Similarity matrices for pairs of graphs

Vincent D. Blondel[1] and Paul Van Dooren[2]

[1] Division of Applied Mathematics, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium, blondel@inma.ucl.ac.be, http://www.inma.ucl.ac.be/~blondel/

[2] Division of Applied Mathematics, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium, vdooren@csam.ucl.ac.be

Abstract. We introduce a concept of similarity between vertices of directed graphs. Let G_A and G_B be two directed graphs with respectively n_A and n_B vertices. We define a $n_A \times n_B$ *similarity matrix* $\mathbf{S}$ whose real entry s_{ij} expresses how similar vertex i (in G_A) is to vertex j (in G_B) : we say that s_{ij} is their *similarity score*. In the special case where $G_A = G_B = G$, the score s_{ij} is the similarity score between the vertices i and j of G and the square similarity matrix $\mathbf{S}$ is the *self-similarity matrix* of the graph G. We point out that Kleinberg's "hub and authority" method to identify web-pages relevant to a given query can be viewed as a special case of our definition in the case where one of the graphs has two vertices and a unique directed edge between them. In analogy to Kleinberg, we show that our similarity scores are given by the components of a dominant vector of a non-negative matrix and we propose a simple iterative method to compute them.

Remark: Due to space limitations we have not been able to include proofs of the results presented in this paper. Interested readers are referred to the full version of the paper [2], and to [3] for a description of an application of our similarity concept to the automatic extraction of synonyms in a dictionary. Both references are available from the first author web-site.

1 Generalizing hubs and authorities

Efficient web search engines such as Google are often based on the idea of characterizing the most important vertices in a graph representing the connections or links between pages on the web. One such method, proposed by Kleinberg [12], identifies in a set of pages relevant to a query search those that are good *hubs* or good *authorities*. For example, for the query "automobile makers", the home-pages of Ford, Toyota and other car makers are good authorities, whereas web pages that list these home-pages are good hubs. Good hubs are those that point to good authorities, and good authorities are those that are pointed to by good hubs. From these implicit relations, Kleinberg derives an iterative method that assigns an "authority score" and a "hub score" to every vertex of a given graph. These scores can be obtained as the limit of a converging iterative process which we now describe.

Let G be a graph with edge set E and let h_j and a_j be the hub and authority scores of the vertex j. We let these scores be initialized by some positive values and then update them simultaneously for all vertices according to the following *mutually reinforcing relation* : the hub score of vertex j is set equal to the sum of the authority scores of all vertices pointed to by j and, similarly, the authority score of vertex j is set equal to the sum of the hub scores of all vertices pointing to j :

$$\begin{cases} h_j \leftarrow \sum_{i:(j,i)\in E} a_i \\ a_j \leftarrow \sum_{i:(i,j)\in E} h_i \end{cases}$$

Let B be the adjacency matrix of G and let h and a be the vectors of hub and authority scores. The above updating equations take the simple form

$$\begin{bmatrix} h \\ a \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} h \\ a \end{bmatrix}_k, \qquad k = 0, 1, \dots$$

which we denote in compact form by

$$x_{k+1} = M\, x_k, \qquad k = 0, 1, \dots$$

where

$$x_k = \begin{bmatrix} h \\ a \end{bmatrix}_k, \quad M = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}.$$

We are only interested in the relative scores and we will therefore consider the *normalized* vector sequence

$$z_0 = x_0, \quad z_{k+1} = \frac{M z_k}{\|M z_k\|_2}, \quad k = 0, 1, \dots$$

Ideally, we would like to take the limit of the sequence z_k as a definition for the hub and authority scores. There are two difficulties with such a definition. Firstly, the sequence does not always converge. In fact, non-negative matrices M with the above block structure always have *two* real eigenvalue of largest magnitude and the resulting sequence z_k almost never converges. Notice however that the matrix M^2 is symmetric and so, even though the sequence z_k may not converge, the even and odd sub-sequences do converge. Let us define

$$z_{even} = \lim_{k\to\infty} z_{2k} \quad \text{and} \quad z_{odd} = \lim_{k\to\infty} z_{2k+1}.$$

and let us consider both limits for the moment. The second difficulty is that the limit vectors z_{even} and z_{odd} do in general depend on the initial vector z_0 and there is no apparent natural choice for z_0. In Theorem 2, we define the set of all limit vectors obtained when starting from a positive initial vector

$$Z = \{z_{even}(z_0), z_{odd}(z_0) : z_0 > 0\}.$$

and prove that the vector z_{even} obtained for $z_0 = \mathbf{1}$ is the vector of largest possible 1-norm among all vectors in Z (throughout this paper we denote by $\mathbf{1}$ the vector, or matrix, whose entries are all equal to 1; the appropriate dimension of $\mathbf{1}$ is always clear from the context). Because of this extremal property, we take the two sub-vectors of $z_{even}(\mathbf{1})$ as definitions for the hub and authority scores. In the case of the above matrix M, we have

$$M^2 = \begin{bmatrix} BB^T & 0 \\ 0 & B^TB \end{bmatrix}$$

and from this it follows that, if the dominant invariant subspaces associated to B^TB and BB^T have dimension one, then the normalized hub and authority scores are simply given by the normalized dominant eigenvectors of B^TB and BB^T, respectively. This is the definition used in [12] for the authority and hub scores of the vertices of G. The arbitrary choice of $z_0 = \mathbf{1}$ made in [12] is given here an extremal norm justification. Notice that when the invariant subspace has dimension one, then there is nothing special about the starting vector $\mathbf{1}$ since any other positive vector z_0 would give the same result.
We now generalize this construction. The authority score of the vertex j of G can be seen as a similarity score between j and the vertex *authority* in the graph

$$\textit{hub} \longrightarrow \textit{authority}$$

and, similarly, the hub score of j can be seen as a similarity score between j and the vertex *hub*. The mutually reinforcing updating iteration used above can be generalized to graphs that are different from the hub-authority structure graph. The idea of this generalization is quite simple; we illustrate it in this introduction on the path graph with three vertices and provide a general definition for arbitrary graphs in Section 3. Let G be a graph with edge set E and adjacency matrix B and consider the *structure graph*

$$1 \longrightarrow 2 \longrightarrow 3.$$

To the vertex j of G we associate three scores x_{j1}, x_{j2} and x_{j3}; one for each vertex of the structure graph. We initialize these scores at some positive value and then update them according to the following mutually reinforcing relations

$$\begin{cases} x_{j1} \leftarrow & & \sum_{i:(j,i)\in E} x_{i2} & \\ x_{j2} \leftarrow & \sum_{i:(i,j)\in E} x_{i1} & & + \sum_{i:(j,i)\in E} x_{i3} \\ x_{j3} \leftarrow & & \sum_{i:(i,j)\in E} x_{i2} & \end{cases}$$

or, in matrix form (we denote by x_i the column vector with entries x_{ji}),

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & B & 0 \\ B^T & 0 & B \\ 0 & B^T & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_k, \qquad k = 0, 1, \ldots$$

which we again denote $x_{k+1} = Mx_k$. The situation is now identical to that of the previous example and all convergence arguments given there apply here as well. The matrix M^2 is symmetric and non-negative, the normalized even and odd iterates converge, and the limit $z_{even}(\mathbf{1})$ is among all possible limits one that has largest possible 1-norm. We take the three components of this extremal limit $z_{even}(\mathbf{1})$ as definition for the similarity scores[3] s_1, s_2 and s_3 and define the similarity matrix by

$$\mathbf{S} = [s_1\ s_2\ s_3].$$

The rest of this paper is organized as follows. In Section 2, we describe some standard Perron-Frobenius results for non-negative matrices that will be useful in the rest of the paper. In Section 3, we give a precise definition of the similarity matrix together with different alternative definitions. The definition immediately translates into an approximation algorithm. In Section 4 we describe similarity matrices for the situation where one of the two graphs is a path graph; path graphs of lengths 2 and 3 are those that are discussed in this introduction. In Section 5, we consider the special case $G_A = G_B = G$ for which the score s_{ij} is the similarity between the vertices i and j in the graph G. Section 6 deals with graphs for which all vertices play the same rôle. We prove that, as expected, the similarity matrix in this case has rank one.

2 Graphs and non-negative matrices

With any directed graph $G = (V, E)$ one can associate a non-negative matrix via an indexation of its vertices. The so-called *adjacency matrix* of G is the matrix $B \in \mathbb{N}^{n \times n}$ whose entry d_{ij} equals the number of edges from vertex i to vertex j. Conversely, a square matrix B whose entries are non-negative integer numbers, defines a directed graph G with d_{ij} edges between i and j. Let B be the adjacency matrix of some graph G; the entry $(B^k)_{ij}$ is equal to the number of paths of length k from vertex i to vertex j. From this it follows that a graph is strongly connected if and only if for every pair of indices i and j there is an integer k such that $(B^k)_{ij} > 0$. Matrices that satisfy this property are said to be *irreducible*.

The Perron-Frobenius theory [10] establishes interesting properties about the eigenvectors and eigenvalues for non-negative and irreducible matrices. Let us denote the spectral radius[4] of the matrix C by $\rho(C)$. The following results follow from [10, **?**].

Theorem 1. *Let C be a non-negative matrix. Then*
(i) the spectral radius $\rho(C)$ is an eigenvalue of C – called the Perron root – and there exists an associated non-negative vector $x \geqslant 0$ $(x \neq 0)$ – called the Perron vector – such that $Cx = \rho x$

[3] In Section 4, we prove that the "central similarity score" s_2 can be obtained more directly from B by computing the dominating eigenvector of the matrix $BB^T + B^T B$.
[4] The spectral radius of a matrix is the largest magnitude of its eigenvalues.

(ii) if C is irreducible, then the algebraic multiplicity of the Perron root ρ is equal to one and there is a positive vector $x > 0$ such that $Cx = \rho x$
(iii) if C is symmetric, then the algebraic and geometric multiplicity of the Perron root ρ are equal and there is a non-negative basis $X \geqslant 0$ associated to the invariant subspace associated to ρ, such that $CX = \rho X$.

In the sequel, we shall also need the notion of orthogonal projection on subspaces. Let $\mathcal{V}$ be a linear subspace of $\mathbb{R}^n$ and let $v \in \mathbb{R}^n$. The *orthogonal projection* of v on $\mathcal{V}$ is the vector in $\mathcal{V}$ with smallest distance to v. The matrix representation of this projection is obtained as follows. Let $\{v_1, \ldots, v_m\}$ be an orthogonal basis for $\mathcal{V}$ and arrange these column vectors in a matrix V. The projection of v on $\mathcal{V}$ is then given by $\Pi v = VV^T v$ and the matrix $\Pi = VV^T$ is the *orthogonal projector* on $\mathcal{V}$. From the previous theorem it follows that, if the matrix C is non-negative and symmetric, then the elements of the orthogonal projector Π on the vector space associated to the Perron root of C are all non-negative.
The next theorem will be used to justify our definition of similarity matrix between two graphs. The result describes the limits points of sequences associated to symmetric non-negative linear transformations.

Theorem 2. *Let M be a symmetric non-negative matrix of spectral radius ρ. Let $z_0 > 0$ and consider the sequence*

$$z_{k+1} = Mz_k / \|Mz_k\|_2, \quad k = 0, \ldots$$

Then the subsequences z_{2k} and z_{2k+1} converge to the limits

$$z_{even}(z_0) = \lim_{k \to \infty} z_{2k} = \frac{\Pi z_0}{\|\Pi z_0\|_2} \quad \text{and} \quad z_{odd}(z_0) = \lim_{k \to \infty} z_{2k+1} = \frac{\Pi M z_0}{\|\Pi M z_0\|_2},$$

where Π is the orthogonal projector on the invariant subspace of M^2 associated to its Perron root ρ^2. In addition to this, the set of all possible limits is given by:

$$Z = \{z_{even}(z_0), z_{odd}(z_0) : z_0 > 0\} = \{\Pi z / \|\Pi z\|_2 : z > 0\}$$

and the vector $z_{even}(\mathbf{1})$ is the unique vector of largest 1-norm in that set.

3 Similarity between vertices in graphs

We now introduce our definition of graph similarity for arbitrary graphs. Let G_A and G_B be two directed graphs with respectively n_A and n_B vertices. We think of G_A as a "structure graph" that plays the role of the graphs *hub* $\longrightarrow$ *authority* and $1 \longrightarrow 2 \longrightarrow 3$ in the introductory examples. Let $\mathrm{pre}(v)$ (respectively $\mathrm{post}(v)$) denote the set of ancestors (respectively descendants) of the vertex v. We consider real scores x_{ij} for $i = 1, \ldots, n_B$ and $j = 1, \ldots, n_A$ and simultaneously update all scores according to the following updating equations

$$[x_{ij}]_{k+1} = \sum_{r \in \mathrm{pre}(i),\ s \in \mathrm{pre}(j)} [x_{rs}]_k + \sum_{r \in \mathrm{post}(i),\ s \in \mathrm{post}(j)} [x_{rs}]_k \qquad (1)$$

These equations coincide with those given in the introduction. The equations can be written in more compact matrix form. Let X_k be the $n_B \times n_A$ matrix of entries $[x_{ij}]_k$. Then (1) takes the form

$$X_{k+1} = BX_kA^T + B^TX_kA, \qquad k = 0, 1, \ldots \tag{2}$$

where A and B are the adjacency matrices of G_A and G_B. In this updating equation, the entries of X_{k+1} depend linearly on those of X_k. We can make this dependance more explicit by using the matrix-to-vector operator that develops a matrix into a vector by taking its columns one by one. This operator, denoted vec, satisfies the elementary property $\mathrm{vec}(CXD) = (D^T \otimes C)\ \mathrm{vec}(X)$ in which $\otimes$ denotes the Kronecker tensorial product (for a proof of this property, see Lemma 4.3.1 in [11]). Applying this property to (2) we immediately obtain

$$x_{k+1} = (A \otimes B + A^T \otimes B^T)\ x_k \tag{3}$$

where $x_k = \mathrm{vec}(X_k)$. This is the format used in the introduction. Combining this observation with Theorem 2 we deduce the following property for the normalized sequence Z_k.

Corollary 1. *Let G_A and G_B be two graphs with adjacency matrices A and B, fix some initial positive matrix $Z_0 > 0$ and define*

$$Z_{k+1} = \frac{BZ_kA^T + B^TZ_kA}{\|BZ_kA^T + B^TZ_kA\|_2} \qquad k = 0, 1, \ldots .$$

Then, the matrix subsequences Z_{2k} and Z_{2k+1} converge to Z_{even} and Z_{odd}. Moreover, among all the matrices in the set

$$\{Z_{even}(Z_0), Z_{odd}(Z_0) : Z_0 > 0\}$$

the matrix $Z_{even}(\mathbf{1})$ is the unique matrix of largest 1-norm.

In order to be consistent with the vector norm appearing in Theorem 2, the matrix norm $\|.\|_2$ we use here is the square root of the sum of all squared entries (this norm is known as the Euclidean or Frobenius norm), and the 1-norm $\|.\|_1$ is the sum of all entries magnitudes. In view of this result, the next definition is now justified.

Definition 1. *Let G_A and G_B be two graphs with adjacency matrices A and B. The similarity matrix between G_A and G_B is the matrix*

$$\mathbf{S} = \lim_{k \to +\infty} Z_{2k}$$

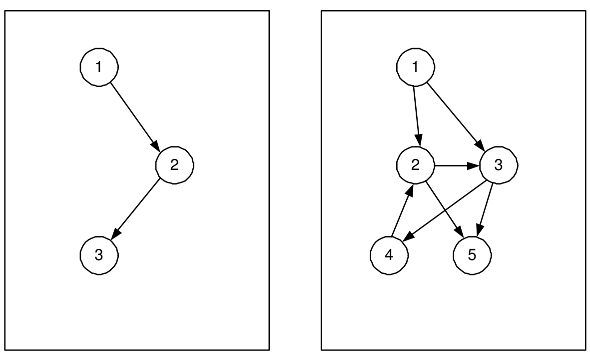
obtained for $Z_0 = \mathbf{1}$ and

$$Z_{k+1} = \frac{BZ_kA^T + B^TZ_kA}{\|BZ_kA^T + B^TZ_kA\|_2}, \qquad k = 0, 1, \ldots$$

A direct algorithmic transcription of the definition leads to an approximation algorithm. An example of a pair of graphs and their corresponding similarity matrix is given in Figure 3. Notice that it follows from the definition that the similarity matrix between G_B and G_A is the transpose of the similarity matrix between G_A and G_B. Similarity matrices can alternatively be defined as the projection of the matrix $\mathbf{1}$ on an invariant subspace associated to the graphs and for particular classes of adjacency matrices, one can compute the similarity matrix $\mathbf{S}$ directly from the dominant invariant subspaces of matrices of the size of A or B; we provide explicit expressions for a few classes in the next sections. Similarity matrices can also be defined by their extremal property.

Corollary 2. *The similarity matrix of the graphs G_A and G_B of adjacency matrices A and B is the unique matrix of largest 1-norm among all matrices X that maximize the expression*

$$\frac{\|BXA^T + B^TXA\|_2}{\|X\|_2}. \tag{4}$$



$$\mathbf{S} = \begin{bmatrix} 0.31 & 0.14 & 0 \\ 0.19 & 0.55 & 0.06 \\ 0.06 & 0.55 & 0.19 \\ 0.15 & 0.06 & 0.15 \\ 0 & 0.14 & 0.31 \end{bmatrix}$$

Fig. 1. Two graphs G_A and G_B and their corresponding similarity matrix $\mathbf{S}$. As an illustration, the similarity score between vertex 2 of graph G_A and vertex 3 of graph G_B is equal to 0.55.

4 Hubs, authorities, central scores and path graphs

As explained in the introduction, the hub and authority scores of a graph G_B can be expressed in terms of the adjacency matrix of G_B.

Theorem 3. *Let B be the adjacency matrix of the graph G_B. The normalized hub and authority scores of the vertices of G_B are given by the normalized dominant eigenvectors of the matrices B^TB and BB^T, provided the corresponding Perron root is of multiplicity 1. Otherwise, it is the normalized projection of the vector $\mathbf{1}$ on the respective dominant invariant subspaces.*

The condition on the multiplicity of the Perron root is not superfluous. Indeed, even for strongly connected graphs, BB^T and B^TB may have multiple dominant roots: for cycle graph for example, both BB^T and B^TB are the identity matrix. Another interesting structure graph is the path graph of length three:

$$1 \longrightarrow 2 \longrightarrow 3$$

Similarly to the hub and authority scores, the resulting similarity score with vertex 2, a score that we call *central score*, can be given an explicit expression.

Theorem 4. *Let B be the adjacency matrix of the graph G_B. The normalized central scores of the vertices of G_B are given by the normalized dominant eigenvector of the matrix*

$$B^TB + BB^T,$$

provided the corresponding Perron root is of multiplicity 1. Otherwise, it is the normalized projection of the vector $\mathbf{1}$ *on the dominant invariant subspace.*

The above structure graphs are path graphs of length 2 and 3. For path graphs of arbitrary length ℓ we have:

Corollary 3. *Let B be the adjacency matrix of the graph G_B. Let G_A be the path graph of length ℓ:*

$$G_A \;:\quad 1 \longrightarrow 2 \longrightarrow \cdots \longrightarrow \ell.$$

Then the odd and even columns of the similarity matrix $\mathbf{S}$ *can be computed independently as the projection of* $\mathbf{1}$ *on the dominant invariant subspaces of EE^T and E^TE where*

$$E = \begin{bmatrix} B & & & \\ B^T & \ddots & & \\ & \ddots & B & \\ & & B^T & B \end{bmatrix} \quad \text{or} \quad E = \begin{bmatrix} B & & \\ B^T & \ddots & \\ & \ddots & B \\ & & B^T \end{bmatrix}$$

for ℓ even and ℓ odd, respectively.

5 Self-similarity matrix of a graph

When we compare two equal graphs $G_A = G_B = G$, the similarity matrix $\mathbf{S}$ is a square matrix whose entries are similarity scores between vertices of G; this matrix is the *self-similarity matrix* of G. Various graphs and their corresponding self-similarity matrices are represented in Figure 2. In general, we expect vertices to have a high similarity score with themselves; that is, we expect the diagonal entries of self-similarity matrices to be large. We prove in the next theorem that the largest entry of a self-similarity matrix always appear on the diagonal and that, except for trivial cases, the diagonal elements of a self-similarity matrix are non-zero. As is shown with the last graph of Figure 2, it is however not true that diagonal elements dominate all elements on the same line and column.

Theorem 5. *The self-similarity matrix of a graph is positive semi-definite. In particular, the largest element of the matrix always appears on diagonal, and if a diagonal entry is equal to zero, then the corresponding line and column are equal to zero.*

For some classes of graphs, similarity matrices can be computed explicitly. We have for example:

Theorem 6. *The self-similarity matrix of the path graph of length ℓ is a diagonal matrix with diagonal elements equal to* $\sin(j\pi/(\ell+1)),\ j = 1, \ldots, \ell$.

When vertices of a graph are similar to each other, such as in cycle graphs, we expect to have a self-similarity matrix whose entries are all equal. This is indeed the case. Let us recall here that a graph is said to be *vertex-transitive* (or *vertex symmetric*) if all vertices play the same rôle in the graph. More formally, a graph G of adjacency matrix A is vertex-transitive if associated to any pair of vertices i, j, there is a permutation matrix T that satisfies $T(i) = j$ and $T^{-1}AT = A$.

Theorem 7. *All entries of the self-similarity matrix of a vertex-transitive graph are equal to* $1/n$.

6 Graphs whose vertices are symmetric to each other

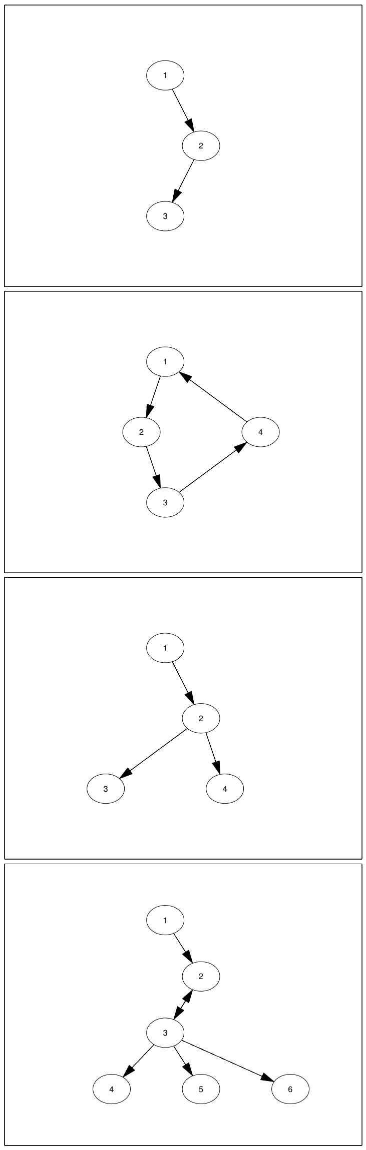
We now analyze properties of the similarity matrix when one of the two graphs has all its vertices symmetric to each other, or has an adjacency matrix that is normal. We prove that in both cases the resulting similarity matrix has rank one.

Theorem 8. *Let G_A, G_B be two graphs and assume that G_A is vertex-transitive. Then the similarity matrix between G_A and G_B is a rank one matrix of the form*

$$\mathbf{S} = \alpha\, \mathbf{1} v^T$$

where $v = \Pi \mathbf{1}$ is the projection of $\mathbf{1}$ on the dominant invariant subspace of $(B + B^T)^2$ and α is the scaling factor $\alpha = 1/\|\mathbf{1}v^T\|$. In particular, if G_A and G_B are both vertex symmetric then the entries of their similarity matrix are all equal to $1/\sqrt{n_A n_B}$.

Cycle graphs have an adjacency matrix A that satisfies $AA^T = I$. This property corresponds to the fact that, in a cycle graph, all forward-backward paths from a vertex return to that vertex. More generally, we consider in the next theorem graphs that have an adjacency matrix A that is normal, i.e., such that $AA^T = A^T A$. In particular, graphs that have a symmetric adjacency matrix satisfy this property. We prove below that when one of the graphs has a normal adjacency matrix, then the similarity matrix has rank one and we provide an explicit expression for this matrix.



$$\begin{bmatrix} 0.408 & 0 & 0 \\ 0 & 0.816 & 0 \\ 0 & 0 & 0.408 \end{bmatrix}$$

$$\begin{bmatrix} 0.250 & 0.250 & 0.250 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.250 \end{bmatrix}$$

$$\begin{bmatrix} 0.182 & 0 & 0 & 0 \\ 0 & 0.912 & 0 & 0 \\ 0 & 0 & 0.182 & 0.182 \\ 0 & 0 & 0.182 & 0.182 \end{bmatrix}$$

$$\begin{bmatrix} 0.103 & 0 & 0.251 & 0 & 0 & 0 \\ 0 & 0.202 & 0 & 0.096 & 0.096 & 0.096 \\ 0.251 & 0 & 0.845 & 0 & 0 & 0 \\ 0 & 0.096 & 0 & 0.074 & 0.074 & 0.074 \\ 0 & 0.096 & 0 & 0.074 & 0.074 & 0.074 \\ 0 & 0.096 & 0 & 0.074 & 0.074 & 0.074 \end{bmatrix}$$

Fig. 2. Graphs and their corresponding self-similarity matrices.

Theorem 9. *Let G_A and G_B be two graphs and assume that A is a normal matrix. Then the similarity matrix between G_A and G_B is a rank one matrix $\mathbf{S} = uv^T$ where*

$$u = \frac{(\Pi_{+\alpha} + \Pi_{-\alpha})\mathbf{1}}{\|(\Pi_{+\alpha} + \Pi_{-\alpha})\mathbf{1}\|_2}, \quad v = \frac{\Pi_\beta \mathbf{1}}{\|\Pi_\beta \mathbf{1}\|_2}.$$

In this expression α is the Perron root of A, $\Pi_{+\alpha}$, $\Pi_{-\alpha}$ are the projectors on its invariant subspaces corresponding to the eigenvalues $+\alpha$ and $-\alpha$, β is the Perron root of $(B + B^T)$, and Π_β is the projector on the invariant subspace of $(B + B^T)^2$ corresponding to the eigenvalue β^2.

When one of the graphs G_A or G_B is vertex-transitive or has a normal adjacency matrix, the resulting similarity matrix $\mathbf{S}$ has rank one. Adjacency matrices of vertex-transitive graphs and normal matrices have the property that the projector $\Pi_{+\alpha}$ on the invariant subspace corresponding to the Perron root of A is also the projector on the subspace of A^T (and similarly for $-\alpha$). We conjecture here that the similarity matrix can only be of rank one if either A or B have this property.

7 Concluding remarks

Investigations of properties and applications of the similarity matrix of graphs can be pursued in several directions. We outline here some possible research directions.

One natural extension of our concept is to consider networks rather than graphs; this amounts to consider adjacency matrices with arbitrary real entries and not just integers. The definitions and results presented in this paper use only the property that the adjacency matrices involved have non-negative entries, and so all results remain valid for networks with non-negative weights. The extension to networks makes a sensitivity analysis possible: How sensitive is the similarity matrix to the weights in the network? Experiments and qualitative arguments show that, for most networks, similarity scores are almost everywhere continuous functions of the network entries. Perhaps this can be analyzed for models for random graphs such as those that appear in [4]? These questions can probably also be related to the large literature on eigenvalues and eigenspaces of graphs; see, e.g., [5], [6] and [7].

More specific questions on the similarity matrix also arise. One open problem is to characterize the pairs of matrices that give rise to a rank one similarity matrix. The structure of these pairs is conjectured at the end of Section 6. Is this conjecture correct? A long-standing graph question also arise when trying to characterize the graphs whose similarity matrices have only positive entries. The positive entries of the similarity matrix between the graphs G_A and G_B can be obtained as follows. One construct the product graph, symmetrize it, and then identify in the resulting graph the connected component(s) of largest possible Perron root. The indices of the vertices in that graph correspond exactly to the nonzero entries in the similarity matrix of G_A and G_B. The entries of the

similarity matrix will thus be all positive if and only if the product graph of G_A and G_B is weakly connected. The problem of characterizing all pairs of graphs that have a weakly connected product was introduced and analyzed in 1966 in [8]. The problem of efficiently characterizing all pairs of graphs that have a weakly connected product is a problem that is still open.
Another topic of interest is to investigate how the concepts proposed here can be used, possibly in modified form, for evaluating the similarity between two graphs, for clustering vertices or graphs, for pattern recognition in graphs or for data mining purposes.

Acknowledgment Three of our students, Maureen Heymans, Anahí Gajardo and Pierre Senellart have provided inputs on several ideas developed in this paper. We are pleased to acknowledge the inputs of all these students.
This paper presents research supported by NSF under Grant No. CCR 99-12415 and by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture.

References

1. Norman Biggs, *Algebraic Graph Theory (Second Edition)*, Cambridge Mathematical Library, Cambridge University Press, 1993.
2. Vincent D. Blondel, Paul Van Dooren, A measure of similarity between graph vertices. With applications to synonym extraction and web searching. Technical Report UCL 02-50, submitted to journal, 2002.
3. Vincent D. Blondel, Pierre P. Senellart, Automatic extraction of synonyms in a dictionary, Technical report 2001-89, Université catholique de Louvain, Louvain-la-Neuve, Belgium. Also: Proceedings of the SIAM Text Mining Workshop, Arlington (Virginia, USA) April 11, 2002.
4. B. Bollobas, Random Graphs, Academic Press, 1985.
5. Fan R. K. Chung, Spectral Graph Theory, American Mathematical Society, 1997.
6. Dragoš Cvetković, Peter Rowlinson, Slobodan Simić, Eigenspaces of graphs, Cambridge University Press, 1997.
7. Dragoš Cvetković, M. Doob, H. Sachs, Spectra of graphs – Theory and applications (Third edition), Johann Ambrosius Barth Verlag, 1995.
8. Frank Harary, C. Trauth, Connectedness of products of two directed graphs, J. SIAM Appl. Math., 14, pp. 150-154, 1966.
9. Hartfiel, D. J. and Maxson, C. J., The chainable matrix, a special combinatorial matrix, Discrete Math., Volume 12, pp. 245–256, 1975.
10. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, London, 1985.
11. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, London, 1991.
12. Jon M. Kleinberg, Authoritative sources in a hyperlinked environment, Journal of the ACM, 46:5, pp. 604–632, 1999.
13. Pierre P. Senellart, Vincent D. Blondel, Automatic discovery of similar words. To appear in "A Comprehensive Survey of Text Mining", Springer-Verlag, 2003.