11th International Conference of the Hellenic Geographical Society
Lavrion Technological and Caltural Park, Lavrion, Greece April 2018

Panagiotis Agourogiannis [1], Charalampos Lepeniotis [2], Giorgos Tataris [3], Dimitris Kavroudakis [4]

[1] PhD Candidate
Department of Geography, University of the Aegean, 81100 Mytilene, geod17010@geo.aegean.gr

[2] MSc Student
MSc Geography and applied Geoinformatics, Department of Geography, University of the Aegean, 81100 Mytilene, geom17032@geo.aegean.gr

[3] PhD Candidate
Department of Geography, University of the Aegean, 81100 Mytilene, tataris@geo.aegean.gr

[4] Assistant Professor
Department of Geography, University of the Aegean, 81100 Mytilene, dimitrisk@aegean.gr

Geographical Analysis of road networks for the identification of similar routes: The case of long race athletic events

Abstract

In its present form, society has developed mechanisms, which, with proper exploitation, can provide solutions to a number of spatial problems. Geography and Geographical Information Systems (GIS), as a "spatial science", by developing the appropriate tools and the necessary decision-making methods, can face problems and analyze policies in scientific fields such as sports, tourism and culture, while promoting local development in all aspects of the economy.
The complexity of modern spatial problems often requires the fragmentation of the original problem into individual smaller chunks, which, by using geoinformatic's and spatial analysis' tools, can lead to realistic results based on the parameters set.
This paper attempts to investigate and examine the process of finding similar routes on large geographic networks. We aim to develop an automative search tools of similar network routes for different geographic networks , by evaluating network's geometric properties and setting constraints. A notable example of this work is the identification of Classical Marathon in another geographical area, by identifying routes with exactly the same geometrical characteristics.
In order to get valid and reliable results, a mathematic algorithm has been developed which uses geographical networks (categorized road networks) with optimization techniques, as well as the use of multivariable statistical methods on variables that are a result from *raster data,* such as the Digital Elevation Model (DEM).
In conclusion, this computational search process of networks, is automated with the use of R programming language and Postgresql (pgRouting, PostGIS) software.

Introduction

The use of contemporary GIS systems can help towards the solution of everyday life problems. The increased availability of spatial data along with the ease of developing complex spatial methodologies also contribute towards the introduction of such systems in day-to-day tasks.The use of spatial sciences in athletic events is somehow limited. There is not much literature about the use of spatial analysis and GIS in general regarding decision-making. Spatial networks, are a relatively new form of spatial data formation which has been used in the past in a number of scientific fields. Athletic events such as running and biking are associated with spatial networks mainly because they take place in roads and routes along known paths in a spatial network (on/off-road). This work presents the use of a novel spatial algorithm which may help athletes identify similar routes between spatial networks in order to conduct training and practicing close to their home residence. This is relatively important considering the comparative advance of athletes conducting training on similar route with the one that the target athletic event will take place. Athletes will be familiar with the geometry of the route and will have developed a strategy necessary to achieve better results without spending resources and traveling. After presenting the proposed spatial algorithm, we present some results from a case study in the islands of the North Aegean regarding a cycling athletic event.

Spatial analysis is a key concept of geographical sciences mainly because it formulates mathematical associations between spatial data and it quantifies aspects of spatial data. One of the key concepts of contemporary spatial data is networks datasets. These spatial structures are abstract associations of spatial data interconnected in a mathematical form similar to the one used in graph theory. Transportation structures such as **road networks** can also be conceptualized in the form of mathematical graphs as they are formed on the basis of nodes and vertices. Nodes could be the mathematical representation of junctions and vertices could be the mathematical conceptualization of road segments. Network analysis has been used in a number of contexts in spatial sciences' **literature**. Notable examples include the works of (Kuby et al., 2007) who investigated fuel station on spatial networks in order to identify candidate sites for siting. Also, the work of (Demšar et al., 2008) illustrate the identification of critical location in a spatial network with the use of classical Graph Theory approaches. On a more **theoretical** basis, the work of (Reggiani et al., 2009) is a very good example of quantifying the notion of complexity in spatial networks as well as the work of (Rozenblat et al., 2013) who focused on a number of methods for multilevel analysis and visualization of geographical networks. Regarding **road networks**, the work of (Okabe et al., 2012) is one of the first works which formulates a number of methodological approaches of spatial analysis along networks, transforming traditional spatial measures to this new form of discrete geographical spatial space.

Spatial networks

Spatial networks are conceptualized in the basis of the scientific field of Graph Theory as sets of Nodes and Vertices. Nodes which represent junctions of streets with single x,y coordinates forming a point in space and connecting one or more vertices. Vertices are connections between two nodes and do not have coordinates. Both nodes and vertices have attributes which describe the characteristics of this specific part of the spatial network.

A sequence of nodes and vertices on a spatial network forms a “**walk**”. This can be conceptualized as a subgraph of the network which is used for the movement from a source node to a destination node. A walk includes the aggregate characteristics of its individual parts. For example that the total length of a walk is the sum of all the vertices it includes and the total altitude range is the range of altitude information in all nodes included in the walk of the network. The number of possible walks in a network depends on the number of nodes and the amount of connections it includes. The number of walks can be potentially infinite as there are many walks that connect node A with node B. It heavily depends on the available connections and the number of steps required from node A to node B. Let A denote the adjacency matrix of a spatial network (graph), then, the **i,j** entry of the matrix. $\mathbf{A_n}$ will denote the number of walks of length n from vertex **i** to vertex **j**. So for example if we have a 3x3 adjacency matrix, with ones in every position, setting $\mathbf{A^4}$ will result to 27 possible walks of length 4. It is clear that for a fairly complex spatial network of thousands vertices and edges, there may be a very big number of possible walks of n steps.

Similarity between walks between spatial networks can be conceptualized after setting necessary similarity thresholds. These thresholds will form ranges of defining two walks as "similar" based on their geometrical or numeric attributes. The more tight the thresholds of similarity, the more difficult to identify two walks as "*similar*". Identifying similar walks on a network can be a time-consuming and computationally expensive task as it requires numerous comparisons of geometrical and numerical attributes between all pairs of possible walks in the network.

In the context of this research there is a **need** for identifying similar walks between two spatial networks. Athletic teams of mountain bikes, before joining an athletic event, conduct training in order to acquire poor performance times and stay in shape for the coming athletic event. This training is conducted mainly close to their residence as there are necessary resources (temporal/financial) for relocation to the are where the athletic event will take place. As a result, usually the teams do not train on the very same terrain where the even will take place. This lead to unfamiliarity with the geometrical features of the actual walk of the event which may have consequences on the overall performance of the team. There is a need for familiarization with the route of the athletic event without spending extra time and funds. Ideally this would also include a training field close to home-town that will be similar to the actual route of the athletic event in order to use it frequently as a training field.

Method

In order to satisfy the need described above, we have developed a spatial algorithm that identifies similar walks between two different spatial networks. This algorithm identifies similar walks in a spatial network based on geographical constrains can be formed as a sequence of steps where each step uses the result of a previous one. The sequence of necessary steps required, are the followed:

Step 1: Preparation of spatial network

The first part of the proposed methodology is to prepare a spatial network in PostGIS database using the PgRouting plug-in which enables the network analysis tools in the database. We use a spatial database as it forms a consistent and efficient way of storing and analyzing large amounts of spatial data with abilities of interconnection between known GIS software. Also, the processing time of analyzing a spatial network in PostgreSQL is considerable faster than any other known spatial file format enabling the analysis of thousand of nodes/vertices in less than a second. This step includes the disassembly of spatial data in tabular format of nodes and vertices along with the creation of an adjacency matrix in the PostgreSQL database holding connections between geometrical features. The results of this step include a number of tables and a routing table of our spatial network.

Step 2: Identification of constrains

The second step is to identify and express in a mathematical form the necessary constrains of similarity. These are mathematical expressions that are applied on the geometrical characteristics of the possible walks, resulting in a binary result (rejecting or keeping a solution). These constrain can be for example similarity on the basis of *total length* between two walks, or the *altitude difference* between first and last node. The constrains could be expressed in a **hard form** which is a very tight formulation without flexible range of accepted values or it can be expressed in a **soft form** with upper and lower acceptance bounds. The number and type of constrains, directly affects the number of solutions identified at the end of the process.

*Step 3: Transformation to an **igraph** network object*

The third step is the transformation of the spatial network to an **igraph** object which is a fietype of the igraph library necessary for mathematical computations on networks. The initial spatial network is transformed in a new format compatible with the igraph computation library in order to facilitate fast computations and ability to use of Graph Theory functions.

Step 4: Recursive Search

The fourth step of the proposed algorithm includes a recursive search approach. This search process evaluates the similarity of each walk of our spatial network with the geometrical features of the target walk. This part is computationally expensive as it requires numerous spatial calculations.

Step 5: Export of the results

The final part of the algorithm exports the accepted solutions to well-known formats such as Shapefiles.

The proposed algorithm is formulated in a way that is consistent with parallel processing operating systems as to minimize the running time and evaluate more candidate solutions in less processing time. We have used R programming language for the development of the above algorithm mainly because of its acceptance in the scientific community and its integration with GIS systems. The algorithm will be published late 2018.

Case Study

In order to illustrate the value of the proposed methodology, we have used a real world case study. The Lesvos Mountain Bike team had the need to identify a training area in Lesvos island that will have similar geometrical characteristics with an athletic event that took place in early 2018 in Chios island. This group of bikers, had the need to find a training field close to their home residence in order to train daily. This will help them to be prepared for the athletic event and help them achieve better running times. The total length of the target athletic event is 11 km has been analyzed in a GIS system in order to obtain its critical geometrical characteristics. This helped us identify and set the necessary constrains for the searching algorithm. The constrains we have developed for this athletic event along with the acceptable solution found during each step of the algorithm are the following:

1. Total length of 11km (with a deviation of +- 500m): found 63 results
2. Start-End altitude diff of 600m: found 55 results (rejected 8 solutions)
3. No hill (100m): found 51 results (rejected 4 solutions)
4. Start close to any city (1000m): found 30 results (rejected 21 solutions)
5. Start altitude 0-40: found 16 results (rejected 14 solutions)
6. Stop altitude 600m (+- 50m): found 5 results (rejected 11 solutions)

The algorithm identified 5 acceptable candidates walks in the spatial network of Lesvos island that fulfill the geometrical characteristics of the athletic event in Chios island. Figure 1 depicts a combined map the two routes (Left: target athletic events, Right: optimal solution in Lesvos island). Figure 2 and 3 depict the altitude profile diagram of the target route and the optimal solution respectively. It is important to note that constrain 4, is set from the athletes in order to facilitate the training process and is not related with the targeting event in Chios island. This illustrates the ability to set spatial queries during the search process which lead to acceptable solution, fulfilling not only similarity constrains but also convenience factors. This approach can also be used when planning an athletic event, by setting convenience constrains such as: proximity to touristic activities and preferability of specific areas (hotels, sightseeings).

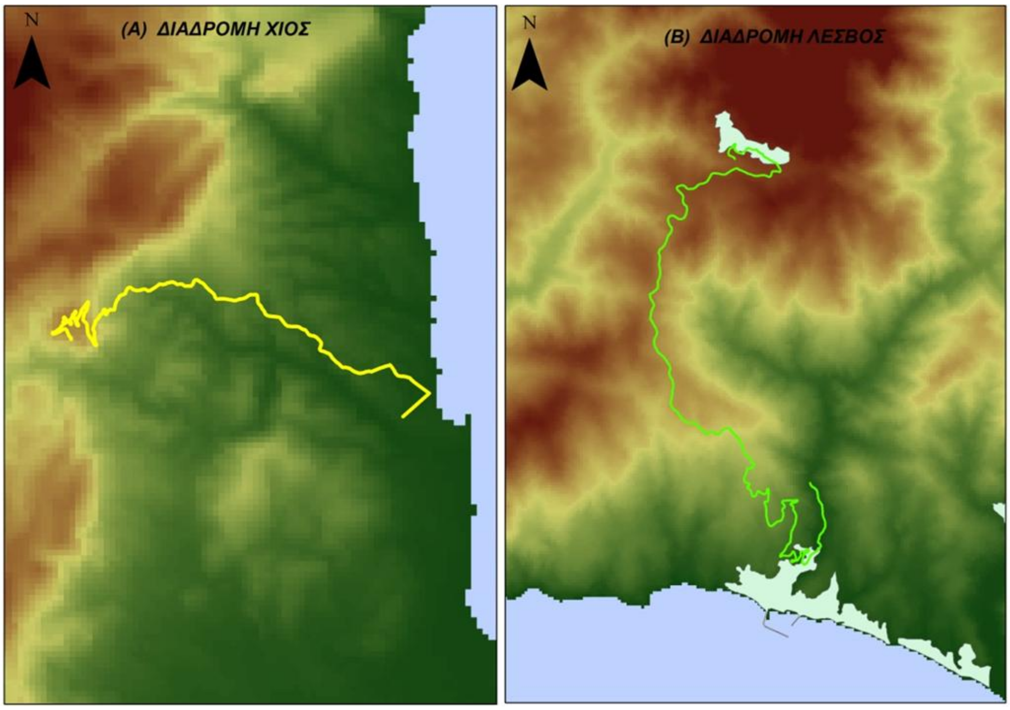

Figure 1: Comparison of initial data (left) and results (right) of searching for a bike event that took place in Chios island

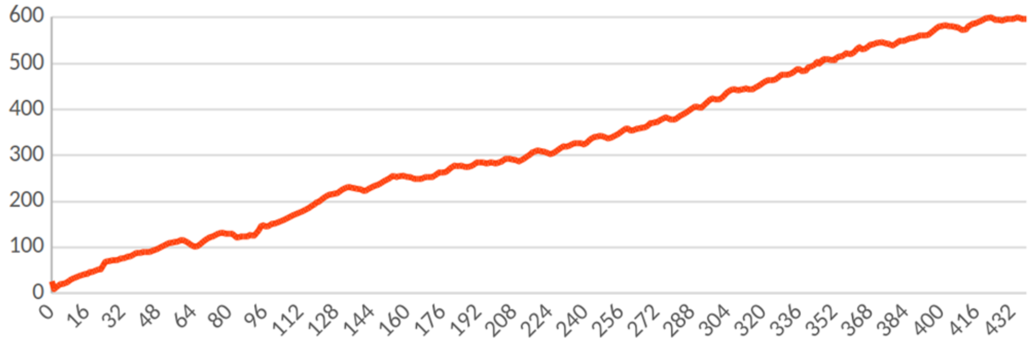

Figure 2: Profile plot of altitude of the target athletic event that took place in Chios island

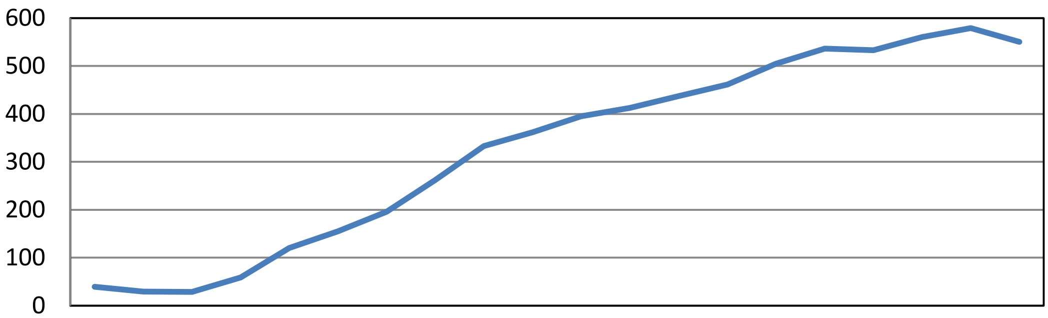

Figure 3: Profile plot of altitude of the result walk found in Lesvos island for training reasons

Concluding remarks

It is clear that the present research has a diverse character with a clear and extrovert style, as it is offered at scientific, athletic and developmental level. The development of a new mathematical algorithm through spatial analysis methods, on geographic networks, makes the science of geography a high-level scientific tool in order to prepare athletes for racing events. The analysis of the geographic characteristics of the networks, such as length, altitude, slope, etc., led to the identification of similar routes, on different geographic networks, which are not in the same area (as it is described in the case study). By identifying such a route, athletes have the ability to improve their preparation, as they are practicing on a route similar to the one they will use at the athletic event. It also reduces their transportation by saving resources, while reducing the difficulties in frequent visits to the venue. It should be noted that the creation of these routes may also contribute to local development, as it may function as a tourism development tool, since many athletes can visit the proposed area for training. A good training route which may be very similar to high level athletic events such as an Olympic Marathon, may attract athletic teams for training. This may have direct impact on local business and touristic activities in the area.

Literature

Demšar, U., Špatenková, O., Virrantaus, K., 2008. Identifying Critical Locations in a Spatial Network with Graph Theory. Transactions in GIS 12, 61–82.

Kuby, M., Lim, S., 2007. Location of Alternative-Fuel Stations Using the Flow-Refueling Location Model and Dispersion of Candidate Sites on Arcs. Networks and Spatial Economics 7, 129–152.

Okabe, A., Sugihara, K., 2012. Spatial Analysis Along Networks: Statistical and Computational Methods, 1 edition. ed. Wiley.

Reggiani, A., Nijkamp, P., 2009. Complexity and spatial networks.

Rozenblat, C., Melançon, G. (Eds.), 2013. Methods for Multilevel Analysis and Visualisation of Geographical Networks. Springer Netherlands, Dordrecht.