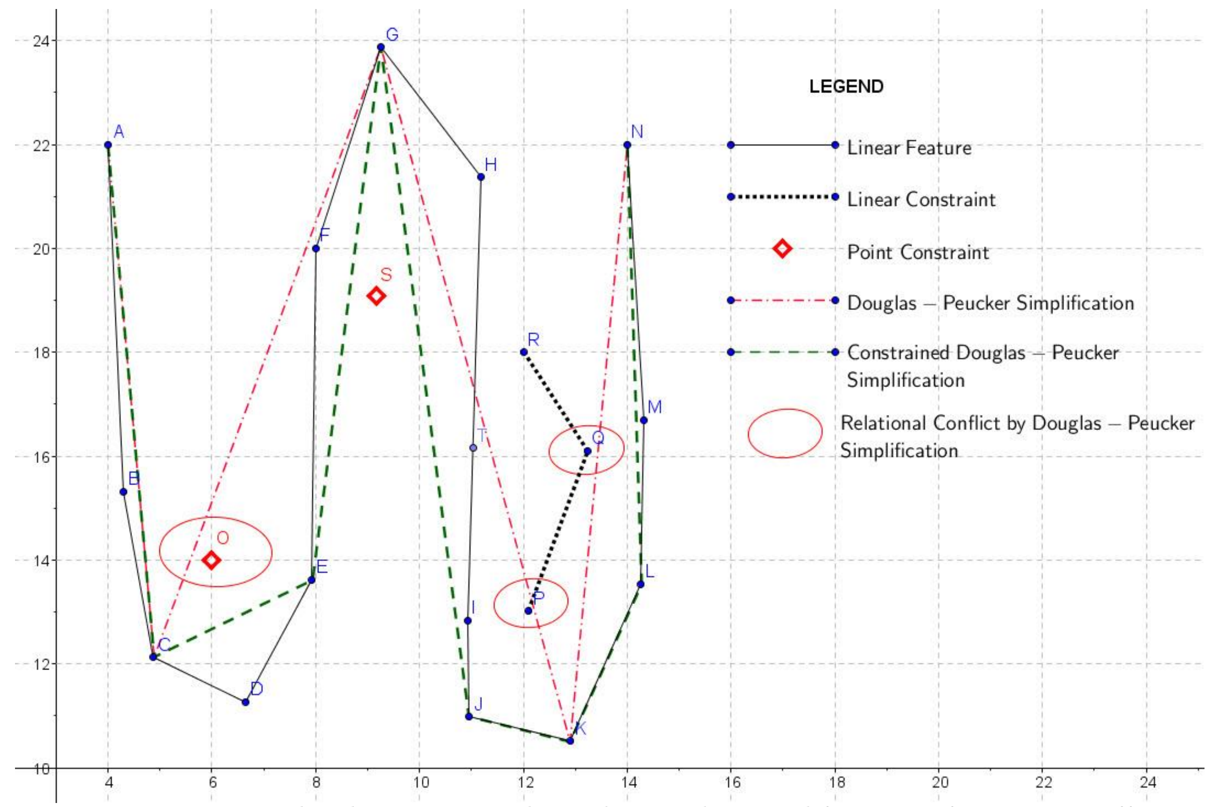

Figure 2: Constrained and unconstrained Douglas-Peucker simplification at five (5) units offset.

Conclusions and Future Work

The contribution of this research is to provide accessible tools for reducing linear features to their essential characteristics in real-time for transmission, pattern mining, storage utilisation, and variable scale representation. An initial prototype demonstrates simplification with Douglas-Peucker algorithm with geometric, metric and direction constraints. Future research will focus on indexing linear and other constraint features using quadtree structures to localize neighbourhood operations. In addition, research efforts will focus on implementing various simplification algorithms with static and moving constraints. A *posteriori* empirical analysis modelled after McMaster (1987) will be performed to compare linear simplification algorithms.

References

Brown, B., Chui, M. and Manyika, J. (2011). Are you ready for the era of 'big data'? *McKinsey Quarterly,* 4, 24-35.

Coleman,D. J., Sabone, B. and N. Nkhwanana (2009). "Volunteering Geographic Information to Authoritative Databases: Linking Contributor Motivations to Program Effectiveness." Geomatica , Vol. 64, No. 2, pp. 27-39.

Douglas D. and Peucker T. (1973), Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *The Canadian Cartographer* 10(2), 112-122.

Estkowski, R.(1998), No steiner point subdivision simplification is NP-complete. ed. *CCCG*.

Estkowski, R. and Mitchell, J. S. B., (2001). Simplifying a polygonal subdivision while keeping it simple. *Proceedings of the seventeenth annual symposium on Computational geometry*. Medford, Massachusetts, USA: ACM, 40-49.

Krishnan, K., (2013). Chapter 1 - Introduction to Big Data. *Data Warehousing in the Age of Big Data*. Boston: Morgan Kaufmann, 3-14.

Madden, S., (2012). Going big on spatial data: A mobile systems perspectiveed.^eds. *ACM SIGSPATIAL GIS 2012*, Redondo Beach, CA, USA.

Manyika, J., *et al*. (2011). Big data: The next frontier for innovation, competition, and productivity.

McMaster, R.B., (1987). The geometric properties of numerical generalization. Geographical Analysis, 19 (4), 330-346.

Muller, J. C. (1990). The Removal of Spatial Conflicts in Line Generalization. Cartography and Geographic Information Science, 17(2), 141-149.

Nickerson, B. G. (1988). Automated cartographic generalization for linear features. Cartographica: The International Journal for Geographic Information and Geovisualization, 25 (3), 15-66.

Saalfeld, A. (1999). Topologically Consistent Line Simplification with the Douglas-Peucker Algorithm. Cartography and Geographic Information Science, 26(1), 7-18.

Saalfeld, A. (2000). Complexity and Intractability: Limitations to Implementation in Analytical Cartography. *Cartography and Geographic Information Science,* 27(3), 239-250.

Samet, H., (2006). *Foundations of multidimensional and metric data structures*. Morgan Kaufmann.

Samet, H., Sankaranarayanan, J. and Auerbach, M., (2013). Indexing methods for moving object databases: games and other applications. *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. New York, New York, USA: ACM, 169-180.

Stefanakis, E., (2012). Trajectory Generalization under Space Constraints. In the Proceedings of the 7th International Conference on Geographic Information Science (GIScience 2012). Columbus, Ohio, Sept. 18-21, 2012.

Weibel, R. (1997). Generalization of spatial data: Principles and selected algorithms. *In* Kreveld, M., Nievergelt, J., Roos, T. & Widmayer, P. eds. *Algorithmic foundations of geographic information systems*. Springer Berlin Heidelberg, 99-152.