The applications of graph theory in the area of networking are of great significance in system analysis of different varieties, including biological systems. In biological systems, the use of networks finds importance in the study of epidemic and its control. A practical example include the evaluation of the spread of disease within human population and the impact of awareness circulating admirist the same population as a result of the infection. Agaba et al. in 2017 proposed a mathematical model that analysed the impact of awareness on the spread of infectious diseases. This was done using the stability analyses of the various steady states of the system of equations and also through the evaluation of some numerical simulations. This paper, with the aid of the system of equations developed by Agaba et al., 2017b, studies the impact of awareness spreading simultaneously with an infectious disease within human population using weighted network.

Key words: Epidemic model, Disease awareness, Graph theory, weighted network
Introduction

Researches on the creation and dissemination of human awareness regarding the spread of an epidemic have been carried out by many scholars in diverse dimensions. Some of these scholars examined the impact of awareness on epidemic models using mean-field models such that evaluate the co-existence of multiple feasible equilibra (Cui et al., 2008; Liu et al., 2007), the occurrence of multiple outbreak of diseases as a result of the spread of awareness or information dissemination (Liu et al., 2007), human behavioural changes caused by the prevalence of infectious diseases (Agaba et al., 2017b; Bauch et al., 2013; d’Onofrio and Manfredi, 2009; Perra et al., 2011; Poletti et al., 2009), optimal disease control programs (Agaba et al., 2017a; Li and Cui 2009; Roy et al., 2015; Tchuenche et al., 2011; Wang et al., 2016) and the role of time delay as regards human response to awareness campaigns on disease dynamics (Agaba et al., 2017b; Agaba et al., 2017c; Greenhalgh et al., 2015; Zhao et al., 2014; Zuo and Liu, 2014; Zuo et al., 15). While other research works that analysed the impact of awareness creation and dissemination on the spread of infectious diseases considered network-based models. Network-based models usually consider the interaction of individuals within the entire population which are often denoted by network nodes and the possible connections between individuals/classes of the population (such as the susceptible, infected and recovered) through which diseases can be transmitted, represented by network edges (Funk et al., 2010; Funk et al., 2009; Gross and Blasius, 2008; Sahneh and Scoglio, 2011; Wang et al., 2013; Wu et al., 2012).

The results from the aforementioned analyses revealed the importance of human awareness in controlling, and in some cases eradicating, infectious diseases from the entire population as portrayed by (Agaba et al., 2017b; Agaba et al., 2017c; Ferguson, 2007; Funk et al., 2010; Jones and Salathé, 2009; Nishiura et al., 2005; Pryut et al., 2015). According to the words of Funk et al. (2010), “the influence of spreading awareness on a disease outbreak depends on how much the individuals at the front of the growing epidemic are aware of its presence”. They also stated that the size of an epidemic outbreak can be potentially reduced by individual behaviour that is responsive to the presence of the infectious disease.

In the same light, it has been observed also that as awareness circulates among human population with respect to the spread of diseases, people of same interest and attitude (which in most cases are inclusive of those whose interest and attitude were adjusted to fit in as a result of peer influence and association) tend to cluster toward the same direction and/or location. Hence, people of similar behaviour and character cluster together to form groups and they tend to exhibit similar ideology and understanding of their activities (Friedkin, 1998; Huberman and Adamic, 2004; Keeling, 1999; Moody, 2001). A good example of this social tie is found in the research carried out by Sade in 1972 as reviewed in Newman (2004), the research studied the social relationship that existed between a selected number of rhesus monkeys by considering their grooming behaviours; those that groomed together and the number of time these were done under due observation for a certain period of time. According to Newman (2004), there exist both strong and weak social ties between individuals in any social network and their strengths are often demonstrated or represented by the strengths of the respective edges within the weighted network.

This aspect of mathematics centred on the study of network interactions or network analyses have over the years find application in several fields of human specialization such as ICT, physics, sociology, biological and biochemical networks (Newman, 2004; Rosen, 2012). In addition, network structure assists in determining the spread of infection transmitted through direct contact of individuals and its control (House and Keeling, 2010). According to the words of Sahneh and Scoglio, it’s clear that “graph theory is widely used for representing the contact topology in an epidemic network” (Sahneh and Scoglio, 2011).

In this paper, weighted network is used to evaluate the impact of awareness spreading simultaneously with an infectious disease within a population. In the analyses, characteristics of weighted network were used to demonstrate how the spread of awareness within a human population affects the dynamics of an infectious disease spreading within the same population. The outline of the paper consists of five Sections. The next Section considered graph theory on the analyses of networks while Section 3 evaluated an epidemic model by studying the interconnections of individuals within their local neighbourhood using the adjacency matrix of the network. In Section 4, the numerical simulations were used to supplement the analytical results of the network as regards the interconnection of individuals with respect to the spread of the disease and awareness springing from global campaigns and direct interactions of the population. The last Section covered the discussion of results.
Epidemic models and graph theory

The analysis of the system of equations modelled by Agaba et al., 2017b focused on the behavioural changes associated with the simultaneous spread of awareness and infectious disease within a population. Two types of awareness were considered: the private awareness which associated with direct contacts between unaware and aware populations, and the public awareness that stems from information campaign. While the disease spread as a result of the interactions of the susceptibles with the infectious individuals within the population. In order to obtain a graphical network representation of the transmission of the spread of infectious disease within the given population, it is pertinent to use spatial structure, contact or social network. The contact network portrait that the probability of infectious disease spreading within a susceptible population through contact with the infected is proportional to the ratio of the infected individuals to the entire given population (Funk et al., 2010; House and Keeling, 2010; Keeling, 1999).

According to the words of Rosen (2012), “graphs are extensively used to model social structures based on different kinds of relationships between people or groups of people”. The graphs representing these social structures are known as social network.

Rosen (2012) and Newman (2004) considered some illustrations of graph models involving individuals and groups of individuals. Among these examples is a transportation network wherein airline networks were modelled by representing each airport by a vertex (node) and all flights by a particular airline moving each day from an airport (departure) to another airport (destination) by directed edges. The graph used in the above model representation is the directed multigraph (Rosen, 2012). Applying the same ideology, this paper considered a network in which the nodes represent groups of individuals determined by their disease and awareness status, while the edges denote the connections between the nodes as a result of the transition of individuals from one node to the other with directed edges indicating the direction of their transitions. The weights on the edges represent the number of possible transition rates. Consequently, the network is assumed to form links connecting \( N \) groups of individuals thus forming a graph, \( G \in \{0,1\}^{N^2} \) where

\[
G = \begin{cases} 
1 & \text{if pairs are connected;} \\
0 & \text{otherwise.}
\end{cases}
\]

According to Keeling (1999), the number of pairs can be obtained as

\[
\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} n_{ij} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} n_{ij},
\]

denoting the sum of all the elements in the matrix, which can also be represented as

\[
\|G\| = nN,
\]

Where \( N \) is the total number of nodes (various groups of individuals according to their disease and awareness status) and \( n \) is the average number of neighbours per node. The value of \( n \) can be generated using

\[
n = N! c_m \Rightarrow n = \left( \frac{N}{m!} \right) = \frac{N!}{m!(N-m)!},
\]

with \( m \) representing the number of connected nodes within the network.

The number of triples is equal to

\[
\|G^2\| = \text{trace} (G^2)
\]

with \( G^2 = G \times G \) , the square of the matrix trace \( (G^2) \) is the sum of the diagonal entries of the matrix \( G^2 \), The measure of how interconnected the local neighbourhoods are, \( \eta \) is defined by the ratio of triangles to triples as follows,

\[
\eta = \frac{\text{number of triangles}}{\text{number of triples}} = \frac{\text{trace} (G^2)}{\|G^2\| - \text{trace} (G^2)}.
\]

In epidemiology network, the measure of interconnection of the local neighbourhood, \( \eta \) is also referred to as the clustering coefficient (House and Keeling, 2010).

The degree of a node in a given network is the total number of connections or edges incident to the node within the network. Whereas, the degree distribution \( P(k) \) of a network is defined as the fraction of nodes in the network with degree \( k \) (Newman, 2004). If there are \( N \) nodes in total and \( mk \) of them have degree \( k \), we have

\[
P(k) = \frac{m_k}{N}.
\]

For instance, the Bernoulli random graph, in which each of \( N \) nodes is connected (or not) with independent probability \( P \) (or \( p-1 \)) (or \( p-1 \).) has a binomial distribution of degree \( k \) is given

\[
P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}
\]

Since this study focused on both the spread of disease and the spread of awareness, it is therefore possible to have a pair state where

1. only disease spreads as a result of the interaction,
2. only awareness is generated by the interaction and
3. both disease and awareness spread concurrently as a result of the interaction.

The spread of disease and awareness through interactions are therefore represented by using the notations \([ ... \]) and \([ ... \]) respectively. If we defined
\(E^d\) to denote the edge on the disease network, then the total number of edges on the disease network would be given as
\[
\|E^d\| = \sum_i E^d_i.
\]

If \(E^a\) denote the edge on the awareness network, we have the total number of edges on the awareness network as
\[
\|E^a\| = \sum_i E^a_i,
\]
and if \(E^b = E^d \cap E^a\) denote the edge that spreads both disease and awareness concurrently, then the total number is given as
\[
\|E^b\| = \sum_i (E^d_i \cap E^a_i).
\]

Thus, similar to Funk et al. (2010), the probability of a selected pair states with a disease edge to also spread awareness is given by
\[
Q_{a|d} = \frac{\|E^d\|}{\|E^a\|}
\]
and the probability of a given pair having an awareness edge to also spread disease is given by
\[
Q_{d|a} = \frac{\|E^b\|}{\|E^a\|}.
\]
Consequently, \(Q_{a|d} = Q_{d|a} = 0\) \(\|E^b\| = 0\). While in contrast, when \(E^d \subseteq E^a\) \(Q_{a|d} = 1\) and when \(E^a \subseteq E^d\) \(Q_{d|a} = 1\) but the value of \(Q_{a|d}\) is not always equal to \(Q_{d|a}\).

Furthermore, if the entire human population within the network is denoted by \(N_p\), then the proportion of each individual on the disease and awareness network is thus given respectively as:
\[
K^d = \frac{\|E^d\|}{N_p} \quad \text{and} \quad K^a = \frac{\|E^a\|}{N_p}.
\]

We note therefore that
\[
K^a = \frac{\|E^d\|}{N_p} = \frac{\|E^b\|}{Q_{a|d}N_p} \Rightarrow Q_{a|d}K^a = \frac{\|E^b\|}{N_p}
\]
and
\[
K^a = \frac{\|E^a\|}{N_p} = \frac{\|E^b\|}{Q_{d|a}N_p} \Rightarrow Q_{d|a}K^a = \frac{\|E^b\|}{N_p}.
\]
Hence,
\[
Q_{a|d}K^d = Q_{d|a}K^a = \frac{\|E^b\|}{N_p}.
\]

In the case of triple states, where ??, ??and ??represent individual node respectively, if ??and ??are connected by the edge, ??and ??and ??are connected by the edge, ??where ??and ??stand for either disease (??) or awareness (??) respectively, then according to Funk et al. (2010) and Keeling (1999), the approximate triple states will be
\[
[XYZ]^{uv} \approx \int_K V^v - Q_{v|u}[XY]^u[YZ]^y\frac{V^v}{Y} = \varphi^{uv}[XY]^u[YZ]^y,
\]
where the correction factor is thus
\[
\varphi^{uv} = \frac{K^v - Q_{v|u}}{K^v}.
\]
By the application of equation 7, we note that \([XYZ]^{uv} = [ZYK]^{vu}\).

**Figure 1:** extracted and modified from Funk et al. (2010), gives the possible connection network between pairs and triples.
Mathematically, weighted networks are often represented using an adjacency matrix whose entries are determined by the weights of the edges, hence they are not always zeros and ones. For instance, considering the graph in Figure 2(a), it generates a weighted network represented with the following adjacency matrix.

\[
\begin{bmatrix}
  h_{1} & h_{2} & h_{3} & h_{4} \\
  0 & 2 & 1 & 0 \\
  2 & 0 & 3 & 2 \\
  1 & 3 & 0 & 0 \\
  0 & 2 & 0 & 0 \\
\end{bmatrix} \implies h_{ij} = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 3 & 2 \\ 1 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}
\]

Where \( h_{ij} \) is the weight of connection from \( i \) to \( j \). Similarly, the adjacency matrix corresponding to Figure 2(c) is obtained as

\[
\begin{bmatrix}
  h_{ij} \\
  0 & 0 & 1 & 0 \\
  2 & 0 & 2 & 1 \\
  0 & 1 & 0 & 0 \\
  0 & 2 & 0 & 0 \\
\end{bmatrix}
\]

Figure 2: (a) and (c) are weighted graphs while (b) and (d) are unweighted multigraphs equivalent to the weighted graphs in (a) and (c) respectively. Note that (c) is a graph with directed edges (directed graph) and (d) is the corresponding directed multigraph.

However, the weighted graphs in Figure 2(a) and (c) can also be represented with the multiple edges graphs captured in Figure 2(b) and (d) respectively. Multiple edges graph is sometimes referred to as multiedges and any network or graph containing multiedges is known as multigraph. While multigraph with directed edges is referred to as directed multigraph. By graph theory, the behaviour of weighted graphs can be obtained by simply mapping them onto unweighted multigraphs. This implies that every edge of weight \( w_{ij} \) connecting the vertices is therefore replaced with equal number of parallel edges of weight 1 each while the adjacency matrix remains the same. Consequently, the multigraphs can be manipulated using all processes and techniques applicable to the unweighted graphs (Newman, 2004, Rosen, 2012).

Model evaluation

The system of equations in Agaba et al.’s model gave the model diagram captured in Figure 3 (adopted from Agaba et al., 2017b) from which the weighted network is generated as the adjacency matrix below representing all possible transitions between the various individual classes within the network associated with the spread of the disease and awareness.
and which implies that

\[ Q_{d|a} = \frac{\| E^b \|}{\| E^d \|} = \frac{2}{4} = 0.5 \]

and

\[ Q_{a|d} = \frac{\| E^b \|}{\| E^a \|} = \frac{2}{8} = 0.25 \]

From the above results, it is obvious that \( Q_{d|a} = 0.5 > Q_{a|d} \) which shows that the probability of individuals on the network with disease edge spreading awareness is higher than the probability of individuals with awareness edge spreading the disease. By implication, this indicates that the presence of awareness, either by direct contact or global campaign, in a disease network tend to curtail the spread of diseases thereby reducing or eliminating the epidemic while on the other hand, as disease is introduced into a given population and allowed to spread, it generates awareness which after an elapsed time interval tend to act as means of controlling the spread of the disease along with other control measures and eventually, eradicates such disease when the spread of awareness become so much. This result agrees with the one obtained in Agaba et al., 2017b.

To determine the measure of the interconnection of the local neighbourhood within the network, the following were obtained from the adjacency matrix \( G_i \) in equation 9 using matrix operations.

\[
G^2_{ij} = \begin{bmatrix}
2 & 0 & 2 & 0 & 8 & 0 \\
1 & 2 & 0 & 0 & 0 & 4 \\
0 & 2 & 2 & 4 & 0 & 0 \\
0 & 4 & 0 & 2 & 0 & 2 \\
0 & 0 & 2 & 1 & 2 & 0 \\
2 & 0 & 0 & 0 & 2 & 2
\end{bmatrix}
\]

which implies that

\[ \| G^2 \| = \sum_{ij} G^2_{ij} = 46, \quad \text{trace}(G^2) = 12 \quad \text{and} \quad \text{trace}(G^3) = 12. \]

Consequently, using equation 3 gives the measure of interconnection of local neighbourhood as

\[
\eta = \frac{\text{trace}(G^3)}{\| G^2 \| - \text{trace}(G^2)} = \frac{12}{46 - 12} = \frac{12}{34} = 0.3529
\]
Numerical study of the network

Numerically, one can also represent the interactions of the total number of individuals within the network with respect to the spread of awareness and the infectious disease graphically to ascertain the movement between the respective classes of population. This in most cases is done using the study of community structure, groups of vertices having dense connections within but scanty between (Newman, 2004). The ability to detect and study communities is central in network analysis. Consequently, some scholars have developed computer algorithms such as the Kernighan-Lin algorithm, Spectral bisection, Girvan and Newman algorithm, Louvain method (Newman, 2004; Gephi, 2010) among others, in order to ease the computation of extracting communities in a network when the topology of the network is inputted into the algorithm. This research used the algorithm of Louvain method embedded in Gephi software application for its numerical computation. In Gephi software, the community detection algorithm, located in the statistics panel, creates a “modularity class” value for each node which the partition module eventually uses to colourize the communities.

![Figure 4: Graph showing the interactions within the network of 50 individuals with (a) $\eta = 0.25$, (b) $\eta = 0.3529$, (c) $\eta = 0.6$, respectively. The colours of the nodes indicate the various six communities.](image1)

Evaluating the information obtained in Section 3 gives the following results represented in Figures 4 and 5 as extracted using Gephi software applications. The first graph was obtained by placing nodes randomly in a network of 50 individuals with the probability of interconnection between the individuals set at $\eta = 0.3529$, as obtained using the adjacency matrix of the network. This process generated the interaction of the individuals graphically and their interconnection produced six (6) communities with the modularity of 0.108 and an average cluster coefficient of 0.183. The diameter which signifies the longest graph distance between any two individuals in the network, that is how far apart the two most distant nodes are, was obtained as 4 while the average path length equals 1.643. The process was repeated using varied values of $\eta$ for a network of 50 and 100 individuals respectively. The overall summary of the results is presented in Table 1 and the pictorial representation in Figures 4 and 5.

![Figure 5: Graph showing the interactions within the network of 100 individuals with (a) $\eta = 0.25$, (b) $\eta = 0.3529$, (c) $\eta = 0.6$, respectively. The colours of the nodes indicate the various six communities.](image2)
Discussion

At the initial state of the interaction in the absence of awareness, the value for awareness edge is zero which implies that the probability of a disease edge spreading awareness at this initial state is also zero. But as awareness is generated, first as a result of the interactions of the unaware susceptibles with the infected which created disease awareness and then by means of global awareness campaign as the number of infections increases, the level of interconnections is increased and consequently, the interactions of the aware individuals with others and also among themselves tend to surpass the interactions between the infected and the susceptibles thus curtailing the spread of the infection with the dissemination of awareness.

The results in Figures 4 and 5 also clearly support the above mentioned. As the value of the measure of interconnection between the individuals within the network \( \eta \) increases, there exist a corresponding increase in the interactions among them and the clustering of these individuals within their respective classes (communities). From the results, it signifies that the susceptibles exhibit the principle of withdrawal or isolation from the infected as a result of their knowledge of the disease and this is one of the measures of controlling infection rate. By this principle, the disease will eventually die out of the population as it connotes that at this point the basic reproductive number, \( R_0 < 1 \).

In the same light, the diameter of the graphs decreases with increase in interconnection showing that the longest graph distance between any two individuals in the network decreases because as they interact, they tend to attract those of same characteristics or social tie towards themselves thereby clustering as communities. Similarly, the average clustering coefficients of the network increases with increase in the level of interconnections among the individuals forming communities of individuals having same characteristics while the average path length decreases with increasing interconnection.

This research work, in addition to existing literatures on the use of weighted network in evaluating epidemic model, has further explored the application of weighted network on epidemic model with an extension on the impact of awareness. Awareness tends to create behavioural changes that often trigger the withdrawal of individuals aware of the disease as means of protecting themselves from being infected. These expressions were illustrated by the clustering of communities (individuals with similar interests or characters) within the weighted network which signifies that people of like manner or status cluster together as a community thereby isolating themselves from other communities.

Consequently, in epidemiology the aforementioned scenario illustrates that the susceptible, infected and recovered population (that is, the aware and unaware population respectively) form different clusters. Whereby indicating that the aware susceptibles thus distant themselves from the infected and by so doing reduces the likelihood of being infected or exposure to infection which in turn minimises the infection rate. Therefore, it connotes that the application of weighted network analysis also aid in evaluating the impact of awareness on an epidemic model. Practically, it implies that the importance of awareness campaign against the spread of infectious diseases cannot be overemphasized. It is thus pertinent for all authorities concern to encourage continuously the dissemination of information, concerning any disease invasion, among the most vulnerable people and also to others within the entire community/country as means of epidemic control and possibly eradication.

### Table 1: Results for the model analyses as obtained using Gephi application

<table>
<thead>
<tr>
<th></th>
<th>Network of 50 individuals</th>
<th>Network of 100 individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure of interconnection, ( \eta )</td>
<td>0.25 0.3529 0.6</td>
<td>0.25 0.3529 0.6</td>
</tr>
<tr>
<td>Modularity</td>
<td>0.189 0.129 0.071</td>
<td>0.140 0.108 0.058</td>
</tr>
<tr>
<td>Average cluster coefficient</td>
<td>0.119 0.174 0.301</td>
<td>0.115 0.183 0.298</td>
</tr>
<tr>
<td>Average path length</td>
<td>1.844 1.662 1.375</td>
<td>1.846 1.643 1.393</td>
</tr>
<tr>
<td>Diagonal</td>
<td>4 4 3</td>
<td>6 4 3</td>
</tr>
<tr>
<td>Community population (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (yellow)</td>
<td>14 12 12</td>
<td>27 11 23</td>
</tr>
<tr>
<td>2 (green)</td>
<td>20 22 14</td>
<td>15 18 13</td>
</tr>
<tr>
<td>3 (purple)</td>
<td>20 14 18</td>
<td>8 15 17</td>
</tr>
<tr>
<td>4 (cyan)</td>
<td>16 16 20</td>
<td>13 17 15</td>
</tr>
<tr>
<td>5 (red)</td>
<td>12 16 12</td>
<td>21 15 17</td>
</tr>
<tr>
<td>6 (blue)</td>
<td>18 20 24</td>
<td>16 24 15</td>
</tr>
</tbody>
</table>
References


