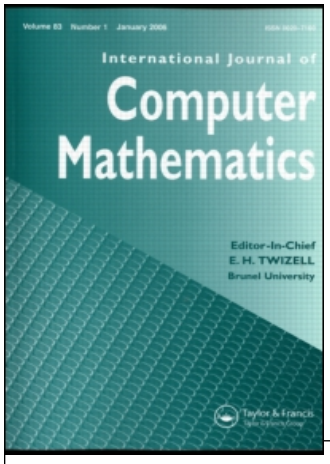

International Journal of Computer Mathematics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713455451

Stochastic simulation of HIV population dynamics through complex network modelling

P. M. A. Sloot [a]; S. V. Ivanov [b]; A. V. Boukhanovsky [b]; D. A. M. C. Van De Vijver [c]; C. A. B. Boucher [c]
[a] Section Computational Science, University of Amsterdam, The Netherlands
[b] St. Petersburg State University, St. Petersburg, Russian Federation
[c] Department of Virology, University Medical Centre, Utrecht, The Netherlands

Online Publication Date: 01 August 2008

To cite this Article: Sloot, P. M. A., Ivanov, S. V., Boukhanovsky, A. V., Van De Vijver, D. A. M. C. and Boucher, C. A. B. (2008) 'Stochastic simulation of HIV population dynamics through complex network modelling', International Journal of Computer Mathematics, 85:8, 1175 — 1187

To link to this article: DOI: 10.1080/00207160701750583
URL: http://dx.doi.org/10.1080/00207160701750583

International Journal of Computer Mathematics
Vol. 85, No. 8, August 2008, 1175–1187

Taylor & Francis
Taylor & Francis Group

Stochastic simulation of HIV population dynamics through complex network modelling

P.M.A. Sloot[a]*, S.V. Ivanov[b], A.V. Boukhanovsky[b], D.A.M.C. van de Vijver[c] and C.A.B. Boucher[c]

[a]*Section Computational Science, University of Amsterdam, The Netherlands;* [b]*St. Petersburg State University, St. Petersburg, Russian Federation;* [c]*Department of Virology, University Medical Centre, Utrecht, The Netherlands*

(*Received 04 July 2007; revised version received 29 September 2007; accepted 15 October 2007*)

We propose a new way to model HIV infection spreading through the use of dynamic complex networks. The heterogeneous population of HIV exposure groups is described through a unique network degree probability distribution. The time evolution of the network nodes is modelled by a Markov process and gives insight in HIV disease progression. The results are validated against historical data of AIDS cases in the USA as recorded by the Center of Disease Control. We find a remarkably good correspondence between the number of simulated and registered HIV cases, indicating that our approach to modelling the dynamics of HIV spreading through a sexual network is a valid approach that opens up completely new ways of reasoning about various medication scenarios.

Keywords: complex networks; HIV; infection spreading; computer simulation; complex systems

1. Introduction

Despite the availability of a large number of mathematical models describing the spreading of HIV, a good understanding of the spreading dynamics through numerical analyses is still a major challenge. It is essential to combine epidemiological processes with sociological models and network sciences. The true incidence of the HIV/AIDS-epidemic is quite uncertain since many people may be unaware of their infection; moreover HIV has a very long asymptotic incubation period which makes study of the actual infection spreading a very complicated task. The various routes of infection and the inhomogeneity of the involved population pose additional problems [27].

Many mathematical models have been suggested to simulate HIV population dynamics. For instance, statistical techniques like the back-calculation method and its related modifications are widely used to estimate the incidence and short-term projection of HIV. Generally, these methods are based on information from annual AIDS cases and incubation periods of the disease [1,2,33]. Popular epidemiological models like SIR models are often used to simulate HIV spreading. New ways to account for homogeneous mixing [8,17,21] and the impact of, for instance, demographic

*Corresponding author. Email: P.M.A.Sloot@uva.nl

ISSN 0020-7160 print/ISSN 1029-0265 online

DOI: 10.1080/00207160701750583
http://www.informaworld.com

effects or drug resistance have been discussed [1,8,16,32]. An SIR model is an epidemiological model that computes the theoretical number of people infected with a contagious illness in a closed population over time. The name of this class of models derives from the fact that they involve coupled differential equations relating the number of susceptible people 'S', number of people infected 'I', and number of people who have recovered 'R'.

One of the key drawbacks of those models is the difficulty to represent population heterogeneity and related risk behaviour. Heterogeneity in these models can only be implemented through different transmission probabilities or different number of contacts for people in risk groups [3,24–26]. An advantage of network-based epidemical models is the natural way in which heterogeneity can be expressed through network degree probability distributions and assortative or disassortative mixing [4,6,13]. The goal of the work presented here is to show that complex network (CN) based modelling techniques provide a universal and natural way to describe any kind of infection spreading and specifically HIV.

The paper describes the infection spreading as a stochastic process taking place over a sociological sexual network. Key ideas like network structure, network generation, and node dynamics are discussed and applied. The models are validated against real historical data.

2. Methods

2.1. *Formal description of the model*

Let us consider a CN-model as a set of the pairs $\{G, \Gamma\}$, where G is a graph, that is, an ordered pair of disjoint sets (V, E) (vortices and edges), and Γ is an evolutionary operator, governing network changes in discrete time steps t:

$$\begin{aligned} \langle V, E\rangle_{t+1} &= \Gamma\langle V, E\rangle_t; \\ \langle V, E\rangle_{t=0} &\stackrel{\text{def}}{=} \langle V_0, E_0\rangle. \end{aligned} \tag{1}$$

The evolutionary operator (1) can be represented as a composition of distinct operators $\Gamma = \otimes_k \Gamma_k$ corresponding to different dynamical aspects, $\otimes$ indicates a 'direct' non-commutative product of the independent operators.

Figure 1 illustrates the effect of the evolution operators Γ on a set of three basic models of epidemiological networks, each of them corresponds to different mechanisms of infection spreading. The simplest one is infection spreading on a static network, *i.e.*, the network structure does not change in time (Figure 1a). At each step an infected node (marked with a cross) can infect its neighbouring nodes with a given probability. An important issue is the impossibility of changing the links. This model of infection spreading corresponds to the so-called 'nodes percolation problem' for the static networks. This model is feasible for infection spreading of geographically segregated agents; this process is represented by the *network percolation operator* Γ_1. A more elaborate model takes into account the network dynamics (Figure 1b). This reflects the fact that network links are not stable and the network structure is evolving over time. The model is fully connected; even in the case of a small number of links (for instance, only one for each node) the possibility to be connected to an infected node cannot be disregarded. This advanced feature makes the model more applicable to infection spreading inside a large society (city or country) and for sexually transmittable diseases (STD), where almost all the people can be infected. This process is controlled by the *network link-dynamics* operator Γ_2. Figure 1(c) represents yet another kind of process occurring on network epidemics models; some epidemics are able to last over a long period of time and people that have been infected at the beginning of the epidemics may be removed from the initial population over

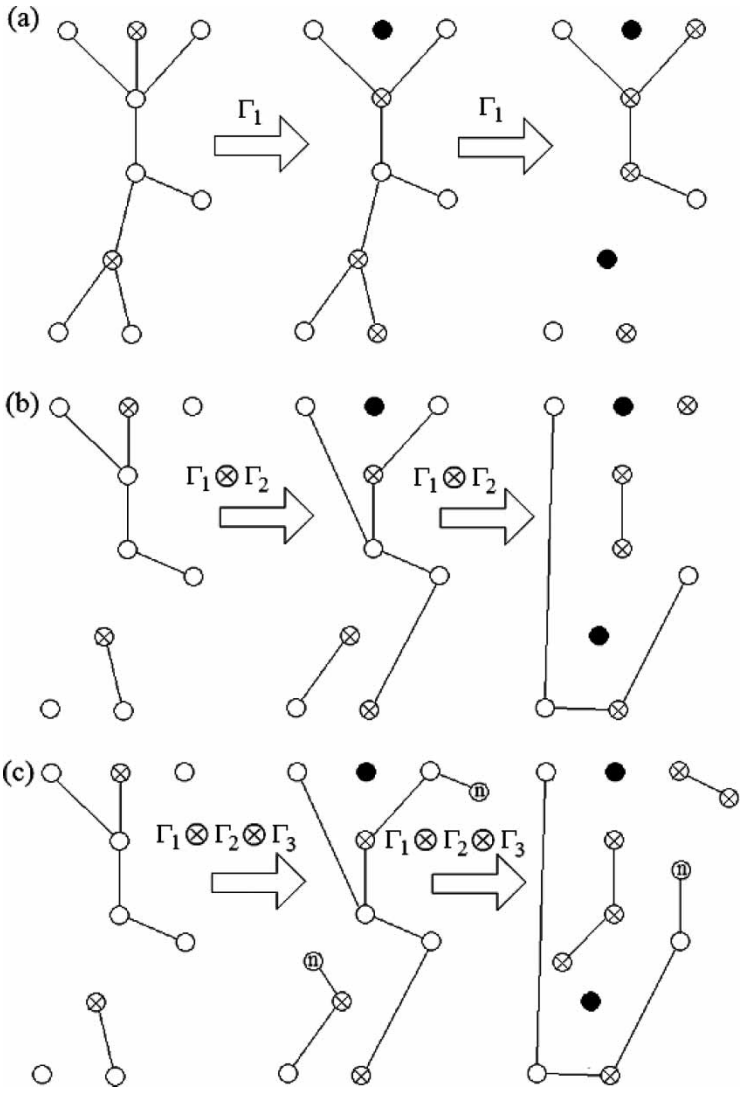

Figure 1. Schematic models of infection spreading through CNs: (a) static graph; (b) dynamic links; (c) dynamic nodes. The circles with crosses indicate infected individuals, the black circles are individuals removed from the network see the text for more details.

time (due to death, age, migration, etc.). This dynamics is captured by the *network demographic* operator Γ_3.

Next we define an operator for HIV spreading as $\Gamma = \Gamma_1 \otimes \Gamma_2 \otimes \Gamma_3$. Note that those operators are non-commutative. In addition, a realistic description of any epidemiological process implies a correct description of the state evolution of the nodes. In our case, one can formalize the time evolution as an additional operator Γ_4, which is defined in the form of changes in the CD4 T-cell count in the blood of infected individuals (Figure 2). The rationale here is that the white blood cells (T-cell lymphocytes) that express CD4 on their membrane are an indication of the health of the individual. If the number of CD4 T-cells drop below $\sim$200 copies per ml the patient is likely to enter the AIDS phase where opportunistic diseases like tuberculoses and cancer will emerge.

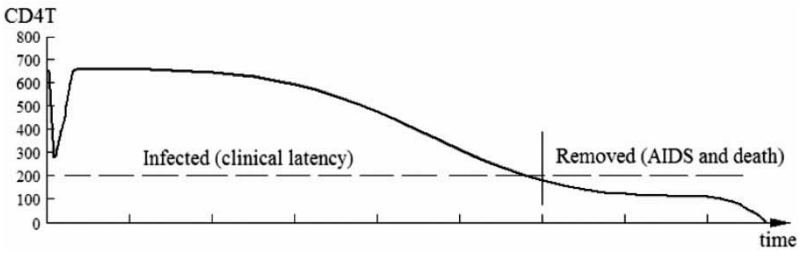

Figure 2. Infected nodes evolution as a function of CD4 T-cell counts.

2.2. *Analytical approximation of CNs dynamics*

Generally, the interplay between the fraction of S, I, and R is described in the form of a standard epidemic SIR model in terms of a system of differential equations. The simplest form can be expressed as:

$$\begin{aligned} \frac{\mathrm{d}s(t)}{\mathrm{d}t} &= -\lambda\rho(t)s(t) \\ \frac{\mathrm{d}\rho(t)}{\mathrm{d}t} &= -\mu\rho(t) + \lambda\rho(t)s(t), \\ \frac{\mathrm{d}r(t)}{\mathrm{d}t} &= \mu\rho(t) \end{aligned} \tag{2}$$

where $s(t) -$ *susceptible*, $\rho(t) -$ *infective*, $r(t) -$ *removed*. This is equivalent to the dynamics described in Figure 1b, with exactly one link per node. Equation (2) can be interpreted as a mean-field approximation to the dynamics shown in Figure 1 [35]. Note that if we disregard demographic effects we can put $s(t) + \rho(t) + r(t) =$ constant. Susceptible individuals could be infected; infective individuals are capable of transmitting the disease; removed individuals have had the disease and are dead. Parameters λ and μ are positive constants representing *infection* and *removal* rates. The value λ is a characteristic of Γ_1 only and is determined by the details of the infection spreading. The value of μ depends on the disease evolution inside an infected node before its death or isolated status in which it does not pass on the infection anymore.

The graph of changing CD4 T-cell counts for a node (Figure 1c) is depicted in Figure 2. The shape of that curve is prototypical for infected individuals without treatment [30]. The mean incubation time is close to 10 years, after which the CD4 T-Cell count drops to a lethal level where the immune system is not capable of fighting against opportunistic infections anymore. This condition is called AIDS and is considered as an isolated state for an infected node, which is consequently removed from the network. The incubation period for an infected individual depends on individual features and social as well as economic factors. The mean value of the incubation time varies from 8.0 to 13.0 years [5,19].

The transition for each individual (per year) from state ρ to state r can be described in the form of an irreversible Markov chain with a matrix of transition probabilities given by:

$$\Pi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-\mu & \mu \\ 0 & 0 & 1 \end{bmatrix}, \tag{3}$$

In particular, for HIV infected nodes μ may vary from $\mu = 0.15$ for individuals without treatment to $\mu = 0.08$ for individuals with treatment [6]. The generalized matrix of transition probabilities Π may be easily extended to any number of intermediate stages of infection.

Combining expressions (2) and (3) one can express the epidemic model with homogeneous mixing in the form of a discrete Markov chain with a corresponding state vector given by: $E = [s, \rho, r]^{\mathrm{T}}$:

$$E_t = \Pi^{\mathrm{T}} E_{t-1} + \lambda(I_2 - I_1)(E_{t-1}I_1)(E_{t-1}I_2)$$
$$E_0 \stackrel{\text{def}}{=} [1 - \rho_0, \rho_0, 0]^{\mathrm{T}} \tag{4}$$

where $I_1 = [1 \quad 0 \quad 0]^{\mathrm{T}}$, $I_2 = [0 \quad 1 \quad 0]^{\mathrm{T}}$.

Expression (4) might be an integral representation of the evolution process (1) in terms of a discrete distribution of s, ρ, r.

The SIR model represented in Equation (4) is too simple to be useful for real applications but is a good starting point for more advanced models, these we will describe in the following paragraphs.

2.3. *Perspectives of HIV modelling*

As a rule, real HIV epidemiological processes are rather complicated. So it makes sense to list some of the procedures, proposed in the epidemiological literature, which make the SIR models more suitable for real situations.

(1) In Equation (2), the population is assumed closed, that is, the population size is fixed ($s(t) + \rho(t) + r(t) = \text{const}$). However, recent results reporting on the spread of HIV indicate the existence of a strong interplay between HIV epidemics and age structures. The demographic impact cannot be neglected. For instance, in Reference [6] the population in SIR models is randomly refreshed to correspond to the age structure for a given country. Another approach has been proposed in Reference [10] where a demographic rate has been introduced into a set of coupled differential equations for the related SIR model.
(2) The SIR models expressed in the form of Equation (2) are only valid for homogeneous mixing, that is, all individuals are able to infect all others with an infectivity λ almost identical for each individual [3]. Recent theoretical results on epidemics show that the rate of the epidemic changes is highly affected by the dynamical behaviour of individuals. The inhomogeneity of behaviour in the exposed groups can be taken into account by introducing a structured sexual contact network. A number of papers show a power-law behaviour for such networks. An interesting result can be found in [23, 24] where a set of coupled differential equations with classes differentiated by connectivity is used for numerical evaluation.
(3) The removal rate μ in Equations (2) and (4) results from a very simple model describing the progression of the time of infection to the final state of the disease. For infections with short incubation time this rate can be described by a single value, but for diseases like HIV the incubation time is comparable with the simulation time. Moreover some changes like diagnosis or treatment [1,32] must be taken into account. This makes the model non-stationary.

Because of these observations we introduce a new method based on direct simulation of CNs in the form given by Equations (1) and (4).

3. The network simulation procedure

HIV is a very special epidemic. Let us list some challenges and approaches arising in HIV modelling:

(1) It is very difficult to build up an HIV model for all the exposure groups simultaneously. Each of these groups has distinctive features and generally they are simulated independently. Some models are used for a general population [22], but others for different exposure groups like heterosexual [6] or homo/bisexual [1] groups only.
(2) The HIV epidemic is inhomogeneous over time. Three epochs pre-ARV (till 1987), antiretroviral therapy (ARV) (1988–1995) and highly active antiretroviral therapy (HAART)-treatment can be distinguished. The system of diagnosis has been improved as well. Those changes may be taken into consideration by non-stationary coefficients of the HIV progression and the probability of infection per partner [1,33].
(3) Despite the improvement of HIV diagnosis only AIDS cases are recorded with some certainty and there is a big uncertainty about early stages of the epidemics. Moreover, some important factors like the size of exposure groups and the interplay with demographic and geographical structures are far from obvious.
(4) With construction of the model in the form given by Equation (1), one should take into account that HIV is primarily an STD. In our model, we identify three ways of HIV infection:
   - men who have sex with men (MSM);
   - heterosexual population;
   - injection drug users (IDU). For this group HIV is not an STD.

Each risk group has its own probability of infection per partner λ and specific demographic factors as well as related contact networks. Obviously sexual contact networks may not be described as a homogeneous graph. Unfortunately the information concerning the structure of contact networks for the third largest exposure group (IDU) is absent. That is why we choose a network model (4) with homogeneous mixing only.

3.1. *Modelling of sexual contact networks*

We construct a network model as a dynamical scale-free network, wherein each individual is represented by a node and the edges are the links between the individuals. The infection may spread along these edges. There are a number of papers which describe scale-free properties of sexual contact networks [18,28]. The scale-free property implies that network has a power-law distribution in the number of outgoing links, with an exponent γ that determines the specific structure of the network for each type of HIV spreading implemented in the model [11]:

$$P(k) \sim k^{-\gamma}, \quad k \leq k_{\max}, \tag{5}$$

where k is the number of sexual partners *per year* and γ is a parameter of the distribution. Small γ denotes more limited sexual contact behaviour and corresponds to a smaller value of $k_{\max}$ indicating the promiscuity and vice versa.

We use for the homosexual contact network the following parameters of the power-law distribution $\gamma_1 = 1.6$ and $250 < k_{\max} < 300$ (see [18] for more details). For the heterosexual network we can take $\gamma_2 = 2.7$ and $60 < k_{\max} < 70$. The cut-off $k_{\max}$ of the distribution is very important as it indicates the possible number of *superspreaders* in the network. Actually, we may use lifetime statistics and build up the model on a static graph but the associated statistics is absent or unreliable.

With respect to Equation (2), we use at each step a *configuration model* for the contact network generation. This flexible approach is based on the generation of a *degree sequence* which allows generating links between any two nodes according to its degrees taken from a specific distribution. In Figure 3 the simulated network patterns for a heterosexual population (a) and MSM (b) are depicted using recorded data.

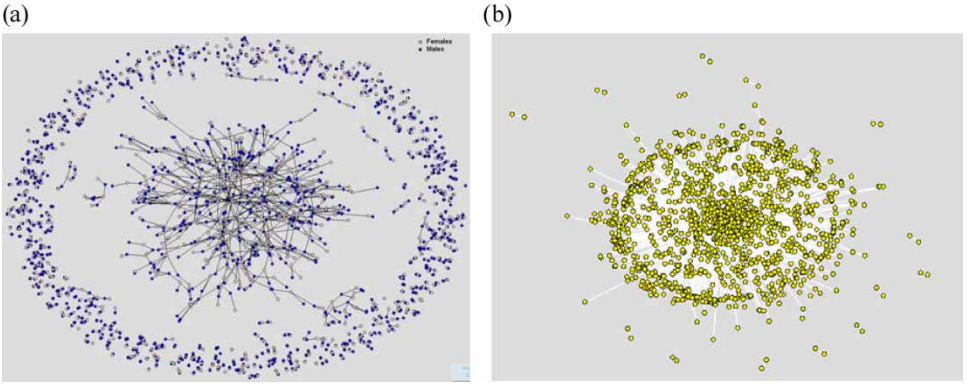

Figure 3. Sexual contact networks over a one-year period. Visualization is done using the Pajek software system [9]. (a) Heterosexual network; (b) MSM network.

From Figure 3 we observe that in the heterosexual contact network there is (at least) one big component and a lot of pairs. For MSM network almost all the nodes belong to one giant component.

3.2. *Modelling the dynamics of the infected nodes*

The Markov chain model with the matrix of transition probabilities given by Equation (3) is a reductive model of infected nodes dynamics, which is equivalent to the model described by Equations (2). Taking into account duration of the incubation period, availability of treatment and the effect of diagnosis (awareness about infection may affect the transmission probability), more advanced models with a larger number of states could be used. In particular, Aalen et al. [1] proposed a multi-state Markov model to represent stages of HIV infection and the diagnosis and treatment. This model is very convenient for back-calculation models [1,2,33], SIR models, and for including SIR models into a network model.

The pre-AIDS stages of HIV infection are defined in terms of the level of CD4+ T lymphocytes (CD4 counts). Four states are specified corresponding to CD4 counts of 500 or higher ($\times 10^6$ cells/l), 350–499, 200–349, and below 200. The mean occupancy times for states 1, 2, 3, and 4 are 5.5, 2, 2, and 1 year, respectively. It reflects a median incubation time of 8.6 years and a mean of 10.5 years. It is assumed that HIV testing became available gradually during 1984, and that general access to testing was available from the beginning of 1985. After years of improvement these tests are assumed to be constant since 1990. Furthermore, we assume that treatment was introduced in 1988 and has been substantially improved after the introduction of combined therapies in 1996. The rates of diagnosis, treatment and other coefficients for this model can be found in a series of papers [1,2].

3.3. *Modelling of demographic effects*

As was shown in Figure 1c, a demographic rule for the network implies permanent changes inside the network that have a strong interplay with the age structure of the population. The easiest way to describe this feature is to define two coefficients d_1 and d_2 which reflect the percentage of the nodes removed (infected and healthy) or embedded (healthy only) in the network at random, respectively. If $d_1 > d_2$ the population (network size) is decreasing, if $d_1 < d_2$ the population is growing. Our assumption is that the population size (*i.e.*, the risk group) is stationary, that is $d_1 = d_2$. This assumption has proven to be quite realistic for last three decades for many developed

countries [34]. For instance, if the network describes the population in the ages of 15–39 years, the coefficient $d = d_1 = d_2$ would be roughly 4%, due to the lack of precise information on the size of this risk group we allow this value be 'tunable'.

3.4. *Modelling of spreading process*

The structure of the operator Γ_1 is quite simple. For each kind of infection an infected node can infect its neighbouring nodes with a given probability λ. The problem is how to estimate this unknown value. Many papers report a probability value of infection per unprotected sexual contact in the range of 0.002% to 40% [20,30]. For this we have chosen to make this value as a free 'tunable' parameter as well. Optimization of λ with a numerical solution $r(t)$ of the model system described in Equation (2) is done through minimization of the following functional:

$$\sum_{Y=1981}^{2005} [r(Y, \lambda) - N_Y]^2 \longrightarrow \min. \tag{6}$$

Here N_Y is the annual officially registered number of AIDS cases in a given country. For instance, data for the USA and Europe can be found in [14,15].

The advantage of this model is the relative scale of modelling, this allows us to avoid a direct approximation of exposure groups' size and to recalculate the scale of the simulation by data reported in literature. For the optimization procedure we use a standard steepest decent method with adaptive step-sizes.

The optimization result for the population in the USA showed a value $\lambda = 0.44$. Of course, this value is much larger than the infection probability per sexual partner which is hard to compare since the number of unprotected sexual contacts is not well known.

Additional 'tunable' parameters are the demographic coefficients d_1, d_2 and the initial size of the infected group for the first years of simulation ρ_0 (for 1975), that define the initial conditions for our model in Equation (4).

3.5. *The direct simulation algorithm*

The basic simulation procedure can be written down in a number of consequent steps:

(1) Generate a network using a given node degree probability distribution with an initial number of randomly infected nodes ρ_0.
(2) Infect nodes with probability for every link (per partner).
(3) For each infected node apply a rule of progression from HIV to AIDS (see paragraph 3.2.). Nodes with AIDS are removed from the network.
(4) Apply the demographic rule.
(5) Store the current nodes state and generate a new random network.
(6) Shift current time to one step forward and repeat (2)–(5).
(7) Repeat (1)–(7) till statistical significance has been obtained.

For simulation of heterosexual spreading the network is represented as a bipartite graph and the transmission probability from men to women is taken to be twice as efficient. Note that operators $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ correspond to Steps (2), (1), (4), and (3), respectively.

The three different types of infection (heterosexual, MSM, IDU), are simulated independently.

4. Results

4.1. *HIV data*

The United States data [14] are used for model identification and tuning mainly because they provide relatively good statistics of AIDS cases, several kinds of infection spreading and a possibility

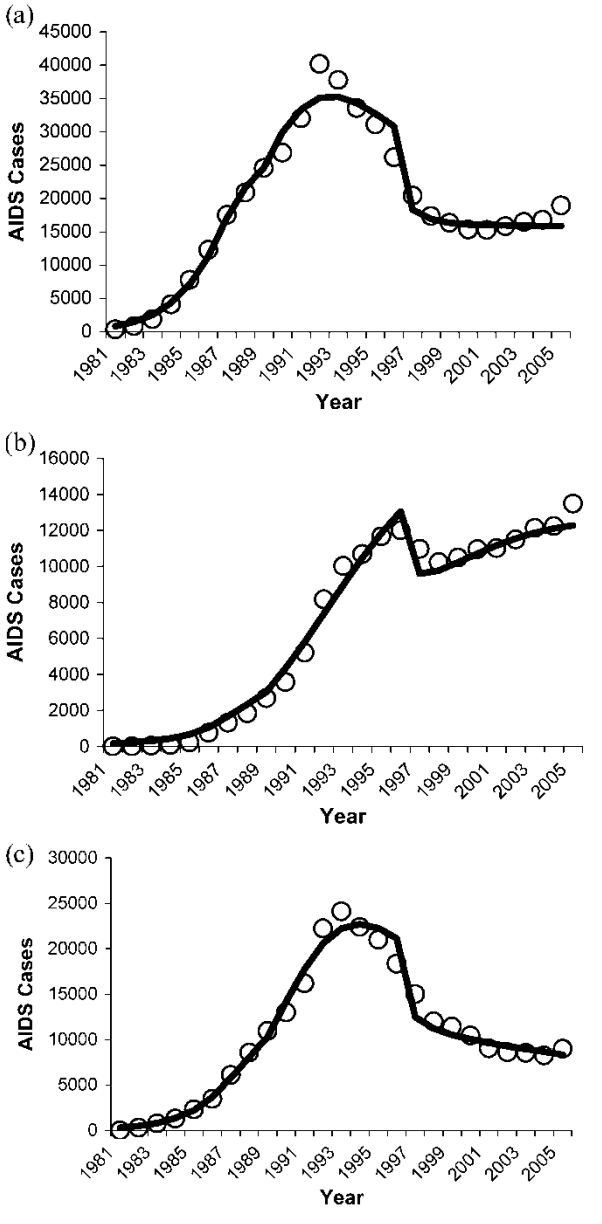

Figure 4. Simulation results and reported data for the AIDS epidemic in USA. (a) MSM exposure groups, power-law distribution with $\gamma = 1.6$ and $k_{\max} = 250$, $I_0 = 0.32\%$, $d = 0.04$, $\lambda = 0.44$; (b) Heterosexual exposure group, power-law distribution with $\gamma = 2.7$ and $k_{\max} = 60$, $I_0 = 0.2\%$, $d = 0.05$, $\lambda = 0.28$; (c) IDU exposure group, homogeneous mixing, simulation parameters are $I_0 = 0.16\%$, $d = 0.025$, $\lambda = 0.72$. Circles are the annual officially registered number of AIDS cases; the solid line indicates the simulation results.

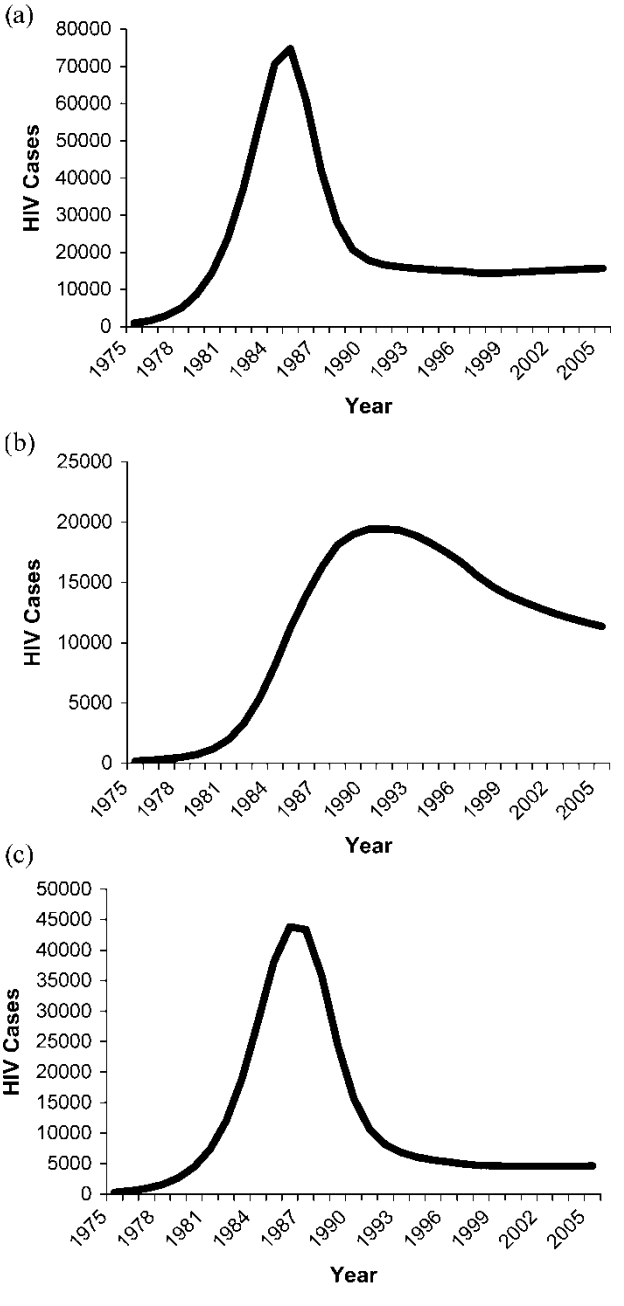

Figure 5. Simulation result for annual HIV cases in USA. (a) MSM exposure groups, power-law distribution with $\gamma = 1.6$ and $k_{\max} = 250$, $I_0 = 0.32\%$, $d = 0.04$, $\lambda = 0.44$; (b) heterosexual exposure group, power-law distribution with $\gamma = 2.7$ and $k_{\max} = 60$, $I_0 = 0.2\%$, $d = 0.05$, $\lambda = 0.28$; (c) IDU exposure group, homogeneous mixing, simulation parameters are $I_0 = 0.16\%$, $d = 0.025$, $\lambda = 0.72$. Measured data are not shown on this figure, because statistic about annual HIV cases is relatively good for past decade only and there is a big uncertainty about early stages of epidemic.

to explore the effects of treatment. These data include at least three epochs of the epidemic evolution: pre-ARV, ARV, and HAART-treatment; and three kind of HIV spreading. For each of these epochs and the three distinct kinds of infection, HIV spreading behaviour was studied.

4.2. *Validation: comparison with annual AIDS data*

As we can observe from Figure 4 the simulation results are very close to the estimation of the officially registered annual AIDS cases. We should, however, consider these result very carefully, since some of model deficiencies may have been diminished by the free 'tunable' parameters. The obvious effect on all the figures shown, is a substantial decline of AIDS cases after introduction of HAART (1996). As can be observed for the homosexual and heterosexual populations, the peaks of the AIDS epidemic are not the same. From a computational point of view the existence of these peaks can be explained as a result of depletion of people with high risk behaviour and a shift of the epidemic to the generic population (with 'normal' risk behaviour), for instance through bisexual contacts. The stable number of AIDS cases in recent years may be explained by the stationary number of HIV infection over a long period of time (see Figure 5).

4.3. *Reconstruction: annual HIV cases*

The simulation result of annual HIV cases reconstruction for different exposure groups is presented in Figure 5. It is encouraging that those estimations are close to the official estimation by the Center for Disease Control [7], which represents approximately 40,000 new transmissions per year all the way from 1990 to the present time, after the peak of 1985 with approximately 150,000 annual HIV cases (for all groups simultaneously).

Some of the effects that we observe from Figure 4 for AIDS cases are present in the graphs for HIV as well. For instance, the stationary flow of HIV epidemics over more than a decade is reflected in the stationary flow of the number of annual AIDS cases. The largest and not obvious difference is a lack of decline of HIV cases after introduction of HAART. It probably means that the process of HIV infection is quite independent from the effect of treatment and that a stationary flow of epidemics may be explained by the balance of rate of infection, incubation period, and demographic factors (supporting the size of risk groups).

We note that the best fit of the demographic coefficient (3–5%) is closely related to the group size (20–40 years) which is generally considered as the size of the exposure groups for different kinds of HIV spreading.

5. Conclusions

We present a parameterized CN model describing the dynamics of HIV spreading. The model has some distinctive features: it takes into consideration all the existing kinds of HIV spreading. Homosexual and heterosexual spreading is described by a scale-free network, drug users spreading is described with the assumption of homogeneous mixing inside the exposure group. All the network parameters have been taken from the medical literature and were fixed during the numerical experiments. The experiments show a promising correspondence between the model results and real demographic historical epidemiological data. This CN model will be included into a generic system of predictive models for HIV infections and associated drug-ranking strategies [29]. As such it will be part of the ViroLab decision support system [31].

- Sexual contact network dynamics is coupled to an inhomogeneous Markov model of HIV infection, diagnosis, and AIDS progression.
- The model was initialized by real historical data. Annual HIV cases were reconstructed by *historical* annual AIDS cases.

The proposed model may be used as a 'back-calculation' [12] model and as a predictive model, because the CN technique mimics the underlying mechanism of HIV transmission.

Acknowledgements

The authors would like to acknowledge the financial support of the European Union through the ViroLab project (www.virolab.org), EU project number: IST-027446. We thank Breanndán Ó Nualláin (University of Amsterdam) and Alex Vespignani (Indiana University) for proofreading the manuscript.

References

[1] O.O. Aalen et al., *New therapy explains the fall in AIDS incidence with a substantial rise in number of persons on treatment expected*, Aids 13 (1999), pp. 103–108.
[2] ———, *A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: application to AIDS prediction in England and Wales*, Stat. Med. 16 (1997), pp. 2191–2210.
[3] R.M. Anderson and R.M. May, *Infectious Diseases in Humans*, Oxford University Press, Oxford, 1992.
[4] R.M. Anderson and A.R. McLean, *Possible demographic consequences of HIV/AIDS epidemics*, Math. Biosci. 90(1–2) 1988, pp. 475–505.
[5] S.R. Arni and R. Srinivasa, *Incubation-Time Distribution in Back-Calculation Applied to HIV/AIDS Data in India*, Math. Biosci. Eng. 2(2) (2005), pp. 263–277.
[6] W.J. Bai, T. Zhou, and B.-H. Wang, *Interplay between HIV/AIDS epidemics and demographic structures based on sexual contact networks*, 2007. Available at arXiv:physics/0602173.
[7] J.G. Bartlett, *Routine opt-out testing for HIV: rationale and obstacles*, 2007. Available at slides/transcript at http://www.medscape.com/viewarticle/551906_2.
[8] F. Baryarama, J.Y.T. Mugisha, and L.S. Luboobi, Computational and mathematical methods in medicine 7(1) (2006), pp. 27–35.
[9] V. Batagelj and A. Mrvar, *Pajek – analysis and visualization of large networks*, Connections 21 (1998), pp. 47–57.
[10] J. Benoit, A. Nunes, and M.T. da Gama, *Pair approximation for models of disease spread*, Eur. Phys. J. B50(1–2) (2006), pp. 177–181.
[11] N. Boccara, *Modeling Complex Systems*, Springer, 2004, p. 410.
[12] R. Brookmeyer and M.H. Gail, *A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic*, J. Am. Stat. Assoc. 83 (1988), pp. 301–308.
[13] S. Gonçalves, M. Kuperman, and M.F. da Costa Gomes, Physica A: Stat. Mech. Appl. 327(1–2) (2003), pp. 6–11.
[14] HIV/AIDS database at Centers for Disease Control and Prevention (CDC). Available at http://www.cdc.gov/hiv/.
[15] EuroHIV (2003). *HIV/AIDS Surveillance in Europe. End-year Report 2003 No. 70.* European Centre for the Epidemiological Monitoring of AIDS. WHO and UNAIDS Collaborating Centre on AIDS. Available at http://www.eurohiv.org/.
[16] J.M. Hyman, J. Li, and E.A. Stanley, Math. Biosci. 181 (2003), p. 17.
[17] W. Kindermann, *HIV Spread among drug addicts in the Federal Republic of Germany*, Addiction 11 (1984), pp. 1372–1373.
[18] F. Liljeros et al., *The web of human sexual contacts*, Nature 411(6840) (2001), pp. 907–908.
[19] I.M. Longini et al., *The dynamics of CD4+ T-lymphocyte decline in HIV-infected individuals: a Markov modeling approach*, J. Acquir. Immune Defic. Syndr. 4(11) (1991), pp. 1141–1147.
[20] D.B. Louria et al., *HIV heterosexual transmission: a hypothesis about an additional potential determinant*, Int. J. Infect. Dis. 4(2) (2000), pp. 110–116.
[21] G. Marks, S. Burris, and T.A. Peterman, Aids 13 (1999), pp. 297–306.
[22] C. Pasqualucci et al., *Estimating the size of the HIV/AIDS epidemic: complementary use of the empirical bayesian back-calculation and the mover-stayer model for gathering the largest amount of information*, Simulation 71(4) (1998), pp. 213–227.
[23] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free networks*, Phys. Rev. Lett. 86 (2001), p. 3200.
[24] ———, *Epidemic dynamics in finite size scale-free networks*, Phys. Rev. E 65 (2002), p. 036104.
[25] M.A. Pedraza et al., Aids 16(7) (2002), pp. 1084–1086.
[26] T.C. Porco et al., Aids 18 (2004), p. 81.
[27] 2006 Report on the global AIDS epidemic, UNAIDS/WHO, 2006.
[28] A. Schneeberger et al., *Scale-free networks and sexually transmitted diseases; a description of observed patterns of sexual contacts in Britain and Zimbabwe*, Sex Transm. Dis. 6 (2004), pp. 380–387.
[29] P.M.A. Sloot et al., *A grid-based HIV expert system*, J. Clin. Monit. Comput. 19(4–5) (2005), pp. 11387–1307.
[30] P.M.A. Sloot, F. Chen, and C.A. Boucher, *Cellular automata model of drug therapy for HIV infection*, in *Lecture Notes in Computer Science*, Vol. 2493, 2002, pp. 282–293.
[31] P.M.A. Sloot et al., *From molecule to man: decision support in individualized E-health*, IEEE Comput. 39(11) (2006), pp. 40–46.

[32] R.J. Smith and L.M. Wah, *Drug resistance in an immunological model of HIV-1 infection with impulsive drug effects*, Bull. Math. Biol. 67(4) (2005), pp. 783–813.
[33] M.J. Sweeting, D. De Angelis, and O.O. Aalen, *Bayesian back-calculation using a multi-state model with application HIV*, Statist. Med. 24 (2005), pp. 3991–4007.
[34] United Nations, Population Division, Department of Economic and Social Affairs. World Population Prospects: The 2004 Revision, 2005.
[35] R. Yang et al., *Epidemic spreading on heterogeneous networks with identical infectivity*, Phys. Lett. A: Gen. At. Solid State Phys. 364(3–4) (2007), pp. 189–193.