AGENT-BASED POPULATION MODELS FOR HOUSEHOLD SIMULATION

Florian Miksch[1], Günther Zauner[1], Niki Popper[1], Felix Breitenecker[2]

[1]dwh Simulation Services, Neustiftgasse 57-59, 1070 Wien, Austria
[2]Vienna University of Technology, Vienna, Austria

florian.miksch@drahtwarenhandlung.at (Florian Miksch)

Abstract

Social contacts are a key element when modeling the spread of epidemics. Classic top-down approaches like ODEs do not consider the social system explicitly because they use abstract parameters that are representing a combination of various real values. By contrast, agent based models provide a more detailed structure and directly use real world values instead of abstract ones. In this paper we first introduce a basic concept of a simple contact system in an agent based SIR model. Such a model corresponds with known top-down approaches very well. Then we propose a method to extend the basic approach with households. Therefore we make use of detailed data that is provided by Statistics Austria. In a validation process we see that, under special presumptions, the household extension is still comparable with the basic approach. This is important because knowledge about behavior of SIR epidemics still has to apply on the extended model. In contrast to the basic model it is now possible so simulate a wide range of new scenarios. Another strength is the ability for more detailed evaluation of simulation runs. In a sample scenario we show the impact of different household sizes on the spread of an SIR epidemic. Furthermore we give ideas for more detailed and complex scenarios.

Keywords: Agent based model, Epidemics, Households, Social contacts.

Presenting Author's biography

Florian Miksch. He studied Technical mathematics with a focus on mathematical methods in computer science at Vienna University of Technology where he earned his master's degree in 2009. His work area is modeling and simulation in health technology assessment (HTA) for *dwh simulation services* and data mining and analysis for the *Main Association of Austrian Social Security Institutions*. Currently he is also pursuing a PhD degree under Prof. Felix Breitenecker at *Vienna University of Technology.*

1 Introduction

Social contacts are a key element when modeling airborne infections. Since a transmission can only happen when two people, one infectious and one susceptible, meet each other, the modeler wants to simulate when, how and where these meetings take place and how he can make use of this information.

Households are considered to play an important role in the spread of airborne diseases because it is an environment where the same people meet each other every day, spend many hours together and may have physically closer contacts than they have with colleagues at work.

At first we give a rough overview about contacts in an agent based model and then we extend this basic concept to implement households. Afterwards we validate that extension and give ideas how to use it.

1.1 An introduction into the simulation of airborne infections

When one considers single persons the basic factors for epidemics are the contacts and the infection probability (Figure 1). This means it starts with a population of single persons. Every person has certain attributes like gender or age and an infection state. Such infection states can be susceptible, infected with or without symptoms, infectious and many more. It is clear that the implemented infection states have to depend on the kind of simulated sickness. During a simulation run people can change their infection state. Recreation (getting healthy from a sickness state) does not necessarily depend on external influences and may happen after a certain time. Against this getting sick from a healthy state requires a transmission. Transmissions can only happen when people meet each other. Therefore we need a submodel that defines contacts. It is crucial to know who meets whom and when the contact happens. Every time when an infectious person meets a susceptible person a transmission happens with a certain probability. This probability can either be set globally for each and every contact or individually depending on attributes of the two involved persons. However an “infectious” as well as a “susceptible” state may be represented by several different states. An infectious person for example can have symptoms or not. Susceptible persons for example can be healthy or sick.

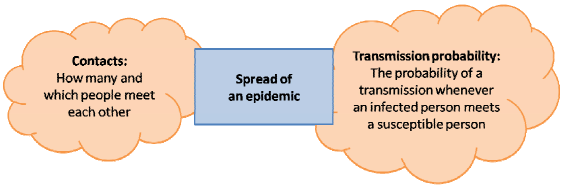

Fig. 1: Basic factors for epidemics

1.2 An introduction into the simulation of contacts

There is no standardized way how to simulate contacts and find out, who meets whom, where and when. The real social system is very complex and does not follow strict rules since people move within their environment more or less without restrictions. They have a family, meet friends, go shopping, do sports, visit bars and clubs and do a lot more social activities where transmissions are possible. People have loose and intense, long and short, regular and unique contacts.

Such a complex contact system that does not follow rules cannot be modeled completely so it has to be simplified and classified.

An approach is to define lists of contacts that are valid for a certain time step until they are regenerated or updated. Another approach is to define a discrete or continuous space where people move around.

Parameters can either be set globally or individually. A global parameter set for the contact submodel might result in the same behavior and contact rates for all people. Individual contact parameters for example can take account of personal attributes like the age. Then it is possible to consider that children play mostly with other children and meet less adults or that young, healthy people have a more active social life than old and sick people. Individual parameters can also lead to contact systems where some people play special roles like nurses in hospitals who meet a lot, sick people. Such individual contact rates for people have been examined in different sociological studies that can be used for calibration of the contact submodel [1][2].

1.3 An introduction into SIR models

SIR models are among the first and best examined models for simulating epidemics. SIR stands for **s**usceptible – **i**nfected – **r**esistant. An SIR model is a simple approach for modeling diseases where all people are considered to be susceptible at first, then all of them can get infected, stay infected for a certain period of time and become resistant to the disease after recovery.

One approach is to use differential equations describing three time-dependent variables: The number of susceptible, the number of infected and the number of resistant people.

Further approaches are cellular automata or agent based models that consider single persons (Figure 2).

This paper deals with an agent based approach for SIR type epidemics.

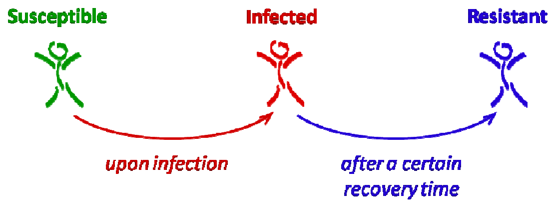

Fig 2: States of an agent in an SIR model

2 The basic model

We start with a dynamic agent based SIR model. The agents are single persons having properties like gender, age, health and infection state. Since the purpose of this task is to implement and examine social contacts the infection part is only modeled as a simple SIR model. The advantage of this strategy is that we can concentrate on the effects of the social system because SIR models are examined very well. The infection part is modeled in a way that single persons are susceptible at first, become infected upon transmission, recover after a certain time and remain resistant after recovery. [3]

2.1 Set up of the model

Before simulation the environment has to be set up.

At first a predefined number of single agents, the persons, are created. These persons have the following attributes:

- Age: The age of all agents is assigned in a way so that it corresponds with Austria's age distribution.
- Gender: Half of the population is considered to be male respectively female.
- Infection state: All people are assigned the state "susceptible" at first. The other possible states are "infected" and "recovered".

After creation of the population a small part of the population is set as "infected" because if everybody is susceptible in the beginning of a simulation no transmissions can ever happen. A typical fraction of infected people in the beginning is about 1% of the whole population.

2.2 Execution of the model

The model is calculated in discrete time steps. For every time step three tasks have to be done:

- Calculate all contacts that happen in this time step.
- Calculate all transmissions that happen upon the contacts.
- Update contact-independent properties like recovery of infected people.

Note: The population remains static in this model while the contacts and the infection part is simulated dynamically. This means that people do not get older, do not die and do not leave the population, that no babies are born and no people enter the model from outside.

2.3 Basic implementation of contacts

For a non-specific implementation of contacts we assume all people to be the same concerning contact rates and distribution of contacts. Calculation of contacts needs 3 steps:

- At first we need to define the number of contacts per person per time step.
- With the number of contacts per person we calculate the number of contacts within the whole population per time step.
- Compute a list with two random persons in every row where the length of the list is the number of contacts for the whole population. Consider two people in a row having a contact. One has to make sure that contacts between the same two persons do not appear more than once in the list and that people do not have contacts with themselves.

2.4 Using the contact list

Using this concept allows to create a contact list in the first time step. For every succeeding time step there are three options available:

- Take the contact list from the previous time step.
- Take a part of the contact list from the previous time step and create a part of the contacts new.
- Create a completely new contact list.

2.5 Results of the basic implementation

The basic implementation with contacts lists and discrete time steps describes a homogenous population that results in the normal, well known behavior of a simple SIR model. Figure 3 shows the results of a run of the basic model with fictional parameters.

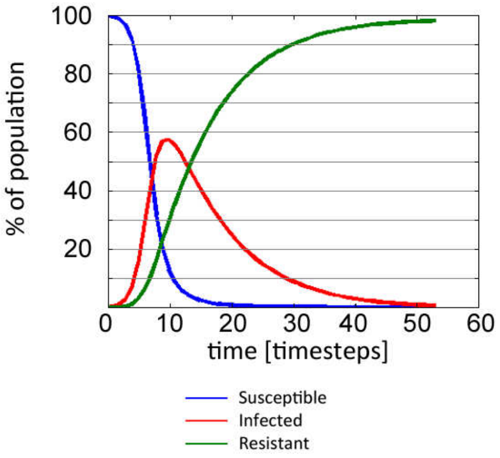

Fig. 3: Results of the basic model

3 The agent based model compared to ODEs

3.1 Advantages of the agent based model

- The model is fully comparable to other implementations of SIR models, for example with ODEs, and therefore it can be used instead of other implementations.
- Since the model is based on single agents it is possible to implement extensions that result in a non-homogenous population just by changing the behavior of some agents. While any extension of ODEs results in a new, extended structure a behavioral change of single agents does not necessarily require restructuring.
- The parameters in the agent based model correspond with real world quantities in contrast to abstract parameters in an ODE model.
- The outcome of the model can easily be calibrated to fit real world data by increasing or lowering the number of contacts or the infection probability for single agents.

3.2 Disadvantages of the agent based model

There are no analytic tools to examine the agent based models. Therefore we need several simulation runs with different parameter settings and compare the results. Against this there exist a wide theory about ODEs and many tools to examine them analytically.

4 Extend the basic concepts to households

4.1 Preliminary remarks

In the basic model we have created a homogenous population with contacts equally distributed over the whole population. But people are not homogeneous, they are individual and act differently.

Concerning the spread of an epidemic there are two parts that can be individualized: the infection part and the contact part.

A change of the infection part is not being discussed here. However it does not require structural changes in the model and is still a powerful tool for bringing the model closer to reality.

Reproduction of the real social contact system seems impossible since every person has an individual social life determining where, when and how often or intense he or she meets other people.

One of the key elements in social life are households. Households are not only key element because they are a place where the same people meet every day closely for a long time but also because it enables the possibility to simulate scenarios like curfew.

In this paper we give an idea how to extend the given structure to households without affecting the existing model. After validation of the new model we show how to use the household structure for scenarios and further extensions.

4.2 Data for households

Statistics Austria [4] provides the following data concerning households that we use for our concept:

- Distribution of household size
- Fraction of men and women living in a household together with their partner
- Fraction of men and women living in a household as a single parent
- Fraction of men and women living in a household as another household member

4.3 Implementation of households

Assign people to households. Assigning people to households can be done in 5 steps:

1. Calculate the number of demanded households and the number of members for each of them. Create these households as empty households.
2. Mark adults (above 18 years) as “partner”, “single parent” or “other household member”. Create couples from male and female partners.

3. Assign couples and single parents to empty households.
4. Assign children (up to 18 years) to households with couples or single parents.
5. Assign other household members to partly filled or empty households.

Create contacts using households. Now we are going to create two contact lists. For the first list we assume that all people within a household have contacts with each other. Then we count all contacts from households and calculate the number of remaining random contacts. Then the list with random contacts can be created as described in the basic concept 2.3.

Now there are two contact lists: The list of household contacts which remain static without change for the whole simulation and the random contact list that can be adjusted and recombined every time step.

Figure 4 shows a visualization of the household structure.

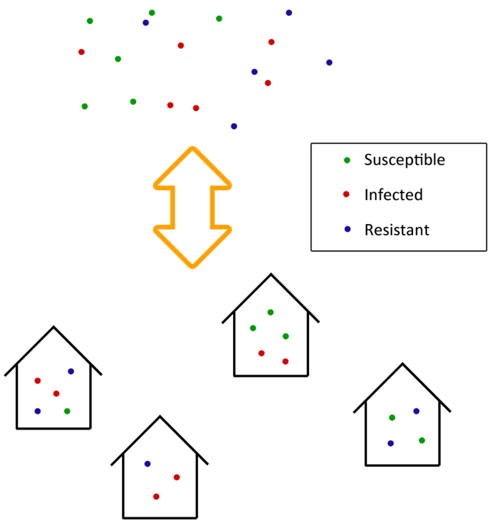

Fig. 4: Visualization of the system with random contacts and households.

4.4 Limits of the household concept

This household scheme is considered to be created in the first time step and remain for the rest of the simulation. Recombination of households caused by either social changes like divorces or marriages as well as changes of the population like immigration or dead people requires a more complex concept that is not discussed here.

4.5 Validation of household structure

For validation we take a simple SIR model and compare a completely random contact structure with a structure that consists of households and random contacts.

At first we need to set up two presumptions:

- The number of contacts in the whole system must be the same for both models.
- The transmission probability is the same for every contact, no matter if it happens within a household or as a random contact.

Parameters: Take a population of 1000 persons with average 15 contacts per person per day and an infection probability of 4% when an infected meets a susceptible person.

The household structure is based upon Austrian data with an average household size of 2.3 persons.

Interpretation of the validation. The simulations show the same behavior and almost the same results. That means that household structure itself does not affect the model but prevails the known and well examined behavior of the standard SIR model (Figure 5).

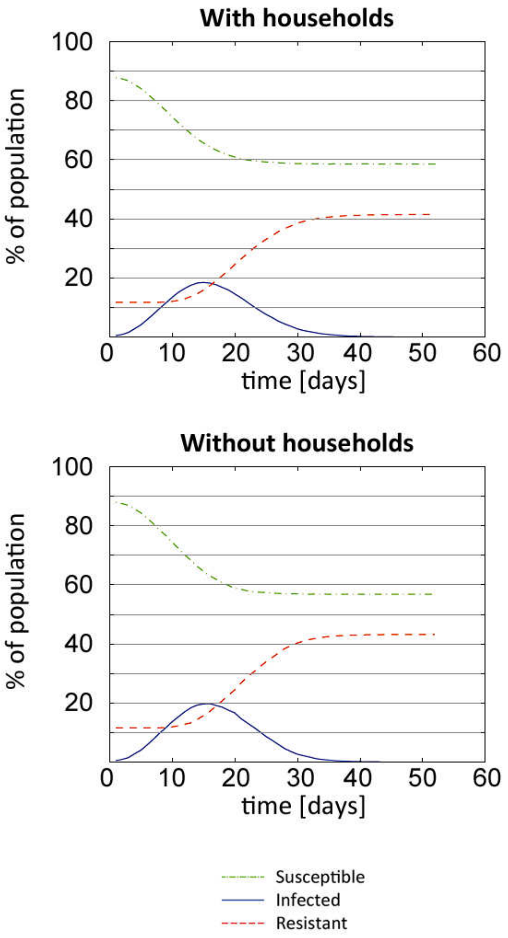

Fig. 5: Validation: Comparing the basic model with the extended model

5 Using households

Why do we need the household structure if it does not affect the model? A model with household structure only provides the same results under the same assumptions as shown above. But the strength of the new structure is the ability to change the assumptions in a way that is not possible in the basic model.

The extended model provides the possibility to simulate a wide range of new scenarios that are not available with simple random contact lists.

5.1 Sample scenario: Varying the household size

This sample scenario should demonstrate the strength of the newly developed extension. The requirements for this scenario are simple: Change the household size and remain the number of contacts outside households the same for all simulation runs.

Parameters: Simulate a population of 1000 persons with an average number of 13.7 random contacts outside households per day and an infection probability of 4% when an infected meets a susceptible person.

Normal household size is 2.3 and double household size is 4.6.

Results of the sample scenario: It is obvious that a bigger household size results in a higher number of contacts within households and therefore in a higher number of overall contacts per person. The fact, that people in bigger households have more contacts is also reported in a wide empirical study among many European countries [2].

So the outbreak of the epidemic is much stronger in case of bigger households and the number of infected people (Figure 6) is much higher.

Interpretation of the sample scenario: The average size of households differs in many countries. Many people in Europe live alone or together with a partner while in other parts of the world, especially in the so called “3rd world”, people live in large households together with many family members. Under the assumptions of the same transmission probabilities and social life outside households we see the impact of bigger households that leads to a significant stronger outbreak of an SIR epidemic.

5.2 Other scenarios

Examples of such scenarios are:

- Higher infection probability within households
- People stay at home if they are infected or as soon as they feel symptoms
- Proclaim a curfew for some time steps if the number of infected people exceeds a certain bound so that everybody has to stay at home.

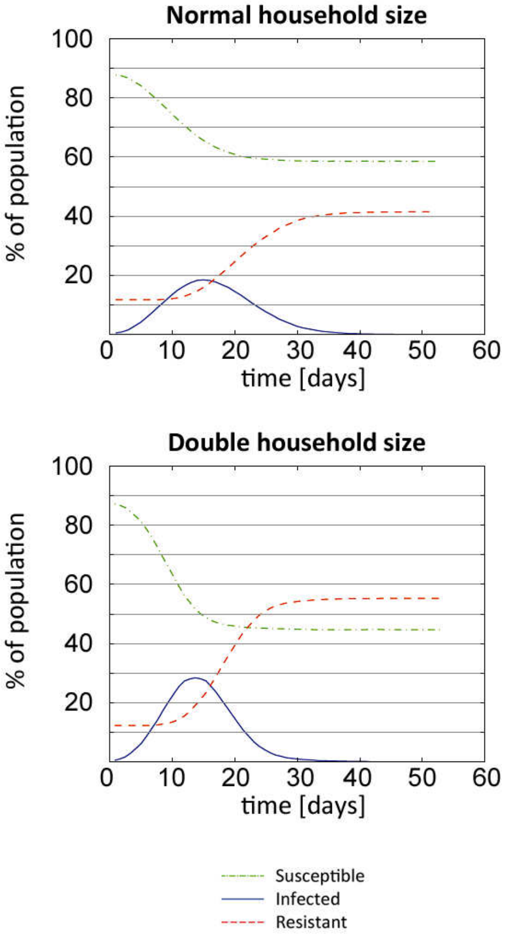

Fig 6: Scenario: Varying the household size

6 Outlook

Based on the concepts for implementation of households it is possible to model other places like workplaces, schools and hospitals.

7 References

[1]R.T. Mikolajczyk, M. Kretzschmar (2008), Collecting social contact data in the context of disease transmission. Social Networks 30, 127–135

[2] Mossong J, Hens N, Jit M, Beutels P, Auranen K, et al. (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3) e74. doi:10. 1371/journal.pmed.0050074

[3] Florian Miksch (2009), “Ein Agentenbasiertes Pneumokokken-Epidemiemodell zur Untersuchung von Langzeiteffekten und Impfstrategien.

[4] Statistics Austria (2010), “Statistisches Jahrbuch Österreichs 2010”, 68-70