This is a previous version of the article published in Applied Mathematics and Computation. 2016, 291: 14-29. doi:10.1016/j.amc.2016.06.036

New highlights and a new centrality measure based on the Adapted PageRank Algorithm for urban networks ☆

Taras Agryzkov[a], Leandro Tortosa[a], Jose F. Vicent[a,*]

[a] *Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante, Campus de San Vicente, Ap. Correos 99, E–03080, Alicante, Spain*

Abstract

The Adapted PageRank Algorithm (APA) proposed by Agryzkov et al. provides us a method to establish a ranking of nodes in an urban network. We can say that it constitutes a centrality measure in urban networks, with the main characteristic that is able to consider the importance of data obtained from the urban networks in the process of computing the centrality of every node. Starting from the basic idea of this model, we modify the construction of the matrix used for the classification of the nodes in order of importance. In the APA model, the data matrix is constructed from the original idea of PageRank vector, given an equal chance to jump from one node to another, regardless of the topological distance between nodes. In the new model this idea is questioned. A new matrix with the data network is constructed so that now the data from neighbouring nodes are considered more likely than data from the nodes that are farther away. In addition, this new algorithm has the characteristic that depends on a parameter α, which allows us to decide the importance attached, in the computation of the centrality, to the topology of the network and the amount of data associated with the node. Various numerical experiments with a network of very small size are performed to test the influence of the data associated with the nodes, depending always on the choice of the parameter α. Also we check the differences between the values produced by the original APA model and the new one. Finally, these measures are applied to a real urban network, in which we perform a visual comparison of the results produced by the various measures calculated from the algorithms studied.

Keywords: Adapted PageRank algorithm; PageRank vector; networks centrality; eigenvector centrality; urban networks.

1. Introduction

One of the fundamental questions in network analysis is to determine the importance of a particular vertex (or edge) in a network. We associate the idea of relative importance of a vertex with the mathematical concept of centrality.

Over the years, network researchers have introduced a large number of centrality indices, measures that give us the importance of the vertices in a network according to a criterion [7]. These indices have proved to be of great value in the analysis and understanding of the roles played by networks like social networks [12, 23, 24], citation networks [17], computer networks [6, 16], urban networks [8, 9, 25], and others.

Depending on the type of the network studied, they are proxies for the structural importance of an element for the overall functioning of the network. Centrality is one of the most studied concepts in network analysis (see [14] for a detailed description of this concept). Numerous measures have been developed, including degree centrality, closeness [11] , betweenness [5, 20], eigenvector centrality [3], information centrality, flow

☆Partially supported by the Spanish Government, Ministerio de Economía y Competividad, grant number TIN2014-53855-P.
*Corresponding author.
Email addresses: taras.agryzkov@ua.es (Taras Agryzkov), tortosa@ua.es (Leandro Tortosa), jvicent@ua.es (Jose F. Vicent)

betweenness, the rush index, the influence measures of Katz [18], Hubbell [15], and so on. In this context, it is not often recognized that the formulas for these different measures make implicit assumptions about the manner in which things flow in a network. For instance, some measures, such as closeness and betweenness, count only geodesic or topological paths, apparently assuming that whatever flows through the network moves only along the shortest possible paths.

Some measures, such as flow betweenness [13], do not assume shortest paths, but do assume proper paths in which no node is visited more than once. Other measures, such as Bonacich ([3, 4]) eigenvector centrality and Katz influence, count walks, which assume that trajectories can not only be circuitous, but also revisit nodes and lines multiple times along the way. Regardless of trajectory, some measures (e.g. betweenness) assume that what flows from node to node is indivisible (like a package) and must take one path or another, whereas other measures (e.g. eigenvector) assume multiple paths simultaneously (like information or infections).

The city is a complex system where a large amount of information and data are generated and this is used as an essential part of the characteristics of the system itself. The source of this vast information can be very diverse: collected from fieldwork, or from Web services supported by social networks, or from existing databases (open or protected). See, for instance, [10] for a detailed study of the structure and behaviour of complex networks.

Nowadays, it is essential to have a fast and reliable classification system for the nodes of a graph to bring order to the data that we can find. As a starting point, we need to define the graph. For instance, the nodes in the graph can represent entities and the directed arcs or edges represent the links. The graph is used to analyse several characteristics of the network and collect statistics, such as the number of incoming and outgoing links of a node, calculation of a strongly connected component of the graph, number of nodes without links, mean diameter of the network and calculation of PageRank. Examples of the calculation of the PageRank is the Adapted PageRank Algorithm (see [1]), in which the authors adapted the concept of PageRank vector to urban networks. In this scenario, the Adapted PageRank Algorithm establishes a ranking of importance of the different nodes in the graph considering external factors to carry out the classification.

This paper presents a new centrality measure based on the Adapted Pagerank Algorithm with a profound modification. In the APA model, the data matrix is constructed from the original idea of PageRank vector, given an equal probability to jump from one node to another, regardless of the distance between nodes. In the new model this idea is modified. A new matrix containing all the data network is constructed so that now the data from neighbouring nodes are considered more likely than data from the nodes that are farther away.

The paper is organized as follows: Section 2 describes the Adapted PageRank Algorithm, that serves as a starting point to establish a new centrality measure. In section 3 we introduce the Adapted PageRank Algorithm Modified (APAM). Section 4 reports the numerical results taking two different networks: first, a small network with 10 nodes and, secondly, a real urban network, in order to analyse differences between the Adapted PageRank Algorithm and the Adapted PageRank Algorithm Modified. Finally, some conclusions are presented in Section 5.

2. The Adapted PageRank Algorithm (APA) and its characteristics

The *PageRank* method [21] was proposed to compute a ranking for every Web page based on the graph of the Web. Therefore, PageRank constitutes a global ranking of all Web pages, regardless of their content, based solely on their location in the Web's graph structure. The purpose of the method is obtaining a vector, called *PageRank vector*, which gives the relative importance of the pages. Since this vector is computed based on the structure of the Web connections, it is said to be independent of the request of the person performing the search. For a more detailed description of the PageRank algorithm, see [22].

In [1], Agryzkov et al. propose an adaptation of the PageRank model to establish a ranking of nodes in an urban network, taking into account the influence of external activities or information. In the following, we refer to this algorithm as the *Adapted Pagerank Algorithm* (APA algorithm). Although the APA algorithm is applied to urban networks, this is perfectly applicable to any network, whenever we want to analyse or represent additional information from the network itself, by means of a numerical assignment of data to the different nodes on the network.

The central idea behind the APA algorithm for ranking the nodes is the construction of a data matrix D, which summarizes the numerical value of the data that we are measuring. This matrix allows us to represent numerically the information of the network that we are going to analyse and measure. This information is placed in columns, where each column represents a specific characteristic or type of information that we want to evaluate or analyse.

It is also important the *weight vector* $\boldsymbol{v}_0$ because it is the key that allows us to establish the importance we assign to any of the type data that we are measuring in the network.

The algorithm proposed in [1] is summarized as:

Algorithm 1. ***APA Algorithm***. *Let us assume that we have a graph $G = (V, E)$ representing an urban network with N nodes n_i, representing squares or streets intersections. We proceed with the following steps.*

Step 1 *Obtain the transition matrix A^* from the graph of the network.*

Step 2 *Consider the different characteristics k_i associated to each of the nodes for the problem studied; evaluate them in each node. With these numerical values, we construct the matrix D.*

Step 3 *Construct a vector $\boldsymbol{v_0}$, according to the importance of each of the characteristics evaluated. This vector represents a multiplicative factor.*

Step 4 *Obtain a vector $\boldsymbol{v}$ by multiplying $D \cdot \boldsymbol{v_0} = \boldsymbol{v}$.*

Step 5 *Transform $\boldsymbol{v}$ into a probability vector $\boldsymbol{v} \longrightarrow \boldsymbol{v^*}$, using the standard method.*

Step 6 *Construct the matrix V, from $\boldsymbol{v^*}$.*

Step 7 *Construct the matrix $M_{APA} = (1-\alpha)A^* + \alpha V$, from A and V.*

Step 8 *Compute the eigenvector $\boldsymbol{x}_1$ associated to the dominant eigenvalue $\lambda_1 = 1$ for the matrix M_G. That is our ranking vector.*

Figure 1 shows an overview of the different steps and the set of mathematical operations involved in the computation of the APA centrality.

The result of applying this algorithm to a network is a ranking vector

$$\boldsymbol{x}_1 = \{x_1(1), x_1(2), \ldots, x_1(N)\}$$

with N components, where the i-th component represents the ranking of the i-th node within the overall network. This vector is the dominant eigenvector $\boldsymbol{x_1}$ associated to the dominant eigenvalue λ_1 of the matrix M_{APA}.

Note that the matrix which is initially taken in Algorithm 1 (step 1) is not exactly the adjacency matrix of the graph. It is a different matrix A^*, that we define as $A^* = (a_{ij})_{i,j=1}^{n} \in R^{n\times n}$, where

$$a_{ij} = \frac{1}{c_j}, \quad \text{if } c_j \neq 0, \tag{1}$$

with c_j, for $j = 1, 2, \ldots, n$, to be the number of connections of the vertex c_j. In other case, we have that $a_{ij} = 0$.

We can say that the mean feature of this algorithm is the construction of the data matrix D and the weight vector $\boldsymbol{v_0}$. Firstly, the matrix D allows us to represent numerically the information we want to study; secondly, the vector $\boldsymbol{v_0}$ allows us to establish the importance of each of the factors or characteristics that have been measured by means of D.

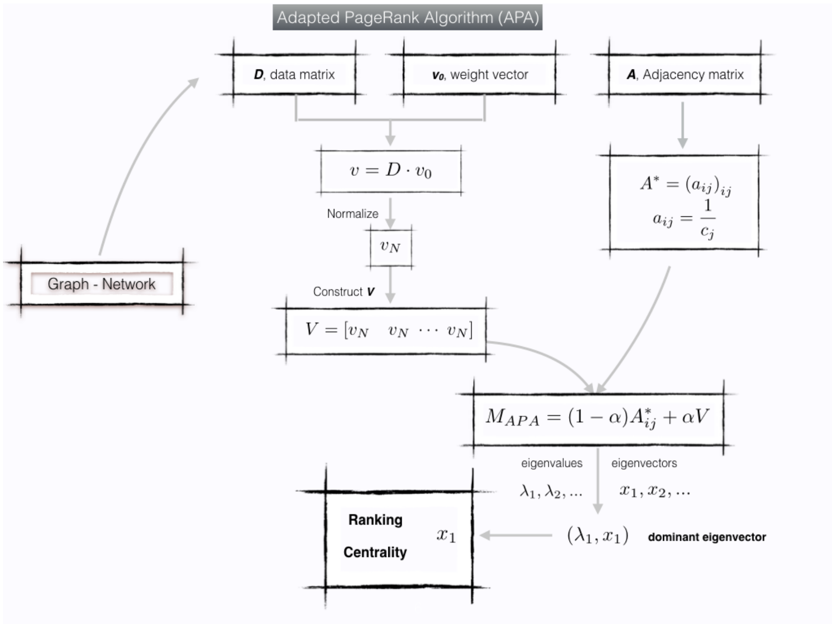

Figure 1: APA theoretical scheme.

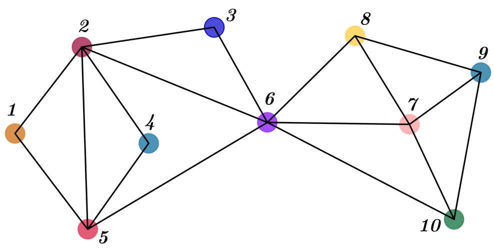

Figure 2: A simple graph with ten nodes to test the centrality measures.

As an example, if we take the graph displayed in Figure 2, we have that

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix},$$

$$A^* = \begin{bmatrix} 0 & 1/5 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 1/2 & 1/4 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/5 & 0 & 1/2 & 0 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 1/5 & 1/2 & 0 & 1/4 & 0 & 1/4 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/4 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/4 & 0 & 1/3 & 0 \end{bmatrix}.$$

Now, let us assume that the data vector $\boldsymbol{v}$ is $\boldsymbol{v} = [1;1;1;1;1;1;1;1;1;1]$, that is, all the nodes in the graph have one data associated to them.

We transform $\boldsymbol{v}$ into a probability vector dividing every component by 10 and construct V as

$$V = \begin{bmatrix} 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \\ 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 & 1/10 \end{bmatrix}.$$

From A^* and V, we construct the matrix M_{APA} as $M_{APA} = (1-\alpha)A^* + \alpha V$, taking $\alpha = 0.15$. The result is:

$$M_{APA} = \begin{bmatrix} 0.0150 & 0.1850 & 0.0150 & 0.0150 & 0.2275 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 \\ 0.4400 & 0.0150 & 0.4400 & 0.4400 & 0.2275 & 0.1567 & 0.0150 & 0.0150 & 0.0150 & 0.0150 \\ 0.0150 & 0.1850 & 0.0150 & 0.0150 & 0.0150 & 0.1567 & 0.0150 & 0.0150 & 0.0150 & 0.0150 \\ 0.0150 & 0.1850 & 0.0150 & 0.0150 & 0.2275 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 \\ 0.4400 & 0.1850 & 0.0150 & 0.4400 & 0.0150 & 0.1517 & 0.0150 & 0.0150 & 0.0150 & 0.0150 \\ 0.0150 & 0.1850 & 0.4400 & 0.0150 & 0.2275 & 0.0150 & 0.2275 & 0.2983 & 0.0150 & 0.2983 \\ 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.1567 & 0.0150 & 0.2983 & 0.2983 & 0.2983 \\ 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.1517 & 0.2275 & 0.0150 & 0.2983 & 0.0150 \\ 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.2275 & 0.2983 & 0.0150 & 0.2983 \\ 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.1567 & 0.2225 & 0.0150 & 0.2983 & 0.0150 \end{bmatrix}.$$

Finally, we compute the eigenvector $\boldsymbol{x}_1$ associated to the dominant eigenvalue $\lambda_1 = 1$ for the matrix M_{APA}. That is,

$$\boldsymbol{x}_1 = [0.1947; 0.4369; 0.1896; 0.1947; 0.3551; 0.4965; 0.3398; 0.2630; 0.2662; 0.2630].$$

From the vector $\boldsymbol{x}_1$, we have a ranking of the nodes according to its importance in the network. In this example node n_6 is the most central one, followed by node n_2. In section 4 we will analyse more numerical results in order to obtain the characteristics of this centrality measure.

In other words, we can say that the APA algorithm constitutes a model to establish a ranking of nodes in a network, with the primary feature that assigns a value to each node according to its significance within the physical network taking into account both the geometry of the network and the data stored in the nodes.

3. The Adapted PageRank Algorithm Modified, APAM

In this section a new algorithm, based on the APA algorithm studied in the previous section, is proposed. It constitutes a new centrality measure applied to urban networks.

3.1. The algorithm

The ranking obtained by the APA algorithm is based on the construction of a matrix M_{APA} which is the sum of two normalized matrices A^* and V, with essential features to understand the philosophy behind the model. For this, we reformulate the meaning of these matrices in probabilistic terms. So,

- The matrix A^* can be interpreted as the probability of going from one node to any of its neighbouring nodes.
- The matrix V represents the probability of jumping from one node to any other in the network, considering the amount of data associated with each of the network nodes. Recall that in the original APA algorithm, following the PageRank vector concept, the term αV was related to the probability that a surfer jumps from a site to another one of the network, without there being a link between them. This probability was the same for every page. In our case, the jump from one node to another is given by the influence of the data present in the network, although the probability of jumping from one node to another does not depend on its topological distance.

However, this feature to jump between nodes with the same probability regardless of the distance between nodes, is certainly debatable when working with urban networks. Recall that the urban networks are a type of complex networks with features that make them very particular when applying traditional algorithms of classical network theory. One of these features, absolutely crucial, is that the degrees of the nodes in urban networks have fairly uniform values and not excessively high. The distribution of the degrees of the nodes on an urban network concentrate their values mostly between 2 and 5. Also, consider equiprobable the fact of jumping from one network node to any other node in the network is at least debatable. In urban networks, the influence that neighbouring nodes exert on the node itself is an aspect that should be taken into consideration.

All these considerations are a fundamental motivation that leads us to propose a new centrality measure based on the idea of PageRank vector, but taking into account the particular characteristics of the networks in which we operate.

A new centrality measure is proposed, denoted by APAM, based on the APA centrality but introducing some modifications with the aim to associate a higher probability to adjacent nodes in the data matrix V, as previously discussed.

Now, the steps we follow are:

Step 1. We start from the graph $G = (V, E)$ and construct the adjacency matrix A and the transition matrix A^* using the expression (1). We also construct the data matrix D and the weight vector $\boldsymbol{v}_0$.

Step 2. We construct the data vector $\boldsymbol{v}$ as $\boldsymbol{v} = D \cdot \boldsymbol{v}_0$. Then, we use the data vector $\boldsymbol{v}$ to construct the matrix V adding the vector $\boldsymbol{v}$ as columns N times. Therefore, V is square of size N.

Step 3. Construct the matrix K, with the idea of collecting the data in the network.

We construct a matrix K as

$$K = A \otimes V + \epsilon F, \tag{2}$$

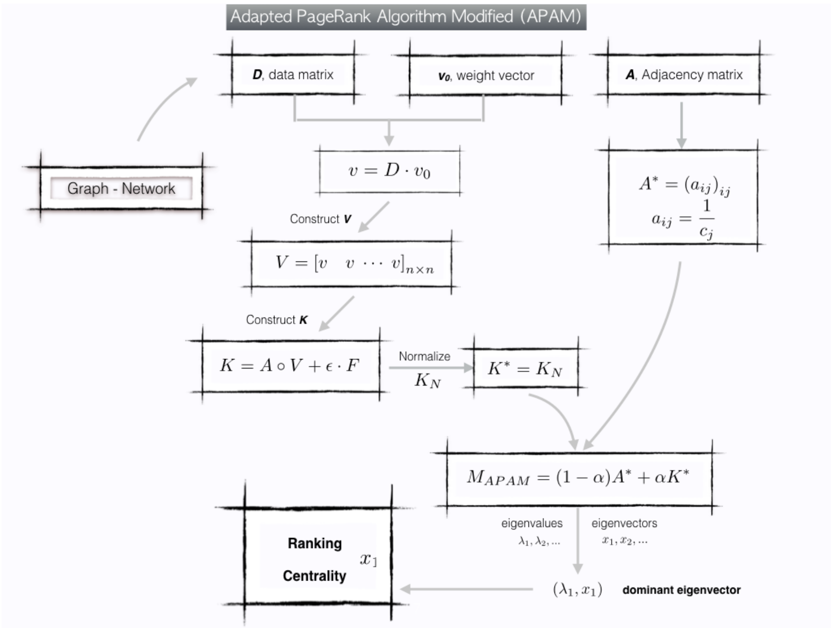

Figure 3: APAM theoretical scheme.

where A is the adjacency matrix, V is the matrix with the data, ϵ represents a small quantity related to the inverse of the total amount of data in the network, and the matrix F is a $N \times N$ matrix whose all its elements are 1. To compute the product of A and V we use the Hadamard product ($\otimes$) of matrices, which multiplies the elements that occupy the same position in both arrays.

Step 4. Normalize K, K_N so its columns are probability vectors. Then, in the following, we denote $K_N = K^*$.

Step 5. Construct the matrix M_{APAM} from A^* and K^* as

$$M_{APAM} = (1-\alpha) \cdot A^* + \alpha \cdot K^*.$$

Step 6. Obtain the dominant eigenvector $\boldsymbol{v_1}$ of M_{APAM}. This vector constitutes the ranking of the nodes of the graph.

According to these steps described above, the algorithm proposed is:

Algorithm 2. ***APA Algorithm modified, APAM.*** *Let us assume that we have a graph representing an urban network with N nodes n_i, representing squares or streets intersections. We proceed with the following steps.*

1 *Obtain the adjacency matrix A from the graph of the network and the matrix A^*.*

2 *Consider the different characteristics k_i associated to each of the nodes for the problem studied; evaluate them in each node. With these numerical values, we construct the matrix D.*

3 *Construct a vector $\boldsymbol{v_0}$, according to the importance of each of the characteristics evaluated. This vector represents a multiplicative factor.*

4 *Obtain a vector $\boldsymbol{v}$ by multiplying $D \cdot \boldsymbol{v_0} = \boldsymbol{v}$.*

5 *Construct the matrix V, from $\boldsymbol{v}$.*

6 *Construct the matrix K according to the expression (2).*

7 *Transform the matrix K in stochastic, $K \longrightarrow K_N = K^*$.*

8 *Construct the matrix $M_{APAM} = (1-\alpha)A^* + \alpha K^*$, from A and V.*

9 *Compute the eigenvector $\boldsymbol{v}_1$ associated to the eigenvalue 1 for the matrix M_{APAM}. That is our ranking vector.*

Figure 3 shows an overview of the steps and mathematical operations involved in the computation of the APAM centrality. We highlight some characteristics of the algorithm.

In this new model two stochastic matrices are constructed. On one hand, the matrix A^* representing, in probabilistic terms, the probability that a node would jump to one of its neighbouring nodes. This matrix is determined by the topology of the urban network studied, without showing anywhere the influence of the data in it. Moreover, we have the stochastic matrix A^*, which itself is related to the data associated with the nodes of the network. On the other hand, we can also make an interpretation of K^* in probabilistic terms, since it represents the probability that a node jump to any of its neighbours, according to the amount of data in each node. This statement leads us to examine in a more detailed way the construction of K^*.

Note that K^* is constructed by normalizing the matrix K, that has been defined as

$$K = A \otimes V + \epsilon F.$$

The matrix K consists of two terms. The term $A \otimes V$ represents the Hadamard product of the adjacency matrix with the data matrix V, which has the data vector $\boldsymbol{v}$ in its columns. The term ϵF represents a small quantity which adds to all the terms of the matrix and which takes into account the existing global data in the urban network. This small amount remains, in a sense, the idea of jumping from one node in the network to any other, at random and without regard to the topological distance between two nodes. This idea is similar to that posed by the PageRank algorithm in the Google search engine.

From the matrices A^* and K^* we construct the matrix M_{APAM} whose dominant eigenvector provides us a ranking of the nodes, according to their importance in the network. It is important to point out that M_{APAM} takes into account, firstly, the network topology and, on the other hand, the importance of the data network in general, and mainly of our data nodes neighbours. This is a fundamental idea behind the design of this new model and is a distinguishing aspect of the original model based on the APA algorithm 1.

Although the idea behind both measures are similar, we can find some differences between Algorithms 1 and 2. In the case of the APAM algorithm 2, the normalization of the matrix K occurs at the end of the process, just before the matrix M_{APAM} is computed. Besides, an important fact is that the data vector $\boldsymbol{v}$ is not normalized, in order to give more importance to the amount of data that is associated with each node.

Following the example given by the graph shown in Figure 2, we can compare the results that we obtain for both algorithms, in order to clarify the differences between them. Although the most interesting is to compare the results obtained in the matrices that reflect the influence of the data on the network. In the APA algorithm, the data matrix is given by V as we have already mentioned, while in the case of the APAM algorithm, the data matrix is K^*, which is the matrix $K = A \otimes V + \epsilon F$ normalized.

In this example, taking $\boldsymbol{v} = [1;1;1;1;1;1;1;1;1;1]$, we have that, after performing the appropriate computations,

$$K^* = \begin{bmatrix} 0.0333 & 0.1833 & 0.0333 & 0.0333 & 0.2200 & 0.0143 & 0.0200 & 0.0250 & 0.0250 & 0.0250 \\ 0.3667 & 0.0167 & 0.4333 & 0.3667 & 0.2200 & 0.1571 & 0.0200 & 0.0250 & 0.0250 & 0.0250 \\ 0.0333 & 0.1833 & 0.0333 & 0.0333 & 0.0200 & 0.1571 & 0.0200 & 0.0250 & 0.0250 & 0.0250 \\ 0.0333 & 0.1833 & 0.0333 & 0.0333 & 0.2200 & 0.0143 & 0.0200 & 0.0250 & 0.0250 & 0.0250 \\ 0.3667 & 0.1833 & 0.0333 & 0.3667 & 0.0200 & 0.1571 & 0.0200 & 0.0250 & 0.0250 & 0.0250 \\ 0.0333 & 0.1833 & 0.4333 & 0.0333 & 0.2200 & 0.0143 & 0.2200 & 0.2750 & 0.0250 & 0.2750 \\ 0.0333 & 0.0167 & 0.0333 & 0.0333 & 0.0200 & 0.1571 & 0.0200 & 0.2750 & 0.2750 & 0.2750 \\ 0.0333 & 0.0167 & 0.0333 & 0.0333 & 0.0200 & 0.1571 & 0.2200 & 0.0250 & 0.2750 & 0.0250 \\ 0.0333 & 0.0167 & 0.0333 & 0.0333 & 0.0200 & 0.0143 & 0.2200 & 0.2750 & 0.0250 & 0.2750 \\ 0.0333 & 0.0167 & 0.0333 & 0.0333 & 0.0200 & 0.1571 & 0.2200 & 0.0250 & 0.2750 & 0.0250 \end{bmatrix}.$$

The comparison of these two matrices is quite revealing. If we think in probabilistic terms and you look at the matrix V, we see that all their entries are equal to 1/10, which indicates that the probability of jumping from one node to another taking into account the data on the network is the same and equal to 1/10. This is normal when you consider the data vector chosen.

But this is not the case when considering the new APAM algorithm. Failure to normalize the vector data before constructing the matrix K^* and its own definition means the probabilities now are not the same. However, we note that the probabilities increase when we jump to a neighbouring node. Looking at, for example, the first column of the matrix K^* note that the most likely jump from node 1 is to nodes 2 and 5, which are precisely its neighbouring nodes. The same happens if you look at the other columns. The influence of the data is greater in our neighbouring nodes, which, as already mentioned above, makes more sense in urban networks.

3.2. The parameter α

The APA algorithm 1 is based on the idea of PageRank model used by the Google search engine to classify websites in order of importance. The PageRank model [2, 19, 21] proposed by the founders of Google, was devised to compute a ranking for every Web page based on the graph of the Web; that is, PageRank constitutes a global ranking of all Web pages, regardless of their content, based solely on their location in the Web's graph structure. The purpose of the method is to obtain a vector, called PageRank vector, which gives the relative importance of the pages.

To perform this, the model is based on the construction of a matrix called Google matrix given by the expression

$$G = (1 - \alpha)A + \alpha B,$$

where A is the transition matrix, B is a matrix defined as $B = (1/n)I_n$, with n the number of nodes (websites). The parameter α is called the *dumping factor* and is related to the idea that a surfer jumps from a website to any other in the web network without any existing link between them. This parameter is usually taken as $\alpha = 0.15$.

If we analyse the Google matrix G we have two terms, $(1 - \alpha)A$ related to the network topology and αB related to the probability that a surfer jumps from a page to another without a link between them. If we take $\alpha = 0.15$ we are giving great importance to the connectivity in the network and considering less relevant the possibility to surf from one page to another with no link between them.

The main contribution of the APA algorithm is that using the basic idea of PageRank model, introduces an important factor in the process of establishing the ranking of the nodes, as it is the presence of data within the network itself. So, the matrix M_{APA} that allows us to establish the ranking for the nodes of the network is given by

$$M_{APA} = (1 - \alpha)A^* + \alpha V, \quad (3)$$

where the first term $(1 - \alpha)A^*$ is related with the network topology and the second term αV is related with the data in the network. The matrix V represents the amount of data associated to each of the nodes. So, we introduce the importance of the data in the network by the term αV. This means that if we consider the same value for the parameter α as in the PageRank model, that is, $\alpha = 0.15$, we are giving much more importance to the network topology than to the weight of the data. In urban networks, this is questionable, due to the characteristics of urban networks, where the great majority of nodes have very low degree values and their distribution is fairly homogeneous. Therefore, it is necessary to perform numerical experiments to study what happens for different parameter values.

The same reasoning can be applied to the algorithm APAM. The matrix from which we get the classification or ranking of the nodes consists of two distinct terms. On one hand the term related to the topology of the network and, on the other hand, the other term relating to the data network. It is multiplied by a parameter alpha, which allows us, in principle, to determine the importance we attach to the network topology or to data. Obviously, if we take $\alpha = 0.5$ would mean that we consider equally important for the classification of the nodes both the aspect of network connectivity as the amount of data present therein.

Therefore, when we study the nodes centrality of an urban network analysing some specific data, we can determine whether we give more or less weight to the amount of data present in each node and its nearest

Nodes	Data	Centrality					
	$\boldsymbol{v}$	APA	APAM				
			$\alpha = 0.15$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
1	1	0.1947	0.1792	0.1841	0.1906	0.1971	0.2037
2	1	0.4369	0.4339	0.4318	0.4280	0.4231	0.4173
3	1	0.1896	0.1788	0.1835	0.1900	0.1966	0.2035
4	1	0.1947	0.1792	0.1841	0.1906	0.1971	0.2037
5	1	0.3551	0.3490	0.3492	0.3488	0.3477	0.3461
6	1	0.4965	0.5166	0.5109	0.5032	0.4953	0.4872
7	1	0.3398	0.3468	0.3459	0.3452	0.3450	0.3451
8	1	0.2630	0.2625	0.2642	0.2671	0.2704	0.2741
9	1	0.2662	0.2624	0.2642	0.2671	0.2705	0.2741
10	1	0.2630	0.2625	0.2642	0.2671	0.2704	0.2741

Table 1: Numerical values of the centrality measures studied for the graph shown in Figure 2 for $\boldsymbol{v} = [1; 1; 1; 1; 1; 1; 1; 1; 1; 1]$.

neighbours. And, what is more important, this can be easily determined by the α parameter. Obviously, when a value closer to 1 is chosen for the α parameter, we are providing greater importance to the data.

4. Numerical results

First, we develop experiments on a small network of 10 nodes, modifying the data values to display its influence on the centrality. After that, we take a real urban network to develop the numerical experiments.

4.1. Numerical examples in a small network

The common denominator, in the measures proposed, is the use of data from the network. For this reason, we show a simple example in which it can be analysed, on a small scale, the differences between both measures and how they depend in greater or lesser extent of data.

In Figure 2 we represent a simple graph with ten nodes and the connections among them. Looking at the graph we realize some characteristics that we must consider. For instance, it is clear that the most relevant node in this graph, from the topological point of view, is the node 6, which presents the highest degree (in this case 6). We also note that if we remove this node of the graph it will split into two unconnected parts. It is evident that if we compute classical centrality measures for this graph, as for example, closeness or betweenness centrality, the most important will be this node 6. Therefore, we can say that it is an essential node when establishing paths between nodes. We can also add that node 4, at first glance, it is the least important in the graph, since it is the one with less connectivity with the rest.

Before computing the centrality measures given by the APA and APAM algorithms to check differences and similarities between them, we must realize that the APAM centrality measure is determined by the value of the parameter α, so we will study this algorithm numerically for different values of the parameter. This means that we run the algorithm giving greater or lesser importance to the presence of network data.

Following with the example presented in the previous section and considering the data vector used, ie, $\boldsymbol{v} = [1; 1; 1; 1; 1; 1; 1; 1; 1; 1]$, we proceed to compute the measures of centrality that the APA and APAM algorithms produce, running the APAM algorithm for different parameter values, as shown in Table 1.

In this table, we note that the most relevant node, for all the cases studied, is the node 6, which is not surprising since it is the node with a higher degree (exactly, 6). We have very few data associated to the nodes of the graph; consequently, the weight of the connections between the elements of the graph significantly predominates over the weight of the data, due to the small quantity of these. However, we can also check how the value of the centrality of the node 6 (most important one), is decreasing when increasing the parameter α in the APAM algorithm. Another consequence of the low number of data in $\boldsymbol{v}$ is that the changes in the centrality values are quite soft.

It is also observed that in this case, with this particular data vector, the centrality values of the nodes when executing the APA algorithm and the APAM algorithm, for $\alpha = 0.7$ are similar. It may be interesting

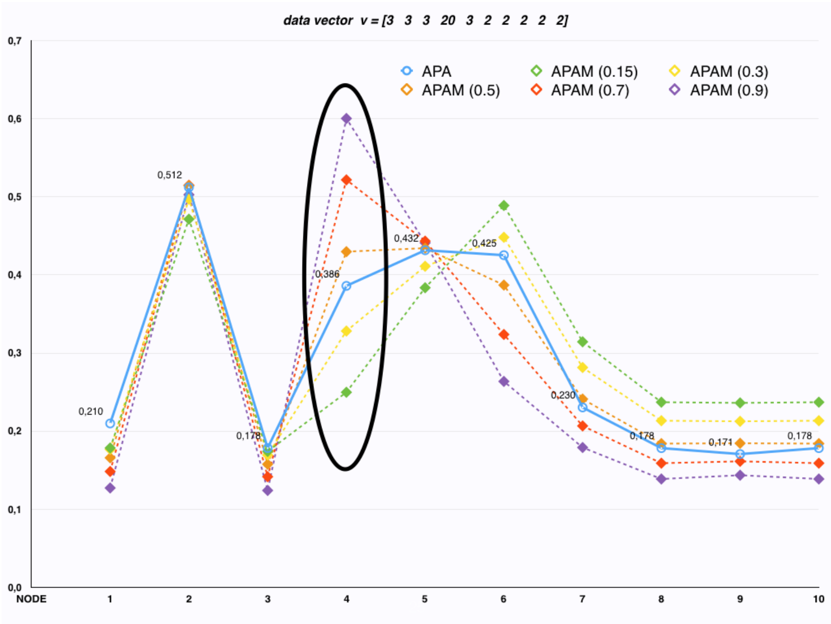

Figure 4: Centrality measures for APA and APAM algorithm taking data vector $\boldsymbol{v_1}$.

to study several numerical examples for this simple network, modifying the data vector to analyse the influence of the data on this network.

Then, let us assume that we have the following set of vector data:

$$\boldsymbol{v_1} = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 20 \\ 3 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} \quad \boldsymbol{v_2} = \begin{bmatrix} 20 \\ 20 \\ 3 \\ 3 \\ 20 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} \quad \boldsymbol{v_3} = \begin{bmatrix} 40 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \boldsymbol{v_4} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$$

The first vector $\boldsymbol{v_1}$ has a set of homogeneous data, with the exception of the node 4 having 20 data associated, a much higher number than the rest. Furthermore, it has been located in the node 4 as it is one of the most isolated in the graph, as has degree equal to 2. In the second data vector $\boldsymbol{v_2}$ we have increased the number of data associated to the nodes $1, 2$ and 5, which are nodes that are connected to each other in the graph (see Figure 2). In data vector $\boldsymbol{v_3}$ we concentrate most of the data existing in the network around the node 1, which is a node with degree 2 and is located in a corner of the graph. Finally, in data vector $\boldsymbol{v_4}$ we have most of the data in the last four nodes in the graph, those located on the right part of the graph.

Table 2 shows the centrality measurement obtained by applying the original APA algorithm 1 and the APAM algorithm 2, taking $\alpha = 0.15$ as usually in the APA algorithm and taking $\alpha = 0.15$, $\alpha = 0.3$, $\alpha = 0.5$, $\alpha = 0.7$, $\alpha = 0.9$ in the new one.

The numerical results in Table 2 show some similar behaviours, for different data vectors, that we will analyze with more detail.

In Figures 4–7, we have graphically displayed the centrality values computed with the APA and APAM algorithms, shown in Table 2, for different values of the parameter α, for data vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$, and $\boldsymbol{v_4}$.

Nodes	Data	Centrality					
	$\boldsymbol{v}$	APA	APAM				
			$\alpha = 0.15$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
1	3	0.2099	0.1785	0.1767	0.1659	0.1484	0.1273
2	3	0.5121	0.4714	0.4970	0.5143	0.5148	0.5027
3	3	0.1784	0.1741	0.1695	0.1577	0.1417	0.1241
4	20	0.3860	0.2498	0.3282	0.4298	0.5217	0.6003
5	3	0.4316	0.3836	0.4111	0.4343	0.4433	0.4411
6	2	0.4253	0.4889	0.4482	0.3870	0.3238	0.2637
7	2	0.2304	0.3144	0.2815	0.2411	0.2067	0.1791
8	2	0.1783	0.2371	0.2134	0.1841	0.1590	0.1388
9	2	0.1707	0.2362	0.2126	0.1845	0.1614	0.1436
10	2	0.1783	0.2371	0.2134	0.1841	0.1590	0.1388
1	20	0.3105	0.2460	0.3115	0.3841	0.4421	0.4899
2	20	0.5685	0.5286	0.5895	0.6280	0.6358	0.6278
3	3	0.1710	0.1666	0.1488	0.1183	0.0877	0.0605
4	3	0.2170	0.1943	0.1947	0.1736	0.1397	0.1012
5	20	0.4886	0.4369	0.5004	0.5497	0.5721	0.5795
6	2	0.4085	0.4608	0.3824	0.2766	0.1850	0.1107
7	2	0.1909	0.2665	0.1911	0.1138	0.0644	0.0366
8	2	0.1478	0.2009	0.1448	0.0869	0.0495	0.0283
9	2	0.1353	0.1990	0.1426	0.0856	0.0498	0.0302
10	2	0.1478	0.2009	0.1448	0.0869	0.0495	0.0283
1	40	0.5167	0.2765	0.3847	0.5248	0.6447	0.7360
2	1	0.5190	0.4615	0.4794	0.4856	0.4714	0.4424
3	1	0.1478	0.1623	0.1459	0.1181	0.0867	0.0560
4	1	0.1920	0.1640	0.1481	0.1186	0.0833	0.0470
5	1	0.4490	0.3796	0.4051	0.4252	0.4276	0.4155
6	1	0.3617	0.4838	0.4334	0.3509	0.2603	0.1732
7	1	0.1643	0.3219	0.2899	0.2402	0.1881	0.1400
8	1	0.1271	0.2420	0.2184	0.1816	0.1427	0.1065
9	1	0.1153	0.2417	0.2183	0.1824	0.1450	0.1106
10	1	0.1271	0.2420	0.2184	0.1816	0.1427	0.1065
1	1	0.0902	0.1412	0.1058	0.0632	0.0327	0.0150
2	1	0.2386	0.3522	0.2629	0.1555	0.0783	0.0335
3	1	0.1137	0.1426	0.1077	0.0645	0.0325	0.0135
4	1	0.0902	0.1412	0.1058	0.0632	0.0327	0.0150
5	1	0.1904	0.2819	0.2106	0.1248	0.0631	0.0272
6	1	0.4519	0.4885	0.4278	0.3219	0.2123	0.1142
7	10	0.4814	0.4288	0.4955	0.5533	0.5806	0.5895
8	10	0.3726	0.3164	0.3600	0.3940	0.4055	0.4042
9	10	0.4049	0.3370	0.4061	0.4773	0.5249	0.5569
10	10	0.3726	0.3164	0.3600	0.3940	0.4055	0.4042

Table 2: Numerical values of the centrality measures studied for the graph shown in Figure 2 for data vectors from $\boldsymbol{v_1}$ to $\boldsymbol{v_4}$.

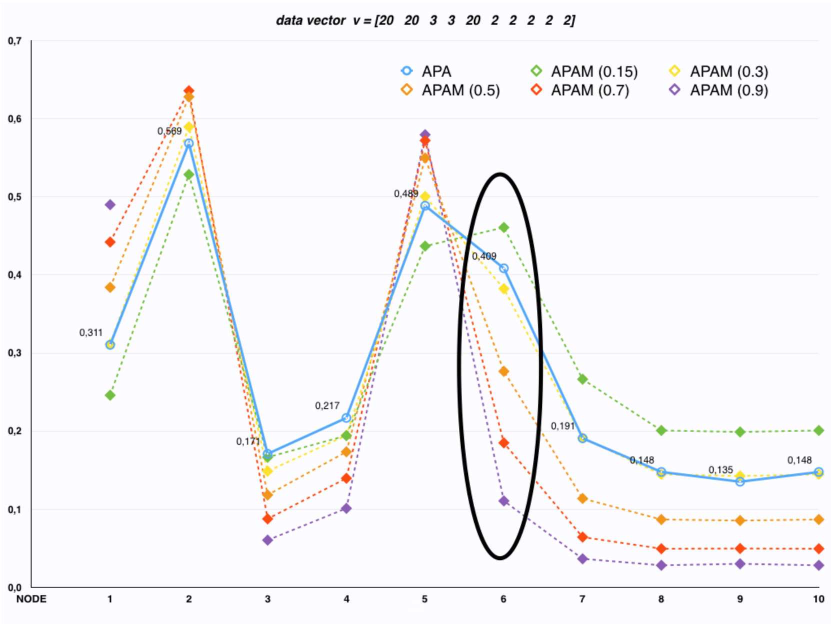

Figure 5: Centrality measures for APA and APAM algorithm taking data vector $\boldsymbol{v_2}$.

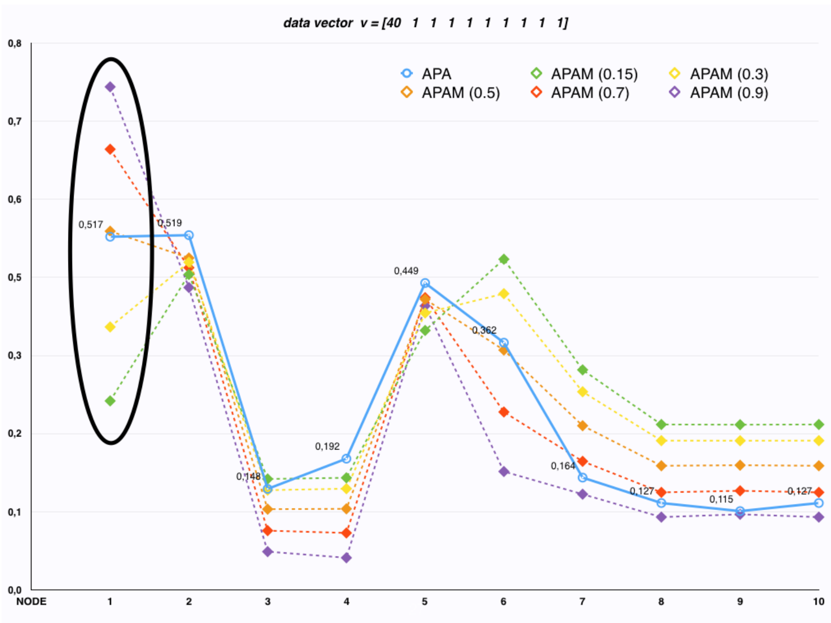

Figure 6: Centrality measures for APA and APAM algorithm taking data vector $\boldsymbol{v_3}$.

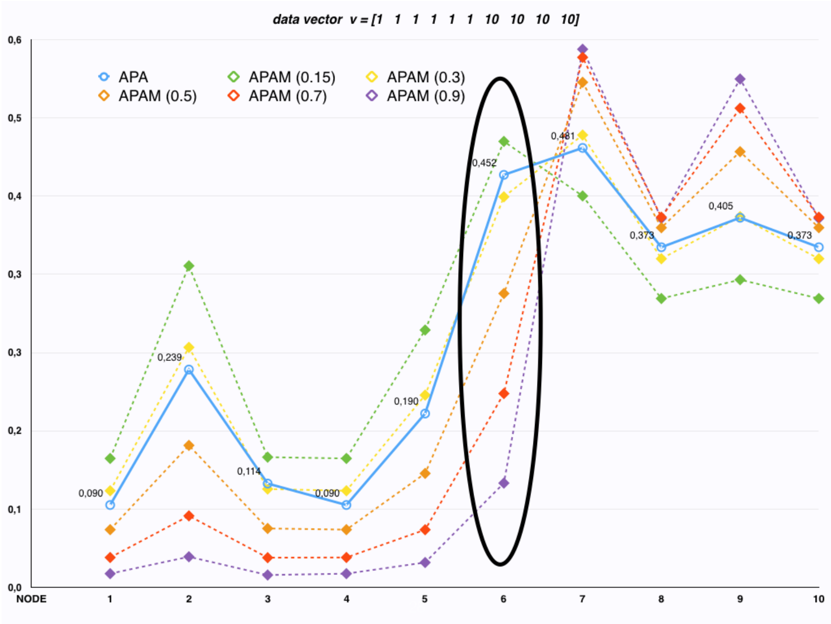

Figure 7: Centrality measures for APA and APAM algorithm taking data vector $\boldsymbol{v_4}$.

It has been indicated in the graphs by an oval, the node where a greater dispersion of the values of the centrality occurs when you execute all the algorithms.

We look at the results for data vector $\boldsymbol{v_1}$, where data are distributed evenly across all the nodes, except for node 4, which is receiving a greater amount of data (20 versus an average of 2.5 per node). After executing the APA algorithm, node 2 is identified as the most central and important in the network. However, when we apply the APAM algorithm, for different values of the α parameter, it is observed that the most important node in the network is changing. So, when $\alpha = 0.15$, that is, we do not give much influence to the data associated to each node, the most important one in the network is the node 6, which is the one with the highest degree. When we increase the value of α, in the range from 0.3 to 0.6, the most important node in the network changes to node 2, the same as when executing the APA algorithm. Only when α takes high values, particularly greater than 0.6, we found that the weight of the data is decisive, which means that the most important node is the one with a greater amount of associated data, which in this case is node 4. So, in Figure 4, we can see that the node that presents a higher dispersion of the centrality values for the different cases studied is the node 4.

We now look at the values of Table 2 corresponding to vector data $\boldsymbol{v_2}$. In this case we choose to place a greater amount of data in the three vertices connected to each other on the left side of the graph. In this example, we observed a very different behaviour in the centrality values of the nodes from previous case. Now, in all cases studied, the most central node is the node 2. The explanation for this fact is that node 2 has two essential characteristics: first, a good connectivity within the graph and, on secondly, a large amount of data associated to it. Logically, give more or less importance to these features, this node in the network is critical. It is also noted that the three nodes with more data associated are the nodes with higher values of centrality.

If we see the centrality values of the data vector $\boldsymbol{v_3}$ we realize that it is similar to the case of data vector $\boldsymbol{v_1}$; the difference is that now, the data are centralized in node 1, which has a small degree and it seems isolated in the left part of the network. Only when α takes high values, particularly greater than 0.6, we found that the weight of the data is decisive and this node is the most important one in the network.

The case of the vector $\boldsymbol{v_4}$ is similar to the case of $\boldsymbol{v_2}$. Now the data are centralized on the right side of the graph (last four nodes of the graph). In all cases except one, the most important node is 7, which has a connectivity a bit higher than the rest of the nodes with the same data.

In view of the graphs shown in Figures 4-7 we can see that the values of the ranking that both models offer have some differences, especially when using extreme values of the parameter alpha domain. Notice a similarity ranking of the APA and APAM models, especially when the alpha parameter takes domain central values, in which case it is given the same importance to the topological connection of the nodes of the network and the data associated with the same.

In the graphs, we have highlighted with oval shapes the nodes that have a higher value of the centrality range. If we look at node 6 in all graphics, we see that it is always among the first or second in the ranking of dispersion values. This is completely logical if we think that the node 6 is the one with the highest degree of the graph, making it the most important network node, from the topological point of view.

Another important aspect to note is related to the deviation of the values given by the APAM ranking model, compared to the values generated by the APA model, when $\alpha = 0.1$ and $\alpha = 0.9$ are used. For $\alpha = 0.9$ we are giving the highest importance to the quantity of data associated to the nodes and barely consider the degree of each node, while for $\alpha = 0.1$ we are giving the highest importance to the connectivity of the graph.

4.2. An example in a real urban network

We take a real urban network as it is the historic center of the city of Murcia, Spain. We create the network (primal graph) from a connected graph where the streets become undirected edges. Nodes usually represent the intersections of the streets, but we can also assign nodes to some points of interest in long streets. The network is composed of 267 nodes and 391 edges.

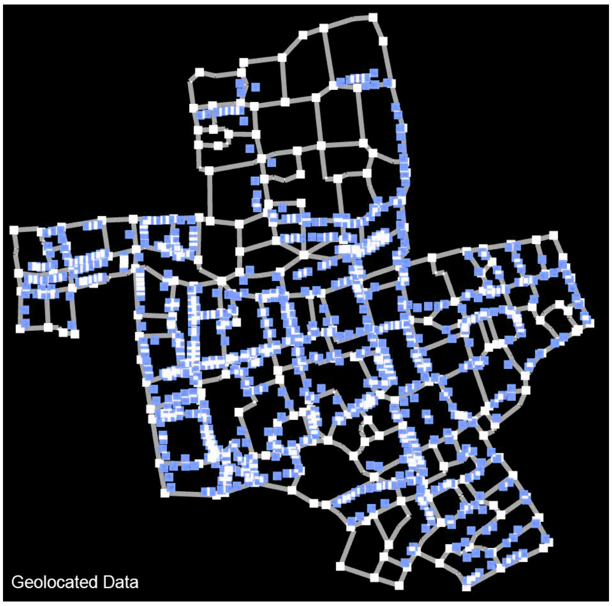

Figure 8: Network of the city area studied and data geolocated on the map.

For this example, we will work only with a part of the city, the historical center and the neighbourhoods that are placed around it. The reason that motivates this limitation lies, on the one hand, in reducing the amount of data to work with and, on the other hand, because the historical center is the most active area of the city and where most activity takes place.

The data collection used for this example starts with fieldwork that consists of collecting the data or information from visual inspection or pictures. These data were assigned to the nodes of the network so that each node has a set of numerical values associated with the information that is being studied. We collected data about existing facilities and commercial activity. Figure 8 displays the primal graph of the city area object of this study and the data geolocated on the map. These data are related to the food-service sector, more exactly with the geolocation of restaurants, bars, coffees, fast food chains, ...

We have developed several numerical experiments with this urban network. The most significant results are shown in Figure 9. Note that the results obtained by running the original APA algorithm are very similar to those obtained when taking into account the overall amount of data on the network. It is therefore necessary to develop a process for the allocation of network data to each node in the graph.

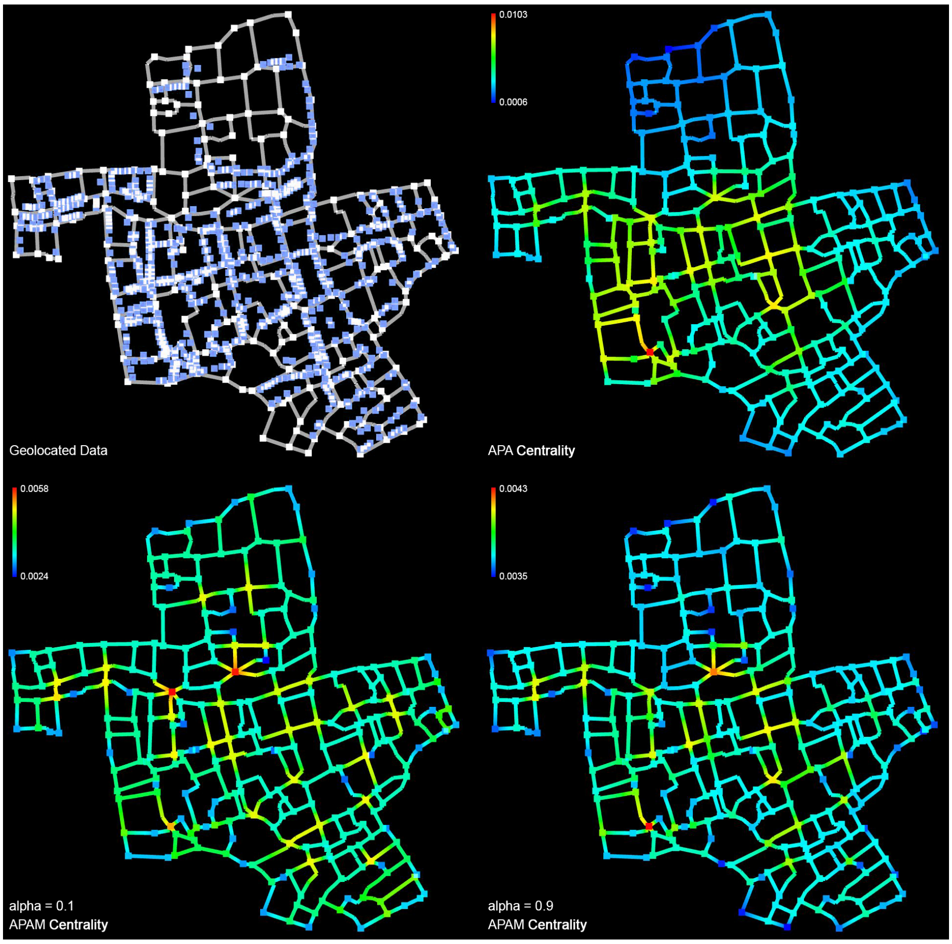

Figure 9: Different centrality measurements, for the urban network shown in Figure 8.

In Figure 9 we display the studied network in four different ways. The upper left image shows the graph associated to the urban network with the geolocated data extracted from it. The top right image shows the display of the network after applying the APA algorithm. The lower left and right images represent the network after applying the APAM algorithm, for $\alpha = 0.1$ (left) and $\alpha = 0.9$ (right).

These quantitative differences in the nodes that constitute the network can be displayed graphically in many different ways. To display the data, we are opting for a chromatic scale of values called Hot-Cold scale, based on the RGB model (Red, Green, Blue). This color domain comprises the following values: Blue[0,0,255], Cyan$[0, 255, 255]$, Green$[0, 255, 0]$, Yellow$[255, 255, 0]$, Red$[255, 0, 0]$.

Consequently, we have two distinct scales: first, the scale of the domain of values that provides quantification of information and, on the other hand, the scale that provides us with the graphic scale. It is necessary to enhance a linear interpolation to set the color that is assigned to each of the nodes, according to the amount of information associated with it.

Looking the graphics of Figure 9 we can appreciate the differences in the values of centrality of some very central points and visualize the centrality in the different areas of the urban network. We note the influence of the data in the computation of centrality.

It is also observed something that has already been discussed in Section 4.1. Regarding the fact that increasing the value of the α parameter the ranges of values of centrality became more dispersed. Accordingly, the visualization for values close to $\alpha = 1$ is much less smooth than when α takes smaller values. No more graphics have been included with other values of the parameter because the differences are not significant.

5. Conclusions

A model based on the centrality APA algorithm, which provides a classification of network nodes according to their importance within the same is proposed. The proposed new model is not intended to correct the APA model but offers an alternative measure of centrality in which the amount of data associated with the node has a greater participation in the stochastic process.

In the process of computation of the APA and APAM centralities, two components are involved: the topology of the network and the data associated with it. The component related to the topological structure of the network in both models remains unchanged, while the component related with the presence of data is different between models.

The idea behind the construction of the data matrix APA algorithm could be interpreted as the probability of making a jump from any node to all other nodes in the network, taking into account the amount of global data network, regardless of the topological distance to nodes of the same. However, in the new model, the data matrix is interpreted as the probability of making a jump from any node to its adjacent nodes, considering the amount of data that each node has associated to it. Thus, the construction of the data matrix takes implicit the topological relationships of the data in the network.

From the studied numerical examples we can conclude that, using the APAM model with extreme values of the alpha parameter, a node, despite having a high degree, if it has few data associated to the same will have little significance in the network. In contrast, if the node has many associated data, regardless of its degree, will have high values of centrality, which does not happen when applying the APA model.

Both models have their practical application in urban networks; each provides a different measure of centrality. For the APA model priority is given the global distribution of data, while the APAM model is focused on the topological data distribution.

References

[1] Agryzkov, T., Oliver, J.L., Tortosa, L., Vicent, J.: An algorithm for ranking the nodes of an urban network based on concept of PageRank vector. Applied Mathematics and Computation **219** (2012) 2186–2193.

[2] Berkhin, P.: A survey on PageRank computing. Internet Mathematics **2(1)** (2005) 73–120.

[3] Bonacich, P.: Power and centrality: a family of measures. American Journal of Sociology **92** (1987) 1170–1182.

[4] Bonacich, P.: Simultaneous group and individual centrality. Social Networks **13** (1991) 155–168.

[5] Brandes, U.: A faster algorithm for betweenness centrality. Journal of Mathematical Sociology **25(2)** (2001) 163–177.

[6] Calabrese, F., Ratti, C., Colonna, M., Lovisolo, P., Parata, D.: Real-time urban monitoring using cell phones: a case study in Rome. IEEE Transactions on Intelligent Transportation Systems **25(2)** (2011) 141–151.

[7] Crucitti, P., Latora, V., Porta, S.: Centrality measures in spatial networks of urban streets. Physical Review E **73** (2006) 036125.

[8] Crucitti, P., Latora, V., Porta, S.: The network analysis of urban streets: a primal approach. Planning and design **33(5)** (2006) 705-725.

[9] Crucitti, P., Latora, V., Porta, S.: The network analysis of urban streets: a dual approach. Physica A: Statistical Mechanics and its Applications **369(2)** (2006) 853–866.

[10] Estrada, E.: The structura of complex networks. Theory and applications. Oxford University Press, Oxford (2012).

[11] Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40(1)** (1977) 35–41.

[12] Freeman, L.C.: Centrality in social networks: Conceptual clarification. Social Networks **1(3)** (1979) 215–239.

[13] Freeman, L.C., Borgatti, S.P., White, D.R.: Centrality in valued graphs: a measure of betweenness based on network flow. Social Networks **13** (1991) 141–154.

[14] Friedkin, N.E.: Theoretical foundations for centrality measures. American Journal of Sociology **96** (1991) 1478–1504.

[15] Hubbell, C.H.: An input output approach to clique identification. Sociometry **28** (1965) 377–399.

[16] Jeong, S.K., Ban, Y.U.: Computational algorithms to evaluate design solutions using Space Syntax. Computer-Aided Design **43(6)** (2011) 664–676.

[17] Jiang, B.: Ranking spaces for predicting human movement in an urban environment. International Journal of Geographical Information Science. **23(7)** (2009) 823–837.

[18] Katz, L.: A new index derived from sociometric data analysis. Psychometrika **18** (1953) 39–43.

[19] Langville, A.N., Mayer, C.D.: Deeper inside PageRank. Internet Mathematics **1(3)** (2005) 335–380.

[20] Newman, M.E.J.: A measure of betweenness centrality based on random walks. (2003) Available via arXiv. http://arXiv:cond-mat/0309045.

[21] Page, L., Brin, S., Motwani, R., Winogrand, T.: The pagerank citation ranking: Bringing order to the web. Technical report 1999-66, Stanford InfoLab (1999).

[22] Pedroche, F.: Metodos de calculo del vector PageRank (in spanish), Bol. Soc. Esp. Mat. Apl. **39** (2007) 7–30.

[23] Stephenson, K. A., Zelen, M.: Rethinking centrality: Methods and examples. Social Networks **11**, 1–37 (1989).

[24] Wasserman, S., Faust, K.: Social Network Anlysis. Cambridge University Press, Cambridge (1994).

[25] Zhu, W.J., Kan, X.D.: Coordination Analysis for Urban Underground Space and Rail Transit Based on Space Syntax. Applied Mechanics and Materials **170-173**, 1388–1391 (2012).