Commute Time for a Gaussian Wave Packet on a Graph

Furqan Aziz, Richard C. Wilson, and Edwin R. Hancock*

Department of Computer Science, University of York, YO10 5GH, UK
{furqan,wilson,erh}@cs.york.ac.uk

Abstract. This paper presents a novel approach to quantifying the information flow on a graph. The proposed approach is based on the solution of a wave equation, which is defined using the edge-based Laplacian of the graph. The initial condition of the wave equation is a Gaussian wave packet on a single edge of the graph. To measure the information flow on the graph, we use the average return time of the Gaussian wave packet, referred to as the wave packet commute time. The advantage of using the edge-based Laplacian of a graph over its vertex-based counterpart is that it translates results from traditional analysis to graph theoretic domain in a more natural way. Therefore it can be useful in applications where distance and speed of propagation are important.

Keywords: Edge-based Laplacian, wave equation, wave commute time, speed of propagation, graph complexity.

1 Introduction

One of the most challenging problems in the study of a complex network is to characterize the topological structure of a network, i.e., the way in which the nodes interact with each other. Each real-world network exhibits certain topological features that characterize its structure. Examples of such features are clustering coefficient, maximum degree, average degree, and average path-length. Over the recent years, researchers have developed different models that have similar properties as the real-world network. These models help us to understand or predict the structure of these systems. Examples of such models are scale-free networks [15] and small-world networks [14].

Recently, spectral methods have been successfully used for quantifying the complexity of a network. Passerini et al. [3] have used the spectrum of the normalized discrete Laplacian to define the von Neumann entropy associated with a graph. They have shown that this quantity can be used as the measure of the regularity of a graph. Han et al. [4] have approximated the von Neumann entropy using quadratic entropy and have shown that the approximate von Neumann entropy is related to the degree statistics of the graph. Escolano et al. [5] have used the diffusion kernel to quantify the intrinsic complexity of the undirected

* Edwin Hancock was supported by a Royal Society Wolfson Research Merit Award.

P. Fränti et al. (Eds.): S+SSPR 2014, LNCS 8621, pp. 374–383, 2014.

networks. They have also extended their work to directed networks [6]. Suau et al. [7] have analyzed the Schrödinger operator for characterizing the structure of a network. Lee et al. [9] have used the spectral methods to discover the genetic ancestry.

The structure of a complex network also plays an important role in the dynamics of information propagation. For this reason the study of a complex networks is becoming increasingly popular in epidemiology, where the goal is to study the mathematical models that can be used to simulate the infectious disease outbreaks in a social contact network. Grenfell [1] has discussed the traveling waves in measles epidemics. Abramson et al. [2] considered traveling waves of infection in the Hantavirus epidemics. Other real-life applications of information propagation over a network include the study of spreading a message over a social network and the study of a computer virus spreading over the internet [10].

While spectral method using discrete Laplacian have been successfully used, they suffer from certain limitations. Since the traditional graph Laplacian is an approximation of the continuous Laplacian to the discrete points, one of its limitations is that it cannot be used to translate most of the continuous results to a graph theoretic domain. For example the wave equation, defined using the discrete Laplacian, does not have finite speed of propagation. This makes it inappropriate for the applications that require spatial analysis or finite speed of propagation; e.g., spread of information in a network. The problem can be overcomed by treating edges of the network as real length intervals. This allows us to define a new kind of Laplacian, the edge-based Laplacian (EBL) of the graph [11][12]. The study of the edge-based Laplacian may be of great interest in appplications where the distance and speed of propagation are important.

In this paper our goal is to study the use of a wave equation, for the purpose of measuring the information flow across the network. The wave equation is defined using the edge-based Laplacian of a graph, where the initial condition is a Gaussian wave packet on a single edge of the graph. We define the wave packet hitting time, i.e., the time required for a wave packet to reach an edge f starting from an edge e, and the wave packet commute time, i.e., the time required for a wave packet to come back to the same edge from where it started. The remaining of this paper is organized as follows: We commence by introducing the edge-based Laplacian of a graph. Next we give a solution of a wave equation defined using the edge-based Laplacian, where the initial condition is a Gaussian wave packet. Based on the solution of wave equation we define wave packet hitting time (WHT) and wave packet commute time (WCT). Finally, in the experiment section, we apply the proposed method to different network models.

2 Edge-Based Laplacian of a Graph

Before introducing the edge-based Laplacian (EBL), in this section we provide some basic definitions and notations that will be used throughout the paper. A *graph* $G = (\mathcal{V}, \mathcal{E})$ consists of a finite nonempty set $\mathcal{V}$ of *vertices* and a finite set $\mathcal{E}$ of unordered pairs of vertices, called *edges*. A *directed graph* or a *digraph*

$D = (\mathcal{V}_D, \mathcal{E}_D)$ consists of a finite nonempty set $\mathcal{V}_D$ of vertices and a finite set $\mathcal{E}_D$ of ordered pairs of vertices, called *arcs*. So a digraph is a graph with an orientation on each edge. A digraph D is called *symmetric* if whenever (u, v) is an arc of D, (v, u) is also an arc of D. There is a one-to-one correspondence between the set of symmetric digraphs and the set of graphs, given by identifying an edge of the graph with an arc and its inverse arc on the digraph on the same vertices. We denote by $D(G)$ the symmetric digraph associated with the graph G. The *oriented line graph* is constructed by replacing each arc of $D(G)$ by a vertex. These vertices are connected if the head of one arc meets the tail of another, except that reverse pairs of arcs are not connected, i.e. $((u, v), (v, u))$ is not an edge.

We now define the EBL of a graph. The eigensystem of the EBL of a graph can be expressed in terms of the normalized adjacency matrix of a graph and the adjacency matrix of the oriented line graph [11][12]. Let $G = (\mathcal{V}, \mathcal{E})$ be a graph with a boundary ∂G. Let $\mathcal{G}$ be the geometric realization of G. The geometric realization is the metric space consisting of vertices $\mathcal{V}$ with a closed interval of length l_e associated with each edge $e \in \mathcal{E}$. We associate an edge variable x_e with each edge that represents the standard coordinate on the edge with $x_e(u) = 0$ and $x_e(v) = 1$. For our work, it will suffice to assume that the graph is finite with empty boundary (i.e., $\partial G = 0$) and $l_e = 1$. The eigenfunctions of the EBL are of two types; vertex-supported eigenfunctions and edge-interior eigenfunctions.

2.1 Vertex Supported Edge-Based Eigenfunctions

The vertex-supported eigenpairs of the EBL can be expressed in terms of the eigenpairs of the normalized adjacency matrix of the graph. Let A be the adjacency matrix of the graph G, and $\tilde{A}$ be the row normalized adjacency matrix. i.e., the $(i, j)th$ entry of $\tilde{A}$ is given as $\tilde{A}(i, j) = A(i, j)/\sum_{(k,j)\in\mathcal{E}} A(k, j)$. Let $(\phi(v), \lambda)$ be an eigenvector-eigenvalue pair for this matrix. Note $\phi(.)$ is defined on vertices and may be extended along each edge to an edge-based eigenfunction. Let ω^2 and $\phi(e, x_e)$ denote the edge-based eigenvalue and eigenfunction. Here $e = (u, v)$ represents an edge and x_e is the standard coordinate on the edge (i.e., $x_e = 0$ at v and $x_e = 1$ at u). Then the vertex-supported eigenpairs of the EBL are given as follows:

1. For each $(\phi(v), \lambda)$ with $\lambda \neq \pm 1$, we have a pair of eigenvalues ω^2 with $\omega = \cos^{-1}\lambda$ and $\omega = 2\pi - \cos^{-1}\lambda$. Since there are multiple solutions to $\omega = \cos^{-1}\lambda$, we obtain an infinite sequence of eigenfunctions; if $\omega_0 \in [0, \pi]$ is the principal solution, the eigenvalues are $\omega = \omega_0 + 2\pi n$ and $\omega = 2\pi - \omega_0 + 2\pi n, n \geq 0$. The eigenfunctions are $\phi(e, x_e) = C(e)\cos(B(e) + \omega x_e)$ where

$$C(e)^2 = \frac{\phi(v)^2 + \phi(u)^2 - 2\phi(v)\phi(u)\cos(\omega)}{\sin^2(\omega)}$$

$$\tan(B(e)) = \frac{\phi(v)\cos(\omega) - \phi(u)}{\phi(v)\sin(\omega)}$$

There are two solutions here, $\{C, B_0\}$ or $\{-C, B_0 + \pi\}$ but both give the same eigenfunction. The sign of $C(e)$ must be chosen correctly to match the phase.

2. $\lambda = 1$ is always an eigenvalue of $\tilde{A}$. We obtain a principle frequency $\omega = 0$, and therefore since $\phi(e, x_e) = C\cos(B)$ and so $\phi(v) = \phi(u) = C\cos(B)$, which is constant on the vertices.
3. If the graph is bipartite then $\lambda = -1$ is an eigenvalue of $\tilde{A}$. We obtain a principle frequency $\omega = \pi$, and therefore since $\phi(e, x_e) = C\cos(B + \pi x_e)$ and so $\phi(v) = -\phi(u)$, implying an alternating sign eigenfunction.

2.2 Edge-Interior Eigenfunctions

The edge-interior eigenfunctions are those eigenfunctions which are zero on vertices and therefore must have a principle frequency of $\omega \in \{\pi, 2\pi\}$. These eigenfunctions can be determined from the eigenvectors of the adjacency matrix of the oriented line graph.

1. The eigenvector corresponding to the eigenvalue $\lambda = 1$ of the oriented line graph provides a solution in the case $\omega = 2\pi$, and we obtain $|\mathcal{E}| - |\mathcal{V}| + 1$ linearly independent solutions.
2. Similarly the eigenvector corresponding to the eigenvalue $\lambda = -1$ of the oriented line graph provides a solution in the case $\omega = \pi$. If the graph is bipartite, then we obtain $|\mathcal{E}| - |\mathcal{V}| + 1$ linearly independent solutions. If the graph is non-bipartite, then we obtain $|\mathcal{E}| - |\mathcal{V}|$ linearly independent solutions.

This comprises all the principal eigenpairs which are only supported on the edges.

Note that although these eigenfunctions are orthogonal, they are not normlaized. To normalize these eigenfunctions we need to find the normalization factor corresponding to each eigenvalue and divide each eigenfunction with the corresponding normalization factor. Once normalized, these eigenfunctions form a complete set of orthonormal bases.

3 Wave Packet Commute Time

Recently, we have solved a wave equation on a graph, where the initial condition is a Gaussian wave packet on a single edge of a graph [8]. The wave equation is a second order partial differential equation, defined as

$$\frac{\partial^2 u}{\partial t^2}(\mathcal{X}, t) = \Delta_E u(\mathcal{X}, t), \tag{1}$$

where Δ_E is the EBL, and $\mathcal{X}$ represents the value of a standard coordinate x on an edge e. Let ω^2 represents the eigenvalue of the EBL with the corresponding eigenfunction $\phi_{\omega,n}(\mathcal{X}) = C(e, \omega)\cos(B(e, \omega) + \omega x + 2\pi n x)$. The complete solution is given as [8]

$$u(\mathcal{X},t) = \sum_{\omega \in \Omega_a} \frac{C(\omega,e)C(\omega,f)}{2} \Big(e^{-a\mathcal{W}(x+t+\mu)^2}$$
$$\cos\left[B(e,\omega) + B(f,\omega) + \omega \left\lfloor x+t+\mu+\frac{1}{2} \right\rfloor \right]$$
$$+ e^{-a\mathcal{W}(x-t-\mu)^2} \cos\left[B(e,\omega) - B(f,\omega) + \omega \left\lfloor x-t-\mu+\frac{1}{2} \right\rfloor \right] \Big)$$
$$+\frac{1}{2|E|} \left(e^{-a\mathcal{W}(x+t+\mu)^2} + e^{-a\mathcal{W}(x-t-\mu)^2} \right)$$
$$+ \sum_{\omega \in \Omega_b} \frac{C(\omega,e)C(\omega,f)}{4} \left(e^{-a\mathcal{W}(x-t-\mu)^2} - e^{-a\mathcal{W}(x+t+\mu)^2} \right)$$
$$+ \sum_{\omega \in \Omega_c} \frac{C(\omega,e)C(\omega,f)}{4} \Big((-1)^{\lfloor x-t-\mu+\frac{1}{2} \rfloor} e^{-a\mathcal{W}(x-t-\mu)^2}$$
$$-(-1)^{\lfloor x+t+\mu+\frac{1}{2} \rfloor} e^{-a\mathcal{W}(x+t+\mu)^2} \Big) . \quad (2)$$

Here $\mathcal{W}(z)$ wraps the value of z to the range $[-\frac{1}{2}, \frac{1}{2})$, and $\lfloor z \rfloor$ is the floor function.

Once we have the solution of the wave equation, we can define a number of interesting invariants to understand the properties of the flow of information across the network. This also helps us to quantify the structure of the network. We commence by defining the wave packet commute time of a graph. Given a graph $G = (\mathcal{E}, \mathcal{V})$ we define the wave packet commute time (WCT) of an edge e as follows. Assume that the initial condition of the wave equation is a Gaussian wave packet on the edge $e \in \mathcal{E}$ and zero elsewhere. Then

$$\text{WCT}(e) = \min_{t>0}\{t : u(e, 0.5) > \delta\}, \quad (3)$$

i.e., the WCT is the time when the wave packet with amplitude at least δ (at the middle of the edge), returns back to the edge e. Figure 4(a) demonstrates the wave commute time for a simple graph with 5 nodes and 7 links. Here the initial condition is a Gaussian wave packet on the edge $e1$ of the graph. The bottom right figure shows the fraction of the wave packet returned back at time $t = 3$. Note that at time $t = 1$, a wave packet with negative amplitude (a trough) returns to the edge $e1$. A trough will always be created when a wave packet is traveling along an edge (u, v) in the directed of v, and the degree of v is at least 3.

Edge-commute time can also be defined in terms of the hitting time of the wave packet. Given two edges $e, f \in \mathcal{E}$, the wave packet hitting time (WHT) can be defined as follows. Assume that the initial condition of the wave equation is a Gaussian wave packet on the edge $e \in \mathcal{E}$ and zero elsewhere. Then

$$\text{WHT}(e, f) = \min_{t>0}\{t : u(f, 0.5) > \delta\}, \quad (4)$$

i.e., the WHT is the time when the wave packet with amplitude at least δ (at the middle of the edge), reaches the edge f, starting from edge e. The edge-commute time can then be defined as:

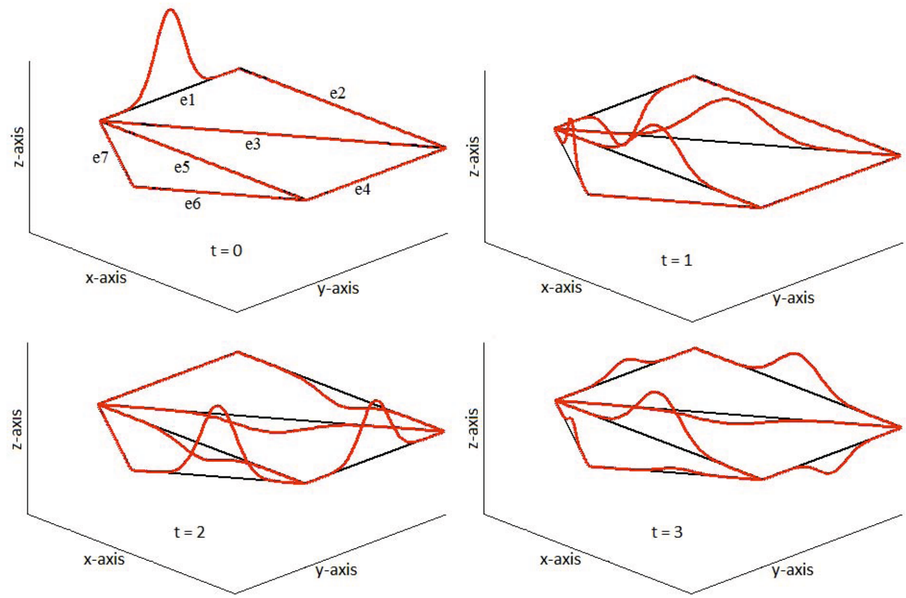


Fig. 1. Commute time of a Gaussian wave packet on a graph

$$\text{WCT}(e) = \frac{1}{|\mathcal{E}|} \sum_{f \in \mathcal{E}} WHT(e, f), \tag{5}$$

i.e., the WCT for the edge e is the average of the WHT over all the edges of the graph. However, the WCT defined using the WHT is computationally more expensive, and therefore in the experiment section we use the WCT defined in Equation 3.

To quantify the complexity of a network, we define a global invariant based on the WCT as:

$$\text{GWCT}(G) = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} WCT(e), \tag{6}$$

i.e., GWCT of a network is the average of the WCT over all the links of the network. In the next section we will show that GWCT provides a good measure for distinguishing graphs with different structures.

4 Experiments

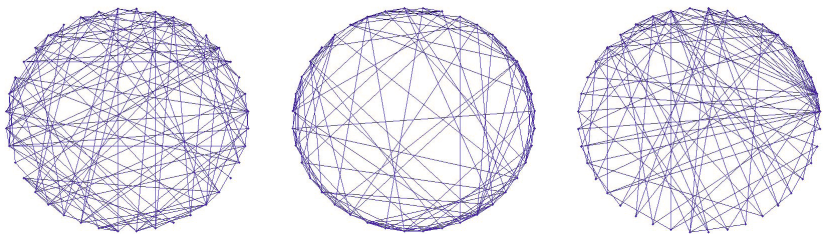
In this section we study the flow of information across a network using WCT and WHT and demonstrate the ability of GWCT to distinguish graphs with different structural properties. We experiment our proposed method on the following three different types of network models.

Erdős-Rényi Model(ER) [13]: An ER graph $G(n, p)$ is constructed by connecting n vertices randomly with probability p. i.e., each edge is included in the graph with probability p independent from every other edge. These models are also called $r$andom networks.

Watts and Strogatz Model(WS) [14]: A WS graph $G(n, k, p)$ is constructed in the following way. First construct a regular ring lattice, a graph with n vertices and each vertex is connected to k nearest vertices, k/2 on each side. Then for every vertex take every edge and rewire it with probability p. These models are also called $s$mall-world networks.

Barabási-Albert Model(BA) [15]: A BA graph $G(n, n_0, m)$ is constructed by an initial fully connected graph with n_0 vertices. New vertices are added to the graph one at a time. Each new vertex is connected to m previous vertices with a probability that is proportional to the number of links that the existing nodes already have. These models are also called $s$cale-free networks.

Figure 2 shows an example of each of these models.



(a) random graphs (b) small-world graph (c) scale-free graph

Fig. 2. Graph models

As mentioned earlier, one of the advantages of the wave equation defined using the EBL is that it has finite speed of propagation [11]. This makes it suitable for applications that require finite speed of propagation. In our first experiment, we demonstrate the ability of edge-based wave equation for identifying infected links in a network. For this purpose, we generate a BA network and a WS network each with 60 nodes and 175 links. We have computed the WHT for each edge of the graph starting from an edge e. The edge $e = (u, v) \in \mathcal{E}$ is selected, such that u is the highest degree vertex in the graph and v is the highest degree vertex in the neighbours of u. Figure 3 shows the cumulative frequencies of infected links for both graphs with different values of δ. As expected, the cumulative number of infected links decreases as δ increases. Note that the links in WS network are infected quickly than links in BA network. This is due to the presence of hub in BA network, which distributes the wave packet with small amplitudes to more links. The WS network, on the other hand, has more regular structure that allows the wave packet to transmit across the network with high amplitudes.

The above experiment shows that the WCT behaves differently on different graphs. This suggests that the WCT can be used to quantify the structure of a

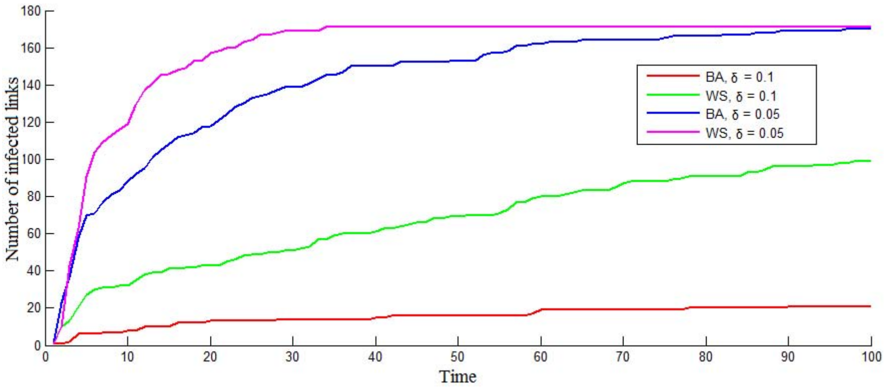


Fig. 3. Number of links infected with time

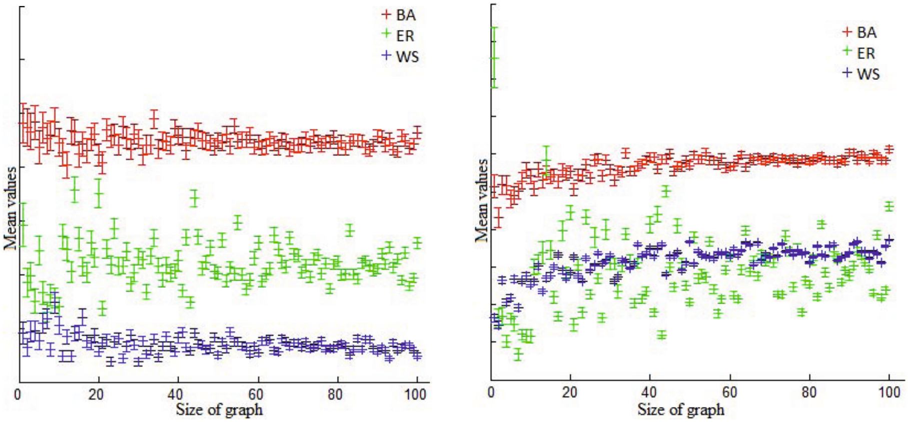
complex network. In our next experiment, we demonstrate the ability of WCT to distinguish networks with different substructures. For this purpose, we generate 100 graphs for each model with $n = 50 + (d-1)k$ with $k = 1, 2, ..., 100$, where n is the number of vertices. We have chosen the other parameters in such a way so that all three types of graphs with the same number of vertices have approximately the same number of edges. For ER models we choose $p = 10/n$, for WS models we choose $p = 0.25$ and $k = 8$, and for BA models we choose $n_0 = 5$ and $k = 4$. For each graph we compute the wave commute time and average it over all the edges. Figure 4(a) shows the average value for the three different types of graphs. Results suggest that the wave commute time is highly robust in distinguishing the graphs with different structures.

Figure 4(b) shows a similar analysis for vertex commute time, which is defined as the expected number of steps for a random walk starting from a vertex u, hits vertex v and then returns to u. The commute time of a vertex u to a vertex v can be computed from the eigenvalues and eigenvectors of the normalized Laplacian. Let (λ, ϕ) be the eigenpair of the normalized Laplacian. Then the commute time is defined as:

$$CT(u,v) = \sum_{i=2}^{|\mathcal{V}|} \left(\sqrt{\frac{vol}{\lambda_i d_u}} \phi_i(u) - \sqrt{\frac{vol}{\lambda_i d_v}} \phi_i(v) \right)^2 , \tag{7}$$

where d_u represents the degree of the vertex u and vol represents the sum of degrees for an unweighted graph. The x-axis in Figure 4(b) shows the average commute time over all vertices.

The mean and the standard error of the edge-commute time depend on the regularity structure of the graph. As the regularity of the graph increases, the value of the standard error decreases. Note that the value of WCT depends on the size of the smallest cycle to which the edge belongs. Figure 5 shows the mean values and the standard errors for the graphs generated in the previous experiment. Since WS networks are more regular as compared to BA networks, they therefore have smaller standard errors. Also, if the probability p of rewiring



(a) Wave commute time (b) Vertex commute time

Fig. 4. Wave commute time vs commute time

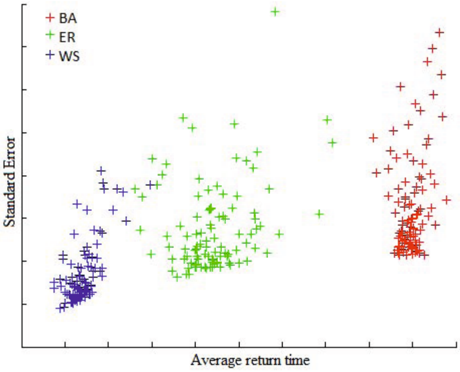


Fig. 5. Mean values and standard errors

is kept low, then the WS network has more small length cycles. Therefore the mean values of WS networks are small as compared to BA networks. Note that ER graphs exhibit more variation in the mean values due to their random structure. Their mean and standard error values lie between that of the BA graphs and the WS graphs.

5 Conclusion

In this paper we have studied the properties of the commute time (WCT) and the hitting time (WHT) of a Gaussian wave packet on a graph. The WCT and WHT are based on the solution of the wave equation defined using the edge-based Laplacian of a graph where the initial condition is a Gaussian wave packet on a single edge of the graph. We have shown the application of WCT and WHT for quantifying the structure and information flow of a network. The advantage of using the edge-based Laplacian (EBL) is that this approach is more closely related to mathematical analysis than the usual discrete Laplacian. This allows us to implement equation on graphs which have finite speed of propagation.

References

1. Grenfell, B.T.: Travelling waves and spatial hierarchies in measles epidemics. Nature, 716-723 (2001)
2. Abramson, G., Kenkre, V.M., Yates, T.L., Parmenter, R.R.: Traveling Waves of Infection in the Hantavirus Epidemics. Bulletin of Mathematical Biology, 519–534 (2003)
3. Passerini, F., Severini, S.: The von neumann entropy of networks. International Journal of Agent Technologies and Systems, 58–67 (2009)
4. Han, L., Escolano, F., Hancock, E.R., Wilson, R.C.: Graph characterizations from von Neumann entropy. Pattern Recognition Letters, 1958–1967 (2102)
5. Escolano, F., Hancock, E.R., Lozano, M.A.: Heat diffusion: Thermodynamic depth complexity of networks. Physics Review E, 036206 (2012)
6. Escolano, F., Bonev, B., Hancock, E.R.: Heat Flow-Thermodynamic Depth Complexity in Directed Networks. In: Gimel'farb, G., Hancock, E., Imiya, A., Kuijper, A., Kudo, M., Omachi, S., Windeatt, T., Yamada, K. (eds.) SSPR&SPR 2012. LNCS, vol. 7626, pp. 190–198. Springer, Heidelberg (2012)
7. Suau, P., Hancock, E.R., Escolano, F.: Analysis of the Schrödinger Operator in the Context of Graph Characterization. In: Hancock, E., Pelillo, M. (eds.) SIMBAD 2013. LNCS, vol. 7953, pp. 190–203. Springer, Heidelberg (2013)
8. Aziz, F., Wilson, R.C., Hancock, E.R.: Gaussian Wave Packet on a Graph. In: Kropatsch, W.G., Artner, N.M., Haxhimusa, Y., Jiang, X. (eds.) GbRPR 2013. LNCS, vol. 7877, pp. 224–233. Springer, Heidelberg (2013)
9. Lee, A.B., Luca, D., Klei, L., Devlin, B., Roeder, K.: Discovering genetic ancestry using spectral graph theory. Genetic Epidemiology, 51-59 (2010)
10. Bradonjic, M., Molloy, M., Yan, G.: Containing Viral Spread on Sparse Random Graphs: Bounds, Algorithms, and Experiments. Internet Mathematics, 406–433 (2013)
11. Friedman, J., Tillich, J.P.: Wave equations for graphs and the edge based Laplacian. Pacific Journal of Mathematics, 229–266 (2004)
12. Wilson, R.C., Aziz, F., Hancock, E.R.: Eigenfunctions of the edge-based Laplacian on a graph. Journal of Linear Algebra and its Applications, 4183–4189 (2013)
13. Erdõs, P., Rényi, A.: On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 17–61 (1960)
14. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature, 440–442 (1998)
15. Barabási, A., Albert, R.: Emergence of Scaling in Random Networks. Science 509–512 (1999)