IEEE Second International Conference on Data Stream Mining & Processing
August 21-25, 2018, Lviv, Ukraine

On Intelligent Multiagent Approach to Viral Hepatitis B Epidemic Processes Simulation

Dmytro Chumachenko
Informatics department
National Aerospace University "Kharkiv Aviation Institute"
Kharkiv, Ukraine
dichumachenko@gmail.com

***Abstract*—Simulation of epidemic processes is actual task, which has high social and economic value. Given research is dedicated to the application of the intellectual multiagent approach to the prediction of the incidence of viral hepatitis B. The structure of the model, agents, environment and rules of agents' interaction has been developed. The multiagent model shows high accuracy.**

Keywords—multiagent simulation, epidemic process, viral Hepatitis B, incidence prediction

I. INTRODUCTION

The modern works of many scientists are devoted to the development of intellectual problem-oriented systems and their application to population dynamics. Managing the epidemic process of population dynamics systems is an important part of this area. The most important tool for studying these systems is adequate mathematical models for predicting the spread of the dynamics of the epidemic process. To date, a significant number of such theoretically valid models of population dynamics have been created. They rely on the mathematical apparatus of statistics and probability theory. A common drawback of existing models is the low accuracy of forecasting, as well as its short-term.

One of the main problems of the epidemic process in the systems of population dynamics is the stochastic nature of its behavior. Also, existing models do not take into account the peculiarities of the internal behavior of the population, the assessment of the external environment by objects, the logical behavior of specimens of the population. In analyzing the epidemic process, in contrast to the behavior of population dynamics systems, the researcher is only interested in one "epidemic-recession" cycle, because the further development of dynamics can vary considerably depending on the consequences of epidemic behavior, as well as the external influences taken.

The analysis [1] of existing models and methods for predicting the dynamics of the epidemic process has shown that the most effective is the use of the multiagent approach, it allows to take into account the intellectual components of the simulated system of population dynamics.

Thus, **the aim of the research** is to develop a multiagent model of the epidemic process, which allows the construction of a prognostic morbidity for the selected population.

II. BACKGROUND ON MORBIDITY BY VIRAL HEPATITIS B

Hepatitis B is dangerous viral infection which has global spread. Annually in the world are nearly 50 million people who suffer only acute form of infection. Up to 600 thousand patients with hepatitis B is dying [2]. The prevalence of this virus varies widely in different parts of the world [3]. Ukraine is a country with a middle prevalence.

Hepatitis B can be transmitted only from people infected with the virus, including those with latent infection [4]. The virus can be transmitted through blood transfusions, in violation of the integrity of the skin and mucous membranes (tattooing, acupuncture, use of shared toothbrushes), through sexual contact. In recent years, the increasing significance of sexual transmission of infection. Also, there is a great risk of transmission during pregnancy. Risk groups include people who inject drugs, having disorderly sexual contacts, hemodialysis patients, medical staff, family members of carriers of the virus, patients with chronic diseases of the skin. The virus has an incubation period with a wide range (from 15 to 180 days). Risk of chronic acute hepatitis B inversely proportional to the age at the time of infection: among adults with a normal immune system is not more than 5%, at the age between 1 and 5 years – 30%, for newborns – 90%. For 25% of patients unable to identify the source of infection [5]. Thus, at the present time it is one of the most widespread and dangerous viral infections that cause anxiety for the health of the population and reducing the average life expectancy of people all over the world. Hepatitis B has an exceptionally high infectivity and the frequent occurrence of severe consequences (including death). Also considering that 65–80% of those who infected with hepatitis B virus disease have no external clinical manifestations, the problem remains relevant.

III. DEVELOPMENT OF STRUCTURE OF SIMULATED SYSTEM

The epidemiological model of Viral Hepatitis B is based on the concept of the epidemic process by Gromashevsky [6], according to which the epidemic process exists with the continuous interaction of the three main components - the source of infection, the mechanism of transmission and the susceptible organism.

A. Structure of agents

The most profitable type of agent in the study of epidemic processes is an emotionally-motivated intellectual agent, for the most complete and accurate model of human behavior. Let's consider the agent as a set of properties:

$$a = \langle s, s_t, c, t_a, l \rangle, \quad a \in A, s \in S, c \in C, t_a \in T_a, \qquad (1)$$

where s_t is time in state s, A is set of all agents, S is set of different agent's states, C is set of working area's cells, T_a is set of possible agent's types, l is length of life.

The set of agent states is predefined and is constant. Depending on the process being studied, the set can be supplemented by different states, the initial set is:

$$S=\{Susceptible, Exposed, Infected, Convalescent, Recovered, Dead\}. \qquad (2)$$

The use of such a set of states is based on the idea of distributing the entire population to subsets, based on their states by epidemic features (the classical model of the SIR type [7]). The proposed set characterizes the model as an analogue of the extended model of the SEIRS type [8].

Fig. 1 shows the transitions between agent states:

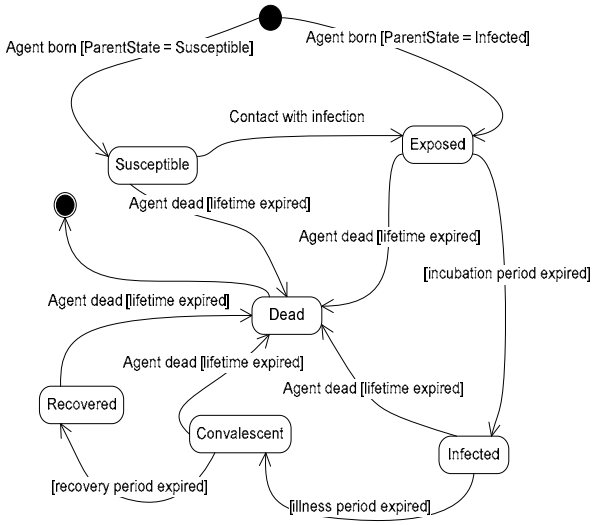

Fig. 1. Agents' states changing

Susceptible – the agent is healthy (may be prone to infection). In this case, a healthy means an agent that is susceptible to the disease of the epidemic process, is modeled.

Exposed – the agent got the disease. This state is an incubation period. During this time, the agent is not yet sick, but already has a chance to transmit the infection.

Infected – agent of the patient. Agents in this condition are the most likely sources of infection for other agents.

Convalescent – the agent recovers. This condition characterizes the period when the clinical symptoms of the disease disappeared, but the agent can still be a carrier of the disease and a source of infection. The presence of this condition is characteristic of certain infectious diseases.

Recovered – the agent recovered (acquired immunity). Agents in this state are no longer subject to the possibility of getting sick.

Dead - the agent is dead from old age or due to illness.

For agents, two types were identified, which can be provisionally called "prudent" and "risky." The characteristics defined for each type are shown in Table I.

TABLE I. DESCRIPTION OF AGENTS' TYPES

Event	Prudent	Risky
Probability of born	***80,0 %***	***20,0 %***
Probability of transition to risk zone in susceptible state	***1,0 %***	***10,0 %***
Probability of transition to hospital in susceptible state	***0,7 %***	***0,1 %***
Probability of transition to risk zone in infected state	***0,5 %***	***5,0 %***
Probability of transition to hospital in infected zone	***80,0 %***	***25,0 %***
Length of staying in home zone	***20 h***	***15 h***
Length of staying in risk zone	***2 h***	***8 h***
Length of staying in Hospital	***2 h***	***1 h***

Taking into account the modeling features of the epidemic process described above, an internal structure of agents has been developed, which includes the following fields (Fig. 2):

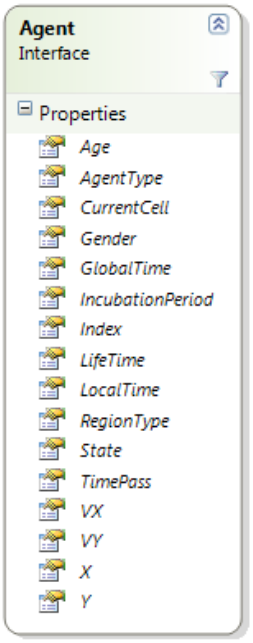

Fig. 2. Internal structure of agent

1. *The index number (Index).* It is intended for the exact determination of the agent among others.

2. *Agent's local time (LocalTime).* It is designed to regulate the order of activity of agents, as well as to determine the need for agent processing by the system.

3. *Global Agent Time (GlobalTime).* Contains the time of the model world, relevant for the agent.

4. *Coordinates of the agent (X, Y).* Designed to determine the physical location in the model world.

5. *Direction of movement (VX, VY).* Contains information about the physical direction of the agent in the modeling world.

6. *State of the agent (State).* Contains the state in which the agent is currently located.

7. *Current cell (CurrentCell).* Contains characteristics inherent in the current location of the agent.

8. *Current Region (RegionType).* Lets you get information about the area in which the agent is currently located.

9. *Time spent in the current area (TimePass).* Contains the moment when the agent has stayed in the current area. At this point in time, the agent moves to another area.

10. *Agent type (AgentType).*

11. *Age of the agent (Age).* It is a value that reflects the model age of the agent.

12. *Sex of the agent (Gender).*

13. *Incubation Period (IncubationPeriod).* Contains the end of the incubation period. Upon reaching this point in time, the agent becomes ill.

14. *Lifetime (LifeTime).* Contains the amount of model time assigned to the agent for life. For each agent this value is different. It is also assumed that this value may change under certain conditions (for example, after the transferred illness).

B. Environment

The composition of the workspace leads to the appearance of a set of cells, as conditional abstract objects. It is assumed that one cell can simultaneously include a number of agents, as well as one object-vector of infection (let's call it an instrument). Therefore, the cell can be described as follows:

$$c = \langle z, \tilde{A}, i \rangle, z \in Z, \tilde{A} \subset A, i \in I, \quad (3)$$

where I is set of all instruments,

Z is set of all working areas.

The working area can be described as follows:

$$z = \langle \tilde{C}, t_z \rangle, \tilde{C} \subset C, t_z \in T_z, \quad (4)$$

where T_Z is set of possible areas types. It is assumed that, depending on the type of area in which the agent is located, the specific nature of the epidemic process changes.

There were 3 working areas defined for experiments:

- *Home area*. In this area only contacts between agents are allowed.
- *Risk zone.* In this area, in addition to contacts between agents, infection from an infected tool is possible.
- *Hospital.* In this area, partial contact between agents and infection from the instrument is possible. When ingested, the agent is treatable, which reduces the duration of the disease.

Under the tool in this model, there are various objects that can somehow transfer the infection from one person to another (syringes, scissors, etc.). They can be subjected to various degrees of decontamination. In order to avoid excessive complication of the model, the features described above are reduced to the tool life expectancy. Thus, the tool can be described as a set:

$$i = \langle s, c, l \rangle \quad (5)$$

Thus, the simulation model can be represented as a function:

$$sim(l_{mean}, |A|_0, \overline{P}, T), \quad (6)$$

where l_{mean} is average life duration,

$|A|_0$ is the initial power of set of agents,

$\overline{P}$ is the vector of probabilities that are taken into account in the simulation model.

C. Interaction of agents

At its core, intelligent multiagent technologies contain elements of a discrete-event approach [9]. This is manifested in the fact that the system has a timeline for the simulation process. The peculiarity is that on this scale, according to certain rules, on the basis of the general situation in the system and the individual characteristics of individuals, there are events of agents that occur and are processed by the system upon reaching the necessary time moment. Among the events are the events-the intersection of the boundaries of cells forming a stream of events of the first kind. Events of this type are represented by an increasing sequence of instants of time, processed as a transition from one cell to another. Also in the system there are events of interaction with other agents, pulling the branching of the results of the event depending on the individual properties of agents interacting with each other. This creates a stream of events of the second kind. The processing of both types of events and the generation of the following for each agent is a difficult task that involves some technical problems, from the decision of which the adequacy and expediency of using the model directly depends.

If we try to consider the interaction of agents as direct physical contact, then when using the event approach, the agent interactions will be events of the second kind [10]. Adding to the second-order event processing system results in a significant reduction in the simulation speed [11]. A similar situation would be acceptable in the case of modeling physical processes. When modeling a phenomenon such as the epidemic process, it is obvious that participants in the process can interact with each other not only through direct physical contact (for example, airborne diseases) [12]. The processing of such events is rather complicated, which leads to a slowing down of the modeling process [13]. In this paper, it is proposed to simplify the consideration of the moment of infection by establishing the belonging of agents to one cell. This allows you to take into account the possibility of interaction of agents and at the same time significantly reduce the loss of productivity.

The program structure of the cell provides a field containing a list of all agents currently in the cell. To simplify the modeling process, the hit in a single cell of

multiple agents is considered to be their interaction. Interaction of agents is processed by the modeling environment, depending on the area to which the cell belongs, as well as the types of agents interacting.

The contact transmission of the disease from agent to agent is realized as follows. To begin with, the possibility of contact infection is checked. It is believed that this is possible if other agents are also processed in the same cell along with the current agent, processed, as well. Each pair of agents (formed from the current agent, is processed with other agents in the current cell) is compared. If at least one of the agents is the carrier of the disease, it is believed that with a certain probability, there is a sufficient contact between the agents for the infection and a healthy agent becomes infected.

Interaction between agents is handled by the modeling environment, depending on the disease, and is modeled. When considering certain epidemic processes, the logic of handling the interaction of agents can be complicated.

In order to increase the population's detail in terms of its membership in conditional social groups, the types of agents were identified, which are presented as follows:

$$t = \left\langle \overline{P}_t, \overline{\tau}_t \right\rangle, \tag{7}$$

where $\overline{P}_t$ is the vector of probabilities that are characteristic for the type of agent, $\overline{\tau}_t$ is the vector of constants of simulation time.

The constructed model admits an extension associated with an increase in the dimensionality of the vectors. In the framework of this task we confine ourselves to the following set:

$$\overline{P}_t = \left(P_b, P_{hr}, P_{hh}, P_{sr}, P_{sh}\right), \tag{8}$$

where P_b is probability of agents with a given type born, P_{hr} is probability of transition to risk zone of agent in susceptible state, P_{hh} is probability of transition to Hospital of agent in susceptible state, P_{sr} probability of transition to risk zone of agent in infected state, P_{sh} probability of transition to Hospital of agent in infected state.

$$\overline{\tau}_t = \left(\tau_h, \tau_r, \tau_m\right), \tag{9}$$

where τ_h is the duration of the agent with current type spending in home area, τ_r is the duration of the agent with current type spending in risk zone, τ_m is the duration of the agent with current type spending in Hospital.

For interaction between agents, the probability vector $\overline{P}$:

$$\overline{P} = \left(P_i, P_r, P_d, P_s, P_a\right), \tag{10}$$

where P_i is probability of transmission of infection from a sick agent or using an infected tool, P_r is the probability of transmission of infection from the agent is at the stage of the incubation period or at the stage of the decline of the disease (let's call it "reduced probability of infection"), P_d is the probability that the agent in the hospital will be diagnosed with a disease that is at the stage of the incubation period, P_s is the probability that in the interaction of the two agents will come into contact, P_a is the probability of accidental transmission of infection to a healthy agent from a carrier agent.

In addition, each type causes the agent to have a length of stay in each of the areas. At the end of the stay in a particular area, the agent passes to a different area according to a given probability. The transition is based on the state (CurrentState) and in which area (CurrentRegion) the agent was located, as well as under the influence of the transition probabilities determined by its type.

IV. Experimantal investigation of intelligent multiagent approach to simulation of viral hepatitis B epidemic process

It is natural to check the model on real statistics about the epidemic process. For the experiments, the incidence of viral hepatitis B and its statistical data for the Kharkiv region from 1994 to 2016 were chosen. Data for conducting experimental studies are provided by the State Institution "Kharkiv Regional Laboratory Center of the Ministry of Health of Ukraine", which monitor the epidemic process and develop and implement preventive and anti-epidemic measures to reduce the incidence rate. The statistical data of the epidemic process of viral hepatitis B included the monthly incidence in the Kharkiv region from 1994 to 2016. The incidence from 1994 to 2009 was divided into the following age groups: children under 2 years, children from 3 to 6 years, children from 7 to 14 years, adolescents from 14 to 18 years, adults from 19 to 30 years, adults from 31 to 40 years , adults from 41 to 50 years, adults from 51 to 60 years, adults over 61 years. Also from 1994 to 2009, the data included the expected pathways of infection: natural, sexual, domestic, vertical, through medical interventions, through drug injections, through manipulations outside treatment and prevention facilities. In turn, data on the parenteral route of infection through medical interventions are divided into injections, blood transfusions, surgical manipulations, gynecological manipulations, blood sampling for laboratory analysis and dental care. On the other hand, all the data were divided into intensive indicators, that is, the real number of patients, and absolute indicators, that is, the incidence of 100,000 people. From 2009 to 2016, statistical data were provided without dividing by age groups and the expected pathways of infection. This is due to changes in the standards of the Ministry of Health of Ukraine on the collection of data on morbidity. Taking into account the structure of the obtained data and the specificity of the spread of the epidemic process, the developed intellectual multiagent model was tuned. Age and social groups were included, it was added that the agents are adults and the interaction does not occur in the "Hospital" area, because the transmission between people is most often performed sexually. To model the epidemic process of viral hepatitis B, the following structure of the population of agents was chosen, the transitions between states are shown in Fig. 3.

The accuracy of the forecast was compared with the statistics on the incidence of viral hepatitis B in the Kharkov region in absolute (Fig. 4) and intensive (Fig. 5) indicators. The abscissa is the year, the ordinate is the indicator. The red

solid line – statistical data, blue intermittent – is calculated by the modeled forecast.

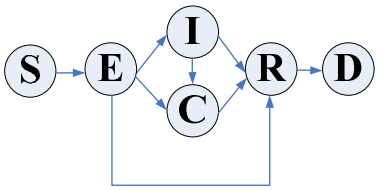

Fig. 3. Transition between states in simulation of Viral Hepatitis B epidemic process

Fig. 4. Comparison of absolute indicators for the incidence of viral hepatitis B in the Kharkov region and the constructed forecast.

Fig. 5. Comparison of intesive values for the incidence of viral hepatitis B in the Kharkov region and the constructed forecast.

The accuracy of the predicted intensive incidence of viral hepatitis B, calculated using the developed intellectual multiagent model, is shown in Table II.

TABLE II. THE ACCURACY OF THE DEVELOPED MULTIAGENT MODEL

Average absolute error	***2,94 %***
Root-mean-square error	***0,017***
Average error	***-0,54***
Standard error	***0,12***
Mean deviation	***1,95 %***

Thus, the accuracy of the constructed forecast is 97.06%.

CONCLUSIONS

A universal intellectual multiagent model of the dynamics of the spread of epidemic processes in population dynamics systems has been developed. The model makes it possible to take into account the types and structure of the population, the features of the spread of the epidemic process are modeled. The described generalized model realizes all the features of the multi-agent approach and is the most universal and susceptible to the type of epidemic process. The advantage of this model construction is the ease in improving and complicating the structure of the simulated system. Without changing already defined modeling patterns, you can add new control parameters, significantly complicate and expand the scope of the subject, and move to higher or lower levels of abstraction. Experimental studies of the universal intellectual multiagent system of epidemic processes have been carried out. The model setting is shown on the example of the incidence of viral hepatitis B, which takes into account the specificity of the epidemic process of viral hepatitis B, the division into age and social groups, the pathways of infection. Calculated with the help of the constructed model, the prognosis was checked on the statistical data on the incidence of viral hepatitis B in the Kharkov region. The accuracy of the forecast is 97.06%.

REFERENCES

[1] D. Chumachenko, K. Bazilevych and E. Meniailov, “Computer simulation of population dynamics epidemic processes,” in Mathematical modeling, optimization, information technologies in technical and social-economic systems, 2018, (in press).

[2] C. Trepo, “Hepatitis B virus infection,” in The Lancet, vol. 384, iss. 9959, pp. 2053–2063, 2014.

[3] L. Rivino, et al, “Hepatitis B virus-specific T cells associate with viral control upon nucleos(t)ide-analogue therapy discontinuation,” Journal of Clinical Investigation, vol. 128(2), pp. 668–681, 2018.

[4] C. Pramoolsinsup, “Management of viral hepatitis B,” Journal of Gastroenterology and Hepatology, vol. 17, iss. S1, pp. 125–145, 2002.

[5] J. MacLachlan and B. Cowie, “Hepatitis B Virus Epidemiology,” in Cold Spring Harbor Perspectives in Medicine, vol. 5, pp. 1–13, 2015.

[6] L. Gromashevsky, General epidemiology, 4th ed., Moscow: Medicine, 1965.

[7] F. Debarre, “SIR models of epidemics” in Modelling course in population and evolutionary biology, Zurich: Institute of Integrative Biology, 2010.

[8] L. Zhang, L. Yingqiu, R. Qingqing and H. Zhenxiang, “Global dynamics of an SEIRS epidemic model with constant immigration and immunity,” in WSEAS transactions on mathematics, vol. 12, iss. 5, pp. 630–640, 2013.

[9] D. Chumachenko, and S. Yakovlev, “Investigation of agent-based simulation of malisious software,” in Econtechmod: an international quarterly journal, vol. 5, iss. 4, Poland: Lublin-Rzeszow, pp. 61–67, 2016.

[10] D. Chumachenko, and T. Chumachenko, “Agent-based model of the epidemic process of diseases with multiple routes of infection transmission development and evaluation,” International Journal of Research Studies in Computer Science and Engineering, vol. 3, iss. 5, India: ARC publications private limited, pp. 61–67, 2016.

[11] A. Chopra, and P. Munindar, “Agent communicatio,” in Multiagent systems, MIT Press, pp. 9–33, 2011.

[12] D. Chumachenko, V. Dobriak, M. Mazorchuk, I. Meniailov, and K. Bazilevych, “On agent-based approach to influenza and acute respiratory virus infection simulation,” 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, pp. 184188, February 20 – 24, 2018.

[13] S. Russel, P. Norwig, Artificial intelligence. A modern approach. New Jersey: Prentice-Hall, 2003.