

Simulation of Swine and Avian Influenza Viruses Recombination Based on Genetic Algorithm

Ria Lestari Moedomo¹, Adang Suwandi Ahmad², Adi Pancoro³, Jorga Ibrahim⁴, and Muhammad Sukrisno Mardiyanto⁵

^{1,2,4,5}School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia

³School of Life Science and Technology, Bandung Institute of Technology, Indonesia

rialmoedomo@students.itb.ac.id

Abstract: This paper describes analysis, design and development of simulation software for Swine (H1N1) and Avian Influenza (H5N1) viruses recombination process. Influenza Pandemics have occurred several times, caused by mutation and recombination of Influenza viruses. Recombination is a sudden change of viruses enabling two different Influenza virus strains combined to become a new virus subtype. The H1N1 recombination has caused Spanish Flu Pandemic in 1918, and the recombination of swine virus, human virus and bird virus has caused Swine Flu Pandemic started in Mexico during April 2009. The main concern raised by WHO is the possibility of recombination between Swine Flu (H1N1), Avian Flu (H5N1) and Human Flu Virus (H3N2) which could trigger Swine Flu becomes more contagious, and have a higher CFR (Case Fatality Rate) causing more people to die. This research's purpose is to define modeling for virus recombination causing Influenza Pandemic phenomena. This research defines several different virus variants which can potentially trigger the Influenza Pandemic. Additionally, this simulation objective is to obtain most possible virus strains formed from the recombination, the scope within this article, which can potentially trigger Influenza Pandemic. New strains could be utilized to support the vaccine planning process. This simulation program was developed based on Genetic Algorithm method, for solving this multi-objective optimization problem. By utilizing Genetic Algorithm approach, the chromosome solution and fitness values/functions of Influenza Pandemic stages are defined and the maximum fitness values are to be searched. The simulation result of H1N1 and H5N1 recombination gave 2 best fitness values as static result and dynamic mean fitness values. Better and more fitness values could be obtained once the database of recombination and mutation virus strains is enhanced.

Keywords: Avian Influenza; Influenza Pandemic; Recombination; (Multi-Objective) Optimization; Genetic Algorithm

1. Introduction

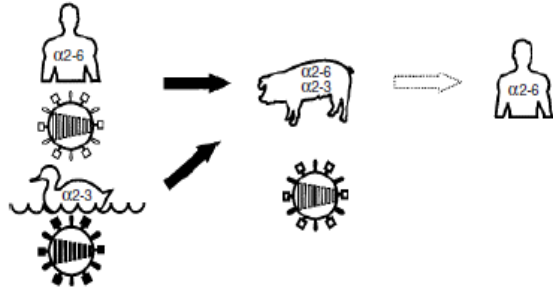
In this world, Influenza pandemic has occurred several times and has caused high number of human population deceased. Spanish Flu Pandemic in 1918, which cause the decease of 40 million people in North America and Europe, was triggered by H1N1 recombination (antigenic shift) process. Other pandemic, Asian Flu pandemic in 1957, Hongkong Flu pandemic in 1968, and Russian Flu pandemic in 1976 were also triggered by recombination process [1]. Pandemic Swine Flu in 2009 has caused almost 10.000 American people deceased (<http://www.tvone.co.id>) while in Indonesia these viruses has caused 948 confirmed cases and 4 deaths until August 2009 (<http://id.wordpress.com/tag/general-healths>) and in Australia there were 34,180 confirmed cases and 138 deaths (<http://ozswineflu.wordpress.com>).

Recombination, a sudden change of RNA viruses, enabling two different Influenza virus subtypes/strains combined or amalgamated to become a new virus subtype. Mutation, a virus change process, occurs in a longer period of time. The recombination and mutation processes can happen within the 8 segments of Influenza virus A: HA (hemagglutinin); NA

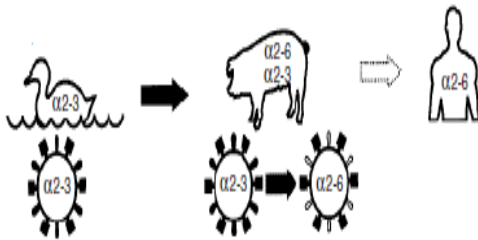
(neuraminidase); NP (nucleoprotein); M (matrix protein); PB1; PB2; PA (polymerase); and NS (non structural proteins) segments, with total 890 – 2341 nucleotide bases, can be represented in ‘G’ (guanine), ‘C’ (cytosine), ‘A’ (adenine) and ‘T’ (thymine) format. Both processes can occur in any virus segments: HA, NA, NP, PB1 and PB2. New virus strains formed from these processes then replace the previous virus strain, which cause the existing vaccine is no longer effective, hence new vaccine production is required.

The Influenza pandemic possibilities can be described as in Figure 1 [2], where process (A) and (C) are the recombination/reassortment process:

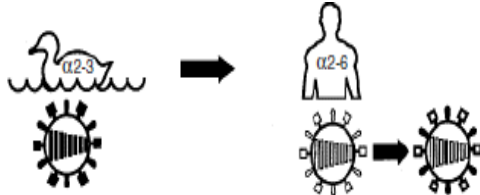
A. Recombination in Pigs



B. Mutation in Pigs



C. Recombination in Humans



D. Mutation in Humans

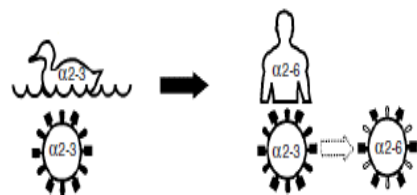


Figure 1. Influenza Pandemic Potential through Antigenic Shift and Drift
 (source: Avian Influenza: Basic Science, Potential for Mutation,
 Transmission, Illness Symptoms & Vaccines [2]).

(A) Reassortment/Recombination in Pigs: both Swine Influenza virus (which can live in mammals and human with $\alpha 2-6$ and $\alpha 2-3$ receptors) from human and Avian Influenza virus (which can live in avian with $\alpha 2-3$ receptors) from avian infecting pigs (with $\alpha 2-6$ and $\alpha 2-3$ receptors) and then recombine and become more infectious and infects more human population.

(C) Reassortment/Recombination in Humans: Avian Influenza virus (which can live in avian with $\alpha 2-3$ receptors and in some cases live in human with $\alpha 2-6$ receptors) somehow infects directly a human being (with $\alpha 2-6$ receptors) already infected by Swine Influenza virus, and these two viruses recombine to become a new virus, which is more infectious and infects more human population.

Another approach is described in [3], stated that pandemic process can be triggered by 3 (three) stages of virus changes which are:

- (I). Increment of virus' infectivity ability/level, from avian/animal to human (this process occurs in avian/animal body)
- (II). Increment of virus' virulence level from low pathogenic to high pathogenic (this process can occur in either avian/animal body or human body)
- (III). Increment of virus' contagiousness ability/level between human (this process occurs in human body)

Each of stage/sub-process above can be recombination/re-assortment (antigenic shift) or mutation (antigenic drift). Possible process flows as in Figure 2 are : (1) Sub-process (I) only; (2) Sub-process (I) followed by (II); (3) Sub-process (I) and (II) occur in parallel, followed by (III); (4) Sub-process (I), (II) and (III) occur in sequence; (5) Sub-process (I) followed by (II) and (III) in parallel; (6) Sub-process (I), (II) and (III) all in parallel.

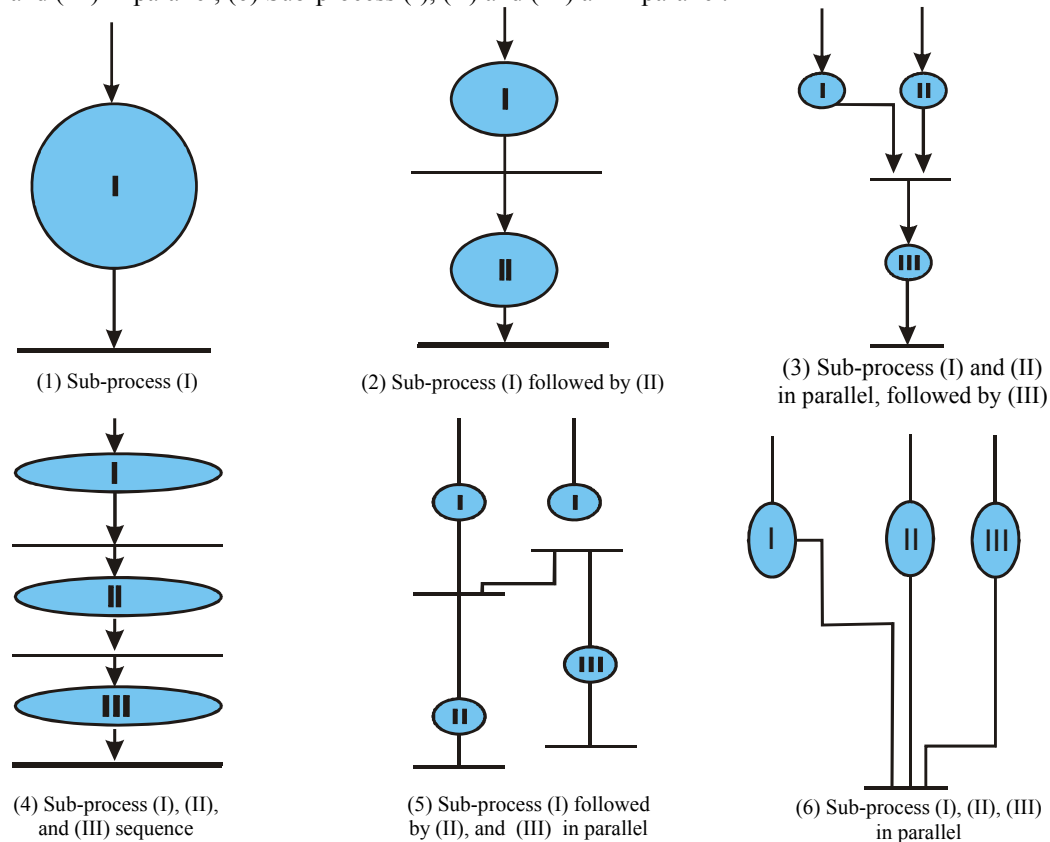


Figure 2. Several possible Influenza pandemic process-flows

For this research, the focus will be on Scenario (4), where sub-process (I), (II) and (III) occur sequentially.

Recombination is a process by which a molecule of DNA and RNA is broken and then joined to a different DNA molecule. Recombination can occur between similar molecules of DNA, namely homologous recombination, or dissimilar molecules of DNA as in non-homologous end joining. If a single host (human, chicken, or other animal) is infected by two different strains of the influenza virus, then it is possible that new assembled strain will be created from segments whose origin is mixed, some coming from one strain and some coming from another. The new recombined strain will share properties of both of its parental lineages. Recombination occurs when two DNA strands come together and exchange information. A recombination example is shown in Figure 3a [4]:

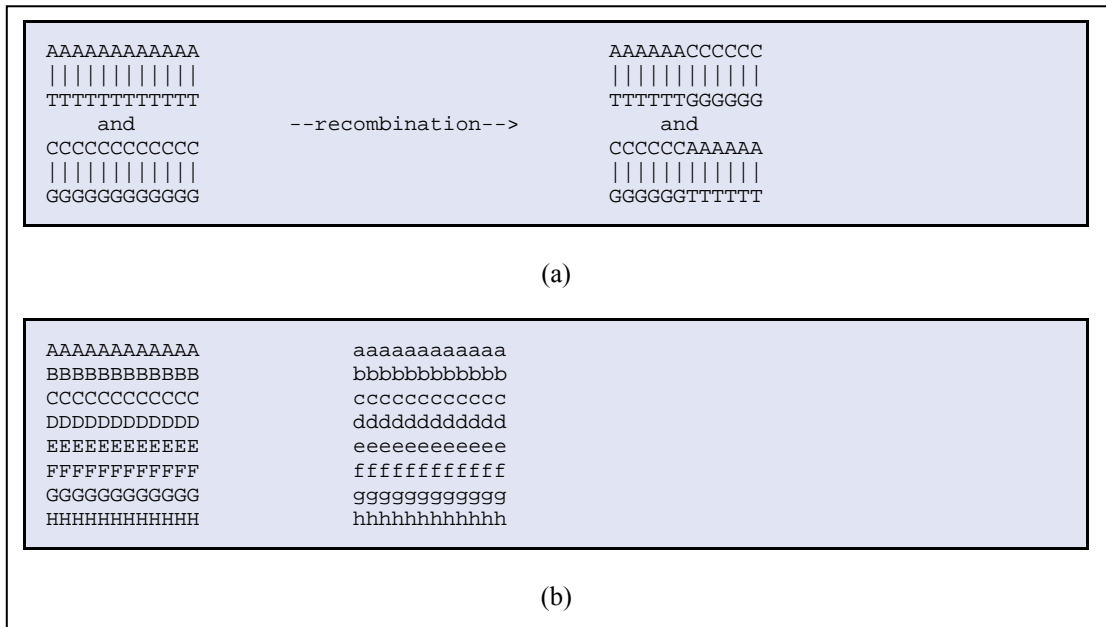


Figure 3. (a) Recombination and (b) Reassortment example

Recombination is different from Reassortment, which is swap of strands between two virus strains. The flu virus genome is comprised of eight different strands, as in Figure 3b:

Figure 3b represents the genomes of two different strains of flu, one in upper case and one in lower case. If a given person/animal is infected with both strains of flu, and more particularly if the same cell in the person/animal is infected with both strains then they can get confused and swap segments between each other, leading to the generation of re-assorted virus particles such as the ones shown below.

Recombination is responsible for some of the major genetic shifts in the history of the influenza virus. The 1957 and 1968 pandemic flu strains were triggered by recombination between an avian virus and a human virus, whereas the H1N1 virus responsible for the 2009 swine flu outbreak has an unusual mix of swine, avian and human influenza genetic. The scope of research in this article is restricted to the analysis of Recombination of Influenza virus, A type, H1N1 and H5N1 subtypes. This research discussion is focused on PB1 segment of these viruses, 1 of 8 segments undergone most frequent recombination.

On the other hand, the benefit of computation as supporting tools in Molecular Biology has been widely known and accepted. In Bioinformatics field, the genetic mutation analysis process and the phylogenetic tree building has been done by multiple alignment and tree

building software tools, for example ones provided by NCBI (National Center for Biotechnology Information) in its website <http://www.ncbi.nlm.nih.gov>. With computer tools, several Molecular Biology research sub processes can be simplified, so that the research process can be done more accurately, efficiently and effectively.

As described above, the Influenza pandemic, triggered by some possible circumstances, as combination of 3 sub-processes with 6 possible process flows, can be categorized as Multi-Objective Optimization Problem. Multi-objective optimization problem can be solved by Genetic Algorithm approach described in Section 3.1 in this article.

2. Related Works

Within the Influenza virus research frame, this research has relationship as preliminary and supporting research for Influenza vaccine researches. From interview with BIOFARMA (SITH ITB, 2009), information about Biofarma's plan to produce the H1N1 vaccine will start on mid of 2010 year, earlier than H5N1 vaccine production.. This is triggered by the fact that global influenza pandemic in Mexico which announced by WHO in April 2009, was started from H1N1, H5N1 and H3N2 viruses recombination process in Mexico.

In [5], the research of H5N1 virus and host interaction investigated the interaction between viruses and human immune system by utilizing the Agent-based-modeling methodology. It simulated 3 simulation components: (a) epithelium, cells which subject to infection by influenza type A (b) IS (immune system) cells and (c) influenza A viruses. It can also simulate both the antigenic shift and antigenic drift phenomena within the model, and it can somehow can integrate smoothly all these biological parameters within the simulation program. This research gave insight to include interaction aspect into this article's research. However, because it started from the human immune system definition, the focus was more on the immune system and not the dynamic changes of virus strains which could cause the Pandemic Flu.

The more recent research article [6] describes the important role of interaction between mutated viruses and their receptors. Previous Avian Influenza mutation researches described only about viruses mutations to increase binding ability from $\alpha 2-3$ receptors (in avian) to $\alpha 2-6$ receptors (in human) per se. This new study suggests that the difference in binding preference between human and avian flu viruses is more complicated than just $\alpha 2-6$ versus $\alpha 2-3$. Rather, it is an affinity for a particular topology, or shape, of $\alpha 2-6$ glycan receptor that characterizes human flu viruses. Specifically, the human flu viruses prefer long $\alpha 2-6$ receptors that occupy an umbrella-shaped space, as opposed to $\alpha 2-3$ and some $\alpha 2-6$ receptors that occupy a cone-shaped space. It also describes about a specific virus A/Texas/36/91's binding specificity to long (umbrella-shaped) $\alpha 2-6$ together with its mixed- $\alpha 2-6$ -and- $\alpha 2-3$ -binding ability would increase the efficient transmission of virus from avian to human.

This research's article was initiated by another research in [7], which investigates the Avian Influenza spread from the geographic/location perspective. Though this research does not include the geographical aspect, this aspect can be included for the future plan, by doing several strain virus changes simulation within a certain geographical scope. This research would observe the interconnection between those virus strains changes and then the chain of these changes is analyzed to identify the virus spread missing links and vectors. As the basis of this research, a preliminary research on H5N1 mutation simulation, another form of virus changes causing the Pandemic Influenza, has been conducted [8]. This preliminary research is then extended to this research which investigates the simulation of H5N1 and H1N1 recombination process.

This research's objective is to define model for virus recombination causing Influenza Pandemic phenomena. This research will define several different virus variants which can potentially cause the Influenza Pandemic. Additionally, this simulation research aims to obtain most possible virus strains formed from the recombination potentially trigger the Influenza Pandemic. New strains could be utilized to support the vaccine planning process which is

conducted by vaccine research and production organization such as Biofarma and Eijkman Institute.

3. Methods

A. Genetic Algorithm Principle

The Pandemic simulation program is developed based on Genetic Algorithm principle. In Genetic Algorithm, several aspects are defined:

- (1). Representation of the genetic solution domain, known as chromosome solution, which is the solution definition with several key parameters, represented by array of bits.
- (2). Fitness values/function to evaluate the solution domain. The fitness values are obtained from the optimization function.

As part of genetic algorithm, there are some processes such as selection, to select some chromosome with highest fitness values as parents; reproduction; crossover; and mutation of chromosomes chosen from the selection process.

In general, the genetic algorithm process flow is described in Figure 4 [9]:

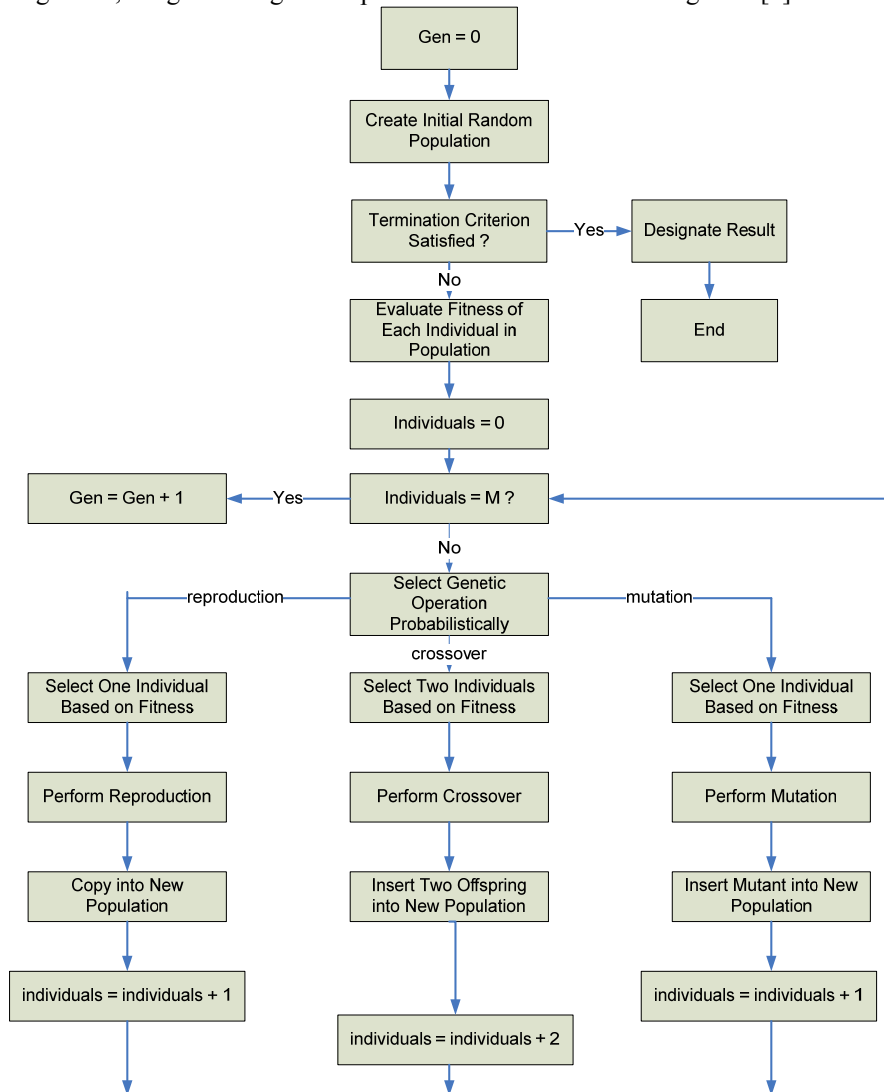


Figure 4. Genetic Algorithm Flowchart

As depicted in the Figure 4, the selection sub-process in chromosome population is conducted first, then the reproduction, crossover and mutation sub-processes are conducted in parallel. In this virus recombination research, the genetic algorithm sub-processes involved are mainly crossover, supported by reproduction and mutation sub-processes. These sub-processes are done iteratively until the termination criteria are achieved.

B. Pandemic Data and Process Modeling

B. 1. Pandemic Data Modeling

The virus data modeling is done by first querying the H1N1 and H5N1 virus strains from virus strains database (NCBI website). There is no H1N1 virus strain from Indonesia in NCBI database, therefore the query is modified to virus strains collected from China. The following is the H1N1 and H5N1 sequences query with several query parameters: only virus sequences located in China and occurred within the 2009 Pandemic years.

The PB1's H1N1 query result from 2009 Pandemic period is 32 nucleotide sequences and the PB1's H5N1 query result between 2007 to 2009 years is 33 nucleotide sequences.

These virus sequences are to be used for multiple alignment and recombination/mutation analysis tasks, to be described later on in Process Modeling section.

Chromosome solution is composed of several information :

1. Virus sequence: obtained from virus sequences database, in this case NCBI. There are 2274 bases, each base has 4 possible values ('G', 'C', 'A', 'T'). Sample of PB1 sequence is the A/Beijing/01/2009(H1N1)'s PB1 sequence information from base position 1-200:

```
atggatgca atccgactct acttttcta aaaattccag cgcaaatgc cataagcacc
acattccctt atactggaga tctccatac agccatggaa caggaacagg atacaccatg
gacacagtaa acagaacaca ccaatactca gaaaagggaa agtggacgac aaacacagag
actggtgcac cccagctcaa cccgattgat
```
2. Virus type: represents the Influenza virus type: H1N1, H3N2, H5N1, H5N2, etc.
3. Segment type: represent the virus strain types: HA (hemagglutinin); NA (neuraminidase); NP (nucleoprotein); M (matrix protein); PB1; PB2; PA (polymerase); and NS (non structural proteins) segments.

B.2. Pandemic Process Modeling

As described in Introduction section above, Pandemic process can be triggered by 3 stages of virus changes, with 6 possible flow possibilities as in Figure 2. These sub-processes are to be modeled as Multi-Objective Optimization Problem. Let's take a look at process flow (4) in Figure 2 (which will be the focus of this article), where each sequential sub-process is represented by an objection function: (I) Infectivity optimization function is represented by Objective Function I, (II) Virulence optimization function is represented by Objective Function II, and (III) Contagiousness optimization function is represented by Objective Function III.

This article describes only Objective function II, the Virulence optimization function because H1N1 virus has reached certain infectivity level to be able to infect human as well as animal, and the global concern is its potential in virulence level increment from low pathogenic to high pathogenic as well its contagiousness level increment. Therefore, the Objective function I, the Infectivity optimization function implemented for H5N1 is not described in this article since this subtopic already done in another research as part of researcher dissertation topic. In the following, the objective function II is described as Virulence Optimization function.

B. 2. 1. Virulence Optimization Function

This function is related to virulence optimization, aims to obtain virus strain with highest Virulence level possible. To evaluate the virulence and pathogenicity level, there are three parameters measured i.e. [10]: the morbidity (measured by weight loss), LD₅₀ titers (amount of virus within certain cell) and virus replication. However, to determine the virulence level, in

this article we consider only the morbidity parameter. This virulence optimization function is implemented as Genetic Algorithm's fitness function.

As in [10], the morbidity impact of H1N1 virus in infected mice is depicted in Figure 5:

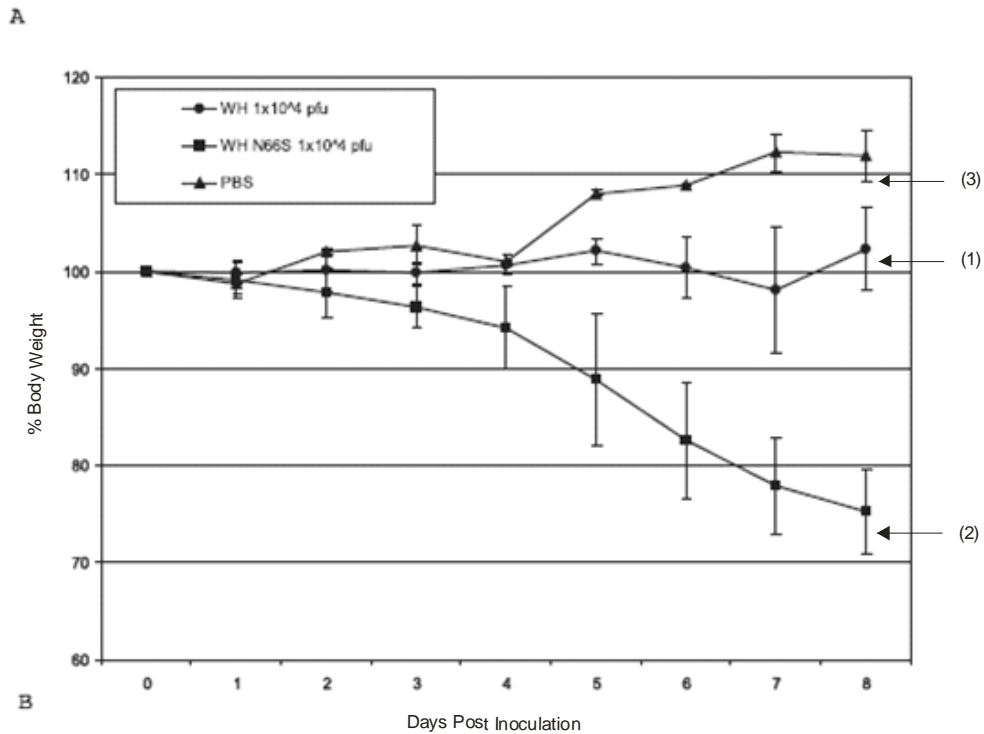


Figure 5. Contribution of PB1-F2 N66S Mutation to Pathogenicity of Recombinant Virus. The mice hosts were inoculated with 1×10^4 PFU of virus or PBS, and their weights were recorded every day after infection

From Figure 5 above, line (1) represents the morbidity impact of WH virus, a virus created in the A/WSN/33 background contained the A/HK/156/97 PB1 gene; line (2) represents the morbidity impact of WH virus with N66S mutation (mutation of N \rightarrow S in base position 66) and line (3) represents the morbidity impact of 1×10^4 PFU of virus (PBS).

Table 1. Numerical interpretation of morbidity impact in Figure 5

The following table depicts the numerical interpretation of line (1), (2) and (3) in Figure 5:

No	Virus type	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
(1)	H1N1 with 1×10^4 PFU	100 %	100%	100%	100%	101%	103%	100%	97%	103%
(2)	H1N1 with N66S mutation	100 %	98% (-2)	96% (-2)	94% (-2)	92% (-2)	88% (-4)	84% (-4)	79% (-5)	75% (-4)
(3)	H1N1 with PBS	100 %	no morbidity							

Table 1 shows that the N66S mutation caused gradual morbidity to H1N1 virus pathogen city, with certain pattern: (1) the weight value decrement for the first 4 days is 2% and (2) the weight value decrement for the second 4 days is 4%.

Virulence optimization function is depicted in Figure 6:

$$F_{\text{all}}(c) = \sum_{s=1} F_j(c)$$

$$F_j(c) = (F_0 - \sum_{s=1} r * \text{interval}) * \text{multiplier}(\text{mutation})$$

F_{all} : virulence optimization function which is the sum morbidity due to all mutations occurred in the virus strain/chromosome (c)
 $F_j(c)$: morbidity function to obtain weight value of infected host due to a certain mutation occurred in virus strain/chromosome (c)
 F_0 : weight value of Day 0 (100%)
r: the function sub-factor: $r = p * i$ for $s > 1$ and $p = 1$; $r = n \bmod (i + 1)$ otherwise
interval: the difference between morbidity values, in this case 2 (the first 4 days) and 4 (the second 4 days)
multiplier(mutation): multiplier factor of a specific mutation (from the following Table 2)
n: number of day/generation, in this case 8
s: number of iteration to obtain F_n , for $n = 1-4 \rightarrow s = 1$; $n = 5-8 \rightarrow s = 1$ and 2
i: n cycles: for $n = 1-4 \rightarrow i = 1$; $n = 5-8 \rightarrow i = 2$
p: number of day in each cycle, in this $p = 1$ to 4 for $i = 1$ and 2

Figure 6. Virulence optimization function implemented as Genetic Algorithm fitness function

The recombination triggers the change of virus strain bases, and the effect of these changes can be analyzed from the change of each individual base. Multiple-alignment is a process to evaluate/compare the changes of existing virus strains related to one consensus/reference virus. We utilized NCBI supporting tool to perform this process.

Table 2 summarizes the effect of changes of PB1 H1N1 virus segment because of recombination/mutation [10]:

Table 2. H1N1 statistic of base position and changes impact

No	Base position	Changes impact	Pandemic stage/sub-process	Multiplier factor to morbidity value
1.	196-198 (amino acid position: 66), mutation from N (AAT, AAC) → S (TCT, TCC, TCA, TCG, AGT, AGC)	Pathogenicity increment	II	1
2.	118-120 (amino acid position: 40), mutation from M (ATG) → I (ATT, ATC, ATA)	Small effect on pathogenicity increment	II	1.25
3.	172-174 (amino acid position: 58), mutation from T (ACT, ACC, ACA, ACG) → I (ATT, ATC, ATA)	Small effect on pathogenicity increment	II	1.25

The initial morbidity value is related to the initial weight value (= 100%, which is the initial fitness value (F_0) in Figure 6), and multiplier factor indicates the effect of such mutation to the weight value/virulence level (multiplier(mutation) in Figure 6).

As an illustration, the following is a chromosome part as input for virulence optimization function, which is the part of FJ789825 virus strain, base position 171-180 and 191-200 with a N \rightarrow S amino acid mutation (AAT or AAC \rightarrow TCT or TCC or TCA or TCG or AGT or AGC bases composition) in base position 196-198. 'G', 'C', 'A', 'T' are represented consecutively as '00', '01', '10', '11'. A specific N \rightarrow S mutation: AAC \rightarrow TCT is written in bold characters:

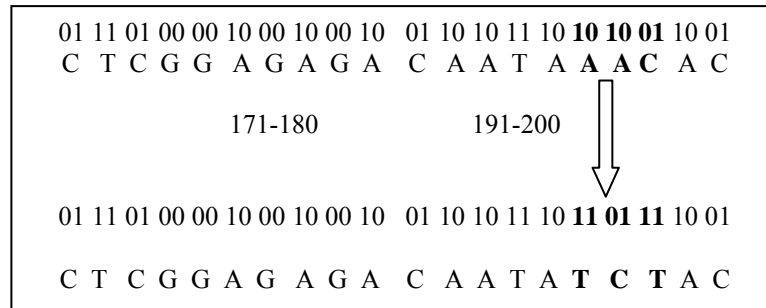


Figure 7. Example of virus strain bases as part of simulation's chromosome

The result of this chromosome input for the virulence optimization function is the virulence fitness value: 76%.

C. Validation Method

The simulation program would generate a number of possible recombined virus strains with their fitness values. The simulation should produce both the newly generated virus strains and their fitness values. Based on the highest fitness values, several virus strains are chosen and determined as virus strains which could potentially trigger the Influenza Pandemic. To validate these virus strains, the Bio-molecular experiment to generate newly mutated virus strains with a similar pattern with this computational (in-silico) process should be conducted, and then the virulence level of these virus strains are measured and validated against the virulence level obtained from this computational process.

4. Result and Discussion

The simulation prototype is developed to identify all potential virus strains with best fitness values, which indicate the highest virulence level and most potential to trigger the Pandemic influenza.

To develop the prototype, MATLAB[®] software with Generic Algorithm toolbox is utilized. The process is as follows:

- (1). To define chromosome: strain virus (2274 base x 2 bits) and its virus type (2 bits), however for this MATLAB[®] prototype, only 20 bases and their virus types are simulated. The segment type is assumed to be PB1 and not defined yet in this prototype.
- (2). To define the fitness function and fitness values, as described in the previous section.
- (3). To perform Genetic Algorithm simulation, to obtain all possible viruses strain, and choose those with highest fitness values.

The virulence optimization function is defined as MATLAB[®] function. The H5N1 and H1N1 virus strains defined are only the PB1 strain within base position 171-180 and 191-200. From Table 2 in section 3.3.2, only 2 changes/mutation defined in the prototype i.e. T \rightarrow I mutation base position 172-174 and N \rightarrow S mutation base position 196-198. The following

example is part of FJ789825 virus strain, base position 171-180 and 191-200 with N → S mutation in human host :

01110100001000100010 01101011101101111001 11

(A) Virus strain part

(B) Host part

These string inputs are then converted to corresponding decimal formats, as inputs to simulation prototype, the above example is converted to 1.995156680167000e+012.

The simulation prototype parameters are set by setting MATLAB's Genetic Algorithm Toolbox, as follows: (1) Fitness function: @ fluburung_rekombinasi; (2) Population size: 20; (3) Initial population: []; (4) Initial score: []; (5) Range: [0 ; 4e+012]; (6) Parameter 'Reproduction' → 'Crossover fraction' is set to 0.593103 to enable crossover occurs in a probabilistic way. The Mutation parameters are: Gaussian; Scale: 1.0; Shrink: 1.0.

The result of this initial prototype is 51 generation with 2 best fitness values with dynamic pattern as well as dynamic mean fitness values, which are: (1) 76% which is obtained from virus strain with N→ S mutation and (2) 95% which is obtained from virus strain with T → I mutation as depicted in Figure 8. This result has answered part of simulation development objective, but not yet give result of mutation combination, i.e. both N→ S and T → I mutations are found in one new virus strain.

5. Conclusions

From the analysis process, the data modeling for several Recombination parameters has been done, with result: the genetic chromosome solution. Additionally, from 3 (three)

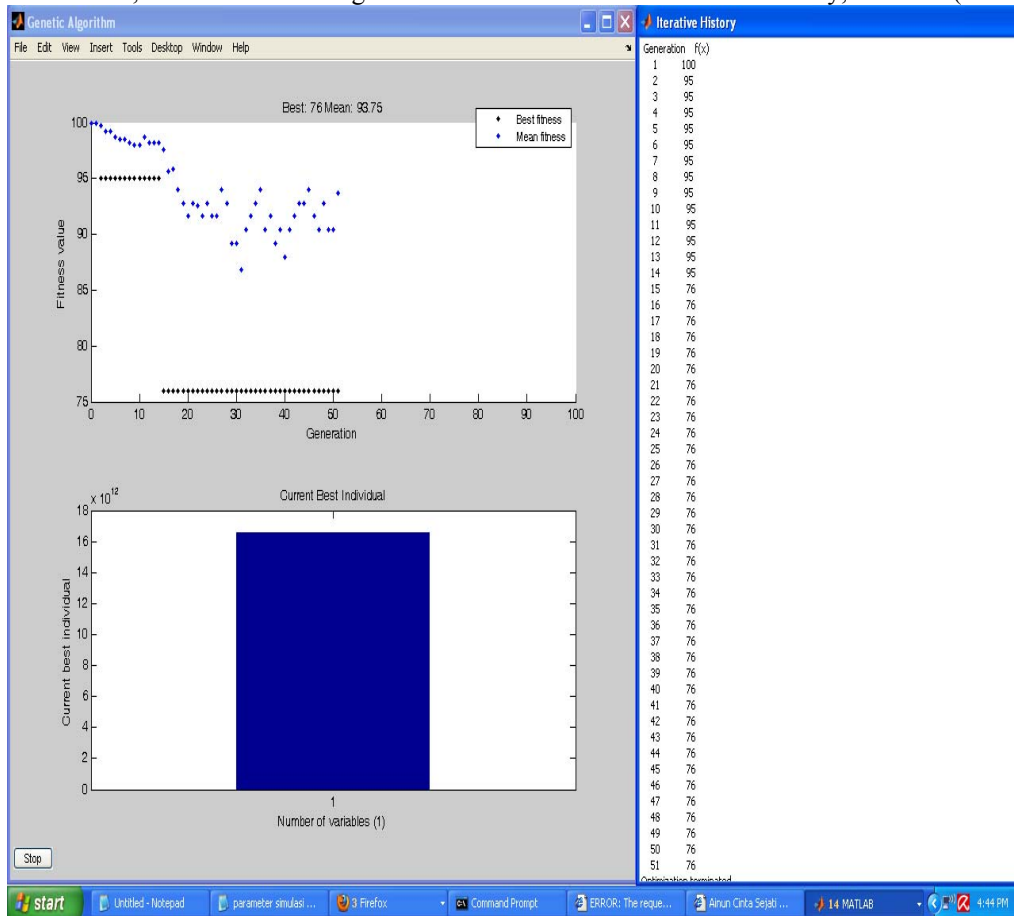


Figure 8. H5N1 and H1N1 recombination simulation result

sub-processes of Influenza pandemic, the process modeling of Virulence function has also been done, with result: objective function definition. With these two inputs, the initial prototype is being developed by utilizing MATLAB[®] software. As in Figure 8, the result of this initial prototype is best fitness values from 51 generations of the virus strains. This simulation result gives 2 best fitness values: 76% and 95% where 76% indicated more weight loss and higher virulence level; and an even more dynamic result for the mean fitness values. Better and more best fitness values could be obtained once the database of recombination and mutation virus strains (Table 2 of Section 3.3.2) is enhanced. Study about MATLAB[®] parameter setting should be done more thoroughly to obtain new virus strains with combination of both $N \rightarrow S$ and $T \rightarrow I$ mutations.

6. Future Work

For future work, we will also complete the infectivity optimization function for H1N1 and H5N1 recombination process and to integrate this virulence optimization function with the infectivity optimization function based on multi-objective optimization principle. We will also integrate the simulation of the H1N1 and H5N1 recombination with the simulation of H5N1 mutation process. To obtain more and better virus strains result, one of the important tasks is to enhance the database of viruses' recombination and mutation impact in Table 2 of this article.

7. Acknowledgements

This article is dedicated to Mrs. Ietje Suwenda Saleh (has passed away in February 8th, 2008), the Biology teacher in Public High School 3, Bandung. She has given a strong Genetic Biology foundation and continuous motivation to her student(s) in Biology learning process.

Sincere thanks and grateful to :

- Prof. Dr. Adang Suwandi Ahmad and Dr. Jorga Ibrahim from STEI ITB as well as Dr. Adi Pancoro from SITH ITB who has given continuous supervision for completing this article as part of PhD. Dissertation research.
- Dr. Bambang Riyanto Trilaksono from STEI who has kindly lend the MATLAB[®] software license to be utilized in this research.
- Dr. Arief Syaichu Rohman from STEI who has kindly give some assistance in understanding MATLAB[®].
- Mr. Deni, PhD Student from Mathematics Dept ITB, and Dr. Agus Yodi Gunawan, Lecture from Mathematics Dept ITB, who give valuable inputs in defining the Multi-objective functions and Genetic Algorithm aspects.
- Prof. Richard Belew, Researcher from University of California in San Diego (USCD), who gives valuable inputs and several papers related to this research.
- Miss Olive, Officer from Health Research and Development, Health Department Republic of Indonesia, who gives valuable input about the importance of H5N1 and H1N1 recombination process.

8. References

- [1] Arora, Dharam Jit S., Letter to the Editor : "Drift in the Nucleoprotein Gene of Swine Influenza virus (H1N1) Causing Respiratory Disease", *Virus Genes* 25:1, 109-111, 2002.
- [2] Chai, Louis Y.-A., "Avian Influenza: Basic Science, Potential for Mutation, Transmission, Illness Symptom. and Vaccines", 2005.
- [3] Shoham, Dany, "Review: Molecular evolution and the feasibility of an avian influenza virus becoming a pandemic strain—a conceptual shift", *Virus Genes* 33:127-132, 2006.
- [4] MedPedia, "Explanations of Mutation, Recombination, Reassortment, Recombination", <http://www.flutrackers.com/forum/showthread.php?t=111165>, June 18th, 2009.
- [5] Nuno Fachada, V. V. L., Agostinho Rosa, "Simulations of Antigenic Variability in Influenza A." *Nature Preceedings* hdl:10101/npre.2008.2303.1, 2008.

- [6] Aarthi Chandrasekaran, et.al., "Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin." *Nature Biotechnology* 26(1), (2007).
- [7] BIC (Business Innovation Center) and MenRisTek, "Track from End to End", 101 Inovasi Indonesia, 2009.
- [8] Moedomo, Ria, "Influenza Pandemic Simulation Program Based on Genetic Algorithm", *Proceeding of 2nd ICMNS, Faculty of Mathematics & Natural Sciences, School of Pharmacy, and School of Life Sciences & Technology*, Institut Teknologi Bandung, 2008.
- [9] Koza, John R, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA, The MIT Press. (Quoted from: The Genetic Program Tutorial website : <http://www.geneticprogramming.com/Tutorial/index.html>)
- [10] Conenello GM, et.al., "A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence", *PLoS Pathog*: 3(10):1414-21, 2007 Oct 5.



Ria Lestari Moedomo was born in Bandung, West Java, Indonesia on April 5, 1966. She graduated from her Bachelor degree program in Informatics Department ITB and Master program from Royal Melbourne Institute of Technology, Melbourne Australia. She is now a PhD student in School of Electrical Engineering and Informatics, Bandung Institute of Technology. Her research interests includes Software engineering, Research Methodology and Bioinformatics. She has won several Research grants and awards including: Wireless Reach Brew Application Funding Program (STEI ITB, Avian Influenza Research Team) in March 2007; Runner Up of Brew Developer Awards 2008 (STEI ITB, Avian Influenza Research Team) in May 2008; and 101 Inovasi Indonesia, Business Innovation Center dan Menristek in August 2009; Fulbright Doctoral Research Program in December 2008 – April 2009.



Adang Suwandi Ahmad was born in Lembang, West Java, Indonesia on November 3, 1948. He graduated from his Bachelor degree program in Electrical Engineering ITB; his Master and Doctoral program from USTL, Montpellier, France. He is now the Dean of School of Electrical Engineering and Informatics, Bandung Institute of Technology. His research interests includes Systems Cognitive Instrumentation. He is a member of several professional organizations: SPIE, IEEE Society of Devices & Aero-space Electronics and Societe Francaise DuVide. He has won several Education and Research awards including: The best Researcher, Research Grant Seminar Donation Research Compete VIII Directorate General Higher Education, National Award of DIKTI DEPDIKNAS in 2002; Satyalancana DWIDYA SISTHA, TNI, Defense Education in 2000; and Satyalancana Karya Satya XXX (30 years), ITB for Education Dedication in 2009.



Adi Pancoro was born in Makassar, South Sulawesi, Indonesia on 19 May 1959. He graduated from his Bachelor degree program in Biology Department ITB; and his PhD degree from Department of Genetics & Biochemistry, Newcastle upon Tyne, England. He is now a Senior Lecturer for Undergraduate and Master graduate programs in School of Life Science and Technology, ITB; Head of Biotechnology Research Center, PAU, ITB, Member of National Research Council- Indonesia (Dewan Riset Nasional) and Research Reviewer Team of Department of National Education. His research interests includes Biotechnology and Bioinformatics. He has won several Research grants and awards including: Competitive Research of ITB in 2010; Riset Unggulan Strategis Nasional-Dikti in 2009-2010; and Riset KKP3T – Departemen Pertanian in 2009-2010.



Jorga Ibrahim was born in Pangkal Pinang, Sumatera, Indonesia on March 10, 1936. He graduated from his Bachelor degree in Astronomy Department ITB; Master of Science degree program from Pure Mathematics in University of Kentucky; and “Doctorat d’Etat es-Science” degree from Universite Paris VI (Universite Pierre et Marie Curie). He is now a Lecturer and Supervisor of PhD students in School of Electrical Engineering and Informatics, ITB. His research interests includes Computer Science and Algorithms. He is a member of several professional organizations: Association des Doctorat d’Etat es-Sciences. International Astronomical Union; South East Asian Mathematical Society; Himpunan Astronomi Indonesia and Himpunan Fisika Indonesia. He has won several Education and Research awards including: “Officier de l’Ordre National de Merite” in 1982 and Achmad Bakrie Award as Best Scientist in 2007.



M. Sukrisno Mardiyanto was born in Cirebon, West Java, Indonesia on May 5, 1953. He graduated from his Bachelor Degree in Informatics from University of Montpellier II in France, 1981; his Master Degree in Informatics from Institut of Applied Mathematic - INPG Grenoble in France and his Doctoral Program from University Montpellier II – France in 1986. He is now an Associate Professor in Software Engineering in School of Electrical Engineering and Informatics, Bandung Institute of Technology. His research interests are in software maintenance and evolutionary software development. He is a member of advisory board of Indonesian National Information Technology and Communication Board (DeTIKNas).