International Journal on Emerging Technologies **11**(2): 245-250(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Mapping of Corona Virus Transmission in India with a Mathematical Approach

R. Varalakshmi
Assistant Professor, Department of Computer Application,
The TamilNadu Dr. Ambedkar Law University, India.

(Corresponding author: Dr. R. Varalakshmi)
(Received 14 March 2020, Revised 04 April 2020, Accepted 06 April 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The corona virus is the new outbreak of virus with symptoms of severe acute respiratory syndrome. The objective of this study is to trace out the existence of local outbreak of virus and its impact to be alleviated through quarantine of symptomatic patients. To analyze this current outbreak of epidemic disease evidences, a mathematical method has been proposed. With emergence of computing tools, mapping mathematical methods and generating numerical solutions with certainty factors have turn out to be achievable. The mathematical mapping has a potential in the study of viral epidemics and framing strategies of global pandemics. In addition to this, the number of person likely to be affected by virus is predicted using publicly reported values. From the result, it is observed that the spread rate is higher. Hence, people should stay at home and follow the instructions given by the Government to avoid the spread of this virus.

Keywords: Coronavirus disease (COVID-19), Emerging viruses, Epidemic diseases, Infectious diseases, Mathematical mapping,

Abbreviations: WHO, World Health Organization; HP, hypothesis; EV, evidence associated; CF, Certainty Factor

I. INTRODUCTION

In December 2019, a new virus was first identified in Wuhan state of China. Wuhan Municipal Health and Health Commission reported that the virus can cause cough, fever and respiratory illness and in critical condition, it result in loss of life. It was believed that the virus was spreading from the animals that are sold in Wuhan seafood market. The virus was started spreading from local seafood market. Initially people who are in direct contact with Wuhan seafood market are affected. Later the people with no direct connection to the market are also affected which shows that It is an infectious disease. The World Health Organization (WHO) named the virus as Coronavirus disease (COVID-19). The Corona virus was spread across the world due to travel from China to different part of the world.

The virus spreads through close contact with an affected person. As the infected person sneeze or cough, the virus spread from the micro droplets. These micro droplets along with virus settled down on various surface. When the other person touches, the virus on it will stick on to the hand and with that hand if he touches his own nose, the virus will be inhaled and it starts affecting the respiratory system.

The spread of virus has four stages:

Stage 1: People who had travel History with an affected Country. In this stage, the diseases doesn't spread locally.

Stage 2: The virus is in the local transmission. The affected people (without evidence of symptoms) are transmitting the virus to the people who are in close contacts. At this stage, the source of infection is identifiable. The affected people can able to be identified and they can be isolated to prevent further spread of virus.

Stage 3: When the source of infection is unidentifiable and the people without travel history, or without close contact of affected people are affected. At this stage, the diseases are difficult to control.

Stage 4: The spread of virus is practically uncontrollable and the infections is across the country.

In India, the first case was reported on 30th January 2020. In order to prevent the spread across the country, the Indian Government has declared Lock down the entire country.

In order to analyze the current outbreaks of epidemic diseases, a mathematical method based on probabilistic reasoning has been mapped. The idea of Probabilistic reasoning is to combine the probabilistic theory to handle uncertainty and deductive logic. One of the major applications of Probabilistic reasoning is Artificial Intelligence and it is applied when there is unsure of the predicates, possibilities of predicates are too large to list down, an error during an experiment.

The sources of uncertainties are as follows:

- When information is obtained from a source rather than obtained personally.
- When the information obtained is misinterpreted or the information is not understood.
- If the measuring equipment is faulty or not calibrated.
- Experimental errors including parallax errors.
- Misinterpretation / wrongly understood / imprecision during conversation or translation.
- Occurrence of a ransom event.
- Noise in the measurements.

- An example is judgement is given based on the benefit of doubt to the accused when there is a lack of certainty or lack of evidence.

When there is such above situation are occurred, probability is widely used. Probability is used to express the confidence of occurring an event. Probability is defined as ratio of number of occurrences to the total number of trials. When an occurrence of an event does not affect the occurrence of other event, then it is called as mutually exclusive events and product theorem is used. When an occurrence of an event depends on another event, then Bayes theorem can be used to calculate the probability. Probabilistic-based reasoning uses Bayes Theorem for tackling uncertainty.

II. LITERATURE REVIEW

The development in the transportation makes easy movement of people, commodities and hence lead to transportation of pathogens around the world. There is more vector borne disease are available in the world and the scientist are doing research to control them as and when they are evolved. The control measures of vector borne disease are very important. Further ports and airports are to be continuously monitored to prevent a novel virus entering into a country [1].

In pharmaceutical industry, the new drugs are developed in three phases viz. Exploratory phase, Drug discovery phase and Clinical development phase. If a new disease is identified, then the Pharmaceutical industry would explore the molecular target in the human body, its mechanism and followed by validation. In human body, for a particular target, the molecules (candidates) will be done in drug delivery phase. In the third stage, the developed drug, first will be tested to healthy human volunteers, then on smaller number of patients and then in larger number of patients with proper permission obtained from the regulatory agency which will take several years to bring a drug to the market for sale [2].

With the help of current and past situation, the future can be predicted. Though machine learning and statistics, knowledge can be extracted by identifying relationships among a pattern. This is widely useful especially in Medical datasets [3].

World Health Organization (WHO), declares as on March 31, 2020, globally 803,011 cases of coronavirus disease 2019 (COVID-19), which includes 591,950 active cases, 172,396 recovered and 39,028 deceased. The virus was first identified in Wuhan state of China during December 2019 and it was started to spread across the globe. The Coronavirus belongs to the family of virus that causes viral pneumonia including fever, breathing difficulty, and bilateral lung infiltrative lesions. Though, initial patients were due to animal-to-person transmission and Sooner more cases were reported are medical staff and others, indicates human-to-human transmission of virus [4].

The virus mainly spread via respiratory droplets from sneezing or cough of infected people during close contact with others. In India, the first case was reported in the last week of January 2020. The outbreak has been declared an epidemic by World Health organization (WHO). To reduce the spread of the virus, the Government of India issued the various guidelines include hand hygiene practices, wearing masks, avoiding public gathering, social distancing and quarantines etc. and through different media, the message has been broadcasted to all the people. In this paper, the context of India where addressed mapping a mathematical methods of Infectious disease transmission.

In order to predict the transmissibility and transmission dynamics, a mathematical method is highly essential. Various study methods were analyzed by the researcher. They computed basic reproduction number using serial intervals and intrinsic growth rate [5-8]. The researcher developed their method using ordinary differential equations and Markov Chain Monte Carlo methods. The Bats-Hosts-Reservoir-People (BHRP) transmission network method was developed assuming that the virus was transmitted among the bats and then it has been transmitted to humans through unknown hosts. The people who were all exposed to the markets and got the risks of infection [9-10].

III. PROPOSED METHOD

With the history of understanding of corona virus infection, it was presumed that all the infections may occur in a probability. This mathematical background allows us to analyze the chance events in a reasonably sound manner. The numerical description shows how likely some conclusion will occur after many repeated evidences or sources. Probability is used to express the confidence of occurring an event. Probability is defined as ratio of number of occurrences to the total number of trials. When an occurrence of an event does not affect the occurrence of other event, then it is called as mutually exclusive events and product theorem is used.

When an occurrence of an event depends on another event, then Bayes theorem [11] can be used to calculate the probability. Probabilistic-based reasoning uses Bayes Theorem for tackling uncertainty. This theorem gives reasoning method about evidences / sources leads to conclusion. Here any source that is occurring or likely to occur is quantified by applying knowledge about the evidence and the probabilistic rules. With reasoning using certainty factors numerical values about sources and conclusion are to be computed.

Fig. 1 depicts the conceptual representation of the proposed system to determine the occurrence of corona virus occurrence based on given evidences or sources.

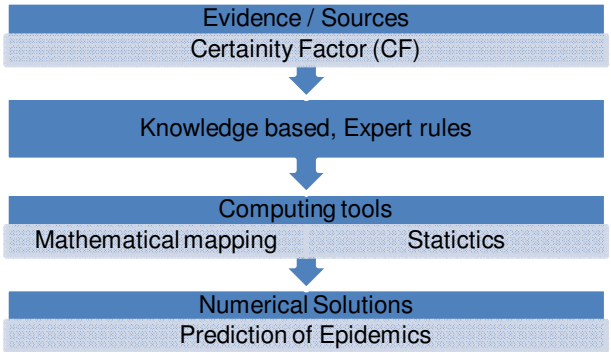

Fig. 1. Conceptual representation of the proposed system.

The mathematical relationship is given as:

$$\mathrm{P(HP|EV)} = \frac{P(HP) * P(EV|HP)}{P(EV)} \quad (1)$$

$$\mathrm{P(\sim HP|EV)} = \frac{P(\sim HP) * P(EV|HP)}{P(EV)} \quad (2)$$

Where $\mathrm{P(\sim HP)} = 1 - P(HP)$

Dividing Eq. (1) by Eq. (2) will result in

$$\frac{P(HP|EV)}{P(\sim HP|EV)} = \frac{P(EV|HP) * P(HP)}{P(EV|\sim HP) * P(\sim HP)} \quad (3)$$

To map Bayes theorem with the proposed implications, consider the following example.

Assume that for a certain medical case, some of the symptoms can be seen as common for some defects.

$$\text{i.e., } \mathrm{P(defect1|symptom1)} = \frac{(P(defect1) * P(symptom1|defect1))}{P(symptom1)} \quad (4)$$

With multiple evidences / sources followed by conclusion can be derived as follows:

Case 1: E1= P(DC) ^ P(S) = P(AP)

$$\mathrm{P(E1|AP)} = \frac{P(E1) * P(AP|E1)}{P(AP)} \quad (5)$$

Where E1 is the evidence of symptom, P(E1 | AP) gives the probability of the Evidence E1 given the probability of conclusion of air pollution denoted as P(AP). The evidence E1 combines the probability of symptom dry cough P(DC) and the probability of symptom sneeze P(S) which leads to the conclusion of air pollution denoted as P(AP).

Case 2: E2= E1 ^ P(M) ^ P(RN) = = P(CC)

$$\mathrm{P(E2|CC)} = \frac{P(E2) * P(CC|E2)}{P(CC)} \quad (6)$$

Where E2 is the evidence of symptom, P(E2 | CC) gives the probability of the Evidence E2 given the probability of conclusion of common cold denoted as P(CC).The evidence E2 combines the evidence E1, the probability of symptom mucus P(M), the probability of symptom roughing nose P(RN) which leads to the conclusion of common cold denoted as P(CC).

Case 3: E3 = E2 ^ P(BA) ^ P(W) ^ P(LF) = P(F)

$$\mathrm{P(E3|F)} = \frac{P(E3) * P(F|E3)}{P(F)} \quad (7)$$

Where E3 is the evidence of symptom, P(E3 | F) gives the probability of the Evidence E3 given the probability of conclusion of Flu denoted as P(F). The evidence E3 combines the evidence E2, the probability of symptom body ache P(BA), the probability of symptom weakness P(W), and the probability of symptom lite fever P(LF), which leads to the conclusion of Flu denoted as P(F).

Case 4: E4 = E3 ^ P(DB) - ~P(LF) = P(CV)

$$\mathrm{P(E4|CV)} = \frac{P(E4) * P(CV|E4)}{P(CV)} \quad (8)$$

Where E4 is the evidence of symptom, P(E4 | CV) gives the probability of the Evidence E4 given the probability of conclusion of Corona Virus denoted as P(CV). The evidence E4 combines the evidence E3, the probability of symptom difficulty in breathing P(DB), and the probability of symptom high fever ~P(LF) or P(HF) which leads to the conclusion of Corona Virus denoted as P(CV).

Once Evidence E4 leads to the conclusion of P(CV), then there comes the existence of mapping the sum of terms in a geometric progression to analyze the consequence of virus spread thru CV.

To apply this theorem estimation of priori and conditional probabilities, certainty factors circumvent with the problem are adopted. A Certainty Factor (CF) is a numerical estimate of the trust or untrust on a conclusion in the presence of a set of evidences. Use a scale measure from 0 to 1 where 0 represents total untrust and 1 stands for total trust. Other values between 0 to 1 represent varying degrees of trust and untrust. The CF values are determined by the knowledge base. In typical expert systems that involves multiple rules relating to the same conclusion. The applied rules must be structured in a manner that it can help the user to enhance the factor as either trust or untrust. In order to achieve this, the two important components Measure of trust and Measure of untrust have to be estimated. Measure of Trust MT[c,e] denoting a value of trust in conclusion 'c' given the evidence 'e'. Measure of Untrust MUT[c,e] denoting a value of untrust in conclusion 'c' given the evidence 'e'.

Using the values of MT and MUT, it is possible to arrive the overall certainty factor (CF) using the formula

$$\mathrm{CF[c,e]} = \mathrm{MT[c,e]} - \mathrm{MUT[c,e]} \quad (9)$$

In most of the expert systems [12], multiple rules using the parameters supporting evidence e_{supp} and against evidence e_{aga} are taken to correlate to a given conclusion. It might either enhance the trust or untrust. MT[c,e_{supp}] denotes the consolidated measure of trust in conclusion c given all the evidences supporting (e_{supp}) it.

For the present scenario, e_{supp} represents a supporting or additional evidence along with the above-mentioned evidences E1 to E4. i.e., a person with other supporting symptoms like sore throat or diarrhea, may be affected by CV.MUT[c,e_{aga}] represents the consolidated measure of untrust in conclusion c given all the evidences against (e_{aga}) it.

And for the present circumstance, e_{aga} represents a person who are against any one of the evidences E1 to E4. i.e, A person who is against the occurrence of all these evidences may likely get affected if he was in contact with CV patient by history. Hence, the Government is insisting to maintain social distance among people to prevent the spread of virus.

From the above discussion, it is represented that if the value of certainty factor is 0, i.e, untrust, the person is not affected by Corona Virus (CV). If the value of certainty factor is 1, i.e, trust, the person is affected by Corona Virus (CV) as explained in Case 4.

Once the person is affected by Corona Virus, knowingly or unknowingly starts spreading the virus to another person. The spread of the virus can be modeled as set of differential equation using classic SEIR model shown in Fig. 2 to simulate SEIR epidemic. The infectious rate, β, controls the rate of spread which represents the probability of transmitting disease between a susceptible and an infectious individual. The incubation rate, σ, is the rate of latent individuals becoming infectious (average duration of incubation is 1/σ).

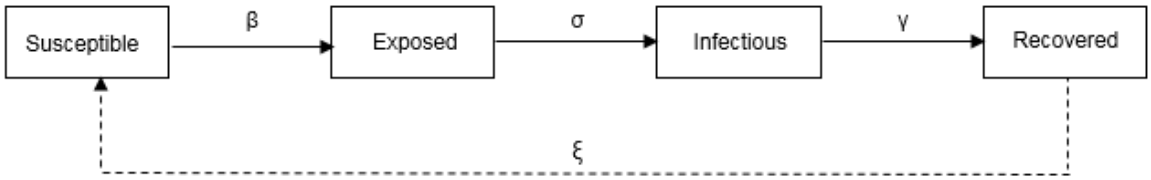

Fig. 2. SEIR Model spread of Virus [13].

Recovery rate, γ = 1/D, is determined by the average duration, D, of infection. For the SEIR model, γ is the rate which recovered individuals return to the susceptible statue due to loss of immunity [13-14]. Many diseases have a latent phase during which the individual is infected but not yet infectious. This delay between the acquisition of infection and the infectious state can be incorporated within the SIR model by adding a latent/exposed population, E, and letting infected (but not yet infectious) individuals move from S to E and from E to I.

IV. RESULTS AND DISCUSSION

Table 1 shows the cumulative number of people affected (Actual) for the period 01st March 2020 to 03rd April 2020 [15]. Fig. 3 illustrates the same.

Table 1: Cumulative number of people affected (Actual) during March 2020 in India.

Date	Cumulative number of people affected (Actual)
01/03/20	3
08/03/20	39
15/03/20	110
22/03/20	396
29/03/20	1024
03/04/20	2547

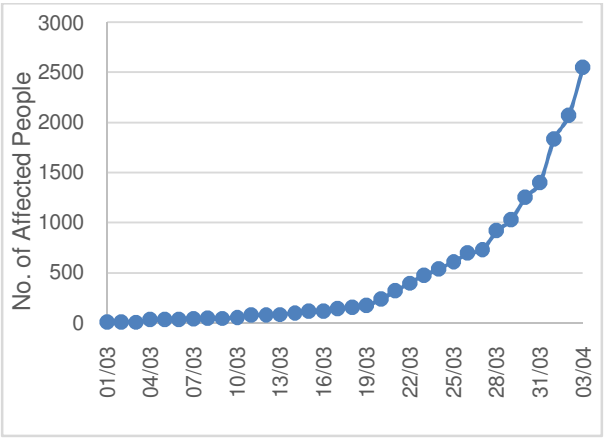

Fig. 3. Total Number of people affected by Corona Virus in India.

In order to predict the number of people likely to be affected in the next 6 days (i.e., from 04th April 2020 to 09th April 2020), simulation has been carried out [16]. In this simulation, the historic data are provided as input and then the prediction is obtained. Table 2 shows the predicted values for these days. In order to evaluate the error %, the actual cases and the predicted cases are compared for the period from 29/03/2020 to 03/04/2020. The graphical representation is shown in Fig. 4.

Table 2: Prediction for the next 6 days.

Date	Cumulative cases			Cases / day		
	Actual	Predicted	*Error (%)	Actual	Predicted	*Error (%)
29/03	1024	1084	5.86	106	165	55.66
30/03	1251	1280	2.32	227	196	13.66
31/03	1397	1510	8.09	146	230	57.53
01/04	1834	1781	2.89	437	271	37.99
02/04	2069	2101	1.55	235	320	36.17
03/04	2547	2478	2.71	478	377	21.13
04/04	-	2923	-	-	445	-
05/04	-	3447	-	-	524	-
06/04	-	4065	-	-	618	-
07/04	-	4792	-	-	727	-
08/04	-	5648	-	-	856	-
09/04	-	6656	-	-	1008	-

Note: * Modulus / Absolute Value

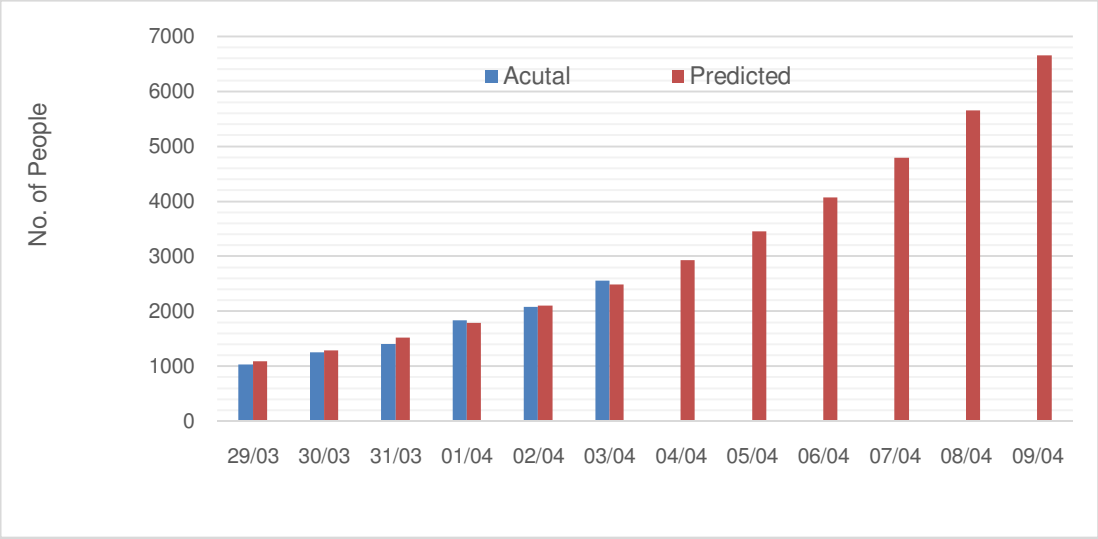

Fig. 4. Forecasting for the next 6 days.

$$\text{Error, } \% = \frac{(\text{Actual Value} - \text{Predicted Value})}{\text{Actual Value}} \times 100\ \% \qquad (10)$$

It is observed that the spread rate is keep on increasing. Hence to reduce the spread of this virus, people has to follow the protective measures as suggested by World Health Organization (WHO) and our Indian Government. If a single person is confirmed CV positive with laboratory tests (as mentioned in probabilistic reasoning) in a country. Then the infected person 'a_1' may contact assuming minimum 3 persons (r=3) in that days. The probability of symptoms can be seen within two weeks (n = 14 days).

In this case, the probability of the conclusion CV spread can be calculated by applying summation of geometric progression as,

$$S_n = \frac{a_1(r^n - 1)}{r-1}, r \neq 1 \qquad (11)$$

$$S_n = \frac{1\,(3^{14} - 1)}{3 - 1} = \frac{4782969\ -\ 1}{2} = 2391484$$

If the number of suspected patients as 500, then within two weeks (n = 14 days),

$$S_n = \frac{500\,(3^{14} - 1)}{3 - 1} = \frac{500\,(4782969\ -\ 1)}{2} = 1195742000$$

Leads to the conclusion that the entire country may get the probability to be infected with CV, which should not be the case to occur.

V. CONCLUSION

This paper analyzes the current outbreaks of epidemic diseases. To analyze the number of evidences or sources which leads to a conclusion that a person affected with corona virus, a mathematical method has been proposed. With emergence of computing tools, mapping mathematical methods and generating numerical solutions leads the prediction of certainty factors. It is observed that day-to-day the spread rate is increasing. Hence, the potential of preventive measures such as practicing personal health and hygiene habits, stay home, maintaining social distancing, isolation, quarantine, etc., should be taken into account in controlling the epidemics.

In this paper, the number of evidences or sources used to arrive at conclusion that a person gets affected by Corona Virus are restricted with limited knowledge. The other evidences or sources such as person age, immune system, any medical history etc. could be taken under consideration as future work.

REFERENCES

[1]. Javid Ahmad Dar and Khursheed Ahmad Wani. (2010). Environmental changes and emerging vector-borne diseases: A review. *Biological Forum — An International Journal*, *2*(1): 78-83.

[2]. Thomas, B., & Chugan, P. K. (2019). Insights from Changing Landscape in New Product Development and Approvals in Pharmaceutical Industry: Recent Trends for Betterment of Humanity. *International Journal on Emerging Technologies*, *10*(4), 85-89.

[3]. Shabia Shabir Khan and S.M.K. Quadri. (2016). Prediction of Angiographic Disease Status using Rule Based Data Mining Techniques. *Biological Forum–An International Journal, 8*(2): 103-107.

[4]. Liu, T., Hu, J., Xiao, J., He, G., Kang, M., Rong, Z., ... & Zeng, W. (2020). Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China. *bioRxiv*. (2020, February 13) Retrieved from bioRxiv 2020.01.25.919787; doi: https://doi.org/10.1101/2020.01.25.919787

[5]. Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... & Xing, X. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa2001316.

[6]. Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, *395*(10225), 689-697.

[7]. Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., ... & Wang, M. H. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases*, *92*, 214-217.

[8]. Zhao, S., Musa, S. S., Lin, Q., Ran, J., Yang, G., Wang, W., ... & Wang, M. H. (2020). Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: a data-driven modelling analysis of the early outbreak. *Journal of Clinical Medicine*, *9*(2), 388.
[9]. Chen, T., Rui, J., Wang, Q., Zhao, Z., Cui, J. A., & Yin, L. (2020). A mathematical model for simulating the transmission of Wuhan novel Coronavirus. *bioRxiv*.
[10]. Chen, T. M., Rui, J., Wang, Q. P., Zhao, Z. Y., Cui, J. A., & Yin, L. (2020). A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infectious Diseases of Poverty*, *9*(1), 1-8.
[11]. W. Feller, An Introduction to Probability Theory and its Applications, Vol. 1, Wiley.
[12] Jones, P. L., & Graham, I. (1988). *Expert systems: knowledge, uncertainty and decision*. Chapman and Hall.
[13]. Retrieved from https://www.idmod.org/docs/hiv/model-seir.html
[14]. Mandal, S., Bhatnagar, T., Arinaminpathy, N., Agarwal, A., Chowdhury, A., Murhekar, M., Gangakhedkar, R. R. & Sarkar, S. (2020). Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach, *Indian J Med Res*, Reterived from DOI: 10.4103/ijmr.IJMR_504_20
[15]. Retrieved from https://www.bloombergquint.com/coronavirus-outbreak/coronavirus-india-updates-india-reports-largest-single-day-death-toll-from-covid-19
[16]. Retrieved from Milan Batista (2020). fitVirus (https://www.mathworks.com/matlabcentral/fileexchange/74411-fitvirus), MATLAB Central File Exchange. Retrieved March 31, 2020.

How to cite this article: Varalakshmi, R. (2020). Mapping of Corona Virus Transmission in India with a Mathematical Approach. *International Journal on Emerging Technologies*, *11*(2): 245–250.