Geometric Characteristics on Road Network for Vehicle Routing With 10 Korean Cities

Da-Young Lee, Bee Geum and Hwan-Gue Cho

Department of Computer Science and Engineering, Pusan National University, Busan, Korea, (e-mail: {schematique, rmaql0106, hgcho}@pusan.ac.kr)

Abstract: The autonomous driving system is an important research topic in the fourth industrial revolution. Finding the shortest path and quickest path under various conditions to maximize the efficiency of autonomous driving cars is the most critical research topic. However, the performance of such driving algorithms heavily depends on the topological characteristics of the road network on which the routing algorithm is exploited. Therefore, prior to predicting the performance of an autonomous driving system, the characteristics analysis of the road network is required and crucial. In this paper, we analyze the topological characteristics of road maps of 10 major Korean cities. And compared lots of topological measures we proposed among 10 cities. We also compare the real road network with the Delaunay Triangulation Graph which guarantees near optimal routing paths for a given point sets in 2D plane to compute how optimally the road network constructed. Also the comparison to Delaunay Graph enable us the find the clusters and bottle necks on road network. Using these comparative measures, it is possible to develop more efficient and reliable routing algorithms for the autonomous vehicles.

Keywords: network topology, routing algorithm, Delaunay triangulation, shortest path, quickest path

1. MOTIVATION

The complexity of a general network can be measured in various ways depending on the specific characteristics that we want to capture. Here are a few possible ways to measure the complexity of a road network.

- Number of nodes and edges: For example, the ratio between the number of vertices and edges can be regarded as the “density of street” in a city.
- Degree distribution: The degree distribution of a network refers to the distribution of the number of edges incident to each node. A more complex network would have a broader degree distribution with a larger variance, indicating that some nodes have many more edges than others.
- Clustering coefficient: The clustering coefficient of a node measures the extent to which its neighbors are also connected to each other.
- Shortest path length: The shortest path length between two nodes in a network is the minimum number of edges that must be traversed to go from one node to the other. A more complex network would have a larger average shortest path length, indicating that it is harder to navigate between different parts of the network.
- Fractal dimension: The fractal dimension of a network is a measure of how its shape changes as we zoom in or out. A more complex network would have a higher fractal dimension, indicating that its shape is more irregular and self-similar at different scales.

The method presented above has been mainly used for social network analysis. However, there are several problems in applying this general abstract graph characteristic analysis method to the road network as it is. First, the road network occupies a specific location unlike an abstract graph. For example, a road network is a planar graph with no intersection of edges, unlike general graphs. Therefore, the degree of most road network graphs is limited in size. For example, most intersections have a branching factor of no more than 8. This is fundamentally different from complex networks with hub nodes of tens or hundreds of degrees. In other words, the typical complex network, social network, bio-interaction network, and road network are fundamentally different. For this reason, analysis indicators that consider geodesic characteristics different from abstract indicators used in social networks should be applied to road network analysis. From this point of view, we propose a new analysis method based on the shortest, quickest distance, and Delaunay triangulation.

2. RELATED WORK

Studies on the graph-theoretic properties of road networks has been going on for a long time. From the perspective of social network characteristics presented in Chapter 1, there are studies on road characteristics in various cities [1,8,9]. As a similar study, the study on how well the road network map reflects the actual topography is the most typical study on the analysis of road network characteristics [3]. In relation to this, there are also studies that analyzed road characteristics based on the traffic volume on the road [7,9]. Another research topic is research on psychological characteristics and preference indexes that general humans find their way, rather than simply some mathematical objective function [2,4]. From this point of view, as an interesting study, there is a study on a simple path that considers the psychological burden felt by the driver, not the shortest distance measure with vehicles [5]. And although there are not many studies that directly use the Delaunay Triangulation Graph, which was used as the main analysis tool in this study, for road network analysis, there is an interesting study that simplifies and generalizes complex road networks using this Delaunay segmentation [6]. However, no study was found that used the relative ratios of the shortest path, quickest path, and Euclidean distance presented in this study as indicators.

3. DISTANCE CHARACTERISTIC

3.1 On Distance Measures

By applying graph theory to the characteristics of road networks in each region, we can evaluate the quality of the road network. In this experiment, Dijkstra's algorithm was used to calculate the shortest path distance, quickest path distance, shortest path time, quickest path time, and hop distance, while the haversine module in Python was used to calculate drone distance. Quickest path refers to the fastest path to reach the end vertex from any starting vertex. To calculate the shortest path time and quickest path time, the equation $M = w + k(d - 2)$ was used. The total waiting time was set as M, the default waiting time at each vertex (intersection) was set as w (20 seconds), and the time multiplier for each degree of the vertex was set as k (30 seconds). If the degree of a vertex was 2 or less, k was set to 0. The vehicle speed was assumed to be 30km/h. For example, consider the graph model in Fig. 1 below. The blue texts represent the length of each edge, and the dotted lines represent the drone distance. In this example, distances are measured in 'm'. Let's say we want to travel from the starting point 's' to the end point 't'. The shortest path is [s, j, i, h, g, t] with a total distance of 110m. However, the quickest path is [s, f, e, d, t] with a total distance of 120m.

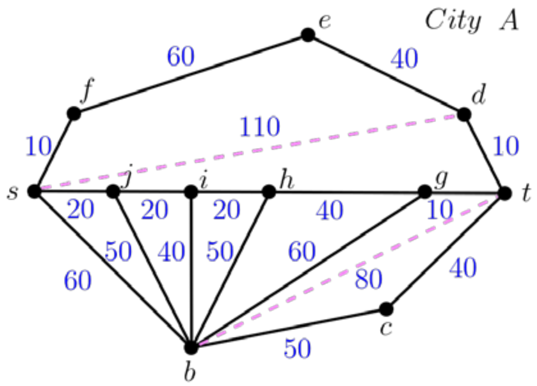

Fig. 1. {quickest, shortest, drone} distance

Here are four examples for Fig 1 listed in the table below. Note that the units for Shortest, Quickest and Drone dist are meters.

Table 1. Three Distance of City A

Pair	Shortest	Quickest	Dist
s, t	110	120	110
s, d	110	130	110
s, g	100	120	100
b, t	70	90	80

In this study, we aim to convert real city data into a graph and conduct shortest path calculations and road network evaluations. We utilized standard node-link data provided by the National Traffic Information Center for the road network data. Intersections were set as vertices and connecting roads were set as edges. The total number of vertices and edges for all ten cities are shown in the Table 2.

3.2 Experiment with 10 Korean Cities

We divided the distances for each city into SHORT (less than 500m), MID (between 500m and 1500m), and LONG (over 1500m). The Table 3 shows the percentage of cases where the ratio of shortest path distance to drone distance is greater than or equal to 1.5. A lower percentage means that the route from the starting point to the destination involves less detours on roads, while a higher percentage means that the route involves more detours.

Table 2. Vertices and edges for each city

City	Vertices	Edges
Seoul	8922	23632
Busan	5740	16402
Daegu	5577	15921
Incheon	5614	16132
Gwangju	5370	15622
Daejeon	3513	9878
Ulsan	5527	15770
Sejong	2056	4823
Jeju	4218	12075
Changwon	563	1494

Table 3. The percentage of cases where the ratio of shortest path distance to drone distance is greater than or equal to 1.5 (%)

City	SHORT	MID	LONG
Seoul	**31.4**	25.3	**1.8**
Busan	25.5	26.4	6.6
Daegu	27.3	26.4	5.0
Incheon	24.9	28.3	19.5
Gwangju	27.7	**29.2**	4.4
Daejeon	23.9	25.3	9.5
Ulsan	18.0	26.2	10.2
Sejong	23.9	21.6	**22.6**
Jeju	**14.4**	**14.2**	6.7
Changwon	17.5	19.3	8.8

For SHORT distances, Jeju Island showed the lowest percentage of values, while Seoul and Gwangju generally had higher percentages. For MID distances, Jeju Island and Sejong City showed relatively low percentages, while Gwangju and Incheon showed relatively high percentages. For LONG distances, Seoul showed an extreme low value, while Sejong City showed a high value.

Now, let's compare the time it takes to travel through the shortest path (shortest path time) and the time difference when taking the quickest path (the path that takes the least amount of time, quickest path time) in each city. The table below shows the percentage of cases where the difference between shortest path time and quickest path time is two minutes or more. The lower the result, the more paths are close to the shortest path in terms of time when traveling through it from the starting point to the destination, and conversely, the higher the result, the fewer paths are close to the shortest path in terms of time.

Now let's look at the results of both distance ratio and time difference together. In the case of LONG distances, Seoul had the best result in distance ratio, but ironically had the worst result in time difference. This means that when moving a LONG distance, Seoul can move closest to the drone distance connecting the starting point and destination, but there are many intersections to pass through.

Table 4. Shortest path time minus Quickest path time for cases where the difference is greater than 2 minutes (%)

City	SHORT	MID	LONG
Seoul	10.1	30.5	**95.7**
Busan	9.5	**25.2**	91.4
Daegu	11.5	31.2	88.2
Incheon	9.3	24.5	86.9
Gwangju	**17.8**	40.7	87.5
Daejeon	8.2	26.7	83.1
Ulsan	15.4	**41.5**	88.2
Sejong	**6.2**	26.2	67.2
Jeju	8.1	26.8	80.5
Changwon	16.8	33.2	**61.4**

Although the distance ratio of Sejong City was not the best in LONG distances, the result of time difference was the second best. Sejong City cannot move closest to the drone distance connecting the starting point and destination, but it has fewer intersections to pass through. We think the reason for this result in Sejong City is that it is a planned city.

Our graph model has straight edges, but the roads that connect all vertices in a city are not necessarily straight. Figure 2 below shows the graph model connecting all vertex pairs for drone distances (edges) of 500m or less.

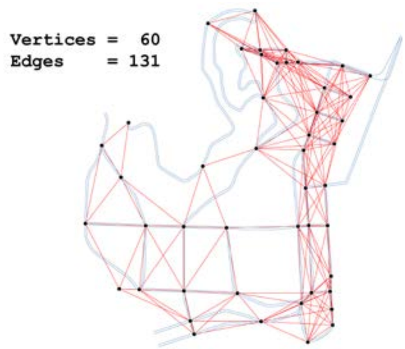

Fig. 2. City Map 2 in Busan Jung-gu, intersection(black dots), street(blue lines), drone path(red lines)

Therefore, there may be cases where an edge that directly connects two vertices is longer than the shortest path that goes through another vertex (in this case, the number of edges is 2). For example, if a direct road connecting two vertices is curvy, the distance will always be longer than that of a straight line. The number of edges where the direct path is longer than the shortest path can be used as a measure of road inefficiency.

In addition, let's compare the ratio of paths where the quickest path time divided by drone path time (time to travel along drone distance with the speed of 30km/h without any traffic) is greater than 5 for LONG distance. The Table 5-(a),(b) show the ratio of edges where the direct path is longer than the shortest path and the ratio of the quickest path time divided by the drone path time. Both tables show that Jeju had the lowest results among all the cities.

Now, let's look at the ratio of the distance traveled when the path is forced to detour due to a broken edge and the ratio of the number of edges to be crossed during the detour for each city. In this case, we measure the ratio of cases where the distance of the detour (measured as the shortest path) is larger than the direct path. We set HOP to 3 for our observations, and for a square-shaped graph, if an edge is missing between a vertex and its neighbor, the detour HOP becomes 3, and the detour distance (HOP distance) becomes 3 as well. Therefore, we set the threshold value to 3. We excluded cases where the direct path is longer than the shortest path.

Table 5. Direct Paths longer than Shortest Paths & The ratio of quickest path time to drone path time in long distances is greater than 5

City	Ratio
Seoul	0.011
Busan	0.006
Daegu	0.013
Incheon	0.017
Gwangju	**0.042**
Daejeon	0.019
Ulsan	0.032
Sejong	0.020
Jeju	**0.003**
Changwon	0.024

(a)

City	Ratio
Seoul	0.025
Busan	0.019
Daegu	**0.026**
Incheon	0.019
Gwangju	0.025
Daejeon	0.017
Ulsan	0.018
Sejong	0.023
Jeju	**0.011**
Changwon	0.019

(b)

Surprisingly, Sejong City, a planned city, had the lowest HOP ratio of 3 when a road was cut off. In addition, the Table 6 below shows the ratio of cases where the HOP distance is greater than 3 and the ratio of cases where the HOP of 10 cities is greater than 3.

Table 6. Hop distance & |Hop| > 3 (%)

City	Ratio
Seoul	**51.0**
Busan	57.5
Daegu	60.0
Incheon	60.1
Gwangju	61.3
Daejeon	60.8
Ulsan	62.5
Sejong	65.6
Jeju	58.9
Changwon	**65.7**

(a)

City	Ratio
Seoul	**46.6**
Busan	47.9
Daegu	50.0
Incheon	52.6
Gwangju	49.8
Daejeon	52.9
Ulsan	51.8
Sejong	**61.7**
Jeju	56.6
Changwon	59.1

(b)

In Table 6(a), Sejong City had the second highest proportion of HOP distance greater than 3, and in Table 6(b), it had the highest proportion of HOP greater than 3. However, in the case of Seoul, both results showed low values.

3.3 Conclusions

- As the size of a city increases (i.e., the number of vertices in the graph increases), the difference between the shortest path time and quickest path time at LONG distances also increases (left of Fig. 3). However, the HOP distance decreases (right of Fig. 3).
  - For example, in the case of Gangseo-gu in Busan Metropolitan City(Fig. 4), it had the highest number of vertices with 1,514, and the difference between the shortest path time and quickest path time at long distances was the largest.
  - In the case of Seoul(Fig. 5), which had the largest number of vertices among the ten cities with 8,922 vertices, the ratio of direct path to HOP distance was the smallest with a ratio of less than 3.

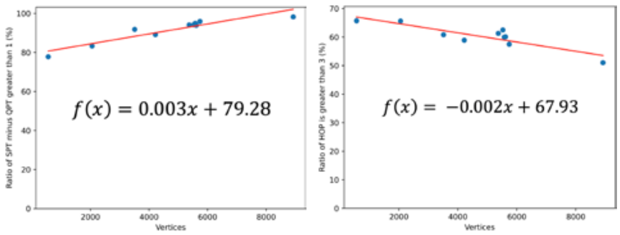

Fig. 3. Relationship between city size and time difference, relation between city size and HOP distance

Fig. 4 Gangseo-gu, intersection(red dots), street(blue lines)

Fig. 5 Seoul, intersection(red dots), street(blue lines)

- If the roads between two adjacent intersections in both planned cities (Sejong and Changwon) cannot be used, the HOP distance was the longest in Changwon and second longest in Sejong. Next, in terms of the number of HOP, Sejong had the highest number and Changwon had the second highest number.
- Excluding two planned cities, Jeju Island showed the smallest difference between shortest path time and quickest path time in long distances (refer to Table 4). In addition, in Tables 5 and Fig. 6, the roads in Jeju Island are relatively even compared to the other 10 cities, and there is little difference between the quickest path time and drone path time in LONG distances.

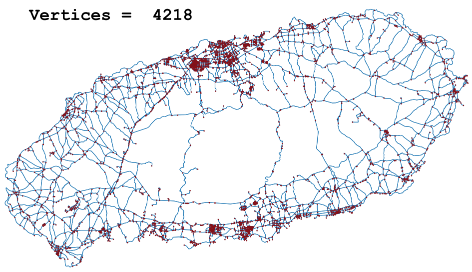

Fig. 6. Jeju, intersection(red dots), streets(blue lines)

4. DELAUNAY BASED NETWORK EVALUATION

4.1 Overview of Delaunay Traiangulation

To build a road network, constructing roads for every pair of vertices is inefficient in terms of land use and cost. On the other hand, allowing only minimal connections like trees can cause traffic congestion due to weak connectivity, which is inefficient in terms of mobility. Therefore, road network must be constructed to ensure that it has an appropriate degree of connectivity. The road network in a grid form has the advantage that the distance is similar no matter which path is chosen to reach the destination, so it is utilized when constructing planned cities. However, most cities construct roads along natural topography, as needed, based on the central area, so it is very difficult for vertex positions or edge configurations to conform to the grid form.

To evaluate the quality of a road network given, we need to check the local and the global connection structure. Delaunay triangulation (DT) is considered a good representation for the point set distribution. DT constructs a triangular mesh for a set of points in a plane such that no point in the set is inside the circumcircle of any triangle in the mesh. DT has a distinct properties. First, DT from a point set guarantees the triangulations that maximizes the minimum angle of all the angles of the triangles in the mesh. And the distance between vertices in the Delaunay triangulation network does not exceed 1.998 times the Euclidean distance. In Fig. 7-(a), for given points set P, when the circumcircle of triangle (x, y, z) is created, since there are no other points inside it, it is included in DT(P) in Fig. 7-(b). DT(P) forms a connected network with more flexible shape than grid-network and limited detour distance, it can be used as a criterion for evaluating the connectivity of a given set of points.

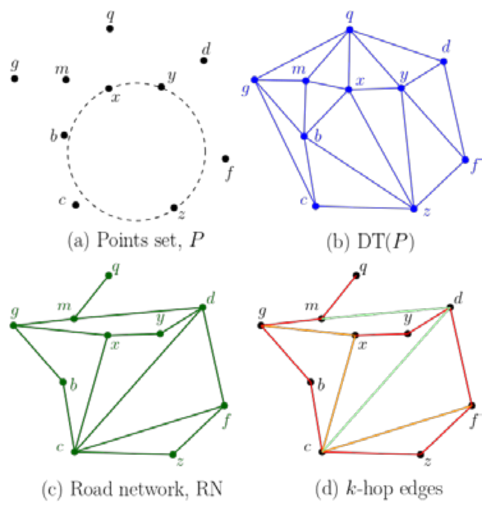

Fig. 7. Point set P (a), Delaunay triangulation DT(P) (b), Road Network RN (c), k-hop edges (d)

4.2 Delaunay Network Similarity Model

The proposed road network evaluation model calculates the edge similarity between the Delaunay network(DT) of the same vertex set as the road network RN. In Fig. 7-(c), it displays road network with same point set P. In Fig. 7-(d), it illustrates one and two hop edges between DT(P) and RN. For example, edge(x,y) exists in both DT(P) and RN, it belongs to 1-hop edges(red). For edge(c,x) in RN, it does not exist in DT(P) and must be reached through the path c-b-x, it is part of

2-hop edges(orange). In the case of edge(d,m) in RN, it is connected through the path d-y-x-m in DT(P), it corresponds to 3-hop edges(light green). Using this method, we devised a model for evaluating the road networks based on the k-hop similarity between RN and DT.

First, $Hop_G^{(k)}(x,y)$ represents whether the shortest path from node x to y on a given graph g requires k hops. In other words, if Hop(shortestPath(x,y)) = k, then $Hop_G^{(k)}(x,y)$=1; else 0. For example, if $Hop_{DT(P)}^{(1)}(x,y)$=1, it indicates that the shortest path from x to y exists as a single edge(x,y) on the DT(P). If $Hop_{DT(P)}^{(2)}(x,y)$=1, it represents that the shortest path from x to y is composed of $< x, v_1, y >$. Similarly, if $Hop_{DT(P)}^{(k)}(x,y)$ is equal to 1, it shows that the shortest path consists of the vertices in the path $< x, v_1, \ldots, v_{k-1}, y >$. Delaunay network similarity $Dsim(P,G,k)$ is defined as follows between a point set P and the network G, and a hop parameter k.

$$Dsim(P,G,k) = \sum_e^{E(RN)} Hop_{DT(P)}^{(k)} e \quad / \quad |E(G)| \qquad (1)$$

Table. 7 Dsim(P,G,k) for example (%)

G	*1*	*2*	*3*
DT(P)	42.86%	14.29%	9.52%
RN	64.29%	21.43%	14.29%

If we take the point set and network shown in Fig.6 as an example, the number of edges in DT(P) and RN is 21 and 14, respectively. For each k value of 1 to 3, there are 9, 3, and 2 hop edges, respectively. Therefore, the results of Dsim are presented in Table 7.

Table 8. Number of Edges for RN, DT

City	E(Road)	E(DT)
Sejong	2,402	5,361
Daejeon	4,988	10,531
Jeju	6,017	12,632
Gwangju	7,862	16,139
Daegu	8,133	16,757
Busan	8,281	16,835
Incheon	8,141	17,213
Seoul	13,980	26,747

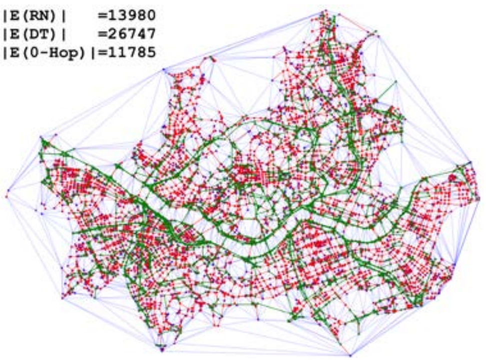

Fig. 8. RN(green) and DT(blue) and 0-HOP(red) of Seoul

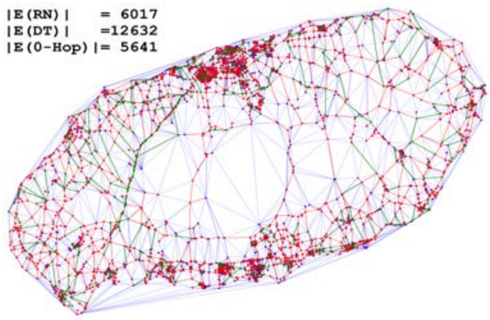

Fig. 9. RN(green) and DT(blue) and 0-HOP(red) of Jeju

4.3 Experiment – Dsim for 8 Korean Cities

First, we constructed DT for each city's road network points. Fig. 8, 9 visualize DT and RN for Seoul and Jeju, respectively. Green line segments depict the street in road network, blue edges depict Delaunay edges, and red edges shows 0-HOP edges. Bidirectional roads were merged into a single edge, resulting in Table 8.

Delaunay similarity are shown in Table 9. When k=0, based on the RN, Jeju's road network matched to DT by 93%, while Seoul matched by 84%. Conversely, based on the DT, similarity was highest for Jeju at 44.66% and lowest for Sejong at 40.57%. By increasing the hop for the Delaunay network to cover the entire road, Jeju required only 4 hops and Seoul required 23 hops. The average Delaunay hop per road being 1.07 for Jeju and 1.23 for Seoul.

Table 9. Delaunay similarity

City	Dsim (RN,0)	Dsim (DT,0)	Max	Avg.
Sejong	90.5	40.57	9	1.12
Daejeon	89.9	42.60	8	1.13
Jeju	93.8	44.66	**4**	**1.07**
Gwangju	90.8	44.24	7	1.12
Daegu	89.5	43.46	11	1.14
Busan	89.1	43.81	8	1.14
Incheon	90.6	42.85	10	1.13
Seoul	84.3	44.06	**23**	**1.23**

A noteworthy point is that the number of edges matching k-HOP follows a power law. We modeled the power function for the number of edges x using equation (2) and conducted curve fitting(Fig. 10). Table 10 shows the results of deriving the optimal p value for each city.

$$f(x) = a(b/x)^p \qquad (2)$$

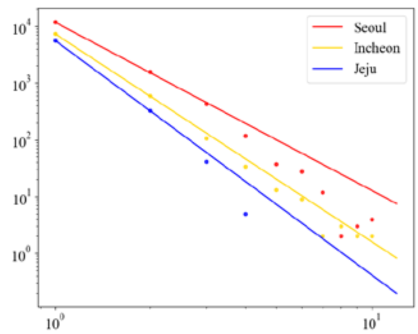

Fig. 10. log-log plot for # of k-Hop edges with k values

Table 10. p values in k-Hop edges power law

City	p	City	p
Sejong	3.59	Daegu	3.48
Daejeon	3.50	Busan	3.42
Jeju	**4.13**	Incheon	3.66
Gwangju	3.66	Seoul	**2.96**

Most cities fall within the range of 3.4~3.7, while Seoul is 2.96 and Jeju is 4.13, which is beyond the average range. This can be explained through visualization. Fig. 11 shows the road network with 0-Hop in grey, 1-3Hop in yellow, and 4 or more hops in red. Roads with 4 or more hops exactly correspond to real urban highways, while roads with 3 or fewer hops correspond to main/subsidiary trunk roads. While Jeju had the only road network without highways in dataset. From this perspective, it can be observed that as the p value decreases, the road network is designed to distribute traffic volume by differentiating the mobility hierarchy.

Fig. 11. Each DT k-Hop edges in RN of Seoul

4.4 Discussion

Based on the experimental results above, the relationship between the Delaunay network and the road network can be seen as follows:

- The main/subsidiary trunk road of cities matches the edge of Delaunay network to a high degree.
- The power law function of the Delaunay multi-hop edge represents the hierarchical structure of the road network.

Table 11. Meta information of each city

City	E(RN) /Node	Area	Node /area	Node /area-f.
Sejong	1.34	465	3.85	8.45
Daejeon	1.41	540	6.51	14.59
Jeju	1.42	1,850	2.28	4.35
Gwangju	1.46	501	10.72	17.40
Ulsan	1.43	1,062	5.20	14.73
Daegu	1.45	884	6.31	14.05
Busan	1.44	770	7.45	13.79
Incheon	1.45	1,066	5.27	8.43
Seoul	1.56	605	14.75	19.82

*Area, area-f : km^2, area-f : The area excluding forest area

Now, we analyze the city's meta-information and road network composition together. In Table 11, Jeju has 2.28 vertices per unit area of 1 km, Sejong has 3.85, which is relatively low. On the other hand, Seoul has 14.75, and Gwangju has 10.72, which are relatively high. This difference becomes even more striking when mountains are excluded from the city area.

It should be noted that in Jeju, the degree per node is similar to other cities at 1.42, and it has a road network that most closely matches the Delaunay network with $Dsim(RN, 0)$ =93.8%, $Dsim(DT, 0)$ =44.66%. Although the vertices are sparse, it can be considered as a road network with appropriate connection strength. On the other hand, Sejong has the lowest degree of 1.34, and while $Dsim(RN, 0)$ is high at 90.5%, $Dsim(DT, 0)$ is the lowest at 40.57%. Not only the vertices, but also the connections are sparse, indicating a road network with weak connection strength. Seoul has a very dense points compared to the area, and the connection between the points is also very high. Moreover, it is a road network that increases accessibility and efficiency of mobility by utilizing roads with Delaunay's multi-hop.

5. CONCLUSIONS

In this paper, we analyzed the topological characteristics of the road networks in 10 major cities in Korea from the two perspectives. The distance-based method evaluated how well the road networks matched the geographical optimal path by using directed, shortest, quickest, and drone paths. The connectivity-based method evaluated how well the connections of the road network matched those of the Delaunay network. The experiments revealed that the road networks in metropolitan cities, planned cities, and islands had distinctive features that could be distinguished from each perspective. In future research, we plan to devise a routing algorithm that considers the width and direction of roads to evaluate routing efficiency based on the characteristics of the road networks.

ACKNOWLEDGEMENT

This work was supported by the Technology development Program(S3104339) funded by the Ministry of SMEs and Startups(MSS, Korea).

REFERENCES

[1] Shang, W. L., Chen, Y., Bi, H., Zhang, H., Ma, C., & Ochieng, W. Y. (2020). Statistical characteristics and community analysis of urban road networks. Complexity, 2020, 1-21.

[2] Denis, M., Pazzaglia, *et al* (1999) "Spatial discourse and navigation," Applied Cognitive Psychology, 13(2), pp.145-174, 1999.

[3] Lyu, H., Sheng, Y., *et al* (2017) "Geometric quality assessment of trajectory-generated VGI road networks based on the symmetric arc similarity," Transactions in GIS, 21(5), pp.984-1009.

[4] Golledge, Reginald G. (1995) "Path selection and route preference in human navigation: A progress report." Spatial Information Theory A Theoretical Basis for GIS.

[5] Duckham *et al* (2003) "“Simplest” paths: automated route selection for navigation." Int. Conf.on Foundations of Geographic Information Science, 2003.

[6] Huang *et al*. (2009)"The detection and removal of conflicts in road network generalization by Delaunay triangulation." Int. Conf. on Geoinformatics. IEEE, 2009.

[7] Shepelev, Vladimir, et al. (2020) "The capacity of the road network: Data collection and statistical analysis of traffic characteristics." Energies 13.7.

[8] Zhang, Yuanyuan, *et al*. (2011) "Centrality characteristics of road network patterns of traffic analysis zones." Transportation research record pp.16-24.

[9] Zhang Q., (2005) "Road network generalization based on connection analysis." Developments in spatial data handling pp.343-353.