A Real-Time Predictive Modelling for Mitigating Contagious Diseases

Shresth Agrawal*, **Samriddhi Vishwakarma****
Guide- Sushil Kumar Pandey
shresthagrawal.31@gmail.com | samriddhi.sv11@gmail.com | sushilpandey0176@gmail.com

Abstract- *In this research we have tried to create a real time system that can predict the behavior and spread of known as well as unknown contagious diseases and finally give policy suggestion on what, where, and how much of vaccination, medication, and isolation is required to stop the outbreak of the disease. This system also tries to come up with an innovative mitigating solution Cyber Quarantine using GPS data points. The system works on the plank of SIR model and tries to use the latest systems and technology to improve the existing SIR Model. It is also linked with a huge Data Base of IDSP (Indian Disease Surveillance Program) and finally the whole system is formulated in form of an Android App so that it can reach the masses and can be easily accessible.*

I. INTRODUCTION

"Contagious diseases is the world's biggest killer which kills approximately 15 million people annually"- **WHO**. There are several mathematical modelling on Epidemics. Ronald Ross in the year 1857, formulated the first mathematical modelling on malaria parasite. Mathematicians generally use the SIR model for forecasting epidemics, this was first founded by A.G. Kendrick and W.O. Kermack in the year 1927 in the research paper- "A CONTRIBUTION TO THE MATHEMATICAL THEORY OF EPIDEMICS', in this paper a simple deterministic model was formulated and was successful in predicting the behavior of outbreaks very similar to that observed in many recorded epidemics. Till date myriad number of statistical model exists but they have certain limitations:

- They fail to explain the most uneven and essential component of a society, migration.
- The transmission and recovery coefficient doesn't yield good accuracy on real time basis as the biological factors which change overtime affecting the prediction accuracy.
- It fails to characterize a disease on the basis of severity:
  for e.g. : thousand cases of malaria in a Topical nation like Africa may not be an epidemic but only a handful cases of malaria in Canada would be an epidemic, that's because thousand cases of malaria in endemic tropical area is considered to be normal.
- Mitigation Methods like vaccination, isolation and medication are not quantified on real time.
- Moreover there exist many android application such as **Healthmaps, Geni-dp, Pro-med** etc. that acts as an alert system for an infectious disease but there exists no application to compute and predict infectious disease using mathematical modelling on real time basis.

Our system solves all the above problems and limitations of the existing SIR model and also creates a system that can be accessible even to the general public.

II. LIME-LIGHT / NOVELTY

- The prediction of this system will be characterized into three levels

  Foremost-prediction: It is often observed that there is a lot of trepidation as soon as any outbreak of an unknown infectious disease takes place, so this system will help to combat the situation by matching its behavior with an existing outbreak.

  Quasi-prediction: As soon as a new disease is identified there may exist mitigating method to it. In this situation quasi prediction will be taken into account to find the most effect policy suggestion.

  Hindmost-prediction: If there is a re-outbreak of a diseases in the same or different area, the system will use the data that was obtained from the earlier outbreak of the same disease to predict the current outbreak and this stage is known as Hindmost prediction.

- In the course of an infectious disease there can be various possible solutions: medication, vaccination, quarantine and isolation. This system will quantify the Mitigating solution using mathematical modelling.
- It is not possible to completely quarantine the susceptible population therefore we have created a method to alert a susceptible about an approaching infection using android app and hence helping people to quarantine themselves. This method is coined as CYBER QUARANTINE.
- This system also includes the migration using GPS data points that will enable the system to track the movement of people in and out of a place using our app installed in their smart phones.
- Government Corner: Hoverer, the government has the resources and capability to provide mitigating solutions but where it lacks is to when, where and how much to provide these solutions. Therefore this system will uses mathematical model to solve this problem as well.

III. Mathematical Modeling

Prediction: In this system we have used the SIR model as a plank for prediction. Let's assume the below block diagram as the world mainly divided into three compartments during the outbreak of an Infectious Disease. Susceptible **(S)**: those who have chances to get infected, Infected **(I)**: those who are infected and can spread the disease, Removed **(R)**: those who are recovered or dead and further can't spread the disease.

We observed that the behavior of a contagious diseases is analogous to unimolecular consecutive reaction **(UCR)**. As in a reaction a molecule of one reactant comes in contact with other reactants to form product molecules, similarly during the spread of a contagious diseases the Susceptible comes in contact with an infected to get infected and simultaneously removed and this analogous behavior was clearly significant in the graphs of both UCR and SIR [fig(v)] which were almost similar. As we derive rate laws for a chemical reaction we have used a similar approach to derive rate laws of contagious diseases. When epidemic controlling steps are taken, the number of factors come into existence that dynamically change the size of the S, I and R parts. So the following factors were included in the SIR system as depicted in fig (i)

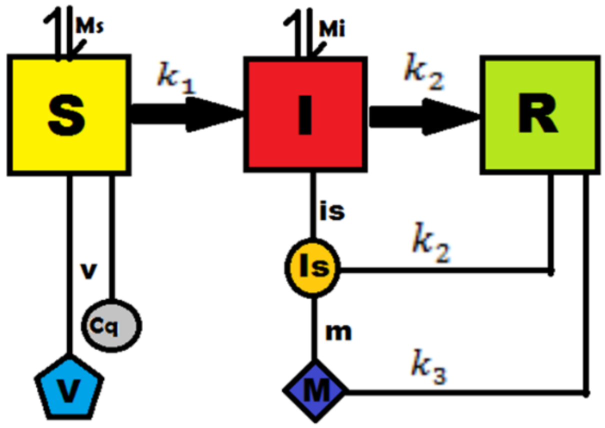


Fig (i): The double sided arrow in Ms and Mi indicate that the change could be both positive and negative

$$\frac{dS}{dt} = -(S - V - Cq)(I - Is)k_1 + Ms \qquad (i)$$

$$\frac{dI}{dt} = (S - V - Cq)(I - Is)k_1 - (I - M)k_2 - Mk_3 + Mi \qquad (ii)$$

$$\frac{dR}{dt} = (I - M)k_2 + Mk_3 \qquad (iii)$$

$$\frac{dV}{dt} = v \qquad (iv)$$

$$\frac{dM}{dt} = m - Mk_3 \qquad (V)$$

Fig (ii): the above equation are the final equations derived for the prediction.

$$\int_{tp}^{s} \frac{dS}{k_a - S - k_b \log s} = \int_0^t dt \qquad (vi)$$

$$S - k_c \log S + k_d = 0 \qquad (vii)$$

Fig (iii): In the equation (vi) and (vii), k represent constant values. We were unable to solve any of the above equation analytically.

Modification done in existing SIR model:

- Mitigating Factors V (vaccination), M (medication), Is (Isolation) and their rates v, is, m respectively.
- Coefficient of recovery k_3 (Rate of Recovery after Medication).
- Mitigating factor Cq (Cyber Quarantine).
- Migration: Ms (migration in susceptible) and Mi (migration in infected).

Derivation of the block diagram and the Equations: [1]

- Let's assume there is an outbreak of a contagious disease, at the initial state **T(0)** the total population will be susceptible to the disease, after some time, some population will get infected at the rate k_1 and some population will get removed at the rate k_2
- At some time **T(n)**, if some mitigating solutions are launched some people will get vaccinated(**V**) at rate **v,** some people will get cyber quarantined(**Cq)** and similarly in infected some people will be hospitalized and hence isolated(**Is**) and further can no more spread; out of those Isolated some will get Medicated(**M**) at rate **m**. People isolated will recover with the same recovery rate k_2 as non-isolated infected population and people who are medicated will recover with a different recovery rate k_3**.**
- During the outbreak there will some migration in both susceptible and infected population which can be either negative or positive. We have included two rates **Ms,** migration in susceptible and **Mi,** migration in infected
- Using the above assumption we derived the equation fig (ii).

Solving of the Equations: [2]

- The equation obtained above are system of linear differential equation of first order. We first used analytical methods to solve this system but the method wasn't successful as the system would end up with an inverse of explicit function having logarithm and linear terms which would further become un-integrable. Check Equation (vi) and (vii).
- Based on our observation of the shape of the curve our second approach was curve-fitting, we tried with different normal distribution curve: Chi-square, Gaussian, and Poison. It yielded good results with infected curve but couldn't explain the behavior of susceptible and removed.
- Finally we moved to the Numerical Integration Method, where we used Runge-Kutta method.

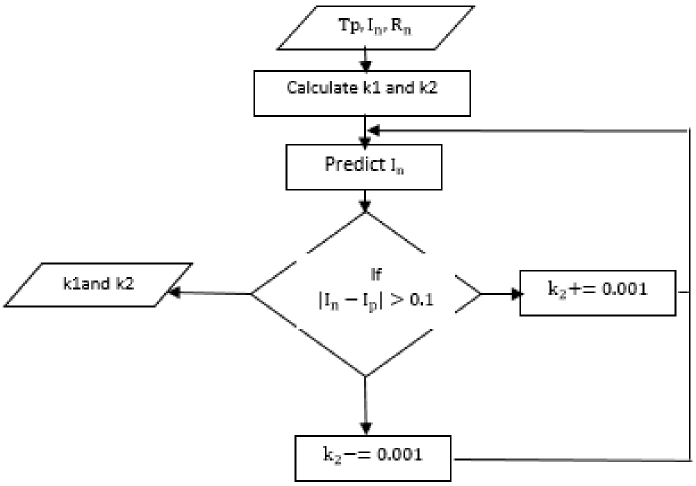


Fig (iv): the above algorithm shows the step by step process involved in finding k1 and k2

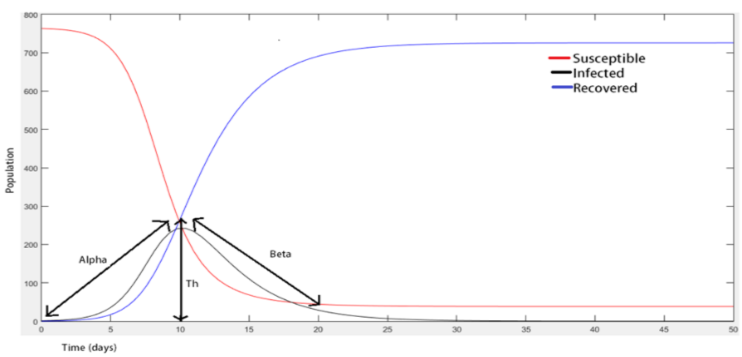


Fig (v): the above graph depicts the susceptible infected and removed population with time

$$\mathrm{Th} = \frac{(I - M)k_2 + Mk_3 - Mi + V + Cq}{k_1(I - Is)}$$

$$\mathrm{beta} = \frac{I_{max} - I_{(2*t_{max})}}{t_{max}}$$

$$\mathrm{alpha} = \frac{I_{max} - I_0}{t_{max}}$$

Fig (vi): the above equation are the final equations derived for the three mathematical factors.

Finding the Coefficients k1 and k2: [3]

$$\frac{\log\left(\frac{Tp}{Tp - I_n - R_n}\right)}{R_n} = \frac{k_1}{k_2}$$

Step1: First the initial parameters **Tp** (Total Population), $\mathbf{I_n}$(Infected at day n) and R_n (Removed at n) are used to calculate the ratio k_1/k_2 using the above equation.

$$k_2 = {}^{1}/_{d}$$

Step2: Initially we assume k_2 as an inverse of d (days required to recover or die) and using the ratio we also find k_1.

Step3: Using $k_1$and k_2 we predict the value of the Infected at day n ($\mathbf{I_p}$) using the above algorithm **[1] and [2]** and check the error difference between $\mathbf{I_p}$ **(**predicted) and $\mathbf{I_n}$ **(**original).

Step4: According to the error, we again modify k_2 and go back to step3 and this is continued until a certain level of accuracy is obtained. And hence both k_1 and k_2 are obtained which could further be used to predict the spread of the contagious disease.

Quantification [4]: There existed no Scale to measure the severity and compare the contagious diseases, hence we have divided the spread of diseases into three major mathematical factors to quantify it.

Threshold (Th) – This is the point when infected population is maximum and the rate of infection is equal to rate of recovery hence **dI/dt** will be equal to zero. Equating equation (ii) we can find **Th**.

Alpha - The rate increase of infected population in per unit time till Threshold is alpha.

Beta - The rate decay of infected population in per unit time after Threshold is beta.

The ratio between the Total Population and the Th gives a relative term to compare severity between different contagious diseases. Alpha and Beta gives the rate of spread of the disease and this could be used to give Mitigations effectively.

Mitigation [5]: During an outbreak there are three possible mitigating solutions Vaccination, Medication, and Isolation. Using the equations derived and the mathematical factors above we can find the effect of any mitigating solution on the severity of the disease. We have also plotted the graph of Alpha and Beta vs the mitigating solutions. These graphs were made using algorithm described in **[3], [4].**

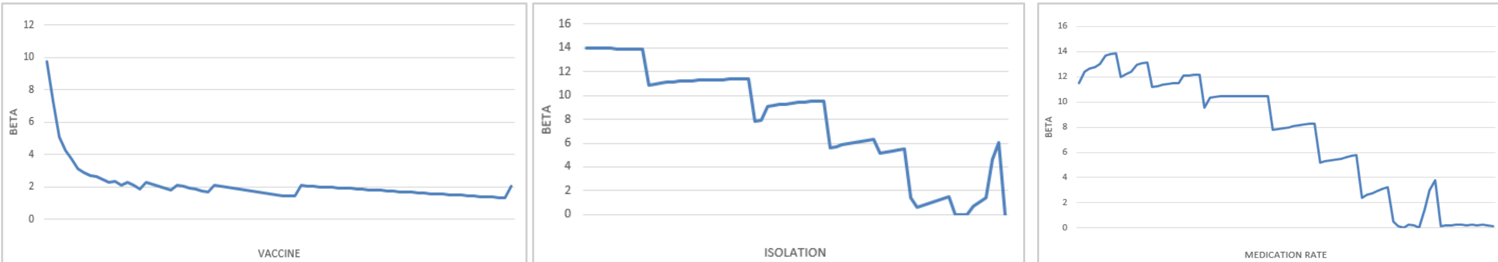


Fig (vii): The above graph represent mod of beta vs vaccination, medication and isolation following

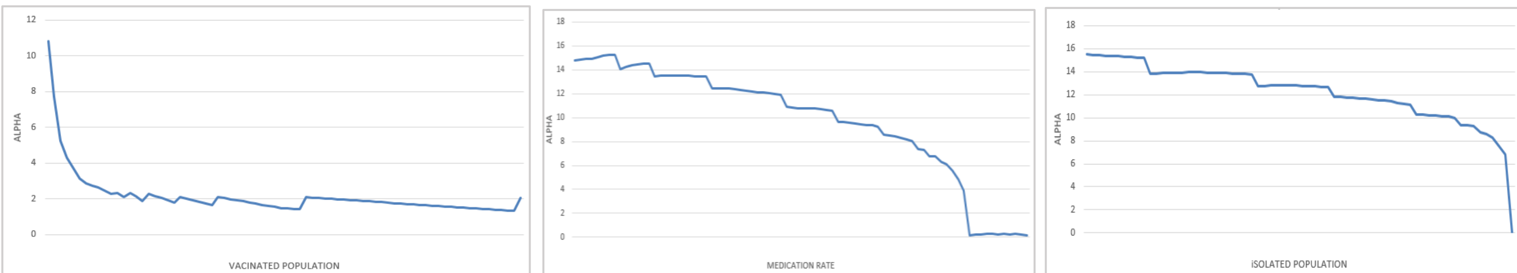


Fig (viii): The above graph represent alpha vs vaccination, medication and isolation following

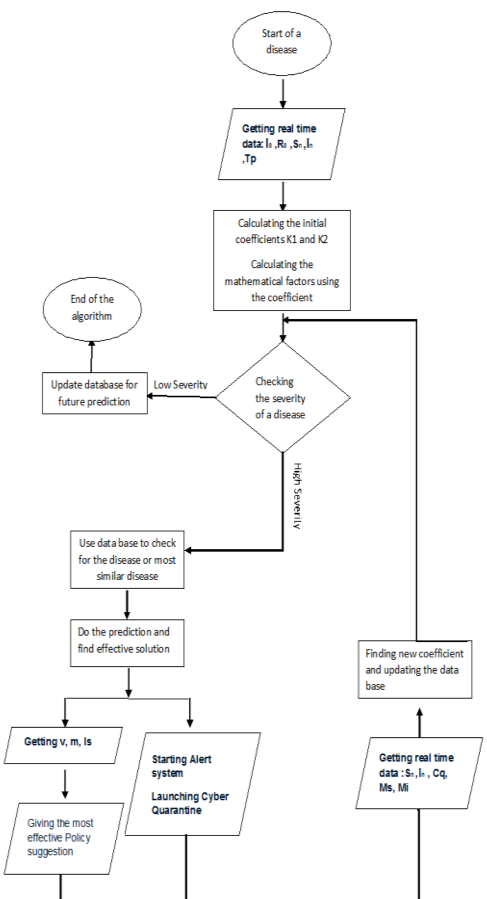


Fig (ix): the above flow chart depicts step by step process followed by the system.

Observation and Conclusion-

From the above graph we observed that in all of the mitigating solution, Vaccination was found to be most effective. In each curve there was a certain value after which the graph decreased significantly which signifies that the diseases will gradually stop. **Hence we concluded for a given set of outbreak there will be a unique set of mitigating solutions to stop the disease. Hence if amount of resources available is given using the above algorithm we can find the best and the most effective policy suggestion.**

IV. Creating Surveillance System

We have created a real-time system which works on our mathematical concepts that were derived above to create a complete contagious disease surveillance system.

Pre-Processing [6]:

We used the database of IDSP (Indian Disease Surveillance Program) which has disease data of more than 5000 diseases from previous 3 years. This data was in PDF format which was first converted into excel and then for each diseases outbreak using Matlab, mathematical factors (alpha, beta, Th) were calculated (**Algorithm described in [3], [4]**) which were finally uploaded in the Virtual Private Server's Database. We also tested our System with three international outbreaks- Ebola, Plague and Influenza.

Data Collection [7]:

General Data: During an outbreak of disease the data is collected in real-time through our Android app created in Android Studio where general public can fill the form and upload basic details about the infection around them. We also have a web portal where hospitals and PHC's (Primary Health Care Centers) can upload information.

Migration**:** Using the android app we will monitor the GPS location and as soon as there is a change in zip code this new zip code data will be triggered to the server.

Government corner**:** In the web portal there is a Government section where government officials can provide the details of the available mitigating solutions.

Data Processing [8]:

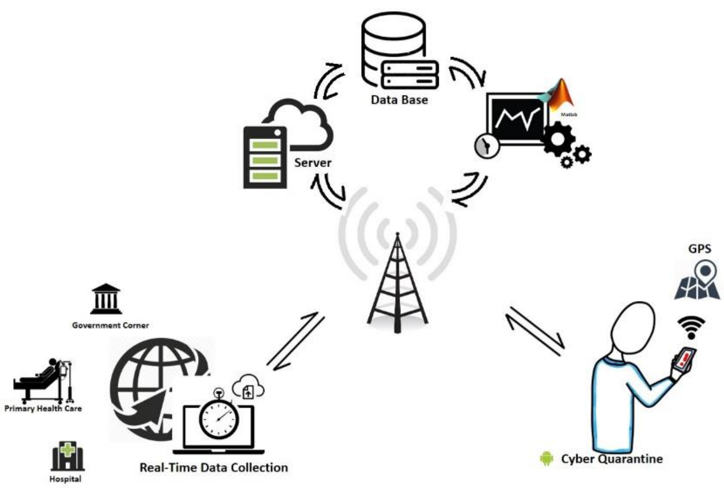
As soon as the data reaches the server it is converted to initial parameters using PHP MySQL. The initial parameters are then used to find the coefficients k1 and k2 (**using algorithm [3]**) and further find the mathematical factors (**using algorithm [4]**). This is done through runtime Matlab program with all the equation coded on the server. After that the severity is calculated and if the severity of a disease is high, then the algorithm proceeds or else the algorithm stops.

Prediction Stages: The Prediction stage and the policy suggestion will differ with different stages and types of outbreak.

•Foremost: In this case, the disease outbreak is totally unknown to the system and no mitigating solution exists. In this case first the factors of original diseases is calculated, then using the existing database the disease most similar to the new outbreak is found and prediction is done on the basis of that. Cyber Quarantine and Isolation are most the effective solutions.

•Quasi: In this case, the disease is known and some mitigating solution exists. Here, first the factors of the original disease is calculated, and prediction is done on the basis of that. Cyber Quarantine, Isolation can be done and the medication, vaccination as per availability

•Hindmost: In this case, the disease and the mitigating solutions are known to the system and there is a re-outbreak, in this case the system uses the exiting mathematical factors to initially scale the disease outbreak and then later change it on real time. All mitigating solutions are possible here.



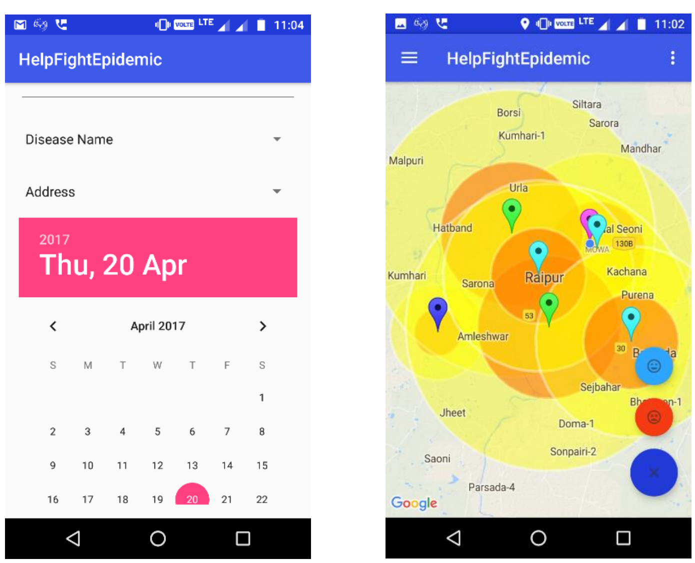


Fig (ii): The above phots are the screen shots of the Android Application

Surveillance System:

Policy Suggestion: If the severity of the disease is high then the using the government corner in our web portal, the amount of available resources like vaccination, medication which is available could be entered by the government officials and the system could then calculate how, where and when of these resources could be used so that the effect is maximum and also how much more is required to stop the Disease. (**Using algorithm [8], [5], [4]**)

Cyber Quarantine: During an outbreak, it is not possible to quarantine susceptible population, and in some cases there are no mitigation solutions available, for such situation we have created an Innovative Method: Cyber Quarantine. People who are susceptible could easily download our app and using there GPS location the system will warn people about the approaching infection around them and hence quarantining them.

End of Algorithm:

After the prediction is done and the mitigation is launched the system updates the database for the future prediction, and hence our system will get better and better with time. It will again move to the data collection stage and again and the severity will be calculated. This will continue on a loop until and unless the severity of the disease is nullified. Calculating factors and coefficients every time will make it ineffective to the change in the environment factors like temperature or climate and thus these changes will not be able to affect the accuracy of the system.

IV. SURVEY

During the research we went on different surveys to collect data about prevailing contagious diseases, some of the major surveys were done in:

- Balagi Multispecialty Hospital Raipur
- AIIMS Medical Hospital Raipur
- Jawaharlal lal Nehru Medical college

V. PILOTING

Currently we have started a campaign "**Kill it B4 it kills u**" to fight with contagious diseases working on this research and under the guidance of the Chief Minister of the State we are piloting this project in state of Chhattisgarh.

VI. CONCLUSION

In this project we have used system of differential equations to predict the rate of spread of contagious diseases, quantify it using mathematical factors and give mitigating solutions accordingly. This method also tries to create an effective real-time working system to for complete surveillance and mitigation of contagious diseases. This system is the need of an hour can help humanity to come out from the clutches of world's biggest killer: contagious diseases.

Link to Campaign: https://salabs.in | Link to Codes: https://github.com/SA-Labs

V. REFERENCE

[1]. www.ncbi.nlm.nih.gov/pmc/articles/PMC3162588/

[2]. www.sciencedirect.com/science/articles/pii/S1198743X14604080

[3]. Mathematical Biology, J.D. Murray, Springer-Verlag, 1989

[4]. Using Real Data in an SIR Model D. Sulsky

[5]. https://www.cdc.gov/about/facts/cdcfastfacts/contagious-diseases-disasters.html

[6]. https://sites.math.washington.edu/~morrow/mcm/mcm15/38725paper.pdf

[7]. https://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/index.html?s_cid=cs_4926
[9]. *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*
[10]. http://web.iitd.ac.in/~pmvs/courses/mel705/curvefitting.pdf
[11]. http://rspb.royalsocietypublishing.org/content/267/1458/2219.short
[12]. https://search.informit.com.au/browseJournalTitle;res=IELAPA;issn=1447-4514
[13]. http://www.sciencedirect.com/science/article/pii/S0895717700000406