

Research Article

Analysis and Characterization of the Spread of COVID-19 in Mexico through Complex Networks and Optimization Approaches

Edwin Montes-Orozco ,¹ Roman-Anselmo Mora-Gutiérrez ,¹ Sergio-Gerardo de-los-Cobos-Silva ,² Eric A. Rincón-García,² Miguel A. Gutiérrez-Andrade ,² and Pedro Lara-Velázquez ²

¹Departamento de Sistemas, Universidad Autónoma Metropolitana Azcapotzalco, Mexico, Mexico ²Departamento de Ingeniería Eléctrica, Universidad Autónoma Metropolitana Iztapalapa, Mexico, Mexico

Correspondence should be addressed to Edwin Montes-Orozco; emontes@azc.uam.mx

Received 13 August 2021; Revised 28 December 2021; Accepted 27 January 2022; Published 23 February 2022

Academic Editor: Samuel V. Scarpino

Copyright © 2022 Edwin Montes-Orozco et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work analyzes and characterizes the spread of the COVID-19 disease in Mexico, using complex networks and optimization approaches. Specifically, we present two methodologies based on the principle of the rupture for the GC and Newton's law of motion to quantify the robustness and identify the Mexican municipalities whose population causes a fast spread of the SARS-CoV-2 virus. Specifically, the first methodology is based on several characteristics of the original version of the Vertex Separator Problem (VSP), and the second is based on a new mathematical model (NLM). By solving VSP, we can find nodes that cause the rupture of the giant component (GC). On the other hand, solving the NLM can find more influential nodes for the entire system's development. Specifically, we present an analysis using a coupled social network model with information about the main characteristics of the contagion and deaths caused by COVID-19 in Mexico for 19 months (January 2020–July 2021). This work aims to show through the approach of complex networks how the spread of the disease behaves, and, thus, researchers from other areas can delve into the characteristics that cause this behavior.

1. Introduction

1.1. Complex Networks, Essential Nodes, and Robustness. In recent years, the study of complex networks and their applications has been an essential topic for research [1–3] because most of the activities can be modeled as networks, for example, the spread of diseases or information, social relationships, transport systems, and electrical or communications networks [4–6].

Then, we can define a complex network as a network with special characteristics that do not occur in simple networks, such as degree distributions, high local cohesion, community formation, and emergent and dynamic properties [4].

The identification and quantification of influential nodes in complex networks are an important activity in several application fields, such as the spread and control of diseases [7], the identification of the most influential members of a criminal group [8], or predicting future commercial relationships [9–12].

In contrast, robustness in complex networks measures the capacity of systems versus specific external or internal events to preserve a relatively normal behavior. Then, we can summarize the robustness of the network as its ability to keep most of the nodes connected after a disturbance (internal or external) [4].

1.2. Coupled Networks. A coupled network is a class of complex networks introduced to better model complex real-world systems. The main characteristic of coupled networks is that two or more independent networks form them.

For example, in Figure 1, we have two interdependent networks (layers). Then, a coupled network is a network

where the nodes of Layer B form a subset of the nodes of Layer 1 (or vice versa), and nodes in common to both networks are coupled. In short, a coupled network comprises all the nodes and all the links that are in each independent layer.

1.3. Approach. In this work, we present two methodologies: the first is to quantify robustness in networks based on the Vertex Separator Problem and the second is based on the information about the shortest paths between each pair of nodes and the degree of each node in the network, where we obtain a quantification of the influence of each node by solving a mathematical programming model based on Newton's law of motion.

Then, we present a set of coupled networks modeled with information on the main information on the spread and deaths caused by COVID-19 in Mexico. In this case, as a case study, it is essential to mention that we have 17 networks for data related to deaths, 18 for data related to infections, and 19 for suspected cases because the information collected is from February 2020 to July 2021.

On the other hand, all the information about COVID-19 used to model the three types of networks was collected from "https://coronavirus.gob.mx/," where the main information studied is the numbers of infections and deaths by state. In this work, we focus on analyzing the information for the 16 municipalities of Mexico City and the 125 municipalities of the State of Mexico.

The idea of working with coupled networks is based on the advantage of analyzing the most important nodes to maintain the connection of the network (robustness) and the identification of the most influential nodes (for the flow of the disease) from the information divided into different periods without leaving aside their conjunction.

Thus, as a result, we can obtain an identification and a quantification based on an analysis that reduces the loss of the dynamic visualization of the complex system.

2. Related Work

Although we defined robustness in complex networks, there is no consensus on the formal definition. However, some works denoted in the specific literature present some properties and characteristics to quantify [13–16].

For example, in [17–19], the authors present some metrics and measurements to quantify the robustness in monolayer networks based on the identification of the most fragile nodes of the network, with respect to the number of direct connections, and the estimation of the importance of these nodes when a disturbance occurs (external attack or internal failure).

In Ren et al.'s work [20], the authors describe that the algorithm for randomly eliminating nodes is ineffective under certain conditions, because removing certain nodes can take a long time to process. Therefore, they propose a link removal strategy called Hierarchical Power Iterative Normalized Cut (HPI-Ncut), in which, instead of focusing on the elimination of nodes, a fraction of connections is eliminated to quantify the impact of nodes that have many or



FIGURE 1: Coupled network example.

few direct connections and to be able to verify their importance to maintain the robustness of the system.

Morone and Makse [21] present a methodology based on the construction and analysis of the optimal percolation problem in random networks, intending to identify the minimum set of nodes that can cause the failure of the system: there is no communication between most of the system elements.

Based on the previous idea, Montes-Orozco et al. [22] present a methodology based on the percolation process and asortality, where certain nodes are eliminated (which have certain special characteristics) to generate disconnection from the network, with which robustness can be quantified. It is known that the elements that belong to a complex system play different roles; therefore, identifying the elements (nodes) that are more influential is very important in real-world issues, since these nodes determine the entire system's behavior.

Identifying the most influential nodes in complex systems has led to the development of different methodologies based on the calculation of the structural information of the network or analogies of statistical physics or mathematical models.

For the methodologies based on structural information, the most representative works are the following:

- (i) Zhong-Ming et al. [23] propose some nodes classification algorithms to identify structural holes. A structural hole is a phenomenon that occurs when a node connected to multiple local bridges (multiple communities) is removed and space is produced.
- (ii) Zhao et al. [24] present an index to calculate the influence of a node based on the number of communities to which it belongs.
- (iii) Wang et al. [25] present a modified efficiency centrality that considers the influence of the average degree of all nodes and the average distance of the network.
- (iv) Li et al. [26] present an improved LeaderRank algorithm, allowing accessible nodes to obtain more values from the base node in random paths.

Regarding the analysis of the disease known as COVID-19, some works are related to the analysis of complex networks. For example, Montes-Orozco et al. [27] present a methodology to identify COVID-19 spreader countries using the inverse percolation approach adapted to work in 5layer multiplex networks. The authors show that the relationship between the sociocultural and economic characteristics caused by COVID-19 greatly impacts the numbers of infections and deaths caused by the rapid spread in countries that do not have good hospital care or social protocols.

Finally, Al-Shargabi and Selmi [28] explore the influence of the COVID-19 pandemic on the tweeting behavior of Saudi users. In particular, the authors show that their research adopts a social media analysis (SMA) for COVID-19 Arab tweets and is based on the analysis of social structures, such as Twitter users and the relationships between them, using networks and graphs theory without the content of the tweets themselves.

3. Materials and Methods

In this section, we present the methodology we use to model each of the networks used to construct the coupled networks and the modification of the Vertex Separator Problem and the mathematical model based on Newton's Law of Motion to identify and quantify the influencing nodes and diffusers of COVID-19.

3.1. Model of Networks about COVID-19 in Mexico. As we mentioned above, we present a case study analyzing the spread of COVID-19 for 2,457 municipalities in Mexico. For these networks, each layer (network) of the coupled network represents a different month, where a node represents each municipality, and the links between the nodes are created using the Euclidean distance between the characteristics to be analyzed (for more information about the process of modeling of networks, see [22, 27, 29, 30]), where mainly we have the following:

- (1) Medical infrastructure.
- (2) Hospital capacity.
- (3) Population.
- (4) Geographic location.
- (5) Distribution of employment, among others.

Then, to better understand the results shown in the following sections, we present the identifiers used for each COVID-19 analysis network in Mexico, where they are in three different branches: number of deaths, number of confirmed infections, and number of suspected cases.

Thus, each network represents a month in February 2020 to July 2021 (in the case of suspected cases, it includes January 2020).

For the death networks, we used the following ID: deaths.2020.*X*, where *X* is the number of the month. For example, deaths.2020.8 represents the network of deaths for August 2020 and deaths.2021.2 represents the network for February 2021.

We used the following ID: cont.2021.*X*, where *X* is the number of months. For example, con.2021.2 represents the network of contagion for February 2021.

We used the following ID: sus.2021.*X*, where *X* is the number of the month. For example, sus.2020.12 represents the network of suspect cases for December 2021.

Finally, the coupled networks are formed by a set of single-layer networks. Then, the indexes are deaths.2020.2-2021.7, cont.2020.2-2021.7, and sus.2019.12-2021.7, which indicates the period to analyze February 2020–July 2021 for deaths and contagion and December 2019–July 2021 for suspect cases.

In Tables 1–3, we present the numbers of nodes and links for each layer which compose the coupled networks to analyze the network size.

Based on the information shown in Table 1, we can see that the number of links for each single-layer network is about 1 and 5 million, where the networks with fewer links are deaths.2021.1 and deaths.2021.2, and the network with the most links is deaths.2020.3. Therefore, it is essential to analyze the relationships between the nodes to obtain information about the deaths caused by COVID-19.

Based on the information shown in Table 2, we can see that the number of links for each single-layer network is about 1 and 6 million, where the network with fewer links is cont.2021.1, and the network with the most links is cont.2020.3. Therefore, as in the previous case, it is essential to analyze the relationships between the elements that compose this network to obtain the characteristics of the contagion dynamics of COVID-19.

For this case, based on the information shown in Table 3, we can see that the number of links for each single-layer network is about 1.5 and 3.7 million, where the network with fewer links is sus.2020.7, and the network with the most links is sus.2020.3.

Therefore, as in the previous cases, it is essential to analyze the relationships between the elements that compose this network to obtain the characteristics of the contagion dynamics and specifically for the suspects' cases of COVID-19 in Mexico.

As mentioned above, we develop three coupled networks formed by the deaths, contagion, and suspect cases networks. Therefore, in Table 4, we present the numbers of nodes and edges that belong to each coupled network.

As in the previous case for the single-layer networks, we can see that the number of links is very high; for this case, it is about 6 million for each coupled network.

Next, we present the main numerical characteristics for each coupled network to analyze (Table 5).

Based on the information presented in Table 5, we can see that the three coupled networks follow the small-world model due to having a high clustering coefficient and small values for the diameter and the average route length. The above is justified because the problem that we are analyzing is a social system, and, therefore, we are working with a social network.

It is essential to mention that we do not calculate more characteristics due to the size of the networks (around 2,500 nodes and 5 million edges (links)).

TABLE 1: Main characteristics for deaths networks.

Network	Nodes	Edges
deaths.2020.3	2,457	5,738,420
deaths.2020.4	2,457	4,006,004
deaths.2020.5	2,457	2,457,056
deaths.2020.6	2,457	1,632,006
deaths.2020.7	2,457	1,257,762
deaths.2020.8	2,457	1,305,306
deaths.2020.9	2,457	1,696,506
deaths.2020.10	2,457	1,903,020
deaths.2020.11	2,457	2,112,662
deaths.2020.12	2,457	1,655,082
deaths.2021.1	2,457	985,056
deaths.2021.2	2,457	985,056
deaths.2021.3	2,457	1,799,622
deaths.2021.4	2,457	2,520,158
deaths.2021.5	2,457	3,227,412
deaths.2021.6	2,457	3,707,550
deaths.2021.7	2,457	2,939,519

TABLE 2: Main characteristics for contagion networks.

Network	Nodes	Edges
cont.2020.2	2,457	5,990,256
cont.2020.3	2,457	4,667,760
cont.2020.4	2,457	2,191,880
cont.2020.5	2,457	967,274
cont.2020.6	2,457	467,172
cont.2020.7	2,457	319,790
cont.2020.8	2,457	375,158
cont.2020.9	2,457	463,080
cont.2020.10	2,457	463,080
cont.2020.11	2,457	587,522
cont.2020.12	2,457	411,522
cont.2021.1	2,457	194,040
cont.2021.2	2,457	373,932
cont.2021.3	2,457	543,906
cont.2021.4	2,457	882,660
cont.2021.5	2,457	1,278,032
cont.2021.6	2,457	1,307,592
cont.2021.7	2,457	444,222

3.2. Robustness through the Adaptation of Vertex Separator Problem (Modified-VSP). As we mentioned above, robustness is the ability to keep the most connected nodes (giant component or GC); therefore, we can quantify the robustness of the network by finding the set of nodes that cause the rupture of the GC.

Formally, the objective of the VSP [31] is to find the minimum set of nodes C that, when they are disconnected from the rest of the nodes, divided the GC into at least two isolated components (A and B), and, formally, the VSP can be modeled as follows:

- (i) *Instance*: A network G (V, E), with |V| = n and a cost ci associated with node i in V.
- (ii) *Problem:* Find a partition of V denoted as (C), where its elimination causes the formation of the disjoint nonempty sets A and B, such that

TABLE 3: Main characteristics for suspect cases networks.

Network	Nodes	Edges
sus.2020.1	2,457	3,688,320
sus.2020.2	2,457	3,703,700
sus.2020.3	2,457	3,095,840
sus.2020.4	2,457	2,221,590
sus.2020.5	2,457	2,324,100
sus.2020.6	2,457	1,859,132
sus.2020.7	2,457	1,576,282
sus.2020.8	2,457	1,804,992
sus.2020.9	2,457	2,180,054
sus.2020.10	2,457	2,318,006
sus.2020.11	2,457	2,314,962
sus.2020.12	2,457	2,072,160
sus.2021.1	2,457	1,746,362
sus.2021.2	2,457	2,460,194
sus.2021.3	2,457	2,780,556
sus.2021.4	2,457	3,046,772
sus.2021.5	2,457	2,980,802
sus.2021.6	2,457	3,198,732
sus.2021.7	2,457	1,655,082

TABLE 4: Main characteristics for coupled networks.

Network	Nodes	Edges
deaths.2020.2-2021.7	2,457	5,893,408
con.2020.2-2021.7	2,457	6,002,542
sus.2019.12-2021.7	2,457	5,068,738

- E does not contain edges (*i*, *j*), with *i* in A and *j* in B.
- (2) max $\{|A|, |B|\}$.
- (3) $P(C_j: j \text{ in } C)$ is minimized subject to 1 and 2.

Based on the above, we can see that, solving VSP on a network, set C that causes the rupture of GC helps quantify the network robustness.

In our case, to obtain homogeneous groups or communities, the modification that we carry out consists of maximizing the sum of weight for the elements that belong to subset C (separator set), while the cardinality of set C is minimized, where it is considered that the weight is the average contagion rate for each of the municipalities to work. Therefore, the sum of the edges that the separating set contains must also be maximized.

Thus, with the adaptation of the VSP to work with coupled networks and considering asortality, we obtain the set of nodes that cause the rupture of GC and that are also chosen homogeneously.

At this point, it is important to mention that we can quantify the robustness of the networks by analyzing the number or percentage of nodes found in the separator set *C* of the GC.

For example, if we use *C*, a high percentage (more than 70%) of nodes indicates that the network is robust, since to cause the rupture of GC, it is necessary to eliminate most of the nodes, while a low percentage (less than 30%) indicates that the network is not very robust, since to cause the rupture of GC, a small number of nodes are needed.

Complexity

TABLE 5: Structural metrics for coupled networks.

Network	Clustering coefficient	Diameter	Average route length
deaths.2020.2-2021.7	0.991948	3	1.02324
con.2020.2-2021.7	0.991948	3	1.05345
sus.2019.12-2021.7	0.99894	3	1.03475

3.3. Simulated Annealing to Solve Modified VSP. The VSP is considered a problem belonging to the NP-Hard class [18, 32]. Therefore, by adding the conditions that cause the communities formed in the components to be homogeneous, the modified version presented in this work is NP-Hard [33]. Therefore, given the nature of our problem, we can solve it using heuristic techniques.

Specifically, for this work, we used a simulated annealing (SA) algorithm [34] developed in C language (for more information, see [30]).

It is important to mention that SA requires 4 control parameters and, to obtain a good performance, it is crucial to find adequate parameter settings for the technique. Then, using the Differential Evolution (DE) algorithm [35] as the calibrator for the main parameters, we obtain the following values (for more information, see [27]):

- (1) Initial temperature $T_i = 1500$.
- (2) Final temperature $T_f = 0.001$.
- (3) Cooling program $T_k \gamma$, with $\gamma = 0.95$.
- (4) Times that a new neighbor is generated $L_k = 20$.

3.4. Identification of Spreader Nodes through Newton's Law of Motion. Because traditional centrality measures focus only on certain aspects (e.g., Degree Centrality focuses only on local information from nodes and ignores global information), we present a methodology to find and quantify the influential nodes in several models and sizes of complex networks.

In this work, each node represents a particle, and we calculate the influence (force) of each node by solving a mathematical programming model that is based on the minimization of the kinetic energy of the system (E_k) , which is defined as the work needed to accelerate a body of a given mass from rest to its stated speed.

We must keep in mind that the forces acting on a particle modify its kinetic energy $E_k = 1/2\text{mv}^2$ and since the aim of our model is the equilibrium state of the system, we have as objective the minimization of the difference between the final Kinect energy (fn) and initial kinetic energy (in) of the system in times T and T + 1, which is equivalent to the scalar sum of the individual kinetic energies:

$$\min z = \frac{1}{2}M_{\rm cm} + \left(V_{\rm cm,fn}^2 V_{\rm cm,in}^2\right) + \frac{1}{2}\sum_{i=1}^{i=N}m_i\left(v_{i,fn}^2 + v_{i,in}^2\right), \quad (1)$$

where $M_{\rm cm} = \sum_{i=1}^{N} m_i$ and $V_{\rm cm} = \sum_{i=1}^{N} v_i$ are the mass and the speed of the center of mass, respectively, N is the total of particles in a system, and v_i is the speed for each particle.

For this, the linear momentum,

$$P_{\rm cm} = \sum_{i=1}^{N} m_i v_i, \qquad (2)$$

of the center of mass must be equal to the sum of the entire center of mass movements.

Therefore, we must consider the constraints.

$$M_{\rm cm}V_{\rm cm,fn} - \sum_{i=1}^{N} m_i v_{i,fn} = 0 \text{ and } M_{\rm cm}V_{\rm cm,in} - \sum_{j=1}^{N} m_i v_{i,in} = 0.$$
(3)

Since the speed is decomposed as the sum of the speed of the center of mass $(V_{\rm cm})$ and the relative initial and final speeds of each particle to the center of mass $v_{i, {\rm in}}$ and $v_{i, {\rm fn}}$, respectively, we present

$$v_{i,\text{fn}} = V_{\text{cm,fn}} + v_{i,\text{fn}}, \forall i = 1, \dots, N \text{ and } v_{i,\text{in}}$$
$$= V_{\text{cm,in}} + v_{i,\text{in}}, \forall i = 1, \dots, N.$$
(4)

This assures that the equilibrium of the system is maintained.

In the center of mass, each particle *i* receives forces from other particles. Then, the total force associated with particle *i* is described as

$$F_{i} = \sum_{k=1}^{N} f_{k,i}, \, \forall \, i = 1, ..., N,$$
(5)

where $f_{k,i}$ is the force from particle k to particle i.

Next, to guarantee the dynamic balance of the system (Third Newton's Law), we present

$$m_i a_i - \left(\sum_{k=1 \ k \neq i}^N f_{k,i}\right) = 0, \ \forall \ i = 1, \dots, N.$$
 (6)

Inspired by the law of universal gravitation, we know that

$$f_{k,i} = G \frac{m_k m_i}{r_{\min(k,i)}^2},$$
(7)

where $r_{\min(k,i)}^2$ is the length of the path between nodes *k* and *i* and *G* is the universal gravitational constant; therefore, to assure that the force exerted and received on any particle is balanced with the gravitational force of the system, we present the inequality

$$f_{k,i} - G\left(\frac{m_k m_i}{r_{\min(k,i)}^2}\right) \le 0, \forall \text{ pair } k, i.$$
(8)

Finally, to avoid infeasible solutions, each variable must be nonnegative $(v'_{i,\text{fn}} > 0 \text{ and } v'_{i,\text{in}} > 0, \forall i = 1, ..., N\varepsilon)$.

Then, the mathematical model based on Newton's Law of Motion (NLM) is denoted as

$$\min z = \frac{1}{2}M_{\rm cm} + \left(V_{\rm cm,fn}^2 V_{\rm cm,in}^2\right) + \frac{1}{2}\sum_{i=1}^{i=N} m_i \left(v_{i,fn}^2 + v_{i,in}^2\right). \tag{9}$$

It is subject to

$$M_{\rm cm}V_{\rm cm,fn} - \sum_{i=1}^{N} m_i v_{i,fn} = 0.$$
 (10)

$$M_{\rm cm}V_{\rm cm,in} - \sum_{j=1}^{N} m_i v_{i,in} = 0.$$
(11)

$$v_{i,\text{fn}} = V_{\text{cm,fn}} + v_{i,\text{fn}}, \, \forall \, i = 1, \dots, N.$$
 (12)

$$v_{i,in} = V_{cm,in} + v_{i,in}, \forall i = 1, ..., N.$$
 (13)

$$m_i a_i - \left(\sum_{k=1 \ k \neq i}^N f_{k,i}\right) = 0, \forall i = 1, \dots, N.$$
 (14)

$$f_{k,i} - G\left(\frac{m_k m_i}{r_{\min(k,i)}^2}\right) \le 0, \forall \text{ pair } k, i.$$
(15)

$$v_{i,\text{in}} > 0 \text{ and } v_{i,\text{in}} > 0, \forall i = 1, \dots, N.$$
 (16)

Equation (1) denotes that the difference in the kinetic energy of the particles must be minimal. Equations (10) and (11) restrict that the center of mass dynamics be equal to the sum of the entire center of mass movements. Equations (12) and (13) ensure that the particle's speed is like the amount of its relative speed and the rate of the center of mass. Equation (14) ensures that the system's dynamic balance is maintained. Equation (15) ensures that the force exerted and received on any particle is balanced with the gravitational force of the system. Finally, (16) ensures that the limits for each variable are respected.

3.5. Differential Evolution to Solve the NLM. To solve the mathematical model presented in equations (1) to (8), for small networks (less than 200 nodes), we use a Differential Evolution (DE) algorithm, which is summarized as follows.

Differential Evolution (DE) is an optimization method belonging to the evolutionary computation category. Like other algorithms in this category, DE maintains a population of candidate solutions, which are recombined and mutated to produce new individuals who will be chosen according to the value of their fitness function.

This algorithm receives as input the values of the connections between any pair of nodes as a vector. The length of each of these vectors is equal to the number of parameters and a space reserved for the fitness value, where the population is composed of NP (number of parents) vectors and each vector is defined as $x_p^g, \forall p = 1, 2, ..., NP$ and g is the corresponding generation.

The DE algorithm consists of 4 steps, which are given as follows:

(1) *Initialization*. The population is initialized (first generation) randomly, considering the minimum and maximum values of each variable:

$$x_{\acute{p},m}^{1} = x_{m}^{\min} + \operatorname{rand}(0,1) \cdot \left(x_{m}^{\max} - x_{m}^{\min}\right), \tag{17}$$

for p = 1, ..., NP, m = 1, ..., n, and rand (0, 1), a random number in the range [0, 1].

(2) *Mutation*. The mutation consists of the construction of NP noisy random vectors created from three randomly chosen individuals, called target vectors, x_a, x_b, x_c . The noisy random vectors (n_p^t) are obtained as follows:

$$n_p^g = x_c + F \cdot \left(x_a - x_b\right),\tag{18}$$

We have p, a, b, and c different from each other, and p = 1, ..., NP and F is a parameter that controls the mutation rate, which is in the range of [0, 2].

(3) *Recombination*. Once the NP noisy random vectors have been obtained, the recombination is carried out randomly, comparing them with the original vectors (*x^g_p*), obtaining the test vectors (trial vectors, *t^p_m*) as follows:

$$t_{p,m}^{g} = \begin{cases} n_{p,m}^{g} \text{ if } \operatorname{rand}(0,1) & <\operatorname{Cr} \\ x_{p,m}^{g} & \text{ in other case} \end{cases},$$
(19)

for p = 1, ..., NP and m = 1, ..., n.

(4) Selection. The selection is made by simply comparing the test vectors with the original ones so that the next-generation vector will be the one with the best value of the fitness function:

$$x_p^{g+1} = \begin{cases} t_p^g \text{ if } \text{fitness}(t_p^g) & > \text{fitness}(x_p^g) \\ x_p^g & \text{ in another case} \end{cases}$$
(20)

The initialization is carried out in the first generation of the algorithm and the processes of mutation, recombination, and selection are carried out for each generation until the maximum number of generations is reached.

Specifically, in this work, we use a mutation factor (F) of 0.55, cross factor (Cr) of 0.83, 200 generations, and 25 individuals in this work.

4. Results and Discussion

In this section, we present the numerical results obtained to quantify robustness and identify influential nodes for the networks belonging to our case study.

4.1. Deaths, Contagion, and Suspect Cases (Coupled Networks). As we can see in the previous sections, the methodologies proposed in this work can cause the rupture

of the giant connected component for several types of models of complex networks and identify the main influence of four of our case study spreaders nodes.

Therefore, in this section, we present the application of both methodologies to our database of COVID-19 in Mexico.

4.1.1. Quantification of Robustness. To quantify the robustness for the COVID-19 networks, we applied our methodology based on VSP. Moreover, due to the dimension of the dataset, we show the information in the following:

- (i) Deaths: https://docs.google.com/spreadsheets/d/ 15snIYBdsKzHnJdk9lozrMv3nZnaoK3uya/edit?us p=sharing&ouid=114121654599377658063&rtpof= =true&sd==true
- (ii) Contagion: https://docs.google.com/spreadsheets/ d/1_IK-FPouxUP7eoFNPxKCH2fO2AXfhDDG/ed it?usp=sharing&ouid=114121654599377658063&rt pof=true&sd=true
- (iii) Suspected cases: https://docs.google.com/ spreadsheets/d/1ThvN1YEBcGQfhhFIIWMTpUM wk4yoILBA/edit?usp=sharing&ouid=11412165459 9377658063&rtpof=true&sd=true

Based on the information shown in previous web links, it is important to mention that the connected components generated by the elimination of the separator set have a high cardinality regarding the separator set; therefore, we can affirm that the COVID-19 networks are robust, but the methodology based on the modification of VSP causes that the identification of the nodes that are more influential for the connectivity be more effectively.

In principle, we can cause the rupture of the network with the elimination of the following:

- (i) 30% of nodes for the coupled network of deaths.
- (ii) 29% of the nodes for the coupled contagion network.
- (iii) 28.5% of the nodes for the coupled network of suspect cases.

Then, our methodology causes the rupture of the network. Therefore, even though it has a high clustering coefficient and a short average rupture diameter and length, it is very effective in quantifying the robustness of complex networks.

Therefore, based on our previous work [29], we can affirm that the modification based on maximizing the sum of weight for the elements that belong to subset C (separator set) while the cardinality of set C is minimized works better than the previously developed methodology.

If we think about the characteristics of health, in principle, we could think that the municipalities that cause the rupture of the giant component are those municipalities where there are many deaths. However, if we analyze the elements that belong to the separating set one by one, we can see that those municipalities present different characteristics. They either have characteristics that make them relate to those municipalities that present few cases, or they have characteristics that make them relate to those municipalities. Therefore, some characteristics of the elements found in the separator set for the coupled networks on infections, deaths, and suspected cases are given below (all the information shown below can be verified at https://gaia.inegi.org.mx/covid19/):

- (i) Some municipalities do not have enough hospitals to treat COVID-19.
- (ii) Some municipalities do not carry out enough COVID-19 detection tests.
- (iii) Some municipalities have many infections.
- (iv) Some municipalities have enough hospitals to treat COVID-19. Therefore, they present many deaths because they serve people from other states/ municipalities.
- (v) Some municipalities conduct many COVID-19 screening tests. Therefore, some people with positive results (coming from other places) decide to stay in the hospitals of that municipality.

Now, to see the results graphically, we present Figures 2–4.

In Figure 2, we can see the entire network for deaths, contagion, and suspect cases. In all three cases, we can see that the networks have a high density, because they have many connections. In addition, we can note that some nodes are on the border; that is, they do not have many connections. Therefore, we could say that they do not have characteristics that make them influential or that are helpful to quantify the robustness of each of the networks.

For example, for the network of deaths (Figure 2(a)), we can see that there are municipalities with little hospital infrastructure (which we can also see at https://gaia.inegi. org.mx/covid19/). Therefore, we can suppose that most of the people infected with COVID-19 are treated in another municipality.

For the case of the contagion network (Figure 3(b)), we observe that there are few nodes with few connections and others with many connections. Therefore, it is important to study those important municipalities for the spread of the virus and the connection of the network by the rupture of the network.

Finally, for the case of suspected cases (Figure 3(c)), we can see that they have similar behavior as the death's cases. This is because, in most of the municipalities or states where there is not a good hospital infrastructure, tests are not carried out to detect the virus.

As we mentioned before, our methodology manages to cause rupture of the giant component (GC) of the network that is made up of the entire network in our three cases.



FIGURE 2: Graphs for the deaths, contagion, and suspect cases networks. (a) Deaths. (b) Contagion. (c) Suspect cases.



FIGURE 3: Graphs for the deaths, contagion, and suspect cases networks (after the rupture of the GC). (a) Deaths. (b) Contagion. (c) Suspect cases.

It is important to mention that the modification of the VSP-based methodology achieves at least three disjoint sets, which present homogeneous characteristics. That is, there are three types of nodes that are of great influence on the network, based on the analyzed characteristics.

On the other hand, given the nature of the public information available to analyze COVID-19 in Mexico, we cannot give accurate results. However, the objective of the manuscript is to show how we can obtain analyses, estimates, and possible results that can help the treatment of the pandemic in Mexico and the world through the complex network approach and the treatment of its dynamic and structural characteristics based on optimization problems.

Now, if we analyze each of the graphs obtained for the division of the networks, we can observe the following.

For the network of deaths (Figure 3(a)), we observe that the nodes that are in the separator set (a component of greater size in the figure) present characteristics of the elements of the other two components. However, we can see that the characteristics are very different for components A and B.

On the one hand, we have the municipalities with many COVID-19 registries, hospital infrastructure, and services to treat the disease (most of the metropolitan area of the Valley of Mexico), while, in the other component, there are some municipalities that do not have many registered cases (such as those mentioned in the description of Figure 3(a)).

Moreover, for both cases (contagion and suspect cases), we can observe that the data show the same behavior as for part (a). This can be justified due to the modification of our technique (which seeks to obtain homogeneous communities with the separation of the giant component (GC)). Then, to analyze the data distribution, we present the maps of the distribution of the contagion, deaths, and suspect cases studied.

In Figure 4, we can see the maps of Mexico for deaths, contagion, and suspect cases, where while the red color is more marked, it means that there is a greater number of cases (as the case may be deaths, infections, and suspected cases).



(b) FIGURE 4: Continued.



FIGURE 4: Maps of Mexico for deaths, contagion, and suspected cases. (a) Deaths. (b) Contagion. (c) Suspected cases.

In Figure 4, we can see that, as mentioned above, most of the data for all three types of analysis follow a trend. In other words, there are more data registered in the main entities or municipalities of the country, whether they have a good hospital infrastructure.

Therefore, based on the numerical results obtained and Figures 2–4, we can see that, in addition to the hospital conditions and capacities to care for people sick with COVID-19, the fact that there is high contact of people in the areas of residence and work are of great influence. Therefore, not only should emphasis be placed on those places with great hospital development, but also measures that help to avoid large crowds and interaction relationships between people should be taken.

4.1.2. Identification of Spreaders. This subsection shows the numerical results in the coupled networks by solving NLM.

As in the previous case (robustness) to identify the most influential or spreaders nodes of COVID-19 in Mexico, we applied our methodology based on NLM and, due to the dimension of the dataset, we show the information in the following:

- (i) Deaths: https://docs.google.com/spreadsheets/d/ 15snIYBdsKzHnJdk9lozrMv3nZnaoK3ya/edit?usp =sharing&ouid=114121654599377658063&rtpof=tr ue&sd=true
- (ii) Contagion: https://docs.google.com/spreadsheets/ d/1_IK-FPouxUP7eoFNPxKCH2fO2AXfhDDG/ed it?usp=sharing&ouid=114121654599377658063&rt pof=true&sd=true
- (iii) Suspected cases: https://docs.google.com/spread sheets/d/1ThvN1YEBcGQfhhFIIWMTpUMwk4yo

ILBA/edit?usp=sharing&ouid=11412165459937765 8063&rtpof=true&sd=true

Based on the information shown in previous web links, it is important to mention that the main influencing or diffusing nodes are mostly in the set of nodes that cause the rupture of the network (because the entire network is the GC). They are a subset of the items found by the modified VSP. Therefore, if we analyze the elements with the greatest influence or diffusion one by one, we can get the following.

Based on the information about the numbers of contagions, suspect cases, and deaths, we can see that some municipalities belonging to Mexico City or State of Mexico are considered the municipalities with the most diseases and deaths and are not in the set of the most spreaders.

On the other hand, some municipalities that are in the spreaders set belong to the next states as Aguascalientes, Veracruz, Tamaulipas, Nuevo León, Puebla, Sonora, San Luis Potosí, Tabasco, Guerrero, Quintana Roo, and Hidalgo. Therefore, with high probability, most of the population (who have contact with many people from different parts of the country) get infected with COVID-19.

However, as we mentioned above, given the nature of the data available to analyze COVID-19, we cannot affirm what phenomena lead to this behavior. Therefore, this work can be of great help so that, in other investigations, they can deepen and analyze the characteristics that cause these municipalities and states to be the most diffusers of the disease.

Finally, regarding the analysis of the complex network, we can verify that our methodologies are easily applicable to different models of complex networks that are related to different systems in the real world and that this work can be the starting point to analyze several phenomena about the spread of COVID-19 as a complex system.

5. Limitations of the Study and Discussion

The analysis presented in this work is based on the identification of the influential nodes in complex networks by the rupture of the GC, the asortality of the nodes, and the quantification of their influence through Newton's Laws of Motion. The idea of using the two ways to analyze the information is that the VSP can find the very connected nodes, and their links can cause a fast spread of the disease.

On the other hand, we can obtain an accurate quantification of the influence or spread for each node using NLM. Then, both methodologies can help obtain a better idea about those states or municipalities where the highest number of infections and spread of the disease occur.

The idea of modeling three different types of coupled networks (deaths, contagion, and suspect cases) is to be able to obtain an identification of the most spreaders nodes, based on a multicriteria analysis that reduces the disadvantages of using only one topic; thus, the technique can identify those municipalities that, based on different characteristics, are the most spreaders of the COVID-19.

Therefore, as the techniques presented in this work depend on the data that report each state or municipality, a limitation of the study is that the drastic change or the lack of data for one or more states or municipalities can cause an erroneous result.

In Mexico, the analysis of deaths, suspected cases, and contagion is carried out from the National Institute of Statistics and Geography (INEGI) information. It is essential to mention that the information is collected from the printed copies of the death certificate, the official source of death statistics.

However, because the compilation and validation of information on deaths take approximately one year to be published, it is possible that the information used in this work does not consider all possible cases.

Finally, another limitation of the database used in this work is that it shows the analysis results of the death certificates issued by the Civil Registry to estimate excess mortality from all causes in Mexico. In addition, this database does not include information on PCR tests applied in private institutions, and, therefore, it has a significant bias.

However, as there is no database to help us analyze data from all institutions in the country, the database available at https://coronavirus.gob.mx and https://gaia.inegi.org.mx/ covid19/ helps to generate a meaningful analysis. Therefore, our work can serve as a basis for future analyses and even analyze various phenomena in different databases.

6. Conclusions

This work presents two methodologies to analyze the contagion and deaths caused by COVID-19 in Mexico in a coupled complex network.

The first is a methodology based on the Vertex Separator Problem (VSP) to quantify robustness in complex networks. For this, since two connected components are formed when applying VSP, we show a social analysis, where the results show that there is excellent communication and interdependence of the elements that share certain types of characteristics.

In addition, it is essential to mention that, with the VSP approach, we can identify which nodes are important and influential for the flow and diffusion of the virus between the people that live in several kinds of municipalities and know which ones have a high rate for the contagion of COVID-19.

Therefore, because the methodology presented in this work is only an indirect process for quantifying the importance or influence of nodes, the second methodology (denoted as NLM) allows identifying and quantifying the most influential nodes in several types of complex networks. NLM is based on the resolution of a mathematical programming method inspired by Newton's three Laws of Motion and universal gravitation law.

Because this methodology only requires the calculation of the degree of each node and the shortest path between any pair of nodes, its adaptation to different models and network sizes compared to traditional methods based on centrality measures (closeness, degree, between, among others) is easy.

Finally, the analysis of COVID-19 networks shows that other characteristics influence the spread of the virus. However, as we mentioned previously, the open-access databases used in this work do not contemplate characteristics outside of health. Therefore, we can see that this work can be the basis for future research related to other areas of study, which can find and classify the characteristics that cause the spread of the virus.

Data Availability

The data used to support the findings of this work are available at the following links: (1) https://docs.google.com/ spreadsheets/d/15snIYBdsKzHnJdk9lozrMv3nZnaoK3ya/edit? usp=sharing&ouid=114121654599377658063&rtpof=true&sd =true. (2) https://docs.google.com/spreadsheets/d/1_IK-FPoux UP7eoFNPxKCH2fO2AXfhDDG/edit?usp=sharing&ouid=11 4121654599%20377658063&rtpof=true&sd=true. (3) https:// docs.google.com/spreadsheets/d/1Usx5aK404EfVUh_kCZxmt jLDCSi02O4E/edit?usp=sharing&ouid=100406467082864939 906&rtpof=true&sd=true.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- A. Aleta, S. Meloni, and Y. Moreno, "A multilayer perspective for the analysis of urban transportation systems," *Scientific Reports*, vol. 7, no. 1, p. 44359, Article ID 44359, 2017.
- [2] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1777–1787, 2012.
- [3] L. Fei, Q. Zhang, and Y. Deng, "Identifying influential nodes in complex networks based on the inverse-square law,"

Physica A: Statistical Mechanics and Its Applications, vol. 512, pp. 1044–1059, 2018.

- [4] W.-B. Du, B.-Y. Liang, C. Hong, and O. Lordan, "Analysis of the Chinese provincial air transportation network," *Physica A: Statistical Mechanics and Its Applications*, vol. 465, pp. 579– 586, 2017.
- [5] J.-H. Liu, J. Wang, J. Shao, and T. Zhou, "Online social activity reflects economic status," *Physica A: Statistical Mechanics and Its Applications*, vol. 457, pp. 581–589, 2016.
- [6] L. Tang, K. Jing, J. He, and H. E. Stanley, "Complex interdependent supply chain networks: cascading failure and robustness," *Physica A: Statistical Mechanics and Its Applications*, vol. 443, pp. 58–69, 2016.
- [7] A. Wahid-Ul-Ashraf, M. Budka, and K. Musial-Gabrys, "Newton's gravitational law for link prediction in social networks," in *Proceedings of the International Workshop on Complex Networks and their Applications*, pp. 93–104, Springer, Lyon, France, 29 November 2017.
- [8] A.-L. Barabási and E. Bonabeau, "Scale-free networks," Scientific American, vol. 288, no. 5, pp. 60–69, 2003.
- [9] L. Liu, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, "Vital nodes identification in complex networks," *Physics Reports*, vol. 650, pp. 1–63, 2016.
- [10] Y. Liu, M. Tang, T. Zhou, and Y. Younghae Do, "Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition," *Scientific Reports*, vol. 5, no. 1, p. 9602, 2015a.
- [11] L.-l. Ma, C. Ma, H.-F. Zhang, and B.-H. Wang, "Identifying influential spreaders in complex networks based on gravity formula," *Physica A: Statistical Mechanics and Its Applications*, vol. 451, pp. 205–212, 2016.
- [12] S.-Y. Tan, J. Wu, L. Lü, M.-J. Li, and X. Lu, "Efficient network disintegration under incomplete information: the comic effect of link prediction," *Scientific Reports*, vol. 6, no. 1, p. 22916, 2016.
- [13] B Min and M Zheng, "Correlated network of networks enhances robustness against catastrophic failures," *PLoS One*, vol. 13, no. 4, p. e0195539, Article ID e0195539, 2018.
- [14] A. Nagurney and Q. Qiang, "Fragile networks: identifying vulnerabilities and synergies in an uncertain age," *International Transactions in Operational Research*, vol. 19, no. 1-2, pp. 123–160, 2012.
- [15] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
- [16] M. Ouyang, L. Hong, Z.-J. Mao, M.-H. Yu, and F. Qi, "A methodological approach to analyze vulnerability of interdependent infrastructures," *Simulation Modelling Practice and Theory*, vol. 17, no. 5, pp. 817–828, 2009.
- [17] L. Liu, Y.-C. Zhang, C. H. Yeung, and T. Zhou, "Leaders in social networks, the delicious case," *PLoS One*, vol. 6, no. 6, Article ID e21202, 2011.
- [18] F. Rendl and R. Sotirov, "The min-cut and vertex separator problem," *Computational Optimization and Applications*, vol. 69, no. 1, pp. 159–187, 2018.
- [19] H. Zhang, E. Fata, and S. Sundaram, "A notion of robustness in complex networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 310–320, 2015.
- [20] X.-L. Ren, N. Gleinig, D. Toli_c, and N. Antulov-Fantulin, "Underestimated cost of targeted attacks on complex networks," *Complexity*, vol. 2018, Article ID 9826243, 15 pages, 2018.
- [21] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, 2015.

- [22] E. Montes-Orozco, R.-A. Mora-Gutiérrez, S.-G. De-Los-Cobos-Silva et al., "Methodology to quantify the robustness in networks:Case study: higher education system in Mexico," *Computing*, vol. 103, 2020b.
- [23] H. Zhong-Ming, W. Yang, T. Xu-Sheng, D. Da-Gao, and Y. Wei-Jie, "Ranking key nodes in complex networks by considering structural holes," *Acta Physica Sinica*, vol. 64, no. 5, 2015.
- [24] Z. Zhao, H. Yu, Z. Zhu, and X. Wang, "Identifying influential spreaders based on network community structure," *Chinese Journal of Computers*, vol. 37, pp. 753–766, 2014.
- [25] W Wang, M Tang, HF Zhang, H Gao, Y Do, and ZH Liu, "Epidemic spreading on complex networks with general degree and weight distributions," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 90, no. 4, Article ID 042803, 2014.
- [26] Q. Li, T. Zhou, L. Lü, and D. Chen, "Identifying influential spreaders by weighted LeaderRank," *Physica A: Statistical Mechanics and Its Applications*, vol. 404, pp. 47–55, 2014.
- [27] E. Montes-Orozco, R.-A. Mora-Gutierrez, S.-G. De-Los-Cobos-Silva et al., "Identification of COVID-19 spreaders using multiplex networks approach," *IEEE Access*, vol. 8, pp. 122874–122883, 2020a.
- [28] A. A. Al-Shargabi and A. Selmi, "Social network analysis and visualization of Arabic tweets during the COVID-19 pandemic," *IEEE Access*, vol. 9, pp. 90616–90630, 2021.
- [29] E. Montes-Orozco, R. A. Mora-Gutiérrez, B. Obregón-Quintana et al., "Mexican university ranking based on maximal clique," in *Lecture Notes in Social Networks Educational Networking*, pp. 327–395, Springer, Cham, 2020c.
- [30] E. Montes-Orozco, R. A. Mora-Gutiérrez, B. Obregón-Quintana et al., "Inverse percolation to quantify robustness in multiplex networks," *Complexity*, vol. 2020, Article ID 8796360, 11 pages, 2020d.
- [31] E. Balas and C. C. d. Souza, "The vertex separator problem: a polyhedral investigation," *Mathematical Programming*, vol. 103, no. 3, pp. 583–608, 2005.
- [32] R. Müller and D. Wagner, "-vertex separator is np-hard even for 3-regular graphs," *Computing*, vol. 46, no. 4, pp. 343–353, 1991.
- [33] M. R. Garey, D. S. Johnson, and R. Sethi, "The Complexity of Flowshop and Jobshop Scheduling," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 117–129, 1976.
- [34] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [35] R. Storn and K. Price, "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.