

Modeling Behavioral Evolution as Social Predictor for the Coronavirus Contagion and Immunization in Nigeria

Arnold Adimabua Ojugo^{a,*}, Andrew Okonji Eboka^b

^a*Department of Computer Science, Federal University of Petroleum Resources, Delta State, Nigeria*

^b*Department of Network Computing, Coventry University, United Kingdom*

Abstract

Since the outbreak of the novel coronavirus (covid-19) pandemic from China in 2019, it has left the world leaders in great confusion due to its fast-paced propagation and spread that has left infected a world population of over Eleven Million persons with over five hundred and thirty four thousand deaths and counting with the United States of America, Brazil, Russia, India and Peru in the lead on these death toll. The pandemic whose increased mortality rate is targeted at 'aged' citizens, patients with low immunology as well as patients with chronic diseases and underlying health conditions. Study models covid-19 pandemic via a susceptible-infect-remove actor-based graph, with covid-19 virus as the innovation diffused within the social graph. We measure the rich connective patterns of the actor-based graph, and explore personal feats as they influence other nodes to adopt or reject an innovation. Results shows current triggers (lifting of inter-intra state migration bans) and shocks (exposure to covid-19 by migrants) will lead to late widespread majority adoption of 23.8-percent. At this, the death toll will climb from between 4.43-to-5.61-percent to over 12%.

© 2021 Author(s).

Keywords: Corona virus, COVID-19, graph-model, SIS, SIR, epidemiology, pandemic

1. Introduction

The novel coronavirus caused by severe acute respiratory syndrome coronavirus emerged in the China in December 2019 and was declared a global pandemic by the World Health Organization (WHO) on the 11 March 2020 [1]. Since then, the disease has quickly spread to all continents with over 11-million cases recorded and with a fatality rate of 6.19% noted on 11 April 2020 [2]. The risk of importing COVID-19 from Europe to Africa is higher than its import from China [3]. Martinez-Alvarez et al. [4] compared its transmission (6-days after first detected cases) in selected countries and observed a more rapid spread of the virus in some West African countries than in Europe [5]. Ojugo and Otakore [6] compared situations in Delta State Nigeria and found that this situation is worsened in Nigeria and also in countries ill-equipped to handle covid-19 disease outbreak due to poor surveillance and response systems, inadequate health infrastructure and services, and unqualified medical personnel to handle the desired and targeted outbreak response. Further investigation by Ojugo and Oyemade [7] notes that current triggers (that is, lifting of inter-

* Corresponding author.

E-mail address: ojugo.arnold@fupre.edu.ng (Arnold Adimabua Ojugo)

intra state migration bans) and shocks (exposure to covid-19 by migrants) will lead to late widespread majority adoption of 23.8-percent. Thus, death toll will climb from between 4.43-to-5.61-percent to over 12%. This value will continue to decrease in terms of percentage of confirmed cases and death when compared to other nations if the health and safety measures are still being implemented by the populace.

Also, [8] African countries with highest importation risk have also been found to possess a high capacity to respond to outbreaks. As of 28 June 2020, Nigeria had a total of 21,317 confirmed cases and 533 deaths from COVID-19. The fast-paced spread and propagation of the COVID-19 contagion has left the world leaders in great confusion. With an infected world population of over Eleven Million patients currently and over five hundred and thirty four thousand deaths with nations such as the United States of America, Brazil, Russia, India and Peru in the lead on these death toll [9], the world is inflicted with a pandemic whose increased mortality rate is targeted at 'aged' citizens, patients with low immunology as well as patients with chronic diseases and underlying health conditions [7].

1.1. The Covid-19 Pandemic

The Coronavirus disease (infection) also called COVID-19, emerged in December 2019 in China in Wuhan city in Hubei province of China [10]. This pandemic has spread across 210 countries and territories around the world and 2 international conveyances. Following this outbreak in China, the virus has recorded new epicentres for its outbreak, the United States of America, Spain, Italy, France and Germany, having confirmed cases of over a hundred thousand [11]. Amidst, the risk posed by this virus on public health, the World Health Organization (WHO) has declared it as a global pandemic and calls for health sectors of the world and government to take it seriously [10]. The spread of the infectious disease is still on the rise despite many efforts people and government of nations to contain it, such as containment, an individual measure of protection, the authorization of the use hydro-chloroquine and other drugs that have not been clinically tested [12].

There are uncertainties as to how COVID-19 is transmitted, although medical researchers (centers) for disease control have noted that it is transmitted via a fluid contact with an infected person [12-15]. They further noted that symptoms of the disease on an individual includes, sore throat, dry cough, fever, shortness of breath, and in the worse stage acute pneumonia and death. According to [14], the incubation period for COVID-19 is between 2-14 days. Also, [15-16] asserted that the reason for the high level of the spread of the virus on individuals across the globe as a result of the symptomatic and asymptomatic nature. Failure to identify and properly manage both the symptomatic and asymptomatic COVID-19 cases by mass and wide COVID-19 testing or screening of the population puts the race at risk [12], also this would significantly affect the way of life of people, as well as businesses across the globe.

1.2. Causatives on The Fast-Paced Spread Propagation in Nigeria

The world is currently dealing with the realities imposed on her by the coronavirus (COVID-19) pandemic, which has led to a huge economic loss for thousands of businesses across the globe. This loss in finance is out-rightly attributed to many of the government's order of shutting down business operations [17-18]. In Nigeria, it is the same scenario as the major economic hub States are on lockdown as a result of the upsurge of the coronavirus. This lockdown features restriction on transportation of people and goods, significantly disrupt outputs and exports, hinders the growth of the private business by undercuts in the investments due to loss of investors' confidence in the market [19]. Against this backdrop of government lockdown and the halt of business activities, private businesses financial and non-financial performances were affected. Same feature befalls the private sector (businesses) and government. As at 9:30 pm 11th April 2020, there were 318 confirmed cases 70 discharged and 10 deaths in Nigeria. The virus tends to be spreading more around and to states in the country with Lagos- 174, FCT-56, Osun-20, Edo-12, Oyo-11, Ogun-7, Bauchi-6, Kaduna-6, Akwa-Ibom-5, Katsina-4, Delta-3, Enugu-2, Ekiti-2, Rivers-2, Kwara-2, Ondo-2, Benue-1, Niger-1, Anambra-1 and Kano-1 respectively [16]. This has led to the federal government taking radical steps like the issuing of an executive order to lock down the major economic hub of the country with a view of the virus.

2. Materials And Methods

2.1. Dataset Gathering

The diffusion study uses as its innovation an attitude change in behavior towards covid-19 spread. The study initiative will sought a convergence in time of final number of adopters – with the singular goal that youth participation, they imbibe idea that behavioral change is a panacea to curb the spread of coronavirus in Nigeria, and (b) they are the right machinery needed to steer such new vehicle to redirect Nigeria. The scheme was designed to convene youths and some select (known infected) cases. All, were kept in a close-contact for the period of eight (8) week – during which time, it is expected that: (a) the youths (inactive or un-infected) will have been exposed to the seedset nodes (in this case, infected cases), (b) seedset cases will be quarantined for the period of 21-days, (c) all nodes (both infected and non-infected) will have been susceptible, infected, removed and made susceptible again. Also, the time in which un-infected nodes (persons) were first exposed to the innovation is known as the time-of-adoption; the time in which the project ends (i.e. 8-weeks) is noted as $t = n$; and, the time before that is $t = t - 1$.

Network data were collected as follows [20] in table 1:

Table 1. Dataset as Pre-Analyzed on Social Network

Network Feats	Youth Innovation
Number of Local Governments	27
Number of Communities	36
Clusters in each community	30wards (clusters)
Targeted Number in Age range	25,000 youth
Seed Sets Phase	5seed-set nodes per Ward
Time of Diffusion / Year	8-weeks
Graph Probability Distribution	$\mu = 0.5346$ and $\delta = 0.34$
Average time of Adoption	3-days
Lowest and Highest Saturation (final number of adopters)	36% and 80% at $t > 0$ and $t = t - 1$ respectively

- a. List ten (10) friends within the age range of 16-24years for the project. The timeline of 8-weeks was chosen. Project is based on: (i) expose actor-nodes to seedset in group without their awareness (so as not to cause panic, indifference and isolation) and we noticed interactions grew with time as they shelved selfish claims and self-centered gains, (ii) seek in time the convergence of a final number of adopters, and (iii) 8weeks timeline will only give time that allows for a sufficient diffusion of the innovation and project. With the change in time, actor-nodes begin to interact and build their personal networks and clusters, which in turn will be tested over time. The innovation (exposure of inactive nodes to the seedset also needs time to be nurtured).
- b. Ten (10) most influential youths in the age range in their immediate community will help reach and further expose the un-infected nodes.

While, the collected data is potentially accurate, its recall event may be erroneous and we hope that these recall errors are normally distributed. The study does not wish to discuss the effect of the network structure on its convergence time nor the effect of clustering on the diffusion model cum process. Many studies note that innovations diffuse faster on highly clustered networks and much slowly on graphs with small degree of clustering. Our experimental model, rather aims at the expected number of final adopters as a function of the convergence task for a specified duration, the effect of behavioural evolution towards the spread of *coronavirus* as the innovation.

2.2. Spectral Clustering (SC)

Spectral clustering [21-22] uses the top eigenvectors of a matrix derived from the input data and transforms the clustering problem into a graph cut problem. The graph cutting approach clusters data points by attribute such that densely packed points are in same cluster, whereas the sparse are in different clusters. The minimum cut formula is given by Eq. 1 as:

$$cut(x, y) = \sum_{i \in A, j \in B} W_{ij} \quad (1)$$

where w_{ij} is the degree of a vertex from i to j . There are usually two balanced cut approaches namely: ratio Cut and Ncut. Also, spectral clustering is a relaxation of the two cut approaches. It is an effective partitioning strategy that includes normalized cut and Ng-Jordan-Weiss algorithms, which are also a relaxations of spectral partitioning problem. Below is SC algorithm with a Laplacian matrix listing described in Algorithm 1.

Step 1: The dataset is divided into two components: training and testing datasets. The training dataset is clustered by SC and this output is regarded as the training subset labelled Training Datasets 1– k . SC centres from the training dataset clustering process are stored to serve as initialisation cluster centre for generating the testing dataset clusters. Because intrusion data features indicate similar attributes of each type in the raw dataset, points in the training dataset with similar features are aligned into groups and regarded as the same subset. In order to improve its performance, different cluster numbers and values of sigma are considered. The number of clusters ranged from 2 to 6 and sigma from 0.1 to 1.0. The samples are assigned to one cluster by similarity. The minimum distance from a data point to each cluster centre is measured by Algorithm 1 and each point is assigned to a cluster.

Listing of the Spectral Clustering Algorithm

Algorithm 1: Spectral Clustering Algorithm

Input: Dataset, clusters k , parameter s and number of iterations $iter$

Output: the set of k clusters

/* Note: symbols “/” and “/” represent comments in the algorithm */

1. Calculate affinity matrix $A \in \mathbb{R}^{n \times n}$ and define $A_{ij} = \exp(-\|s_i - s_j\|^2 / 2\sigma^2)$ in which s_i and s_j are the original data points i and j , respectively */
 2. If $i \neq j$, then the $A_{ii} = 0$
 3. D is diagonal degree matrix and computed with elements: $d_i = \sum_j^n A_{ij}$. So that given the graph G with n -input nodes or vertices, the Laplacian matrix $L(n \times n) = D - A$
 4. Find k -largest eigenvectors of matrix L and $[x_1 x_2 \dots x_k] \in \mathbb{R}^{(n \times k)}$
 5. Generate the matrix Y by renormalizing each row of x as thus: $Y_{ij} = X_{ij} / (\sum_j X_{ij}^2)^{\frac{1}{2}}$
 6. Minimize the distortion of each row Y to regard as the point in clustering term \mathbb{R}^k using any clustering algorithm, such as a distance-based clustering approach.
 7. Finally, the original point s_i is assigned to cluster j when the row of y_i belongs to the cluster j .
-
8. return the set of k clusters and cluster centre
-

Step 2: The testing dataset, which has been divided from the raw dataset, is used to generate k datasets. The previous cluster centres obtained from SC cluster in Step 1, are regarded as the initialization cluster centres of the SC algorithm in this step. The test dataset which are divided by SC cluster, are regarded testing subsets. These subsets are denoted as Test 1 via Test k .

Step 3: The k testing data subsets are fed into k clusters, which were completed by the k training data subsets in Step 1. The output of each is integrated as the final output and employed to analyse positive detection rates. Then, a confusion matrix is used to analyse the data-point classification performance.

2.3. Susceptible-Infect-Susceptible (SIS) Model

Here, [7] at discrete time-step at $t = 0$, the graph is inserted with actors exposed to the innovation – so that if an actor x is exposed and has adopted the innovation at time t , it has single chance to convince its immediate neighbour y that is currently active but yet-to-be-exposed. Probability x succeeds with y is P_{xy} . If x succeeds, y is exposed and has adopted the innovation at time at $t+1$; Else, x tries again in the future (even if y adopts via another actor’s neighbour). Also, actor y can decide to reject the innovation at another time-step (due to path dependence, morphing of ties, external influences and other feats). This process continues and stops after n -steps ($t - 1$), at which point there are no more actors to be exposed to innovation. It requires an actor to be exposed more than once. The graph can be of size M , with M_d subset of nodes and d copies of seed-set actor nodes placed on the network. With propagation complete, $S(M_d, G)$ is the expected number of final adopters. Thus, the process continues to evolve in time to either propagate or eventually die. An actor x adopts based on the rate of $\frac{\beta}{\delta}$ and probability β . At same time, an actor may reject the

adoption after having previously adopted probability δ . With an adjacency matrix T , $\lambda_1(T)$ is the largest eigen-value of T . Thus, $\frac{\beta}{\delta} < \frac{1}{\lambda_1(T)}$ is true as threshold and is sufficient for quick recovery as easily proven [20].

2.4. Susceptible-Infect-Susceptible (SIS) Model

Here, [7] at discrete-time at $t = 0$, the graph is inserted with actors exposed to the innovation – so that if an actor x is exposed and has adopted the innovation at time t , it has single chance to convince its immediate neighbour y that is currently active but yet-to-be-exposed. The probability that x succeeds with y is P_{xy} . If x succeeds, y is exposed and has adopted the innovation at time $t+1$; Else, x tries again in the future (even if y adopts the innovation via another actor’s personal network or neighbour). This process continues and stops after n -steps (at time $t - 1$) – at which point there are no more active actors to be exposed to the innovation. It requires an actor to be exposed exactly once. The graph can be of size M , with M_d subset of nodes and d copies of seed-set actor nodes placed on the network. With propagation complete, $S(M_d, G)$ is the expected number of final adopters. Expectation can exceed all random choices depending on the diffusion model and nature of graph structure in use. Eq. 2 describes the maximum expected number of final adopters, which exceeds all possible initial seed-set choice adopter placements [20].

$$S_d(G) = \max_{M_d} S(M_d, G) \quad (2)$$

Subset $A_d = \arg \max_{M_d} S(M_d, G)$ corresponds to choices made by an adaptive adversary. $S_d(G)$ is epidemic spread in G and a similar definition of epidemic spread of randomize adversary as in Eq. 3 in which case, define it to be the expected epidemic spread where the expectation takes over all possible positions of the d viruses placed on the network and given by:

$$S'_d(G) = E_{M_d}[S(M_d, G)] \quad (3)$$

2.5. Behavioral Evolution

Events modelled on a social graphs will never treat all actors as equal. Thus, it distinguishes between trusted friends from total strangers (in that continuum) via relationship spectrum as provided by *ties* leading up to improved *behaviour* (behavioural evolution). Loose acquaintances that an actor interacts with on demand are said to have *weak ties* with between such actors. Though weak, they can help such actors generate ideas, adopt innovations as well as expedite knowledge transfer across the network or groups. Conversely, interaction between trusted friends and family are said to be *stronger-ties*, which affects the actor’s emotional health and can lead a society to safe heavens in time of crisis. Whether a link between actors exists or not, relationships have their own few properties that helps define tie-strength that exists between agents as they form clusters, cliques and communities [23].

Dyads are all pair of interactions to measure relationships of n actors with m ties – resulting in an $m \times n$ binary matrix of elements in G . The ordered pair (i, j) is 1 if it is sampled and 0 if not. Both directed/undirected graphs have a set of observed dyads incident in at least one sampled node. Behavior then, can be measured from 7-dimensions that manifest in various forms to include: (a) time structure, (b) emotional intensity, (c) mutual intimacy and (d) reciprocal services [24]. Also, [25] notes that structural variables and factors like (e) topology and (f) emotional support from informal circles indicate a stronger tie. Lin et al (1981) note that: (g) social distance as embodied in socio-economic status, education, race, political affiliation and gender of an actor – all which influences tie-strength. In practice, the structural variables are substituted with simple proxies such as communication recency, interaction reciprocity, mutual friendship and frequency of the interaction.

Thus, [20] in extending [26] with the potential of a feedback into the society in ways that benefit users as in Eq. 3. Thus, tie-strength is modeled as a linear combiner, where R_i represents number of predictive variables used in task, e_i is error term, D_i is dyads pairs, N_i is network structure and EI_i are external influences.

$$S_i = \alpha + \beta R_i + \gamma D_i + N_i + EI_i + e_i \quad (3)$$

A cluster is a strong candidate of a cohesive set or clique by itself due to large number of ties amongst members. As clusters increase, actors form more cliques with others in close proximity, which also decreases the set cardinality and

increase the number of sets in graph. High-clustered graphs have short-path length, large number of cliques with small cardinality. It results in large expected number of final adopters, introduces a close knit relationship between clustering coefficient and bound for number of final adopters. Such graphs are quite beneficial for complex diffusion process (contagion) to reinforce adoption of the innovation as they are more likely to be exposed to multiple adopters and overlapping influences via such short path-lengths during diffusion. Its merit is, in the existence of a seed set adopter inside the cluster. In contrast, with no *seed-set* adopter in a locally dense cluster, it is highly stable and may resist adoption of innovation. Thus, while clustering reinforces adoption if the innovation penetrates, it also weakens adoption by making penetration more difficult for small cluster coefficient k . Graphs with smaller clusters have long path-lengths to diffuse innovation more and further [27].

N_i encodes an idea that its structure is dependent on the structural predictive variable in tie-strengths with respect to: (a) tie history, (b) reciprocity, and (c) cohesive clusters and its coefficients. A major reason in using clustering includes that: (i) there is a direct relationship between cohesiveness and clustering, and (ii) clustering has been used extensively in many studies to capture graph network structure as it also allows for easy comparison of result – is given as thus:

$$N_i(G) = P_{\theta,G} + \lambda_0 \mu_L + \lambda_1 Med_L + \sum_{\tau=0}^{\tau-1} \sum_{i \in L} \lambda_{\tau} (s - \mu_L)^{\tau} + \lambda_5 Min_L + \lambda_6 Max_L \quad (4)$$

where $L = S_j$ and i, j are mutual friends. N_i uses the following parameters: (a) $P_{\theta,G}$ to encode the system/graph’s probability distribution of actors via defined upper and lower bound of agent dispositions in the system, (b) μ and Med are actor’s cum system threshold respectively of graph, (c) expression $\sum_{\tau=0}^{\tau-1} \sum_{i \in L} \lambda_{\tau} (s - \mu_L)^{\tau}$ encodes learning with momentum as a function of its convergence in time/outcome of the diffusion process, and (d) Max/Min values are upper/lower bounds of final adopters [28].

Since each node i ’s neighbour has potentially unique set of mutual friends, the model uses 5-descriptors of tie-strength distribution to describe the graph’s structure thus: (a) mean number of ties to cater for cohesion among the nodes of the graph, (b) variance in the number of ties to cater for nodal and joint degree of separation in the graph, (c) kurtosis of the ties to cater for clusters and its coefficient, (d) minimum number of ties in the graph, which caters for local density and degree of distribution, and (e) maximum number of ties in graph, which caters for reciprocity – all belonging to the structural dimension of ties as associated to clustering and cohesiveness of the graph system; while it also introduces a dependency that tie-strength also depends on other tie-strengths [26, 20].

2.6. Dynamic SIS and SIR Implementation

As diffusion process continues at $t \geq 0$, actors are first exposed (to seed-set) to allow them form their perception and behaviour about the innovation. In time, as more of the actors are exposed to innovation, they make preferences – and based on their threshold (behaviour), adopt the innovation and form clusters and cliques via learning outcome. These strengthen their ties as well as improves the actor’s personal network and consequently, helps them better retain information within their memory in time as the system continues in its search for optimality [29-31]. The random exchange in an actor’s personal network allows knowledge swap – so as to yield in time agents with a new set of disposition (consider adoption and behaviour change cum evolution). As more adoption is encountered, more agents continue to learn/retain contents within their memories that better their personal network via community-influences. The listing in Algorithm 2 shows the SI (susceptible-infect) model

Listing of the SI Algorithm

Algorithm 2: Spectral Clustering Algorithm

Input: Dataset, clusters k , parameter s and number of iterations $iter$

Output: the set of k clusters

1. Initialize number of nodes n ; number of Ties m
 2. Set Initial Tie Strength = 10+, Cluster Structure = 25+;
 3. Set Network Structure as function of clusters
 4. Initialize Graph via PD of $P_{\theta,G}(G = g) = \frac{\exp[\theta^T u(g)]}{c(\theta,G)}$
 5. Set Agent position with Max/Min bounds of expected number of final adopters as a function of Motion or Movement M .
-

6. Randomly select nodes for *seedset choice*
7. While Node are yet to be Exposed
8. Choose agents position in *seedset* as best position in graph
9. Initialize current Agent position = $\{M_{\min} + \text{rand}(P_{\max} - P_{\min})\}$
10. Compute Threshold cum Path dependence for all nodes in Graph
11. Set Pairwise Variables Interactions (+D_i)
12. Set Predictive Variables for task (+R_i)
13. Compute Agent's *Tie Strength* as function $\{+R_i, +D_i, +N_iG, +El_i\}$
14. For Each Agent, Node or Actor i
15. Do {
16. If *seedset* is member of an agent's personal network
17. Then *node.list.append(seedset node)*
18. End If }
19. End For Each
20. Compute Network Structure $N_i(G) = P_{\theta,G} + \lambda_0\mu_L + \lambda_1Med_L + \sum_{t=0}^{t-1} \sum_{i \in L} \lambda_t (s - \mu_L)^t + \lambda_5Min_L + \lambda_6Max_L$
21. Compute actors change in behaviour from exposure at $t \geq 0$ as: $M_{\text{new}} = w * M_{\text{old}} + c_1 * \text{rand}() * [(P_i)/T] + c_2 * \text{rand}() * [(P_n)/T]$
22. Updating agents' positions as $P_{\text{new}} = (P_{\text{old}} - M_{\text{new}})$
23. // continue till all nodes are exposed, implies stop criterion reached

Each exposure yields an updated number of final adopters as its optimal solution in time via recomputed threshold value for exposed actors. We note that: (a) the agent position range is normalized between [0-1] dividing it by maximum range of agents, (b) each position randomly determines swap type needed for adoption rate, and (c) positions are reset and these recomputed new values will eventually reflect system threshold. With each solution found, model restarts with another randomly selected point for the planted seed-set choice in the graph space [32].

Agents with threshold value above 0.5 are chosen. Process continues till all agents are exposed time $\geq t - 1$ at which all agents will have a threshold of 1 for the diffusion process or the nodes are continuously re-evaluated till an agent is found of threshold lesser than or equal to start-off threshold value (these form the stopping criterion for the model). At which point the solution is reached [33-35].

3. Findings and Discussion(s)

3.1. Model Evaluation

Model performance is evaluated via computed values of mean square error (MSE), mean absolute error (MAE) and mean relative error (MRE), coefficient of efficiency (COE) and coefficient of determination (R^2). MSE, MRE, MAE have an ideal value 0; while COE and R^2 have an ideal value 1 [26]. Validation is more of a scientific discussion that ambiguous results as improperly applied, often impedes. Our study aims to minimize confusion in social graph diffusion using agent based modeling as we aim to establish parameters for its measurement [36].

Table 2. Model Performance of Parameters

Strategy	MAE	MRE	MSE	COE	R^2
SIS	0.012	0.110	0.036	0.753	0.821
SIR	0.01	0.192	0.029	0.688	0.812
Spectral	0.10	0.110	0.032	0.871	0.901

Model's performance aim to exploits predictive variables in network structure, tie strength and pairwise interactions – as the model's achieves its coefficient of determination R^2 with MSE, MRE and MAE at $p < 0.1$ on a continuous 0–1 scale, where 0 is weakest and 1 is strongest. On the average, R^2 for the parameters as predicted shows seventh-tenth of its true-value. Its error interval tightens at end of the continuum to suggest a strong evidence of interaction between all these dimensions at $p < 0.1$. Parameters as structural dimension, plays minor role (in its linear form factor) but has an important modulating role to imply that relationship matter; Thus, cursory look via clusters impacts on ties, network structure and eventually, behaviour change.

3.2. Model Convergence

Figure 1 shows the time convergence of the graph models with profile hidden Markov model (PHMM) as the benchmark model. Spectral is red, SIS is blue, SIR is purple; while, the benchmark (PHMM) model is green. Figure 1 shows mean convergence time on the effectiveness cum efficiency of proposed model-based solution.

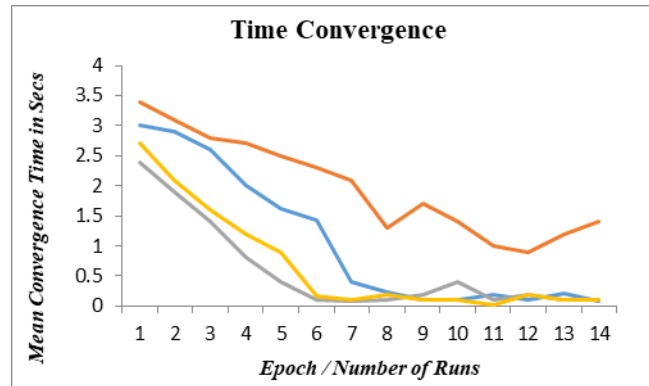


Figure 1. Convergence Time Using Proposed Model

Figure 2 shows the prediction accuracy in simulating the effectiveness and efficiency of proposed models.

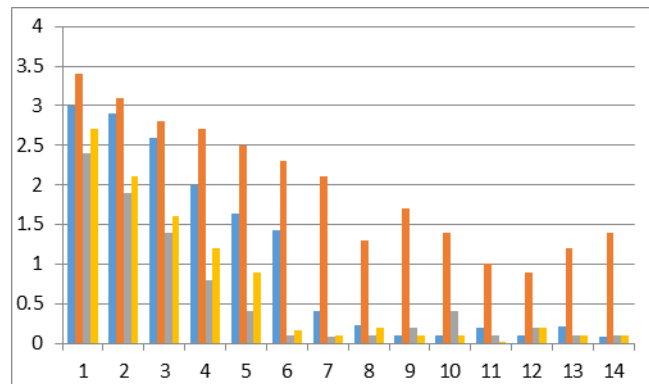


Figure 2. Prediction Accuracy Using Proposed Models

3.3. Result Presentation and Discussion

The adoption of a linear combiner model for behavioural evolution and change with ties prediction allows us to leverage on the benefits of the dataset and explain the results once built. s_i is ties between actors i , R_i is number of predictive variables, e_i is the error term in ties, $N_i(G)$ is network structure. D_i all pairwise between the predictive variables with 90% or greater completion rate given the 7-dimensions as thus: reciprocity, intensity, social distance, intimacy, emotional support, and structural distance. Of these, we note that there are 3-major variables for the 7-dimensions namely: reciprocity, intimacy and emotional support. Using a linear combiner also force more variables than required data points into the model with a 90% system-threshold as in table 3, to ensure every dimension is adequately represented to explore interactions between dimensions of behavioural evolution, as its novel approach

Table 3. Model Performance of Parameters

Strategy	μ	$+R_i$	$+N_i$	$+D_i$	EL_i
SIS	0.47	0.87	0.78	0.89	0.10
SIR	0.23	0.42	0.34	0.43	0.19
Spectral	0.43	0.90	0.92	0.95	0.21

The results shows that where *seedset* of agents that can influence others exists, and are uniformly distributed over the graph, the study explores the relationship of the expected number of final adopter (our metrics), clustering effects, threshold values with path-dependence, network structure and seed cardinality – to suggest that the highly clustered networks with more seed-sets allowed for easier diffusion of innovation. While clustering and high clustering coefficients promotes diffusion where there exists seed node inside an agent’s personal networks and social system, such cohesive and highly clustered sets are also quite difficult to penetrate when they are not targeted during initial seeding phase. In some cases, sampling strategy used to generate the agents for the diffusion process can create some errors in the time of adoption, alongside the external influences that acts on the system etc. These external influences such as finance, tribalism, race and gender etc – can also act as a motivator during each transaction that involves corruption ethic, especially with such actors not being monitored. This will alter the outcome of the diffusion process significantly. A major reason why project was kept at 8weeks is to seek time convergence and the acceptability of the initiative that will help change actor perception, disposition and behaviour and consequently allow complete diffusion as actors create stronger ties. External influence and path dependence only accounts for about 11.8% of deflection from the true purpose (as in table 2).

4. Conclusions

With graph-based model, there are still a lot more variance to understand such as predictive variables and “behind-the-scenes” data. Though, the addition of more data to the task at hand may not solve it – but, we have aimed to define some parameters that helps predict social graph. There is the need to standardize parameters for social-graph models, which in turn raises new questions for the theory. In modelling ties, it is important to know: (a) what feats and parameters are necessary predictors to be used in predicting threshold, and (b) what limits are to be set for behavioural evolution. We believe our work makes some important contributions to the theory as thus: (a) extending tie-strength dimension as manifested in all social-graphs, (b) defining network structure dimension as a function of probability distribution of agents in problem space, (c) all dimensions modelled as a continuous value, (d) our result extends the realization of how structural dimension in predictive variables used in task help modulates other dimensions within, by filtering agent relationships via cliques and clusters, and (e) previous works assumed either the presence or absence of a link in the graph without recourse to the properties of the link itself.

References

- [1] World Health Organization. Coronavirus Disease (COVID-2019) Situation Reports. Available online: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports> (accessed on 2March 2020).
- [2] A.A. Ojugo., O.D. Otakore., (2020). *Forging an optimized Bayesian network model with selected parameter for detection of the Coronavirus in Delta State of Nigeria*, J. of Applied Sci., Engr., Tech. & Edu., 3(1): pp37–45, 2020, doi: 10.35877/454RI.asci2163
- [3] Adegboye, O.; Adegunle, A.; Pak, A.; Gayawan, E.; Leung, D.; Rojas, D.; Elfaki, F.; McBryde, E.; Eisen, D. Change in outbreak and impact on importation risks of COVID-19 progression: modelling study. medRxiv **2020**. [CrossRef]
- [4] Martinez-Alvarez,M.; Jarde, A.; Usuf, E.; Brotherton,H.; Bittaye,M.; Samateh, A.L.; Antonio,M.; Vives-Tomas, J.; D’Alessandro, U.; Roca, A. COVID-19 pandemic in west Africa. Lancet Glob. Health **2020**. [CrossRef]
- [5] Gilbert, M.; Pullano, G.; Pinotti, F.; Valdano, E.; Poletto, C.; Boëlle, P.-Y.; et al., *Preparedness and vulnerability of African countries against importations of COVID-19: A modelling study*. Lancet **2020**, 395, 871–877. [CrossRef]
- [6] Ojugo, A.A., O.D. Otakore., *Forging an optimized Bayesian network model with selected parameters for detection of corona-virus in Delta State Nigeria*, J. Applied Sci., Eng, Tech & Edu. 3(2): pp128-138
- [7] Ojugo, A.A., D.A. Oyemade., *Predicting potential spread of corona-virus in Nigeria through ties threshold for dynamic SIR graph model*, submitted to Nigerian Journal of Research and Technology. 2020.
- [8] Adegboye, O., Adegunle, A, Gayawan, E., Early transmission dynamics of novel COVID-19, Int. J. Environmental Research and Public Health, medRxiv **2020**. [CrossRef]
- [9] J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell et al. *Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts*, 2013, doi: 10.1371/journal.pone.0072168
- [10] McKibbin, W. & Fernando, R. (2020). The Global Macroeconomic Impacts of COVID-19: Seven Scenarios. Retrieved from <https://researchgate.net>

- [11] Worldometers (2020). Global Statistics on COVID-19 Pandemic Outbreak. Retrieved from, <https://worldometers.info/coronavirus>
- [12] Addi, R. A., Benksim, A., Amine, M., & Cherkaoui, M. (2020). Asymptomatic COVID-19 infection management: The key to stopping COVID-19. *Journal of Clinical and Experimental Investigations*, 11(3), 1-2
- [13] Kim, K. H. (2020). COVID-19. *International Neurological Journal*, 24(1), 1-1. doi.org/10.5213/inj.2020edi.001
- [14] Nigeria Centre for Disease Control [NCDC] (2020). COVID-19 Pandemic Outbreak. Retrieved from, <https://covid19.ncdc.gov.ng>
- [15] Bai, Y., Yao, L., Wei, T., Tian, F., Jin, D, Y., Chen, L. & Wang, M. (2020). Presumed asymptomatic carrier transmission of COVID-19. *JAMA*.
- [16] Nigeria Centre for Disease Control Twitter Handle [NCDC] (2020). COVID-19 Pandemic Outbreak of Statistics. Retrieved from, <https://twitter.com/ncdc.gov.ng>
- [17] Tashanova, D., Sekerbay, A., Chen, D., Luo, Y., Zhao, S. & Zhang, Q. (2020). Investment opportunities and strategies in an era of coronavirus pandemic. Retrieved from <https://ssrn.com/abstract=3567445>
- [18] Kim, J., Kim, J., Lee, S. K., & Tang, L. R. (2020). Effects of epidemic disease outbreaks on the financial performance of restaurants: Event study method approach. *Journal of Hospitality and Tourism Management*, 43, 32–41.
- [19] Jung, H., Park, M., Hong, K & Hyun, E. (2016). The impact of an epidemic outbreak on consumer expenditures: An empirical assessment for MERS Korea. *Sustainability*, 8(454), 1-15. DOI:10.3390/su8050454
- [20] A.A. Ojugo., F. Aghware., R. Yoro., M. Yerokun., A. Eboka., C. Anujeonye., F. Efozia., *Predict behavioral evolution on graph model*, Advances in Networks, 3(2): pp8-21, 2015
- [21] Aggarwal CC, Reddy CK. Data Clustering: Algorithms and Applications. Chapman and Hall, 2013.
- [22] M. Azarafza, M. Azarafza, H. Akgunc, *Clustering method of spread pattern analysis of corona-virus (covid-19) infection in Iran*, J. Applied Science, Engineering Tech. & Education, 3(1) pp1-6, doi: 10.35877/454ARI.asci31109, 2020.
- [23] David, P.C., (2007). *Path dependence – a foundational concept for historical social sciences*, *Climetrica – Journal of Historical Economics and Econometric History*, 1(2).
- [24] Granovetter, M. (1978). *Threshold Models of Collective Behavior*. *American J. Sociology*, 83(6), p1420–1443. doi:10.1086/226707, (<http://dx.doi.org/10.1086%2F226707>). JSTOR 2778111
- [25] Sala, A., Cao, N., Wilson, C., Zablith, R., Zheng, H and Zhao, B.Y., (2010). *Measurement calibrated graph models for social network experiments*, IW3C2, ACM 978-1-60558-799-8/10/04.
- [26] Gilbert, E and Karahalios, K., (2009). *Predicting tie strengths with social media*, *Journal of computer and Human Interface*, 15, p76-97, ACM 978-1-60558-246-7/09/04.
- [27] Golbeck, J. (2013). *Analyzing the Social Web*, Morgan Kaufmann, ISBN: 0-12-405856-6.
- [28] Granovetter, M. (1985). *Economic action and social structure: The problem of embeddedness*, *American J. Sociology*, 91(3), p481–510, doi:10.1086/228311 (<http://dx.doi.org/10.1086%2F228311>).
- [29] Handcock, M.S and Gile, K.J., (2009). *Modeling social networks from sampled data**, *Annals of Applied Statistics*, arXiv: math.PR/00000.
- [30] Schnettler, S., (2009). *A small world on feet of clay? A comparison of empirical small-world studies against best-practice criteria*, *Social Networks*, 31(3), p179-189, doi:10.1016/j.socnet.2008.12.005.
- [31] Scott, J.P., (2000). *Social network analysis: A Handbook* (2nd Ed). Thousand Oaks, CA: Sage Publications.
- [32] Smith T.S and Stevens, G.T., (1999). *The architecture of small networks: strong interaction and dynamic organization in small social systems*, *American Sociological Review*, 64, p403–20
- [33] Toivonen, R., Kovanena, L., Kivelä, M., Onnela, J.K., Saramäki, J and Kaskia, K., (2009). *A comparative study of social networks models: network evolution and nodal attributes models*, *Social Networks*, 31, p240-254, doi:10.1016/j.socnet.2009.06.004.
- [34] Valente, T.W., (1996). *Social network thresholds in the diffusion of innovation*, *Social Networks*, 18, p69-89, SSDI 0378-8733(95)00256
- [35] Wilson, C., Boe, B., Sala, A., Puttaswamy, K.P and Zhao, B.Y., (2009). *User interactions in social networks and their implications*. In *Proc. of EuroSys* (April 2009).
- [36] Ojugo, A.A., A. Eboka., E. Okonta., et al, *Genetic algorithm rule-based intrusion detection system*, *J. of Emerging Trends in Computing Information Sys.*, 3(8): pp1182-1194, 2012