4th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Exploring the effect of Steiner points on the simplification algorithms

Amir Gholami[1], Abolfazl Ghanbari *[1], Neda Kaffash Charandabi [2]

[1] University of Tabriz, Planning & Environmental Sciences, Remote Sensing and Geographical Information System (GIS), Tabriz, Iran
[2] University of Tabriz, Marand Technical College, Geomatics Engineering, Tabriz, Iran

Keywords
Least Squares
Douglas Poker
Simplification
Steiner points
Trajectory data

Abstract

With the increasing volume of spatial data generated by a variety of spatial data recording tools such as smartphones, the importance of geometric simplification approaches has become more and more over time. The goal of geometric simplification is to achieve more summarized and less complex features. It provides an algorithm that results in terms of geometric properties such as area, perimeter, and angles being more similar to the primary feature. Algorithms with lower accuracy select consecutive subsets of primary points. As a result, some points of the geometric shape are completely ignored. While the results of methods such as least squares (LS) are more accurate in geometric simplification. Also, most geometric simplification algorithms of linear features focus on points in their processes and ignore the edges. Therefore, in this study, to improve the accuracy of geometric simplification accuracy, the effect of Steiner points on Douglas Poker (DP), LS, and a combination of them (DP-LS) was investigated. For this purpose, the trajectory recorded in Einali Mountain of Tabriz was used. The results showed that the use of Steiner points on average led to an improvement of 3.46% angle changes, 914941 m^2 area difference, 2.66% curvature similarity, and 0.36% node reduction in DP-LS and LS methods.

1. Introduction

With the advancement of technology and the equipping of smartphones with microchips of the Global Positioning System (GPS), huge volumes of data begin to be generated (Laurila et al., 2012). Many users use them in various applications such as finding places, events, restaurants and shops, and so on. The popularity of these devices has led to an increasing amount of trajectory data (Muckell et al., 2011), which can be used to determine the mobility of people, traffic network zoning, traffic detection, social anxiety, extracting interesting and scenic places for the tourism industry (Williams & Kemp, 2020). In addition, it is used in updating the roads and obtaining the boundaries of different areas and buildings (Shu et al., 2020).

Trajectory data is so large in the initial state that, according to the research findings, if trajectory constituents are collected at 10-second intervals, one gigabyte of storage capacity is needed to store more than 4,000 objects in a given day without data compression (Meratnia & Rolf, 2004). By increasing the volume of data storage, the transfer of this information becomes very costly and reduces user satisfaction and reduce their use.

Different methods of data simplification are used to reduce data volume and thus increase the speed of processing (Sun et al., 2016). One type of simplification method that can be applied to trajectories data, is a geometric simplification, which is usually applied to the geometric properties of linear and polygonal features (Ying et al., 2003). Linear and polygonal features are a set of nodes that are connected to each other by the edges that enable each connection between the two nodes, respectively.

During the process of geometric simplification of these features, the number of nodes decreases and therefore the volume of data is reduced, which can be in two main ways, including selecting nodes from the primary nodes of the features or producing nodes with coordinates different from the original shapes (Renjian et al., 2009).

* Corresponding Author

(amirp5576@gmail.com) ORCID ID 0000 – 0001 – 7749 – 3534
*(a_ghanbari@tabrizu.ac.ir) ORCID ID 0000 – 0001 – 6225 – 0433
(n_kaffash@tabrizu.ac.ir) ORCID ID 0000-0002-6281-9009

Cite this study

Gholami, A., Ghanbari, A., & Charandabi, N. K. (2022). Exploring the effect of Steiner points on the simplification algorithms. 4th Intercontinental Geoinformation Days (IGD), 314-318, Tabriz, Iran

Different methods are used to reduce the volume of data with different accuracy. Accuracy refers to the degree of proximity of the simplified feature to the main feature in the real world, which can be examined from various perspectives such as geometric similarity, semantic, and spatial relationships (Wang et al., 2015).

It is important that most geometric simplification algorithms try to select a subset of the primary nodes (Song & Miao, 2016) so the final geometry of the algorithm output does not have any effect from the deleted nodes. However, a number of algorithms, implemented mainly on the basis of least squares (LS), examined the effect of all primary nodes on the final geometry of the simplified feature. Thus, more accurate estimates of the original geometry of the features have been obtained (Tong et al., 2015). However, it can be seen that these algorithms in the processing are mostly focused on nodes whereas edges are also a major part of the features.

Therefore, in this study, in addition to nodes, edges were also considered and this was done with the help of Steiner points. Steiner points are known as points that are not part of the primary feature. These points could be added to solving a geometric optimization problem, to create a better solution. For example, Steiner points have been used to construct triangulations with better angles and total line length. In the context of feature simplification, the offered Steiner points can allow for less displacement than popular algorithms that do not have Steiner points (Kronenfeld et al., 2020).

In this study, the least squares (LS), Douglas Poker (DP), and the least squares and Douglas Poker (DP-LS) algorithms were compared and the effect of Steiner points on the simplification process was evaluated.

LS method is the most well-known regression analysis technique. It is utilized to solve problems in which the quantity of observations is more than the quantity of unknowns. One of the most important applications of the least squares process is fitting lines and curves to points, so it was used in this study to fit a line to each of the features' segments (Ghilani & Wolf, 2006).

The main goal of DP method is to select and maintain a few mainline points. In other words, the deleted points are completely ignored in DP and have no effect on the simplified features (Douglas & Peucker, 1973).

The DP-LS simplification method is based on DP and line-fitting algorithm, and the area and length of the feature lines are kept constant before and after feature simplification by using LS method. In this model, first, the nodes are selected by DP, then the points between them (unselected points) are not deleted. Rather, lines with the LS are fitted to these points. By doing this, the effect of the arcs is more than the DP method in the simplified features. According to the geometry features of these lines, three condition equations are calculated including area, length, end-vertexes identical, and fixed-point condition equation. The results of the research show that the DP-LS model is feasible to ensure the data quality in the simplification process (Xiaohua & Gusheng, 2004).

To implement the idea of this research, the trajectories recorded in the Einali Mountain (in north of Tabriz, Iran), which were long and complex trajectories.

2. Method

To evaluate the proposed model of this research, two trajectories were selected from the Einali Mountain, Tabriz, East Azarbaijan Province, Iran (Figure 1). The reason for choosing this region was the long and complex trajectories of this area.

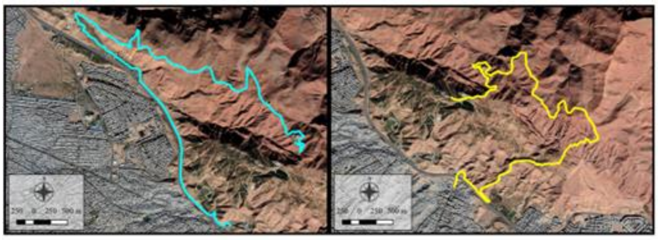

Figure 1. The study area of research, trajectory 1 (right) and trajectory 2 (left)

To investigate the effect of Steiner points on the linear features simplification, DP, DP-LS, and LS algorithms were applied to two different trajectories with three different thresholds of 1, 10, and 50 meters each. Each of the models was implemented once in the presence of Steiner points and again without them. Finally, 18 different models are obtained for each trajectory. The following describes how to create Steiner points on linear effects.

First, duplicate nodes must be removed from the trajectory. This will remove the zero edges. Steiner points are created on the edges of the trajectory at intervals of half the size of the shortest edge (Lmin). Therefore, it can be ensured that there is at least one Steiner point on each edge. If the edge length (Li) is not divisible by Lmin / 2, then the distance of the Steiner points (D_s) for that edge is calculated according to Equations (1), (2).

$$n = \left\lfloor \frac{2 \times L_i}{L_{min}} \right\rfloor \quad (1)$$

$$D_s = \frac{L_i}{n} \quad (2)$$

According to this approach, creating Steiner points is a kind of preprocessing to perform the geometric simplification of linear effects, because the algorithms themselves remain unchanged and only the inputs change. In this study, the results are evaluated using five different indicators. These indicators are described below.

1. Area difference (I_1): By connecting the start and end of linear features, a polygon is obtained whose area can be calculated. The absolute value of the area difference of a linear feature, before (A0) and after simplification (A), is used as the area difference index. The lower the value of this index, the better the simplification.

2. Reduction percentage of vertices number (I_2): Simplification of Polylines geometry is done with the aim of reducing the number of vertices and thus reducing the storage volume, and the higher the percentage of vertex reduction, the better the simplification is considered.

3. Percentage of average curvature similarity (I_3): The value of average curvature is obtained by dividing the sum of the feature geometry angles by the sum of the lengths of all its edges. The percentage similarity of the mean curvature is calculated according to Equations (3).

$$I_3 = 100 \times {|MC - MC_0|}/{MC_0} \quad (3)$$

The mean curvature of the primary feature (MC0) and simplified feature (MC) are obtained by dividing the sum of the angles by the sum of the lengths of their edges.

4. Percentage of changes in angles (I_4): This index is equal to the ratio of the sum of the angles of the simplified feature (Ang) to the sum of the angles of the primary feature (Ang_0) multiplied by 100. The higher the value of this index, the better the simplification.

5. Median Housdroff distance similarity (I_5): In this index, the minimum distance of each point of the primary polyline from the simplified polyline ($D_{p\text{-}p0}$) and vice versa ($D_{p0\text{-}p}$) is calculated. The smaller the distance, the better the simplification.

These five indicators and simplification algorithms DP, DP-LS, and LS, as well as the function required to create Steiner points, were implemented in the Python programming environment using the QGIS software API. In the next section, the implementation results are presented.

3. Results

After applying DP, DP-LS, and LS algorithms, simplified trajectory outputs were generated for different models. For example, trajectory 1 output with a threshold of 10 meters and trajectory 2 output with a threshold of 50 meters are shown in Figure 2 and Figure 3, respectively.

The similarity of all simplified shapes is that the distance of any point from them to the original feature is not more than the specified threshold. Therefore, by considering the threshold for all three algorithms equally, their performance can be evaluated by indicators. The evaluation results based on the five mentioned indicators are shown in Table 1.

According to Table 1, the values of the five evaluation indicators for the DP algorithm did not change before and after the use of Steiner points. Therefore, it can be concluded that the use of Steiner points has no effect on the performance of the DP algorithm. As mentioned, the DP algorithm only uses distance to simplification. In each iteration, it selects and holds the point farthest from the line connecting the other two points. For each edge, the start or end point is always selected and the Steiner points between them are eliminated. In the DP method, the summarized nodes are always a subset of the primary nodes, so the I_5 index for this method will always be zero.

Since the DP-LS method uses the DP method at the start of the operation, the reduction percentage of its points is similar to the DP method, according to Table 1. Therefore, whether or not Steiner points are used, combining LS with DP has no effect on improving results in terms of reducing the number of points. But according to other evaluation indicators, the use of Steiner points has improved the performance of the DP-LS algorithm. LS algorithm implemented based on least squares. The results show that the use of Steiner points in simplification with LS algorithm, in most cases, has improved the performance of this algorithm. This result can be inferred based on all 5 implemented indicators.

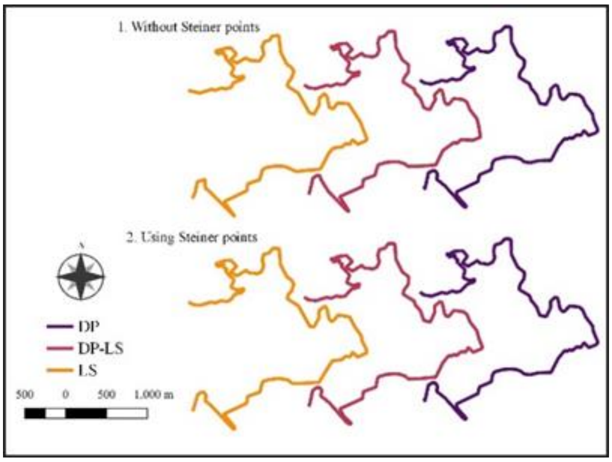

Figure 2. DP, DP-LS, and LS results for trajectory 1

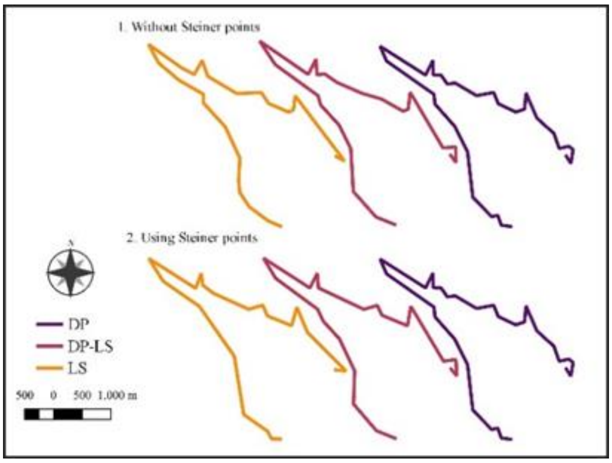

Figure 3. DP, DP-LS, and LS results for trajectory 2

Each Polyline consists of a set of points and edges. Edge lengths are ignored when fitting-based algorithms are used. Because only the points in the fitting process are used and the length of the edges is not affected. According to the results, creating Steiner points can be a good way to apply the effect of edges on the fit of the line. Thus, the longer the edge, the more points Steiner will have along it. This will affect the line fit and the output edge of the algorithm.

4. Conclusion

In this study, the effect of Steiner points on the geometric simplification of linear features was investigated. For this purpose, three algorithms DP, DP-LS, and LS with different tolerances of 1, 10, and 50 meters were applied on two different trajectories. In each model, the influence of Steiner points was evaluated by five different indicators. Based on these results, Steiner points have no effect on DP simplification.

Steiner points also have no effect on the results obtained by the DP-LS method in terms of reducing the number of points. But the LS method has improved the simplification process about 0.72%.

The use of Steiner points based on the indices of area difference, average curvature similarity percentage, similarity percentage of changes in angles, and middle Hassdorf distance, for the DP-LS method has improved 1829830 square meters, 4.05%, 4.09%, and 5 meters, respectively. The same improvement was 53.22 m², 1.26%, 2.82%, and 0.04 m for the LS method, respectively.

In future studies, it is suggested that other methods such as the weighted least squares (taking into account the weight in proportion to the length of the edges) be used and compared.

Table 1. The results of indicators for trajectory 1 and 2 in different model

Trajectory	Algorithm	Steiner Points	Threshold	I_1	I_2	I_3	I_4	I_5
1	DP	No	1	1603676	68.00	54.72	54.27	0.00
1	DP	Yes	1	1603676	68.00	54.72	54.27	0.00
1	DPLS	No	1	1536	68.00	53.36	52.71	0.01
1	DPLS	Yes	1	141	68.00	61.92	61.60	0.01
1	LS	No	1	213	71.61	69.59	69.40	0.05
1	LS	Yes	1	164	71.81	73.56	73.00	0.02
1	DP	No	10	3769018	93.45	24.99	24.31	0.00
1	DP	Yes	10	3769018	93.45	24.99	24.31	0.00
1	DPLS	No	10	2708556	93.45	24.16	23.27	0.13
1	DPLS	Yes	10	2410490	93.45	27.44	27.10	0.11
1	LS	No	10	3161	93.65	27.82	27.00	0.15
1	LS	Yes	10	3113	94.12	28.40	28.10	0.15
1	DP	No	50	21945175	97.53	12.67	11.55	0.00
1	DP	Yes	50	21945175	97.53	12.67	11.55	0.00
1	DPLS	No	50	59911014	97.53	14.81	13.96	0.42
1	DPLS	Yes	50	53258062	97.53	14.94	14.11	0.18
1	LS	No	50	12822	98.00	15.57	14.30	0.17
1	LS	Yes	50	12773	98.20	16.89	14.30	0.08
2	DP	No	1	272269	73.23	48.57	30.61	0.00
2	DP	Yes	1	272269	73.23	48.57	30.61	0.00
2	DPLS	No	1	440	73.23	49.59	48.20	0.02
2	DPLS	Yes	1	396	73.23	59.83	55.55	0.02
2	LS	No	1	490	75.89	65.79	65.67	0.04
2	LS	Yes	1	432	78.29	65.83	65.67	0.04
2	DP	No	10	2956864	93.65	26.60	9.38	0.00
2	DP	Yes	10	2956864	93.65	26.60	9.38	0.00
2	DPLS	No	10	4525456	93.65	30.67	21.62	0.10
2	DPLS	Yes	10	1507018	93.65	32.29	25.96	0.08
2	LS	No	10	2794	94.62	31.22	49.21	0.14
2	LS	Yes	10	2736	95.40	32.36	59.59	0.08
2	DP	No	50	17134916	97.99	10.46	10.36	0.00
2	DP	Yes	50	17134916	97.99	10.46	10.36	0.00
2	DPLS	No	50	6056549	97.99	11.09	14.51	0.19
2	DPLS	Yes	50	5048459	97.99	11.54	14.51	0.17
2	LS	No	50	44751	98.38	15.46	30.22	0.11
2	LS	Yes	50	44693	98.64	15.98	32.05	0.03

References

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: the international journal for geographic information and geovisualization, 10(2), 112-122.

Ghilani, Ch. D., & Wolf, P. R. (2006). Adjustment Computations: Spatial Data Analysis, Wiely, Fourth Edition.

Kronenfeld, B. J., Stanislawski, L. V., Buttenfield, B. P., & Brockmeyer, T. (2020). Simplification of polylines by segment collapse: Minimizing areal displacement while preserving area. International Journal of Cartography, 6(1), 22-46.

Laurila J K, Gatica-Perez D, Aad I, Bornet O, Do T M T, Dousse O, ... & Miettinen, M. (2012). The mobile data challenge: Big data for mobile computing research (No. CONF).

Meratnia N & Rolf A (2004) Spatiotemporal compression techniques for moving point objects. In International Conference on Extending Database Technology (pp. 765-782). Springer, Berlin, Heidelberg.

Muckell, J., Hwang, J. H., Patil, V., Lawson, C. T., Ping, F., & Ravi, S. S. (2011). SQUISH: an online approach for GPS trajectory compression. In Proceedings of the 2nd international

conference on computing for geospatial research & applications, pp. 1-8.
Renjian, Z. H. A. I., Fang, W., Li, Z., & Lei, G. (2009). Line simplification based on geographic-feature constraint. ICC, Santiago.
Shu, J., Wang, S., Jia, X., Zhang, W., Xie, R., & Huang, H. (2020). Efficient Lane-Level Map Building via Vehicle-Based Crowdsourcing. IEEE Transactions on Intelligent Transportation Systems.
Song, J., & Miao, R. (2016). A novel evaluation approach for line simplification algorithms towards vector map visualization. ISPRS International Journal of Geo-Information, 5(12), 223.
Sun, P., Xia, S., Yuan, G., & Li, D. (2016). An overview of moving object trajectory compression algorithms. Mathematical Problems in Engineering.
Tong, X., Jin, Y., Li, L., & Ai, T. (2015). Area-preservation Simplification of Polygonal Boundaries by the Use of the Structured Total Least Squares Method with Constraints. Transactions in GIS, 19(5), 780-799.
Wang, Y., Lv, H., Chen, X., & Du, Q. (2015). A PSO-neural network-based feature matching approach in data integration. In Cartography-Maps connecting the world (pp. 189-219). Springer, Cham.
Williams, R., & Kemp, V. (2020). Principles for designing and delivering psychosocial and mental healthcare. BMJ Mil Health, 166(2), 105-110.
Xiaohua, T., & Gusheng, X. (2004). A new least squares method based line generalization in GIS. In IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium (Vol. 5, pp. 2912-2915). IEEE.