Ranking nodes in directed networks via continuous-time quantum walks

Paola Boito[1] and Roberto Grena[2*]

[1]Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, Pisa, 56127, Italy.
[2*]C. R. ENEA Casaccia, via Anguillarese 301, Roma, 00123, Italy.

*Corresponding author(s). E-mail(s): roberto.grena@enea.it;
Contributing authors: paola.boito@unipi.it;

Abstract

Four new centrality measures for directed networks based on unitary, continuous-time quantum walks (CTQW) in n dimensions – where n is the number of nodes – are presented, tested and discussed. The main idea behind these methods consists in re-casting the classical HITS and PageRank algorithms as eigenvector problems for symmetric matrices, and using these symmetric matrices as Hamiltonians for CTQWs, in order to obtain a unitary evolution operator. The choice of the initial state is also crucial. Two options were tested: a vector with uniform occupation and a vector weighted w.r.t. in- or out-degrees (for authority and hub centrality, respectively). Two methods are based on a HITS-derived Hamiltonian, and two use a PageRank-derived Hamiltonian. Centrality scores for the nodes are defined as the average occupation values. All the methods have been tested on a set of small, simple graphs in order to spot possible evident drawbacks, and then on a larger number of artificially generated larger-sized graphs, in order to draw a comparison with classical HITS and PageRank. Numerical results show that, despite some pathologies found in three of the methods when analyzing small graphs, all the methods are effective in finding the first and top ten nodes in larger graphs. We comment on the results and offer some insight into the good accordance between classical and quantum approaches.

Keywords: Quantum walks, Network analysis, Node centrality

1 Introduction

Centrality measures for networks [1, 2] can be loosely defined as measures of importance of nodes. Notions of centrality have received considerable interest in the last decades, especially boosted by the important task of finding efficient algorithms for ranking web pages in search engines. Two classical examples of algorithms for computing centrality scores are HITS [3] and PageRank [4].

In directed networks, nodes can be ranked according to their importance as *hubs* or *authorities*. Roughly speaking, a strong hub is a node that points to many important nodes and a strong authority is a node that is pointed to by many important nodes. The notion of "important nodes" depends on the method: in HITS, which simultaneously provides hubs and authority scores, good hubs point to good authorities and good authorities are pointed to by good hubs, whereas in PageRank a node has a higher authority centrality score if it is pointed to by strong authorities.

Algorithms for webpage ranking usually compute authority centrality, but any authority method can be easily converted to a hub method by applying it to the reverse graph obtained by inverting the direction of each link. In matrix terms, if A is the adjacency matrix of the original graph, the reverse graph has adjacency matrix A^T. For instance, PageRank is an authority ranking algorithm, but it can be easily converted to hub ranking (Reverse PageRank, see e.g. [5, 6]) by applying it to A^T instead of A. HITS, on the other hand, is fundamentally a power method that alternates between A and A^T and yields both hub and authority rankings.

An interesting feature of PageRank is the fact that it can be seen as a random-walk problem: the Google matrix G built from A is a stochastic matrix, thus representing a random walk, and the ranking score of a node is the asymptotic probability of finding the walker in that node. This is an example of the important role that random walks play in many ranking methods.

The increasing development of research in quantum computation has sparked an interest in finding quantum-formulated ranking algorithms for networks. The relation between node-ranking algorithms and classical random walks suggests that it may be useful to explore the theory of quantum walks on networks, for ranking purposes. On the theory of quantum walks, we mention, among others, the seminal papers [7–10] and the review [11]. A good introduction can be found, for instance, in the recent book [12]. Among other features, quantum walks exhibit faster diffusion w.r.t. classical random walks, thus allowing for algorithmic speedup [13, 14].

In contrast to classical random walks, unitary quantum walks typically do not converge to a stationary distribution, because of unitary evolution. Therefore, time-averaging is expected to be required for the extraction of centrality scores, unless one resorts to the use of mixed quantum-classical evolution [15] or tunable time-dependent Hamiltonian for adiabatic computations [16]. Another remarkable feature is that the average occupation of each node may depend on the initial state, i.e., the prescription of the initial state is part of the ranking method.

Like classical random walks, quantum walks can be defined either in a discrete-time (DTQW) or a continuous-time (CTQW) framework [17, 18]. Here we focus on the continuous-time case, which has the merit of taking place in a Hilbert space of dimension equal to the number of nodes, whereas DTQW typically require a dimension of the order of the number of edges.

The formulation of a quantum method for centrality measures based on DTQW or CTQW has been explored by several authors [19–27]. [20] proposes a method, called *Quantum PageRank*, based on a DTQW that evolves in a Hilbert space of dimension n^2, where n is the number of nodes. In CTQW-based methods such as [21, 22], the quantum walk takes place in a Hilbert space of dimension n; however, such methods are applicable to undirected graphs only. The authors of [23] have circumvented the difficulty of obtaining an evolution operator from a non-symmetric A by forgoing unitarity, and adopting instead non-standard PT-symmetric Hamiltonians.

In a recent work [28], three unitary-CTQW-based centrality measures for directed networks were introduced and discussed. Unitarity was obtained by defining the CTQW on the associated bipartite graph, following the idea behind [29], hence with a number of degrees of freedom equal to $2n$. One of the methods proved to be especially robust and its rankings were very well related to HITS rankings; however, no satisfactory results were obtained when trying to design an algorithm well-related to PageRank. Moreover, the doubling of the dimension of the Hilbert space is a drawback, only partially mitigated by the fact that two of the three proposed methods compute hub- and authority-centrality in a single run.

However, doubling the dimension of the Hilbert space could be unnecessary. Classical HITS and PageRank methods can both be reframed as eigenvector problems on $n \times n$ symmetric matrices. This fact suggests the idea of using such matrices as Hamiltonians to define a unitary n-dimensional CTQW on the network. Rankings obtained by such a CTQW can be expected to yield results similar to HITS and PageRank, respectively.

In this paper we explore this possibility, proposing two evolution operators for CTQW in n dimensions, based on a (slightly modified) HITS-derived Hamiltonian and on a PageRank-derived Hamiltonian. As mentioned above, the initial state must be supplied as well, in order to properly define the method. Two choices were studied: a uniformly occupied initial state and an initial state with occupation numbers weighted w.r.t. node degrees (in-degrees for authority centrality or out-degrees for hub centrality). On the whole, we propose four new algorithms, combining the two choices for the Hamiltonian and the two initial states.

Section 2 provides the background to understand the problem and recalls the HITS and PageRank algorithms, formulating them as eigenvector problems on symmetric $n \times n$ matrices. It also outlines the main ideas behind quantum centrality algorithms. Our CTQW-based methods are presented in Section 3. Section 4 is devoted to numerical experiments: tests on simple toy models, which allow us to discuss certain features of the methods, are followed by tests

on more than 5000 larger graphs of different sizes (from 128 to 1024 nodes). These tests are designed to check whether the obtained rankings are in good accordance with the results of the classical methods they are derived from. Three indicators are used: the probability of finding the same top-ranked node, the number of common nodes in the top-ten ranking, and Kendall's τ [30]. Sections 5 and 6 contain a discussion of the results and concluding remarks.

2 Background

A *(directed) graph* $\mathcal{G}$ is a pair (V, E), where V is the set of nodes, labeled from 1 to n, and $E = \{(i, j) | i, j \in V\}$ is the set of edges. The graph is called *undirected* if $(i, j) \in E$ implies $(j, i) \in E$. The graphs we consider are weakly connected, unweighted and contain no loops (i.e., edges from one node to itself) or multiple edges.

The *adjacency matrix* associated with $\mathcal{G}$ is the $n \times n$ matrix $A = (A_{ij})$ such that $A_{ij} = 1$ if t$(i, j) \in E$, and $A_{ij} = 0$ otherwise. Clearly A is symmetric if and only if $\mathcal{G}$ is undirected.

The *in-degree* $\deg_{\text{in}}(i)$ of node i is the number of nodes pointing to node i, whereas its *out-degree* $\deg_{\text{out}}(i)$ is the number of nodes pointed to by node i. It holds $\deg_{\text{in}}(i) = (\mathbb{1}^T A)_i$ and $\deg_{\text{out}}(i) = (A\mathbb{1})_i$, where $\mathbb{1} \in \mathbb{R}^n$ is the vector with all entries equal to 1.

2.1 Classical centrality measures for directed graphs

Among the many centrality measures proposed in the literature for directed graphs, we recall two popular ones that inspired the present work. Both of them were originally developed for ranking web pages.

HITS (Hyperlink-Induced Topic Search) [3].
This is an iterative scheme that computes an authority score and a hub score for each node. Each iteration takes the form

$$x^{(k)} = A^T y^{(k-1)}, \qquad y^{(k)} = A x^{(k)}, \tag{1}$$

followed by normalization in 2-norm, where $x^{(k)}$ and $y^{(k)}$ are the vectors of authority and hub scores, respectively, at the k-th step. Authority and hub centralities are defined as the limit values of the corresponding scores for $k \to \infty$. The initial vector is usually chosen as $\frac{1}{\sqrt{n}}\mathbb{1}$.

This process can also be reframed as a power method, yielding the dominant eigenvector of the symmetric matrices $A^T A$ (for authorities) and AA^T (for hubs).

PageRank [4, 31].
Here we recall the PageRank method for authority scores, as originally conceived for web page ranking. Hub centralities can be obtained via *Reverse*

PageRank [5, 32], that is, PageRank applied to the "reversed graph" with adjacency matrix A^T.

In the PageRank algorithm, the adjacency matrix A of the graph is modified to obtain the so-called patched adjacency matrix $\tilde{A}$, which is row-stochastic:

$$\tilde{A}_{ij} = \begin{cases} A_{ij}/\deg_{\text{out}}(i) & \text{if } \deg_{\text{out}}(i) > 0, \\ 1/n & \text{otherwise.} \end{cases}$$

This can be seen as the transition matrix of a random walk: each entry $\tilde{A}_{ij}$ is the probability, for a walker placed at node i, to reach node j, with the assumption that from each dangling node − i.e., a node with zero out-degree − the walker can reach all other nodes with uniform probability. Next, one introduces a teleportation effect by adding a rank-one correction:

$$G = \alpha\tilde{A} + \frac{1-\alpha}{n}\mathbb{1}\mathbb{1}^T, \tag{2}$$

where $\alpha \in [0, 1]$ is a parameter, usually chosen as $\alpha = 0.85$. The Perron-Frobenius theorem ensures that 1 is the (simple) eigenvalue of maximum modulus for G^T. The associated eigenvector $x > 0$ such that $G^T x = x$ yields the PageRank scores, and it can be computed efficiently via the power method.

The Google matrix G as defined above yields authority rankings. For hub rankings, one can define a matrix G_h in the same way, starting from A^T instead of A. Note that $G_h \neq G^T$.

The PageRank method can be rewritten in terms of the symmetric matrix

$$H_G = (I - G)(I - G)^T.$$

Indeed, it is easily seen that the problem of finding the dominant eigenvector of G^T is equivalent to the problem of finding a vector in the null space of H_G. Obviously the dominant eigenvector of G^T belongs to the null space of H_G, since it has eigenvalue 1. On the other hand, if $x \in \text{Ker}(H_G)$, then $0 = x^T H_G x = ((I - G)^T x)^T (I - G)^T x = \|(I - G)^T x\|_2^2$, which implies $(I - G)^T x = 0$, i.e., x is an eigenvector of G with eigenvalue 1; therefore it is the dominant eigenvector. This fact is used, e.g., by [16] in order to redefine the PageRank method as a ground-state method suitable for a quantum adiabatic computation.

2.2 Quantum centrality

Much of the literature on quantum centrality methods relies on discrete-time quantum walks. One notable example is the (discrete) Quantum PageRank, here abbreviated as DQPR [20]. For a graph with n nodes, these methods typically work in an n^2-dimensional state space, as required by the coined or Szegedy formulations of the DTQW, although it can be proved that in DQPR the non-trivial component of the dynamics actually takes place in a subspace of dimension at most $2n$.

However, continuous-time quantum walks [33–35] have also been used to define notions of graph centrality, mostly for undirected networks [22]. A CTQW on a graph $\mathcal{G}$ with n vertices takes place in a complex Hilbert space of dimension n; a generic state of the system takes the form

$$|\psi\rangle = \sum_{j=1}^{n} a_j |j\rangle,$$

where $|j\rangle$, with $j = 1, \ldots, n$, is the basis state associated with a walker localized at node j, and the amplitudes $a_1, \ldots, a_n \in \mathbb{C}$ are such that $\sum_{j=1}^{n} |a_j|^2 = 1$. Note that it holds $a_j = \langle j | \psi \rangle$. The square modulus $|a_j|^2$ is the *occupation* of node j, i.e., the probability of finding the system in the state $|j\rangle$ after a measurement.

The time evolution of a CTQW on a graph is described by the Schrödinger equation

$$i\frac{\partial |\psi(t)\rangle}{\partial t} = H|\psi(t)\rangle, \tag{3}$$

where $|\psi(t)\rangle$ is the state of the system at time t. The Hamiltonian operator H encodes the structure of the graph and is often chosen as the graph Laplacian or the adjacency matrix, which are symmetric when $\mathcal{G}$ is undirected; see [36] for motivation and a comparison.

If H does not depend explicitly on time, the solution of (3) is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle, \qquad U(t) = \mathrm{e}^{-iHt}. \tag{4}$$

Here $U(t)$ is the unitary evolution operator of the system and $|\psi(0)\rangle$ is the initial state. Quantum properties of the CTQW include superposition of states and time reversibility; the latter is a consequence of the unitarity of the evolution operator and implies that there is no limit state.

For this reason, one typically resorts to limits of time averages (*limiting distribution*) in order to define centrality and communicability scores. For instance, in [22] the authors define a CTQW-based centrality measure for undirected graphs where the initial state $|\psi(0)\rangle$ is a uniform superposition of all vertex states, the evolution of the walk is modeled by equation (3) with $H = A$, and the centrality score of node j is

$$C_j = \lim_{T \to \infty} \frac{1}{T} \int_0^T |\langle j | \psi(t) \rangle|^2 dt. \tag{5}$$

The following result, recalled from [28] and adapted from [8], ensures the well-posedness of the above definition and gives an explicit, closed-form characterization of the time-average limit.

Theorem 1. *Let $U(t) = \mathrm{e}^{-iHt}$ be a unitary evolution operator in a n-dimensional Hilbert space, and denote as $\{\theta_j, |\phi_j\rangle\}_{j=1,\ldots,n}$, the eigenvalues and*

eigenstates of the time-independent Hamiltonian H. Recall that the eigenvalues and eigenstates of $U(t)$ are $\{\lambda_j(t) = \mathrm{e}^{-it\theta_j}, |\phi_j\rangle\}_{j=1,\dots,n}$. Let $|\psi(0)\rangle = \sum_{i=1}^{n} a_j|\phi_j\rangle$, with $a_1, \dots, a_n \in \mathbb{C}$, be the initial state of the system, written in the eigenstate basis, and denote as $|\psi(t)\rangle$ the state of the system at time t. Then for any state $|\xi\rangle$ it holds

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |\langle\xi|\psi(t)\rangle|^2 dt = \sum_{j,k \text{ with } \theta_j = \theta_k} a_j a_k^* \langle\xi|\phi_j\rangle\langle\phi_k|\xi\rangle, \qquad (6)$$

where the asterisk * denotes the complex conjugate.

Note that the limit depends on the initial state and on the eigenstates of H, but not on the eigenvalues of H, except for their multiplicity structure.

3 Centrality for directed networks based on unitary CTQW

It was mentioned in Section 2.1 that both HITS and PageRank for directed networks can be seen as eigenvector problems on certain symmetric matrices. This suggests that we can define unitary CTQW-based centrality scores derived from classical HITS and PageRank, by using suitable modifications of these symmetric matrices as Hamiltonians.

The new methods proposed here follow a similar general scheme as in [22]:

1. Prepare the initial state (the "walker") in a state $|\psi(0)\rangle$.
2. Propagate the walker for a sufficiently large time t:

$$|\psi(t)\rangle = \mathrm{e}^{-iHt}|\psi(0)\rangle. \qquad (7)$$

3. Compute the time-average probability distribution of finding the walker at each vertex j; this is the centrality score of node j:

$$C_j = \lim_{T \to \infty} \frac{1}{T} \int_0^T |\langle j|\psi(t)\rangle|^2 dt. \qquad (8)$$

These centrality measures are well-defined, as a consequence of Theorem 1. Differently from [22], however, H here is not the adjacency matrix; moreover, the initial state can be different from the equal superposition of all vertex states.

3.1 HITS–derived CTQW

HITS looks for the dominant eigenvector of $A^T A$ (authority) or AA^T (hubs). It is therefore quite natural to investigate the use of such matrices as Hamiltonians to define evolution operators for hub and authority rankings. We will

describe the method for authority rankings; for hub rankings, simply exchange A and A^T.

In fact, some difficulties arise if the matrix $A^T A$ is chosen as a Hamiltonian for authority rankings. Indeed, $A^T A$ might have a block-diagonal structure; for instance, it is diagonal for a path graph. In this case, the resulting CTQW will be "trapped" in separate subspaces. This means that the quantum walk cannot explore the whole space and the resulting rakings might not be meaningful. In the example of a path graph, with a diagonal Hamiltonian, there is no evolution at all of the occupation of the nodes, because the states $|j\rangle$ are eigenvectors of the Hamiltonian, and the resulting ranking vector will be equal to the initial state. This is clearly not a desirable behaviour. For this reason we borrow from PageRank the idea of adding a rank-one correction to the adjacency matrix:

$$\tilde{A} = \alpha A + \frac{1 - \alpha}{n} \mathbb{1}\mathbb{1}^T, \tag{9}$$

where $\alpha \in [0, 1]$ is a suitably chosen parameter. Now define the following $n \times n$ symmetric matrix:

$$H_H = \tilde{A}^T \tilde{A}. \tag{10}$$

The matrix H_H represents our Hamiltonian operator, which will be used to model a CTQW in a Hilbert space of dimension n according to (7). In our experiments we take $\alpha = 0.85$ as is often done for classical PageRank.

3.2 PageRank–derived CTQW

Direct PageRank (for authorities) seeks a vector in the 1-dimensional null space of $H_G = (I - G)(I - G)^T$, where G is defined as in Section 2.1. As above, we will describe our PageRank-like method for authorities, and hubs can be found replacing G with G_h.

Matrix G already contains corrections that ensure the irreducibility of H_G; so, no further corrections are introduced. The Hamiltonian is

$$H_G = (I - G)(I - G)^T. \tag{11}$$

The Hamiltonian H_G will be used to model a CTQW in a Hilbert space of dimension n according to (7). The parameter α that defines the matrix G is set to 0.85.

3.3 Initial state

As already mentioned, the choice of the initial state is required to properly define a quantum walk and its limiting distribution. In [28] it was shown that the same evolution operator may give remarkably different rankings, depending on the choice of the initial state. Two options have been considered here for

the initial state. The first one is simply a uniformly occupied state:

$$|\psi_u(0)\rangle = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} |k\rangle. \tag{12}$$

The second choice is an initial state where each node has occupation proportional to its in-degree (for authority centrality) or to its out-degree (for hub centrality):

$$|\psi_w(0)\rangle = \frac{1}{\sqrt{\sum_{k=1}^{n} \deg_{\text{in}}(k)}} \sum_{k=1}^{n} \sqrt{\deg_{\text{in}}(k)} |k\rangle \quad \text{for authority,}$$

$$|\psi_w(0)\rangle = \frac{1}{\sqrt{\sum_{k=1}^{n} \deg_{\text{out}}(k)}} \sum_{k=1}^{n} \sqrt{\deg_{\text{out}}(k)} |k\rangle \quad \text{for hub.} \tag{13}$$

This choice of initial state is mainly motivated by the fact that $|\psi_w(0)\rangle$ is at the same time easy to compute and strongly correlated with the HITS score vector. Numerical experiments [28] also confirm its effectiveness.

Combining the two Hamiltonians defined above with these two choices of the initial state, we obtain four quantum centrality methods:

- CQHITSu: Hamiltonian H_H (10), initial state $|\psi_u(0)\rangle$ (12);
- CQHITSw: Hamiltonian H_H (10), initial state $|\psi_w(0)\rangle$ (13);
- CQPRu: Hamiltonian H_G (11), initial state $|\psi_u(0)\rangle$ (12);
- CQPRw: Hamiltonian H_G (11), initial state $|\psi_w(0)\rangle$ (13).

4 Numerical experiments

The four proposed methods have been tested on small graphs, where clear hub / authority results can be expected, in order to spot possible drawbacks of the methods, and then tested on a large sample of larger graphs generated using the Python package NetworkX [37]. All the methods were implemented in Octave and run on a laptop equipped with a 4-core Intel i7-7500U 2.70GHz processor, with 16 GB of RAM. In this paper we do not attempt a quantum implementation of the proposed centrality measures: all the numerical tests involve classical computations. The limit in (8) is computed via Theorem 1.

This section summarizes the main results obtained for the various methods.

4.1 Small graphs

The small graphs chosen for the initial test are:

Example 1: path graph, i.e., a simple chain of nodes.
Example 2: diamond graph, i.e., a graph with a "main hub" (node 1) with $n-2$ outgoing edges directed towards nodes $2, 3, ..., n-1$, and a "main authority" (node n) with $n-2$ incoming edges from nodes $2, 3, ...n-1$.
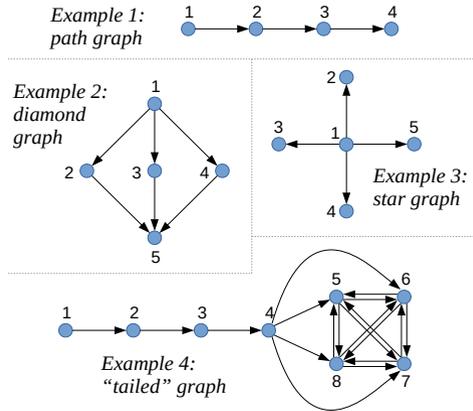
Fig. 1 Simple graphs used for testing the ranking methods (examples 1-4).

Example 3: star graph, i.e., a central hub (node 1) connected to all the other nodes, and no other edge present.

Example 4: tailed graph, i.e., n_1 initial nodes connected as in a path graph, followed by n_2 nodes that form a complete subgraph. Node n_1 is connected to all the n_2 nodes in the complete subgraph.

These graphs are shown in Figure 1. The computed centrality scores for Examples 1–3 and the rankings for Example 4 are shown in Tables 1–6. The rankings provided by the quantum methods are typically similar to the rankings given by their classical counterparts. However, some details are worth discussing.

Table 1 Hub scores for Example 1, with 4 nodes. Authority scores can be read by inverting the order of the nodes.

Node	CQHITSu	CQHITSw	HITS	CQPRu	CQPRw	PR
1	0.2683	0.3301	0.5774	0.4541	0.4479	0.3701
2	0.2683	0.3301	0.5774	0.2795	0.3147	0.2988
3	0.2683	0.3301	0.5774	0.1820	0.1636	0.2149
4	0.1952	0.0097	0.0000	0.0844	0.0737	0.1161

In [28], a "quantum pathology" affecting almost all the tested quantum methods (including DQPR) was spotted in small examples similar to these ones. When a graph contains "no-hub" nodes, i.e., nodes with zero out-degree, or "no-authority" nodes, any sensible ranking should place them at the bottom of the respective rankings. However, this is not always the case for quantum methods, and this problem sometimes also appears in the four centrality measures presented here. CQHITSu and CQPRu fail in Example 2: node 5, which has no outgoing edges, is placed second in hub rankings instead of being in the

Table 2 Hub scores for Example 2, with 5 nodes. Authority scores can be read by exchanging node 1 and node 5.

Node	CQHITSu	CQHITSw	HITS	CQPRu	CQPRw	PR
1	0.4055	0.4886	0.5000	0.5606	0.6787	0.4683
2	0.1400	0.1695	0.5000	0.0955	0.0879	0.1407
3	0.1400	0.1695	0.5000	0.0955	0.0879	0.1407
4	0.1400	0.1695	0.5000	0.0955	0.0879	0.1407
5	0.1746	0.0028	0.0000	0.1528	0.0578	0.1096

Table 3 Hub scores for Example 3, with 5 nodes.

Node	CQHITSu	CQHITSw	HITS	CQPRu	CQPRw	PR
1	0.2599	0.9906	1.00000	0.5685	0.7162	0.5238
2	0.1850	0.0023	0.00000	0.1079	0.0710	0.1190
3	0.1850	0.0023	0.00000	0.1079	0.0710	0.1190
4	0.1850	0.0023	0.00000	0.1079	0.0710	0.1190
5	0.1850	0.0023	0.00000	0.1079	0.0710	0.1190

Table 4 Authority scores for Example 3, with 5 nodes.

Node	CQHITSu	CQHITSw	HITS	CQPRu	CQPRw	PR
1	0.1850	0.0007	0.0000	0.1491	0.2484	0.1709
2	0.2037	0.2498	0.5000	0.2127	0.1879	0.2073
3	0.2037	0.2498	0.5000	0.2127	0.1879	0.2073
4	0.2037	0.2498	0.5000	0.2127	0.1879	0.2073
5	0.2037	0.2498	0.5000	0.2127	0.1879	0.2073

last position, and the same happens to node 1 for authority rankings. CQPRw fails in the authority ranking in Example 3: node 1, the central node of the star (with no incoming edges) is ranked first instead of last. No evident pathologies emerge from Examples 1 and 4.

The only method that seems to be immune to "quantum pathologies" is CQHITSw. In fact, the introduction of the initial weighted state was aimed at correcting this pathology; it is effective for HITS-derived methods, but the behaviour of PR-derived methods seems to be more erratic, since in Example 3 the pathology unexpectedly *appears* as an effect of the weighted initial vector.

We add some more remarks:

- In path graphs, different behaviours are observed for HITS, which isolates the last node (hub) or the first node (authority) placing the other ex-aequo, and for PR, which outputs descending hub scores and ascending authority scores, without ties. Both behaviours are acceptable; it is interesting to see which one is reproduced by the new methods. It appears that the two HITS-derived methods have the same behaviour as HITS, and the two PR-derived methods have the same behaviour as PR, as expected.
- Classical HITS shows a notable failure on Example 2 (it does not isolate the first-ranked node, both in hub and authority rankings) and a less significant

Table 5 Hub rankings for Example 4, with $n_1 = n_2 = 4$.

CQHITSu	4	1,2,3			5,6,7,8	
CQHITSw	4	5,6,7,8				1,2,3
HITS	4	5,6,7,8				1,2,3
CQPRu	1	2	3	4	5,6,7,8	
CQPRw	1	2	3	4	5,6,7,8	
PR	1	2	3	4	5,6,7,8	

Table 6 Authority rankings for Example 4, with $n_1 = n_2 = 4$.

CQHITSu	5,6,7,8	2,3,4			1
CQHITSw	5,6,7,8	2,3,4			1
HITS	5,6,7,8	1,2,3,4			
CQPRu	5,6,7,8	3	4	2	1
CQPRw	5,6,7,8	4	3	2	1
PR	5,6,7,8	4	3	2	1

Table 7 Number of scale-free graphs used in numerical tests and corresponding number of nodes.

Number of nodes	128	256	384	512	640	768	896	1024
Number of graphs	800	400	267	200	160	133	114	100

failure in Example 4 (node 1 is given the same authority as nodes 2, 3 and 4). In Example 2, these problems are likely linked to the degeneration of the dominant eigenvector. Note that the two CQHITS method do not suffer from these drawbacks and correctly spot the first-ranked node in Example 2 (and also the last-ranked node in Example 4), isolating them from the others.

- An interesting feature of Example 4 is the fact that classical PR and HITS output different hub rankings: PR places node 1 first, followed by the other nodes of the "tail" – thus emphasizing the fact that all the following nodes can be reached from them – while HITS privileges node 4, emphasizing the number of direct connections. It is noteworthy that the quantum versions reproduce this different behaviour.
- In Example 4, CQHITSu identifies the same top hub as HITS, but it places nodes 1-3 above nodes 5-8, contrarily to HITS (and CQHITSw); in the same example, there is a curious inversion of nodes 3 and 4 in the authority ranking for CQPRu.

4.2 Tests on larger graphs

4.2.1 Scale-free graphs

Here we test ranking algorithms on 2174 larger scale-free graphs authomatically generated by the `scale_free_graph` command in NetworkX [37]. These are directed graphs built according to a preferential attachment process. The number of graphs generated for each choice of the number of nodes is given in Table 7. Tests are repeated for hub and authority rankings, so each graph is used twice.

Table 8 F_1, F_{10} and Kendall's τ for scale-free graphs.

Nodes	CQHITSu			CQHITSw			CQPRu			CQPRw		
	F_1	F_{10}	τ	F_1	F_{10}	τ	F_1	F_{10}	τ	F_1	F_{10}	τ
128	0.841	6.22	0.0288	0.900	7.68	0.7000	0.917	8.27	0.3198	0.978	9.23	0.5658
256	0.804	6.69	0.0026	0.893	8.09	0.6940	0.933	8.29	0.2652	0.984	9.19	0.5225
384	0.768	6.54	0.0019	0.869	7.85	0.6697	0.901	8.26	0.2718	0.963	9.15	0.5118
512	0.758	6.67	-0.0205	0.855	7.87	0.6754	0.910	8.23	0.2321	0.970	9.16	0.4894
640	0.731	6.68	-0.0088	0.872	7.87	0.6671	0.922	8.07	0.2363	0.972	9.05	0.4850
768	0.695	6.42	-0.0024	0.831	7.56	0.6508	0.914	7.95	0.2520	0.970	8.96	0.4878
896	0.667	6.55	-0.0123	0.851	7.80	0.6523	0.882	7.89	0.2406	0.961	8.95	0.4750
1024	0.685	6.51	-0.0088	0.865	7.87	0.6545	0.920	7.78	0.2296	0.975	8.91	0.4705

Table 8 shows the comparison of the four centrality methods with their corresponding classical counterpart (HITS for CQHITSu and CQHITSw, PR for CQPRu and CQPRw), showing

- the fraction F_1 of tests in which the first node is the same as found in the corresponding classical algorithm,
- the average number F_{10} of the top 10 nodes that appear in both the quantum and classical rankings, and
- the Kendall's τ [30] of the overall ranking, w.r.t. the ranking obtained by the classical algorithm.

We see that HITS-derived algorithms perform quite well, with a significant improvement of CQHITSw over CQHITSu. CQHITSw finds the same first node as HITS in 83–90% of the tests, with an average coincidence of more than 7 nodes in the top 10. Results for CQHITSu are still acceptable but clearly worse. Note the marked difference in Kendall's τ parameter for the two methods: about 0 for CQHITSu, and above 0.65 for CQHITSw. The τ value for CQHITSu indicates that, although the method is generally able to find top-ranked nodes, the overall rankings are weakly related to HITS rankings. This is almost certainly due to the above-mentioned quantum pathology, which disrupts rankings in lower positions; see [28] for an extensive discussion. CQHITSw, on the contrary, exhibits the highest values of τ among the four methods.

PR-derived methods perform even better than HITS-derived methods in finding top-ranked nodes: these are correctly identified in about 90% of the cases for CQPRu and in more than 96% of the cases for CQPRw. The number of "correct" nodes among the top ten is also remarkable: about 8 for CQPRu and about 9 for CQPRw. Kendall's τ, however, is not high (usually below 0.3 for CQPRu and around 0.5 for CQPRw). The improvement achieved adopting a weighted initial vector is significant, but it is less marked than for HITS-derived methods. A possible explanation might come from the fact that the HITS ranking is known to be strongly correlated to the simple degree ranking, while the correlation is weaker for PR.

A regular trend can be spotted in the F_1 parameter for CQHITSu, as it appears to decrease when increasing the size of the graphs. Such a behaviour is not observed in other methods.

Table 9 F_1, F_{10} and Kendall's τ for random graphs.

	CQHITSu			CQHITSw			CQPRu			CQPRw		
	F_1	F_{10}	τ	F_1	F_{10}	τ	F_1	F_{10}	τ	F_1	F_{10}	τ
Hub	0.923	9.43	0.8877	0.923	9.42	0.8854	0.958	9.59	0.8955	0.959	9.60	0.9007
Auth.	0.972	7.19	-0.0640	0.991	9.12	0.8136	0.785	7.67	0.4174	0.988	9.72	0.7944

4.2.2 Random *k*-out graphs

Another test was performed on random graphs built via preferential attachment, with the constraint that all nodes have (nearly) the same outdegree. These graphs were generated by the `random_k_out_graph` command in NetworkX [37] and then modified by removing multiple edges, which is the reason for the slight deviation from uniform outdegree. A sample of 3000 graphs of 128 nodes each was considered. The same quantities F_1, F_{10} and τ are computed for the four methods, as in 4.2.1. Since the roles of hubs and authorities are asymmetrical due to the outdegree constraint, separate averages are computed. Results are shown in Table 9.

For hub rankings, results for weighted methods are almost identical to results for a uniform initial vector, as can be expected. The correspondence with classical methods seems to be very good, even − quite unexpectedly − for Kendall's τ (around 0.9). In authority ranking, uniform methods give good results − except for Kendall's τ, as expected − but the improvement from uniform to weighted methods is significant: the weighted methods show a striking 99% probability of spotting the same first-ranked node as their classical counterparts, an average coincidence of more than 9 among the first 10 nodes and a Kendall's τ around 0.8.

5 Discussion

One remarkable feature of the methods presented here is their overall good (sometimes *very* good) accordance with classical methods such as HITS or PageRank. One may ask why, since the classical methods seek the eigenvector associated with a specific eigenvalue (the dominant or the null eigenvalue), whereas from Theorem 1 it is clear that eigenvalues have no role at all in the quantum rankings, except possibly for their multiplicity structure. Indeed, suppose for simplicity that the Hamiltonian matrix H has distinct eigenvalues: then the limiting distribution of the associated CTQW only depends on the eigenvectors of H and on the initial state. Why should CTQW-based centrality give similar results to the classical case, if CTQWs do not "see" eigenvalues?

The explanation we found to be most convincing is the following. The classical ranking vector, say ϕ_1, is positive by the Perron-Frobenius theorem. Now, recall that coefficient a_1 from Theorem 1 is the (possibly weighted) sum of the elements of ϕ_1. Then one can typically expect that $|a_1| \gg |a_j|$ for $j > 1$, thus amplifying the contribution of eigenvector ϕ_1 in the quantum ranking. Figures 2 and 3 show the behaviour of the coefficients $|a_j|^2$ for a scale-free and for a k-out graph, respectively, for the CQPRu method. In both cases a_1

Fig. 2 Coefficients $|a_j|^2$ for the CQPRu authority ranking of a scale-free graph with 128 nodes and parameters $\alpha = 0.557$, $\beta = 0.136$, $\gamma = 0.307$. Here $|a_1|^2 = 0.5078$, whereas the other coefficients are in the $[0, 0.0608]$ range.
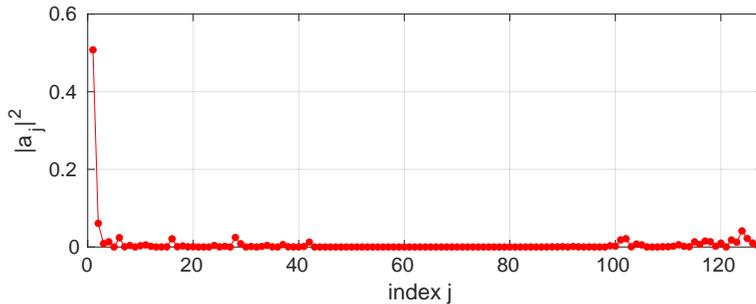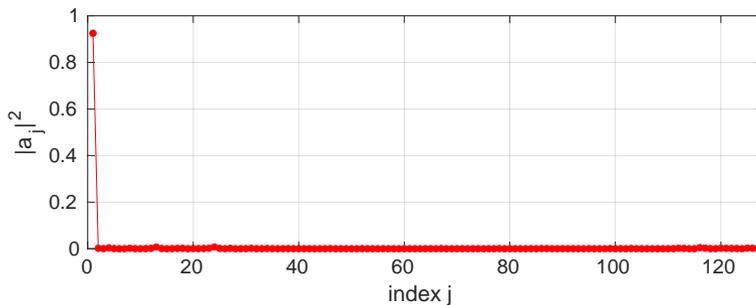

Fig. 3 Coefficients $|a_j|^2$ for the CQPRu hub ranking of a k-out graph with 128 nodes and parameters $\alpha = 0.3$, $k = 5$. Here $|a_1|^2 = 0.9249$, whereas the other coefficients are in the $[0, 0.0077]$ range.

is the coefficient associated with the classical PageRank eigenvector and it is clearly larger than the others. Note that both rankings are quite faithful to their classical counterparts: they identify the same top-ranked node and the same set of top-ten nodes.

This remark may also help explain why hub rankings for k-out graphs are in extremely good accordance with classical PageRank: since the PageRank vector is close to being uniform (as shown in Figure 4), the value of a_1 is nearly maximised, so in Theorem 1 the term associated with the PageRank eigenvector is strongly dominant.

Fig. 4 Entries of a few eigenvectors of the Hamiltonian matrix for the CQPRu hub ranking of a k-out graph with 128 nodes and parameters $\alpha = 0.3$, $k = 5$. The first eigenvector (thick blue line) is the classical PageRank vector, i.e., it is associated with eigenvalue 0. Also depicted are the eigenvectors associated with eigenvalues 0.2908, 0.9517 and 1.4143, that is, columns of indices 3, 50 and 100 in the eigenvector matrix computed by MATLAB. Clearly the first eigenvector has a comparatively uniform behaviour, whereas the other eigenvectors are much more localized.
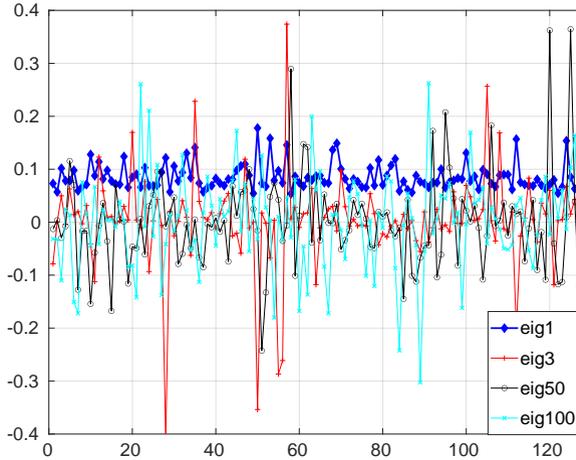
One may also ask whether the behaviour of the quantum methods is related to localization properties of eigenvectors. One may conjecture, for instance, that quantum approaches amplify the contribution of localized eigenvectors. Localization of eigenvector centrality is a well-known phenomenon: see, e,g., [38–40]. However, in our experiments the classical ranking eigenvector does not appear to be remarkably more localized than others. Moreover, the various eigenvectors tend to be localized over different sets of indices, so they do not necessarily reinforce the contribution of the classical ranking eigenvector. Hub rankings for k-out graphs are a good counterexample for this conjecture: the classical ranking eigenvector is typically much less localized than other eigenvectors (see Figure 4), but, as noted above, the quantum rankings are in very good accordance with classical rankings.

Another notable phenomenon is the fact that, for the CQPRu method, many coefficients a_i may be zero, and therefore the associated eigenvectors of the Hamiltonian do not contribute to the ranking. For instance, in the example of Figure 2, a_j is numerically zero for $j = 21, 22, 43$ to $80, 98, 107, 108$. More precisely, the Hamiltonian has many eigenvectors with eigenvalue 1 that are orthogonal to the initial uniform state. In fact, in such cases G and G^T both have large-dimensional null spaces, so it is not surprising that the intersection of these null spaces also has a comparatively large dimension. Now $\mathrm{Ker}(G) \cap \mathrm{Ker}(G^T)$ is an eigenspace of H associated with the eigenvalue 1, and each

vector in this subspace has zero average, i.e., it is orthogonal to the uniform vector. Indeed, let $\mathbf{e}$ be the vector with all entries equal to 1. Since G is row-stochastic we have $\mathbf{e}^T G^T = \mathbf{e}^T$. Then for any $\mathbf{v} \in \mathrm{Ker}(G) \cap \mathrm{Ker}(G^T)$ it holds $\mathbf{e}^T G^T \mathbf{v} = \mathbf{e}^T \mathbf{v}$. But $G^T \mathbf{v} = 0$ and therefore $\mathbf{e}^T \mathbf{v} = 0$.

6 Conclusions

We tested node-ranking methods based on continuous-time quantum walks, obtained by reframing two well-known classical methods (HITS and PageRank) as eigenvector problems on symmetric matrices, and using such matrices (with a slight modification for HITS) as Hamiltonians for the quantum walk. Each Hamiltonian was tested on two initial vectors (uniform and degree-weighted), yielding four quantum ranking methods. Numerical results are in good agreement with their classical counterparts in finding the top-ranked nodes, with a significant improvement in weighted methods w.r.t. the uniform ones. For instance, CQPRw finds the same top-ranked node as classical PageRank in more than 95% of the cases, and the average number of common top-10 nodes is always above 8.9, for all sets of test graphs. CQHITSw seems to be immune to the anomalous rankings of no-authority or no-hub nodes that are frequently encountered when using quantum methods, and has the best performance for the overall agreement of the rankings, as measured by Kendall's τ. Analysis of the results shows that the reason for this agreement is not related to localization of the dominant (or null) eigenvector found in the classical case. It is more likely due to the Perron-Frobenius theorem, which ensures that the initial positive state (uniform or weighted) has a large component along the eigenvector corresponding to the classical ranking vector.

Declarations

Funding

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Data availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

References

[1] L. C. Freeman, Centrality in social networks conceptual clarification, Soc. Netw. 1, 215-239 (1978).

[2] F. A. Rodrigues, Network Centrality: An Introduction, In E. Macau (eds), *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems. Nonlinear Systems and Complexity, Vol 22* (Springer, Cham, 2019) pp. 177-196.

[3] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, J. ACM 46, 604-632 (1999).

[4] S. Brin, and L. Page, The anatomy of a large-scale hypertextual web search engine, Comput. netw. ISDN syst. 30, 107-117 (1998).

[5] D. Fogaras, Where to start browsing the web?, in *Innovative Internet Community Systems 2003*, Lecture Notes in Computer Science Vol. 2877 (Springer, Berlin, Heidelberg, 2003) pp. 65-79.

[6] A. O. Zhirov, O. V. Zhirov, and D. L. Shepelyansky, Two-dimensional ranking of Wikipedia articles, Eur. Phys. J. B 77, 523-531 (2010).

[7] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, Phys. Rev. A 48, 1687 (1993).

[8] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani, (2001, July) Quantum walks on graphs in *STOC '01: Proceedings of the thirty-third annual ACM symposium on Theory of computing* (ACM, New York, USA, 2001) pp. 50-59.

[9] J. Kempe, Quantum random walks: An introductory overview, Contemp. Phys. 44, 303-327 (2003).

[10] M. Szegedy, Quantum speed-up of Markov chain based algorithms, in *45th Annual IEEE Symposium on Foundations of Computer Science* (IEEE, 2004) pp. 32-41.

[11] S. E. Venegas-Andraca, Quantum walks: a comprehensive review, Quantum Inf. Process. 11, 1015-1106 (2012).

[12] R. Portugal, *Quantum Walks and Search Algorithms*, 2nd edition (Springer, 2018).

[13] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, D. A. Spielman, Exponential algorithmic speedup by a quantum walk, in *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing* (ACM, New York, USA, 2003) pp. 59-68.

[14] J. Kempe, Discrete Quantum Walks Hit Exponentially Faster, Probab. Theory Relat. Fields 133, 215-235 (2005).

[15] E. Sánchez-Burillo, J. Duch, J. Gómez-Gardeñes, and D. Zueco, Quantum Navigation and Ranking in Complex Networks, Sci. Rep. 2, 605 (2012).

[16] S. Garnerone, P. Zanardi, and D. A. Lidar, Adiabatic Quantum Algorithm for Search Engine Ranking, Phys. Rev. Lett. 108, 230506 (2012).

[17] A. M. Childs, On the Relationship Between Continuous- and Discrete-Time Quantum Walk, Commun. Math. Phys. 294, 581-603 (2010).

[18] O. Mülken, A. Blumen, Continuous-time quantum walks: Models for coherent transport on complex networks, Physics Reports 502, 37-87 (2011).

[19] S. D. Berry, and J. B. Wang, Quantum-walk-based search and centrality, Phys. Rev. A 82, 042333 (2010)

[20] G. D. Paparo, and M. A. Martin-Delgado, Google in a quantum network, Sci. Rep. 2, 444 (2012).

[21] L. Rossi, A. Torsello, and E. R. Hancock, Node Centrality for Continuous-Time Quantum Walks, in *S+SSPR 2014: Proceedings of the Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition*, Lecture Notes in Computer Science Vol. 8621 (Springer, Berlin, Heidelberg, 2014) pp. 103-112.

[22] J. A. Izaac, X. Zhan, Z. Bian, K. Wang, J. Li, J. B. Wang, and P. Xue, Centrality measure based on continuous-time quantum walks and experimental realization. Phys. Rev. A 95, 032318 (2017).

[23] J. A. Izaac, J. B. Wang, P. C. Abbott, and X. S. Ma, Quantum centrality testing on directed graphs via PT-symmetric quantum walks. Phys. Rev. A 96, 032305 (2017).

[24] T. Loke, J. M. Tang, J. Rodriguez, M. Small, and J. B. Wang, Comparing classical and quantum PageRanks. Quantum Inf. Process. 16, 1-22 (2017).

[25] P. Chawla, R. Mangal, and C. M. Chandrashekar, Discrete-time quantum walk algorithm for ranking nodes on a network, Quantum Inf. Process. 19, 158 (2020).

[26] T. Wu, J. A. Izaac, Z.-X. Li, K. Wang, Z.-Z. Chen, S. Zhu, J. B. Wang, and X.-S. Ma, Experimental parity-time symmetric quantum walks for centrality ranking on directed graphs. Phys. Rev. Lett. 125, 240501 (2020).

[27] L. Böttcher, and M. A. Porter, Classical and quantum random-walk centrality measures in multilayer networks, SIAM J. Appl. Math. 81, 2704-2724 (2021).

[28] P. Boito, and R. Grena, Quantum hub and authority centrality measures for directed networks based on continuous-time quantum walks, J. Complex Netw. 9, cnab038 (2021).

[29] M. Benzi, E. Estrada, and C. Klymko, Ranking hubs and authorities using matrix functions, Linear Algebra Appl. 438, 2447-2474 (2013).

[30] M. G. Kendall, A new measure of rank correlation, Biometrika 30, 81–93 (1938).

[31] A. N. Langville, and C. D. Meyer, Deeper inside pagerank, Internet Math. 1, 335-380 (2004).

[32] A. D. Chepelianskii, Towards physical laws for software architecture, arXiv preprint arXiv:1003.5455 (2010).

[33] R. P. Feynman, Quantum mechanical computers, Found. Phys. 16, 507-531 (1986).

[34] E. Farhi, and S. Gutmann, Quantum computation and decision trees, Phys. Rev. A 58, 915 (1998).

[35] A. M. Childs, E. Farhi, and S. Gutmann, An example of the difference between quantum and classical random walks, Quantum Inf. Process. 1, 35-43 (2002).

[36] T. G. Wong, L. Tarrataca, and N. Nahimov, Laplacian versus adjacency matrix in quantum walk search. Quantum Inf. Process. 15, 4029-4048 (2016).

[37] A. A. Hagberg, D. A. Schult, and P. J. Swart, Exploring network structure, dynamics, and function using NetworkX, in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, G. Varoquaux, T. Vaught, and J. Millman (Eds) (Pasadena, CA, USA, 2008) pp. 11–15.

[38] K.-I. Goh, B. Kahng, and D. Kim, Spectra and eigenvectors of scale-free networks, Phys. Rev. E 64, 051903 (2001).

[39] T. Martin, X. Zhang, and M. E. J. Newman, Localization and centrality in networks, Phys. Rev. E 90, 052808 (2014).

[40] F. L. Metz, and I. Neri, Localization and universality of eigenvectors in directed random graphs, Phys. Rev. Lett. 126, 040604 (2021).