Understanding and Modeling the HPV Virus behaviour in the society

Manoel A. Freitas Neto

Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco (UFRPE)
Caixa Postal 52171 – 900 – Recife – PE – Brazil

mbetmann@gmail.com, manoel.freitasnt@ufrpe.br

Abstract. *This article describes a work developed to the academic discipline Mathematical-computational modeling applied to epidemiology, on the system of information course - UFRPE. The project was built with the assistance of the professors Jones Albuquerque and Silvana Bocanegra. The objective of the study was to understand and model the behaviour of the different types of HPV virus, in the woman organism, based on statistical data and using cellular automata and compartmental models as support tools to assist on the understanding and on the simulation of the conduct of the pathology in a closed society.*

Resumo. *Este artigo descreve o trabalho desenvolvido na cadeira de Modelagem matemático-computacional aplicada à epidemiologia, no curso de Sistemas de informação - UFRPE. O projeto foi construído com o auxílio dos professores Jones ALbuquerque e Silvana Bocanegra. O objetivo do estudo foi, compreender e modelar o comportamento das variedades do vírus HPV, no organismo de mulheres, baseado em dados estatísticos e utilizando autômatos celulares e modelos compartimentais como ferramentas de apoio para auxiliar no entendimento e na simulação da conduta da patologia em uma sociedade fechada.*

1. Introduction

Using compartmental models and cellular automata, this work tried to understand and modeling the behaviour of the HPV. This is possible because the infection follow some rules to infect people and other laws when it is installed on humans bodies. The objective to model, is to simulate or simplifying how something works in the real world but using finite resources.

Our world is a combination of a non deterministic, probabilistic, dynamic and exponential machines, thats why is so difficult to simulate that behaviour. So, we introduce the cellular automata, a discrete model, used to simulate complex behaviors and expose this in a visual way.

Looking to the health side, the amount of the cases of STD (Sexually Transmitted Disease) has been increasing, according to the OMS (2018), this growth, is mainly on the younger part of the population, 15 to 25 years old. The HPV is a virus with easy transmission because only needs dermal contact, and keeps on the latent state

or without showing symptoms for months or even years, thereby, turning a person into a vector and spreading the infection.

2. Cellular Automata and Structure.

Cellular automaton is a spatial model of temporal evolution, they can have one, two or three dimensions and are used to display and simulate natural behaviors based on predetermined rules. Is based on turing machines, that is, state machines that evolve according to transitions.

One of the most known cellular automata is the Game of Life, developed by John Conway in 1970. The model evolution follow four simple rules: Any live cell with fewer than 2 neighbors dies by underpopulation; any live cell with more than 3 neighbors dies by overpopulation; any dead cell with exactly 3 neighbors become a live cell; any live cell with 2 or 3 neighbors, unchanched, to the next generation. This laws generate cellular structures which simulate social behavior.

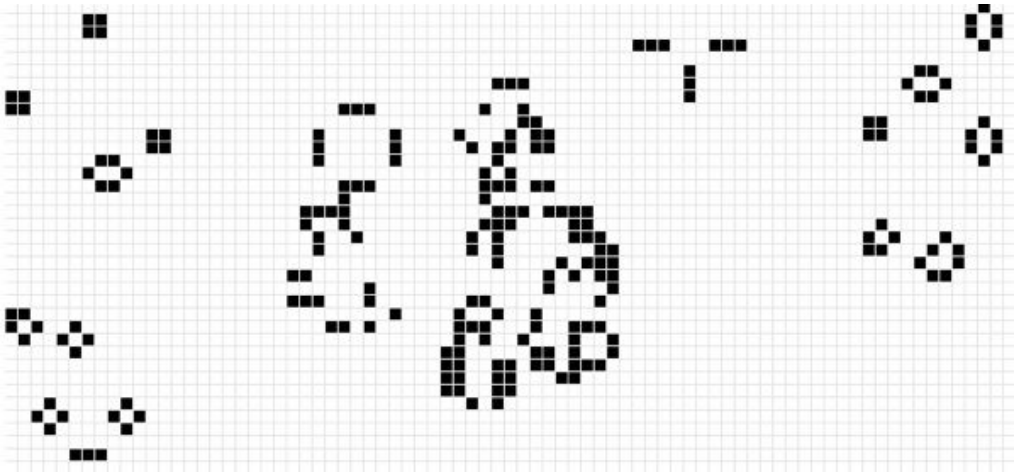

Figure 1. Conway's Game of Life

2.1. Compartimental Models

When we work with epidemiologic models, one known technique is used to classify the people in categories, or compartments, so comes up the compartment models. The main idea is to separate each being, based on their state on the instant t.

Inside this approach, the more basic deterministic model is called SIR (susceptible-infected-recovered), this let us model a viral disease like flu. Walking to the complexity, we face with infections which does not fit in this model because does not have the recovered state or just never get better. For this, we just adapt the number of compartments to the pathological conditions, for example, in measles, the individual gets out of an infected state to a recovered and after immune, in others viral spreads, the person become susceptible after recovered.

Still talking about structures, we have to work with the idea that in a society, people born and dies, so, the number N of the individuals is not constant and follow the

birth rate and mortality rate. To simplify the model, we can set this value. The idea in the HPV project is to have a non variant N, but we can eliminate and create new cells, generating new individuals with new characteristics but preserving the amount to keeps on control, but knowing that the better case is to have a society with independent rates.

3. HPV Virus

HPV (Human Papillomavirus) is a virus from the family papillomaviridae with a double stranded DNA, when it is on symptomatic state, the disease develop warts, more known as condylomas, in mucous membranes like mouth, vagina and anus. The virus has more than 150 cataloged types. Classified as an STD, the virus is acquired frequently throughout sexual relations without protection, but also can be transmited by mother to fetus or sharing of contaminated objects.

The virus has three states of life in an organism, it can keeps latent, in this case the people are not vectors, and it is the smallest installment of the cases, the other two was used on the research, is the infection state with symptoms and the infection state without symptoms, both very important to understand the disease infection capacity.

According to the statistics given by Dr. Waldemir Washington Rezende, part of the central institute of medicine college from USP, 75% of the women, during their life, will get in touch with some variant of the HPV virus. Fortunately, 95% of these cases will get better by spontaneous resolution of the organism, the remaining, who can not take care can developer cervical cancer in case of the wounds advance over time.

For prevention, we have the same contraceptive methods used against the others STD and the vaccines given by the health system, in Brazil, to the girls who has 9 until 13 years old, this prevents the types 16, 18, 11 and 6, which more commonly cause cancer. Furthermore, there is the pap smear exam, which can detect the condylomes and the presence of the virus.

4. The Model

The experiment was made based on compartmental models and structured like a cellular automaton with two dimensions, a model which uses month as the time unit (each transition), working with the Moore neighborhood, it means, we consider the 8 neighbors around a cell. With predetermined rules, the automata can change the state of the individuals. The complexity and precision of the model increases proportionally the amount of the variable and laws applied to the cells.

The first step was delimit the size of the matrix, that is, the quantity of the cells which would be part of the grid. The people' rating also needed to respect statistics real data, therefore, was used real data from the census IBGE (Brazil - 2010). This information was extracted to delimit the amount of people in the age ranges, moreover, the number of individuals who are classified as virus carries, was taken from the research from Ministry of health (Brazil - 2018).

Going deeper into the code and the construction, was used the programming language Python with the GUI (graphical user interface) package named Tkinter to present the visual. A class Cell was made to instance objects with the attributes: time of

infection, state, age and gender. With this individual characteristics, the transitions rules are applied to change the states.

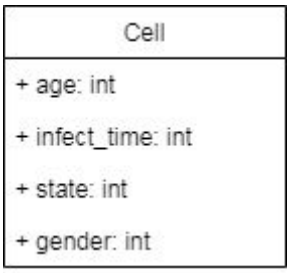

Figure 2. UML diagram from a cell

The cell population was divided em age ranges like 15 to 25, 25 to 49 and 50 to 70 years old, each one with the quantities based on the census of the brasilian people, also accorind to the biological sex male and female. Focusing on the young population, to observe the results of the study, we consider that about 55% of the young (15 to 25 years old) had contact with the disease, this data was taken from the research from the Ministry of health (Brazil- 2018).

Inside the model, the possibility to create new cells is given by the death of old cells, so, we have an N (number of individual) static, like was said before, but with the generating of new individuals, making the birth rate equals the mortality rate in the system. Is understandable that this harms the simulation of reality, but currently the simplicity of the model was preserved and after we can sophisticate and create new conditions on the research.

4.1. States

With all the system based on compartmental models, was created boxes to classify the cells according to the state which they are. We made this to identify the individuals and make accounts, during the model execution, aiming at the follow amd compare with real database. The amount of the people in each state was previously determined statistically based on health data which contains age ranges about how many people got in touch with the virus. Initially, was generated 20 distinct cells, with mixed attributes according to the statistics. Was created three compartimental states represented by 0, 1 and 2 in the code.

The first state is the susceptibility state, where most of the cells are, this means that those who are in this phase is fit to get the infection and it is up to the rules. From the array with 20 cells, 15 was initially set on the state 0, that is 75% of the closed population was in the state. This quantity includes half of the young population and a part of adult and old people, of both sexes.

The second state already identified as infection is reached when the individual get in touch with the virus (the vector) but not developer symptoms yet, nevertheless, the person has the capacity to infect other, this will get a better approach in the next topic. The idea is that the automaton follow the contamination statistics from the young Brazilians.

On the third state fit the people who start to manifest the symptoms, is who present the condylomes on the mucous region. This was the more difficult classification because there is no exact time or medium time to people began to manifest anything, some organisms take 2 months while others can take 20 years to show any indication of infection.

Knowing that HPV is a disease that has no cure, which the individual one contaminated keeps with any viral charge, we do not have a recovered state, like the basic model SIR suggest, to treat this case and to avoid that all cells in the grid get infected, was created a mechanism, which if one cell stay more than 12 months infected, the cells is excluded from the matrix and the blank space is filled with a new cell. In the end, the compartmental model of the states of the automaton is represented by the flowchart below:

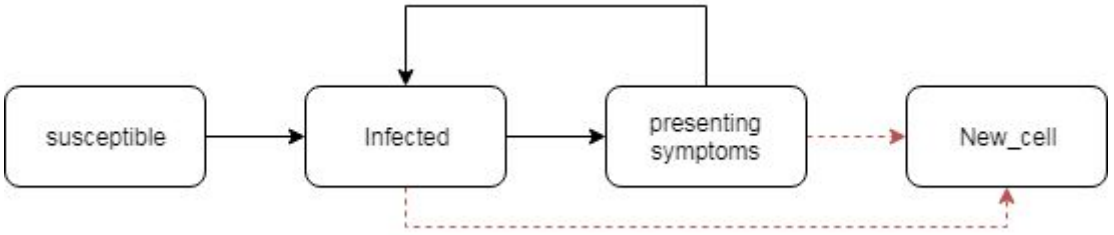

Figure 3. Compartimental model flow chart

4.2. Transition Functions

The rules of the model were developed to the cells change the states. For each state, there are some rules which need to be respected to the individual advance. They were created based on the natural condition of the acting of the pathology and use the attributes of each object (cell), the goal was to keep more closer as possible the reality.

An individual located on the state 0, is susceptible to get the virus, to this, he needs to get in touch with another individual who has the disease. The big question is that, in the model, there are two types of cells which are infected, the ones which are on the state 1 (infected) and the ones which are on the state 2 (infected and symptomatic), according to the INCA (Instituto Nacional de Cancer - Brazil), a person who has condylomas has more chances to infect a susceptible individual than other who has not the wounds, but we do not have an exact number of how bigger is that capacity, to cover this question, it was determined that cells in state 2 can infect more. We follow the rules below:

- If the cell has more than 3 neighbors on state 1 or more than 2 neighbors on state 2 AND the age of the cell is between 16 and 25, it will get infected and goes to the state 1.
- If the cell has more than 4 neighbors on the state 1 or more than 3 neighbors on the state 2, it will get infected and goes to the state 1.

The same way, the rules to the transition from the state 1 to state 2 was determined based on real events, like was said before, the difficult was determined the time which an organism take to developer symptoms. As we do not have an exact time, was set that:

- If an organism remain infected for 5 months, it will begin to present symptoms e goes to the states 2.

One more condition is the way back, because a person can discover that she has the virus and start to treat the condylomas, so she will not present symptoms anymore, and then the cell needs to go back to the state 1. The rule is:

- If the infect time of the cell is more than 8, it goes back to the state 1.

It will keep on this transition until the infect time get in 12, when the cell will be excluded and created a new one.

5. Conclusions

After de automaton complete 60 transition, considering that the time of one transition is equal to one month, after 5 years, was observed some data about the amount of people infected compared to the whole, and compared with the non infected. But the focus here is to watch the number of female youths on the research. Bellow is the visual result of the automaton in four stages: 1, 12, 23 and 36 months.

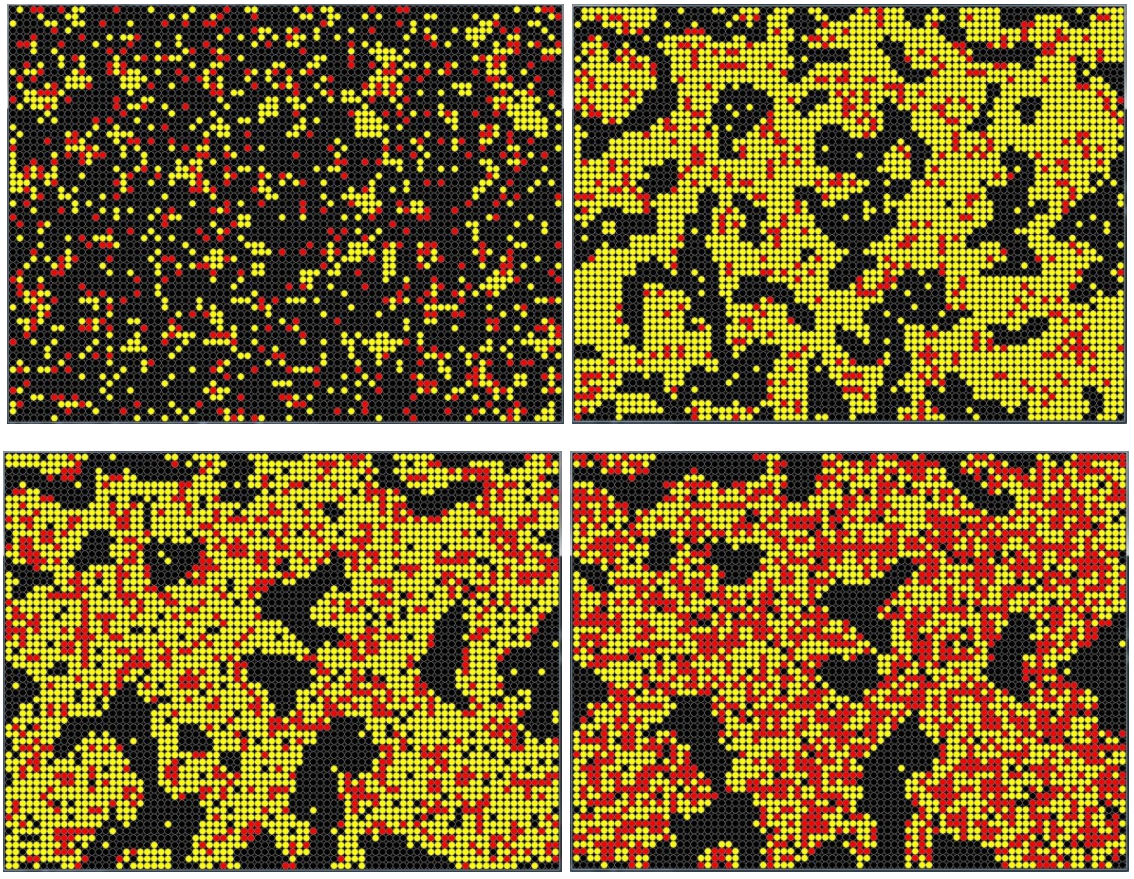

Figure 4. Temporal evolution in 4 years

In the image, we can observe that on the first moment, about 75% of the cells are susceptibles (black), and the other 25% are divided on the most on infected (yellow) and a minority that are presenting symptoms (red), like in the real world, many people get the virus and do not know because they do not produce wound.

Over the years, the amount of people infected, in the yellow and red compartmentals tends to increase with the dissemination of the virus, follow the trend of the other STD actually, mainly in the young part of the population, according to the data extracted from the simulation, about 50% of the young were infected, matching the reality, corresponding the research of the Ministry of health.

6. What We Can Improve

As we know, the more characteristics the cells have, more specific and precise the is de model, but do this make the process more costly computationally speaking. Thus, as a goal, the work will be reviewed and each individual will get more attributes which are more relevant to the system, for example: civil state, socioeconomic class and average of the number of partners.

Another point to be study better, is the delimitation of the neighborhood and the size N of the population, because we know that the reality is not exact and the birth rate and mortality rate vary according to the region which is being study.

The goal to be achieved is to reach women who are susceptible to get cervical cancer and decrease the statistics of fourth leading cause of death due to cancer factor in Brazil, for this do happen, the model needs to the more delimited and focus on the types which cause cancer, 70% of the cases are caused by the 6, 18, 16 and 11 type.

7. References

J.L Schiff (2008) Introduction to cellular automata, Wiley-Interscience, 1th edition.

Nascimento, B. J. (2013) “Um modelo epidemiológico baseado em autômatos celulares para análise comportamental de incubadas, micro e pequenas empresas ”, http://www.ppgia.ufrpe.br/sites/ww4.ppgia.ufrpe.br/files/documentos/2013_-_um_modelo_epidemiologico_baseado_em_automatos_celulares_para_analise_comportamental_de_incubadas_micro_e_pequenas_empresas.pdf, June.

Palefsky, J. M. (2011) “Human Papillomavirus-Related disease in men: not just a woman issue ”, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2871537/, June.

Ault, K. A. (2006) “Epidemiology and Natural History of Human Papillomavirus Infection In the Female Genital Tract ”, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1581465/, June.

Alves, D., Gagliardi, H. (2006) Técnicas de Modelagens de Processos Epidêmicos e Evolucionários, 1th edition.

Holton, M. and Alexander, S. (1995) “Soft Cellular Modeling: A Technique for the Simulation of Non-rigid Materials”, Computer Graphics: Developments in Virtual Environments, R. A. Earnshaw and J. A. Vince, England, Academic Press Ltd., p. 449-460.

CDC (2006) Principles of epidemiology in public health practice, CDC, 3th edition.