Fluid Model-Checking in UPPAAL for Covid-19*

Peter G. Jensen, Kenneth Y. Jørgensen, Kim G. Larsen, Marius Mikučionis, Marco Muñiz, and Danny B. Poulsen

Department of Computer Science, Aalborg University, Denmark

Abstract. During the spring of 2020, the BEOCOVID project has been funded to investigate the use of stochastic hybrid models, statistical model checking and machine learning to anlyse, predict and control the rapid spreading of Covid-19. In this paper we focus on the SEIHR epidemiological model instance of Covid-19 pandemics and show how the risk of viral exposure, the impact of super-spreader events as well as other scenarios can be modelled, estimated and controlled using the tool UPPAAL SMC.

1 Introduction

Epidemic modeling has gained tremendous interest in both news and research communities in 2020 due to the rapid spread of Covid-19. In the news most of the interest has been to use modeling as a way to explain the spread of Covid-19 while much research is about using models to predict and control the spread.

In Denmark three (collaborating) initiates on combating Covid-19 using mathematical models has been made:

- In March a expert group headed by Statens Serum Insitut (SSI[1]) was established with the task of developing mathematical models to predict the impact of Covid-19 spread in the Danish society and to evaluate the effect of preventive measures.
- In early April researchers at Danmarks Tekniske Universitet (DTU) and Aalborg Universitet (AAU) started a research project funded by Novo Nordisk Fonden (NNF) to develop and improve modeling tools of Covid-19 to assist decision makers to evaluate the effectiveness and impact of preventive measures. The project has been carried out in collaboration with SSI.
- April 20th researchers from the Distributed, Embedded and Intelligent systems group at Department of Computer Science, Aalborg University (AAU) received a grant by Poul Due Jensens Foundation (PDJ) to futher aid development of Covid-19 models. The PDJ project has been working in close collaboration with the NNF projekt.

The key research question addressed in the NNF and PDJ projects has been to identify the best strategy for social restrictions dependent on age and region

* The project was funded by Poul Due Jensens Foundation grant.

[1] www.ssi.dk

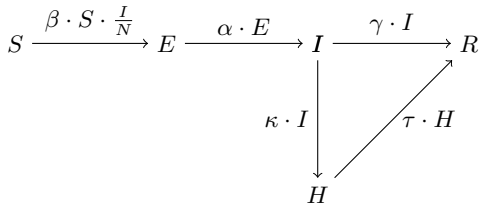


Fig. 1: Rate diagram of the basic compartmental SEIHR model.

in Denmark for protecting the population and society from Covid-19 mortalities caused by exceeding the intensive care capacity in the Danish hospital system.

An ambition of the two projects has been to provide a strategic decision tool for the Danish authorities. The projects has involved dessimination to Statens Serum Institute (SSI), through a number of scheduled meetings. These has not only included the concepts behind the models, but also effects of changing the underlying assumptions and the uncertainties inherent in the different input data and estimations, as well as the predictive power of the analysis methods.

Classically epidemics are modeled using so-called *compartmental* models where a population is divided into a number of different compartments, e.g. the following five compartments of the so-called SEIHR model:

- *susceptible* (S) being those that can affected by the disease,
- *exposed* (E) being those that has the disease but not yet infectious,
- *infectious* (I) being those that has the disease but not yet infectious,
- *recovered/removed* (R) being those that has had the disease and either recovered, in quarantine are died from the disease, and
- *hospitalised* (H) being those that are hospitalised.

The dynamic change of the distribution of a population over compartments may be described using a rate diagram such as Fig. 1. The expression above the arrows in the rate diagram describes the rate of the flow between different compartments, e.g. the arrow $E \xrightarrow{\alpha \cdot E} I$ means a conversion from E to I with a rate α multiplied by the number of E elements. Similarly $S \xrightarrow{\beta \cdot S \cdot \frac{I}{N}} E$ is a conversion from S to E with a rate β multiplied by the number of S, except this conversion is facilitated by infectious I elements where the probability of meeting one is I/N, therefore the overall conversion rate is scaled with this probability. As we shall see later, the rate diagram may be analysed using a number of different mathematical models.

In this paper we focus on the PDJ project, with the purpose of illustrating how statistical model checking in the tool Uppaal SMC [3] has been used to model, analyse and synthesize a variety of scenarios relevant for Covid-19 [4], ranging from abstract (continuous) population models to detailed (stochastic) agent-based models as well as (fluid) mixtures of these, allowing to reason about health risks of selected individuals in the setting of particular populations.

2 SEIHR Models in UPPAAL SMC

The rate diagram in Fig.1 describing how individuals move between the different compartments of the SEIHR model can be captured by a number of different mathematical models. Traditionally, rate diagrams are most often interpreted as ordinary differential equations (ODE) which are deterministic. However, the diagrams can also be viewed as stochastic models, where the rates of the reactions are used as parameters of exponential distritubions. The stochastic models may either be aggregated or be agent-based. In the latter, the health-status (i.e. compartment) for each individual is faithfully reflected, whereas in the former only the number of individuals in each compartment is maintained. The stochastic models are more realistic than the ODE models, but also come with a significantly increased complexity in their analysis: the analysis of agent-based models are exponentially[2] more complex than the analysis of aggregate models. A well-known fact about the aggregate stochastic model is that it can be easily translated into a set of ODEs capturing the expected behaviour of the model in the limit. Furthermore, the aggregated stochastic model can be proven to be a correct abstraction of the agent-based stochastic model using the notion of probabilistic bisimulation. To mediate between the accuracy of modelling and the complexity of analysis, it is possible to have mixed models – so-called fluid models – where selected individuals are modeled as agents, whereas the remaing population are modelled using either ODE or aggregated models.

In the remainder of this section we will show how the tool Uppaal SMC [3] can easily model and analyse all the three (four) above types of models for reaction networks.

2.1 Ordinary Differential Equation Models

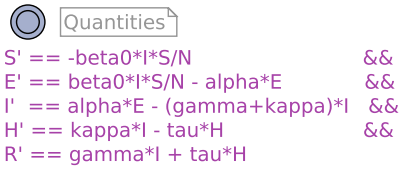
Figure 2 shows the ODE model of the SEIHR reaction network in Uppaal. The main ingredience of the model are the five continous state-variables `S`, `E`, `I`, `H` and `R` declared as (initialized) clocks in the declaration part Fig. 2a. The declaration part also sets a number of constants for the various rates of the reaction network to fit the evolution of Covid-19 in Denmark. Here we are looking at small subpopulation of Denmark with 10,000 people and with `1`% being exposed initially. The behavioural part of the model is given in Fig. 2b, being a one-location automata, with an invariant describing the behaviour of the state-variables as a system of ODEs. The ODEs are derived from the SEIHR reaction network in a very simple manner: for any state-variable X there is an ODE expressing that the derivative of X equals the difference between the total rate of the incoming edges and the total rate of outgoing edges, i.e.:

$$X' = \sum_{Y \xrightarrow{E} X} E - \sum_{X \xrightarrow{E} Z} E$$

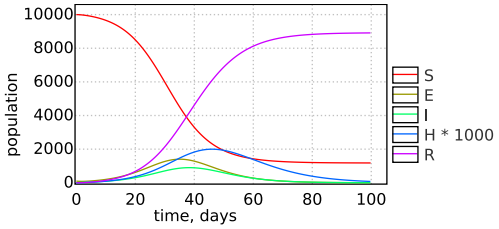
[2] Assuming a fixed number of compartments.

```
typedef int[0,1<<31-1] int32_t;
const int32_t N = 10000;
const double eps = 0.01;
const double BRN = 2.4;
const double alpha = 1.0/5.1;
const double gamma = 1.0/3.4;
const double beta0 = BRN * gamma;
const double pH = 0.9e-3;
const double kappa =
    gamma*pH/(1.0-pH);
const double tau = 1.0/10.12;
clock S = (1.0-eps)*N;
clock E = eps * N;
clock I = 0.0;
clock H = 0.0;
clock R = 0.0;
```

(a) Declarations.



(b) SEIHR hybrid automaton.



(c) Simulation.

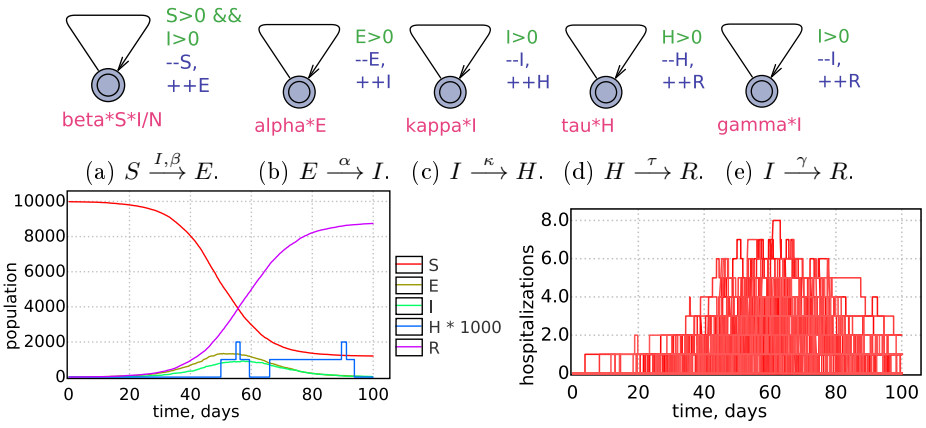
Fig. 2: ODE Model of SEIHR Rate Diagram.

Finally, we see the evolution of the state-variables over a period of 100 days in Fig. 2c. In particular, we note that out the total population of 10,000 some 1,403 will get exposed, 900 infected and 1,98 hospitalized. The time for the simulation was 0.077s.

2.2 Aggregated Stochastic Models

In Fig. 3(a)-(e), we show the aggregated stochastic model of the SEIHR rate diagram. In this model the different compartments are represented as integer variables (counters `S`, `E`, `I`, `H` and `R`) representing at any given point in time the number of individuals being in that state. A key assumption is that $\mathtt{S}+\mathtt{E}+\mathtt{I}+\mathtt{H}+\mathtt{R}=N$, where N is the number of individuals. As for the rate-expressions of the SEIHR rate diagram, these are used as rates of exponentially distributed transitions incrementing/decrementing the relevant counters. E.g. the (looping) transition of Fig. 3(b) indicates that one individual is transferred from `E` to `I` with rate `kappa*I` – of course provided that `E` is larger than 0 as expressed by the guard `E>0`. The resulting aggregated model is a continuous time Markov chain (CTMC) with states being vectors $(\mathtt{S},\mathtt{E},\mathtt{I},\mathtt{H},\mathtt{R})$ and where the five transitions of Fig. 3(a)-(e) are racing against each other.

In Fig. 3(f) we see the evolution of the state counters resulting from a single simulation over a period of 100 days. Despite the randomness of the simulation, the evolution of `S`, `E`, `I` and `R` seems indistinguishable from that of the ODE model Fig. 2. However, considering the variable `H`, we see a variation between 0, 1 and 2 over the 100 day period. Fig. 3(g) visualizes 100 random simulations with



(f) One simulation plot. (g) 100 simulations.

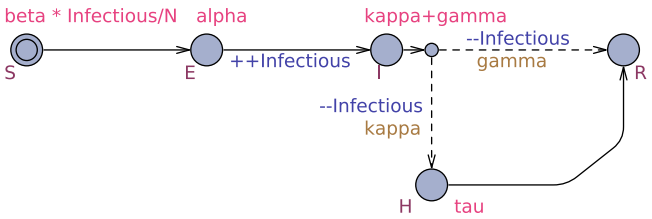
Fig. 3: Aggregated CTMC Model of SEIHR Rate Diagram.

H ranging between 0 and 8. In fact, based on the 100 random simulations the expected value of H in the aggregated CTMC model is found to be 3.82 ± 0.24. Moreover, using 291 random simulations the probability that H will exceed 4 is found to be in the confidence interval $[0.183, 0.282]$ with 5% confidence. These stochastic analyses significantly refines the expected behaviour analysis provided by the deterministic ODE model, where H was below 2. For this aggregated CTMC model the time to perform a single simulation is approximately 0.702s (a factor of 10 more than the ODE model).

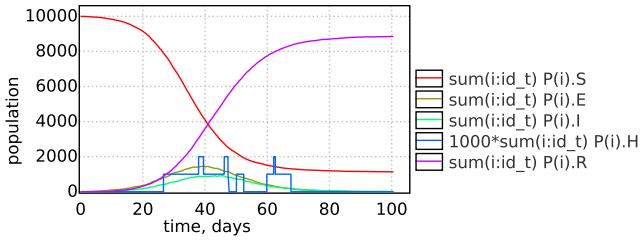
2.3 Agent-based Stochastic Models

Both the ODE model and the aggregated CTMC model provide sufficient information to address[3] the key question as to whether the capacity at hospitals will be exceeded within a given period. However, questions such as "how many different individuals will be hospitalized" and "what is the expected time before a given individual becomes exposed" cannot be readily answered by these two models. To answer such questions, we need an agent-based model, where the healthiness status of each individual is accounted for. Fig. 4(a) provides a SEIHR agent automaton (template) to be instantiated for each individual of the population. In the automaton, the five locations S, E, I, H and R are used to represent healthiness status. The time of transitions between the last four states are exponentially distributed with rates alpha, kappa and gamma respectively The rate of the transition between S and E has rate beta*infectious/N, where infectious keeps count on the total number of infected individuals, i.e. infectious/N is the probability that a random individual is infected.

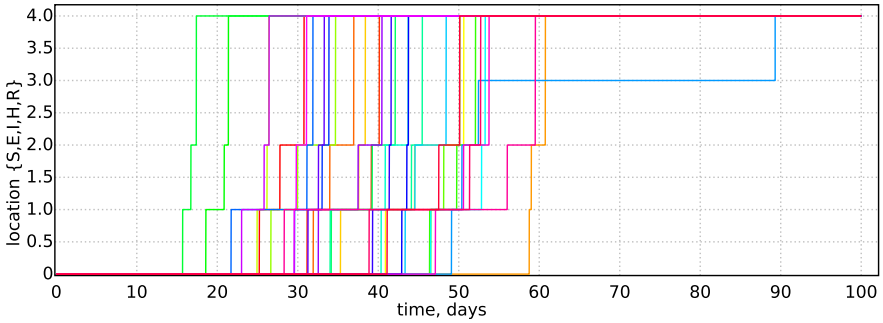
[3] Assuming, of course, that the given model is valid with respect to reality.



(a) SEIHR agent template `P`.



(b) Simulation.

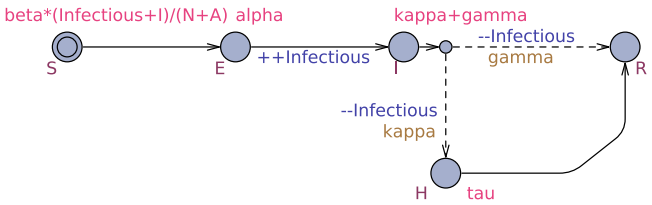


(c) Location changes of 30 individuals: `S` at `0`, `E` at `1`, `I` at `2`, `H` at `3`, `R` at `4`.

Fig. 4: Agent-based CTMC Model of SEIHR Rate Diagram.

Instantiating the template 10,000 times we see in Fig. 4(b) the evolution of the number of individuals in the different states resulting from a single simulation over a period of 100 days. During the simulation we are tracking expressions such as `sum(i:id_t) P(i).S`. Here `P(i)` refer to the `i`'th instance of the template `P`, and `P(i).S` is a Boolean indicating whether `P(i)` is in state `S`. Summing over all instances `i` the overall expression `sum(i:id_t) P(i).S` provides the total number of individuals in state `S`. We see that the evolution matches that of the aggregated CTMC.

Now we select 30 individuals out of the total population of 10,000. In Fig. 4(c) we track the state of these selected 30 individuals during one simulation. In particular, we note a wide variation in the time of becoming exposed as well as in the length of time being in the various states. Out of the 30 randomly selected individuals one a single person becomes hospitalized. The time for the



(a) Fluid SEIHR agent template P.

Fig. 5: Fluid Model of SEIHR Rate Diagram.

simulation for 10,000 individuals was 407.49s (several orders of magnitude larger than the aggregated CTMC model).

2.4 Fluid Models

For agent-based models – as described in the previous section – simulations involve *all* individuals of the total population resulting in significant increase in simulation time. However, if the properties of interests only refer to a limited number of selected individuals it may be advantageous to apply the method of *fluid model checking*.

In [2] a potential use of fluid approximation techniques in the context of stochastic model checking has been investigated. Here the focus is on properties describing the behaviour of a single agent in a (large) population of agents, exploiting a limit result known also as fast simulation. In particular, the behaviour of the single agent is approximated with a time-inhomogeneous CTMC, which depends on the environment and on the other agents only through the solution of the ODE. This approach has been proven assymptotically correct in terms of satisfiability of logical properties including reachability probabilities.

In Fig. 5(a) we revise the agent-based CTMC model from the previous section with the purpose of exploiting fluid model checking. For each of the 30 selected individuals, the template `P` will be instantiated. To model the behaviour of the remaining 9,970 individuals we will use the ODE model of Fig. 2 (with $N = 9,970$). Importantly, the template `P` of Fig. 5(a) describes a time-inhomogeneous CTMC as the rate of the transition of leaving `S` given by the expression `beta * (infectious+I)/(N+A)` is time dependent. Here `infectious` is the number of individuals infected out of the 30 selected ones, and the `I` one of the five state-variables of the ODE describing the number of infected individuals out of the 9,970 large population. Finally, `A` respectively `N` is the number of selected individuals (30) respectively the amount of individuals of the ODE model (9,970). The time for a single simulation of the resulting fluid model is 0.164s being several orders of magnitude faster than simulation of the corresponding agent-based model (407.49s).

3 Covid-19 in Denmark

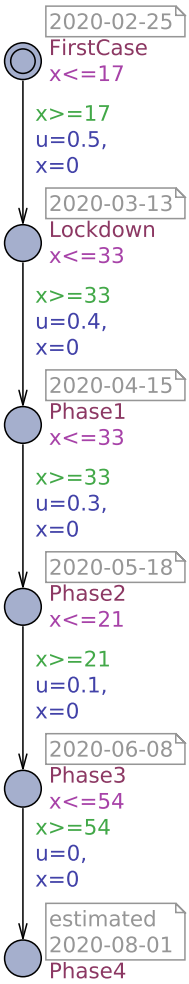


Fig. 6: Phases.

Covid-19 was first identified at December 2019 in Wuhan, China and from there quickly spread throughout the world. At February 27th the first case of Covid-19 was confirmed in Denmark.

On March 11th 2020 the Danish prime minister Mette Frederiksen announced a countrywide quarantine (lockdown) taking effect from March 13. The order closed down all non-essential public services, including daycare, primary- and secondary schools, upper secondary schools and universities. All non-essential public sector works was required to stay and work from home, the order urged the private sector to follow the same procedure. Closely follow on March 16th an order restring public gatherings of more than 10 people, as well as closing down shopping centers and stores where people are in close proximity, including bars, restaurants, fitness centers, hairdressers, dentists and shopping centers.

The lockdown was kept into effect until April 15th when a gradual reopening of the country started. The reopening was planned and approved by politicians in collaboration with the government, assisted by modeling and expert input by SSI. The plan consisted of 4 phases gradually lifting the quarantine:

Phase 1: daycare and primary schools (1.-5 grade) as well as hairdressers, dentists.

Phase 2: staring May 18th, included opening shopping centers, bars and restaurants (with reduced opening hours), secondary school, upper secondary schools, outdoer sports clubs churches and professional sports and athletics.

Phase 3: staring June 8th, included universities, public swimming pool, gyms, sports, tourist attractions, parties and larger gatherings (up to 500).

Phase 4: everything else including lifting the ban on public gatherings on more than 500 people. The phase is planned to start in August.

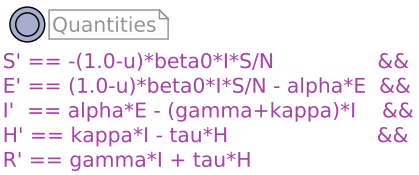
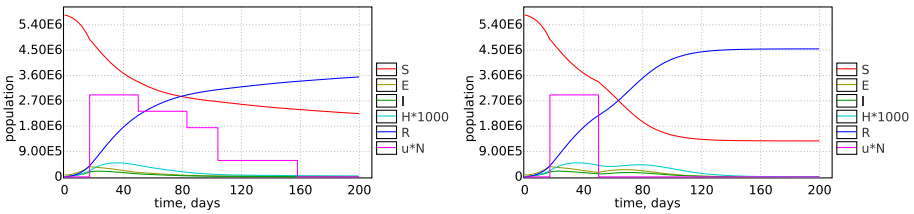


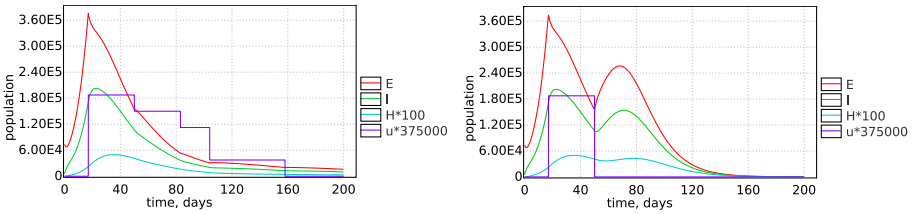
Fig. 7: SEIHR model for Denmark with quarantine.

Figure 6 is a timed automaton modeling the four phase of the gradual lifting of the quarantine. Here `u` is a variable between 0 and 1 giving the degree of quarantine, i.e. 0 corresponds to the complete lifting of quarantine. Figure 7 is a slighty modified version of the ODE SEIHR model, taking into account the degree of quarantine at any given point in time.



(a) Danish quarantine. (b) Early reopening.

Fig. 8: SEIHR trajectories.



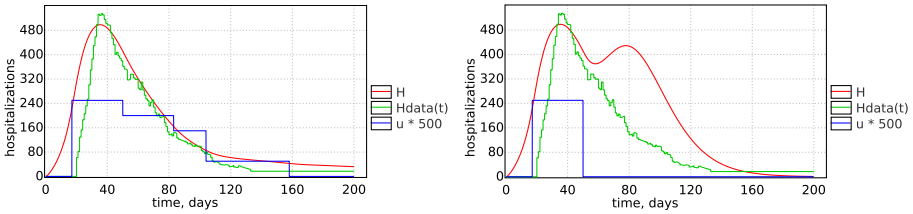
(a) Danish quarantine. (b) Early reopening.

Fig. 9: EIH trajectories.

Figures 8, 9 and 10 compare the planned lifting of quarantine with a hypothetical plan, where quarantine is completely lifted after phase 1. The three Figures focus on different subsets of `S`, `E`, `I`, `H` and `R`. In the first two figures the values of `H` as well as `u` are so small that a scaling has been used. From Fig. 8 we see that the planned lifting of quarantine slowly brings `R` close to the level needed for heard immunity in Denmark at approximately 3,267,000 (at the time of writing this paper we are in the middle of Phase 3). In contrast in the alternative plan with early reoponing heard immunity would have been achieved already now.

In Fig. 10 focus is on the number of hospitalized individuals. Here the predictions of the two models are compared to the actual Covid-19 hospitalized numbers as published by SST (Sundhedsstyrelsen)[4]. We see that the trajectory

[4] `https://www.sst.dk/da/corona/tal-og-overvaagning`



(a) Danish quarantine. (b) Early reopening.

Fig. 10: Hospitalizations.

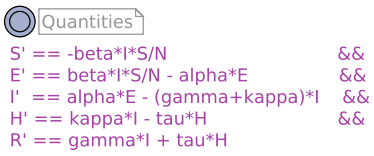
in Fig. 10(a) obtained from the ODE model under the planned quarantine phases is extremely close to the actual observed data. Most importantly we see that that maximum number of hospitalized individuals at any given point in time is less than 520 well below the capacity of Danish hospitals. In Fig. 10(b) we see that the early complete lifting of quarantine results in a small temporary increase in number of hospitalizations.

4 Family Routines in Cities

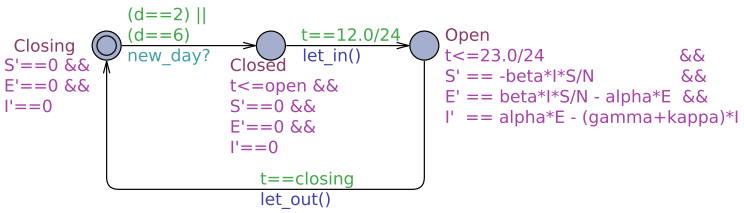
In this section, we consider a scenario focusing on the healthiness of a family with three members, a mother, a father and a son, living in Copenhagen. Besides living in Copenhagen (613,288 inhabitants) the members of the family spend considerable time at work and at school and occasionally enjoy some leisure activity. More precisely, the mother works at Maersk (estimated 2,000 employees at Esplanaden), the father works at ITU (estimated 2,300 employees) and the son goes at Vesterbro Ny school (752 pupils). As for leisure, the father is fanatic about FCK (FC Copenhagen) and has season tickets for home matches at Parken (capacity of 38,000 spectators) twice a week. During weekend the mother and father enjoy a dinner at one of restaurants in Nyhavn (some 5,000 people may gather there). In this scenario the son enjoys no leisure activities.

Now the city of Copenhagen as well as the 5 locations relevant for this particular family, i.e. Parken, ITU, Maersk, Vesterbro Ny School and Nyhavn, will have their own SEIHR ODE-based model. Each location has a population-size as well as a specific transition-rate for flow between susceptible (`S`) and exposed (`E`) reflecting the differences in being exposed at various locations. Fig. 11(a) is an instantiation of the ODE SEIHR model for Copenhagen, where `KBH_N` is the number of inhabitants in Copenhagen and `beta` is an array with a distinct value for each location e.g. `beta[kbh]` is the exposure rate for Copenhagen.

The SEIHR model for Parken is essentially the product of an ODE SEIHR model with a timed automaton [1] indicating the opening hours of Parken. In Fig. 11(b) we see that the opening hours is on Tuesdays (`d==2`) or Saturdays (`d==60`) between hour 12.00-23.00. Only in the `Open` location, the ODE for Parken is activated. The function (not shown here) `let_in()` (`let_out()`) will "transfer" a random number of spectators from (to) Copenhagen into (from) Parken upon



(a) ODE model for Copenhagen.



(b) Timed ODE model for Parken.

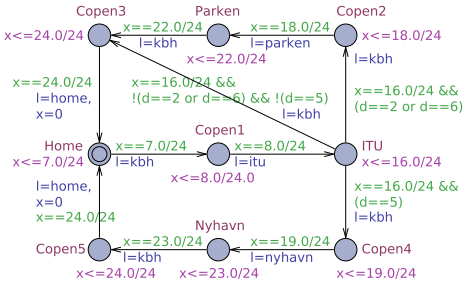
Fig. 11: SEIHR models for locations.

opening (closing). The SEIHR models of the remaining four locations are similar to that of Parken taking opening hours into account.

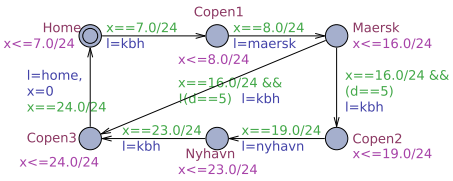
As for the three members of the family we will use two components: an agent-based model for recording the healthiness status and a timed auomtata describing the weekly itinerary reflecting work-hours and leisure activities. As an example Fig. 12(a) is timed automata describing the where-abouts of the Father over a week. Here `x` is a clock used to determine the precise timing of the various location-visits. As such we see that he leaves `Home` at 7 o'clock in the morning and reaches ITU at 8 o'clock. At 16 o'clock he leaves either for `Home` in order to go to `Nyhavn` or to `Parken`. Fig. 12(b,c) describes the itineraries for the mother and son. Fig. 12(d) is the agent-based SEIHR model used to describe the health status for each family member. Here we note that the rate for leaving `S` is a composite expression essentially picking the index of the array `beta` corresponding to current location of the family member (given by the expression `l[id]`, where `l` is an array holding the location of all three family members). Note here the final case, where the family member is at home (potentially the location with highest exposure), where the integer variable `Home_I` counts the number of family members being infected.

In Fig. 13 we see the result of a single simulation of the SEIHR model Copenhagen. In Fig. 13(a) we notice that twice week a significant part of the Copenhagen population going to Parken. Also, in Fig. 13(b) we note that the number of hospitalized peaks on day 42 at approximately 125 people.

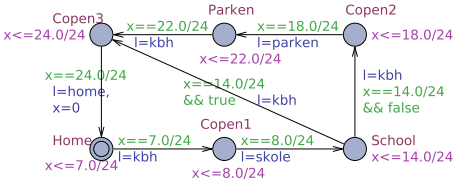
In Fig. 14 we estimate for each family member the probability that this person becomes exposed during a duration of 300 days. For the father (similar for the mother) the returned 95% confidence intervals are 0.738 ± 0.025, whereas the confidence interval for the son is only slightly below being 0.731 ± 0.025. Note that in all cases the exposure happens within the first 100 days – after this



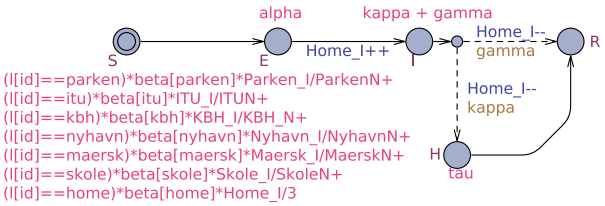
(a) TA itinerary for Father.



(b) TA itinerary for Mother.



(c) TA itinerary for Son.



(d) Agent model of the Health Status per each member.

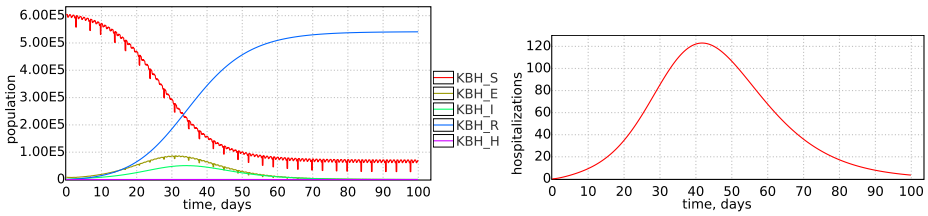
Fig. 12: Models for family members.

the number of infected in the various locations relevant for the family becomes too low.

The very marginal difference between the exposure of the son and the parents may seem strange as the son in this scenario do not enjoy any leisure activities (where the `beta` has been set substantially higher than at work-places). However, the son meet his father and mother regularly and for substantial amount of time at their home. Thus, we investigated the alternative scenario, where the son lives alone (still not enjoying leisure activities). The result becomes significantly different as show in Fig. 14(c), where the estimated 95% confidence interval becomes 0.661 ± 0.025. Thus the lesson is: it is not enough that you yourself stay away from highly exposed places, you should avoid spending long time-periods with others having this behaviour.

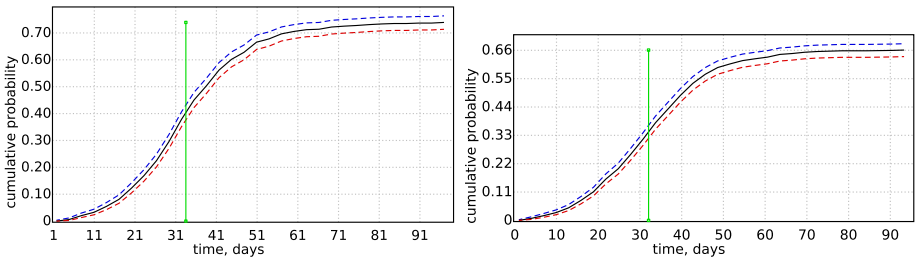
5 Super-Spreading and Bars

In this section we consider a scenario of a super-spreader, being a single individual who is infected and has a personal extremely high rate for spreading the virus to other people at the same location. For locations, we consider five bars (say at Nyhavn in Copenhagen) each with a capacity of 300 persons (out of



(a) SEIHR (b) Number of hospitalized.

Fig. 13: Copenhagen.

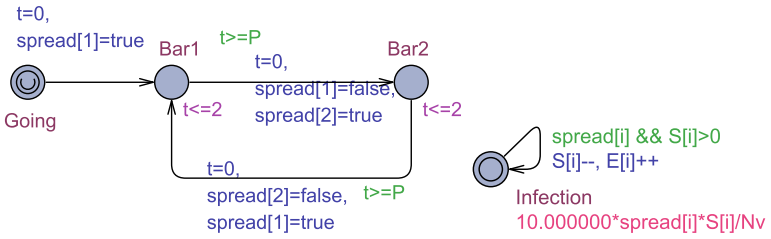


(a) Father: 0.738 ± 0.025 (95% CI). (b) Son *living alone*: 0.661 ± 0.025.

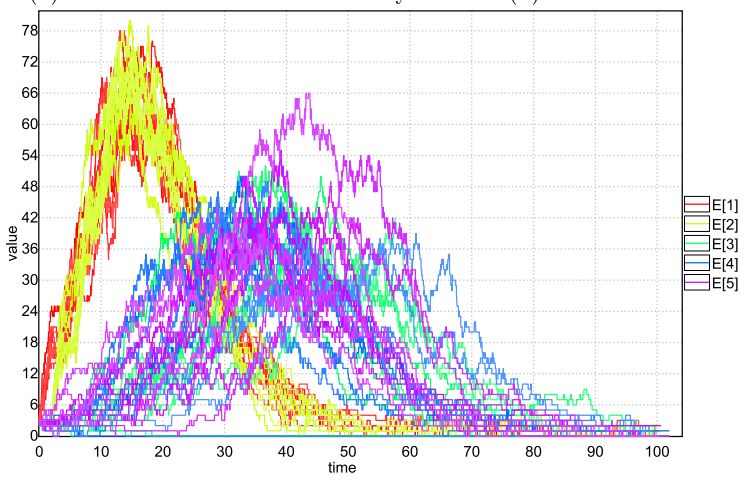
Fig. 14: Probabilities of becoming exposed within 300 days.

which 3 are assumed to be exposed already). For each bar we instantiate the aggregate CTMC model of Fig. 3(a). Thus a complete state will be captured by five arrays of counters `S`, `E`, `I`, `H` and `R`, e.g. `E[2]` will be the number of exposed people in bar number 2, `Bar[2]`.

In this scenario the super-spreader walks between bars in a periodic manner. In fact, to demonstrate the damage of the super-spreader only `Bar[1]` and `Bar[2]` are visited. The periodic behaviour of the super-spreader is given as a timed automaton in Fig. 15(a), where we note that the period is 2 days. `spread` is a Boolean array where `spread[i]` is **`true`** when the super-spreader is in `Bar[i]`. In Fig. 15(b) we see the extra reaction rule added to `Bar[i]`. We see that the rule cause a susceptible person to be exposed with extremely high rate (`beta`=10) but only if the super-spreader is in `Bar[i]`. In Fig. 15(c) we see 10 simulations tracking the number exposed people in each of the five bars over a period of 100 days. Clearly, there are many more people being exposed in the bars visited by the super-spreader in comparison to the other bars. Also the exposure happens much faster in these bars with a peak around day 13 compared to day 40. In fact, the expected number of people becoming exposed in `Bar[1]` is 72.58 ± 1.38 compared to that of `Bar[5]` being 48.04 ± 2.43.



(a) Timed Automaton for Itinerary (b) Infection Rule



(c) Evolution of number of exposed people in the five bars.

Fig. 15: Super-Spreader

6 Tracing Covid-19

One strategy for limiting the spread of epidemic diseases is containment: isolate infected people and rapidly determine who they infected and also isolate them. This strategy works very well if

1. we can discover initially infected people through testing, and
2. we can trace their interactions.

The smitte|stop smartphone application[5], mandated by the danish government [6], has recently made such a trace-and-isolate strategy possible. The application is build atop frameworks of Apple and Google for their respective smartphone platforms.

After being installed smitte|stop constantly emits unique IDs to nearby phones running smitte|stop. It also stores IDs of phones it has been in close contact with

[5] https://smittestop.dk/

[6] https://sum.dk/Aktuelt/Nyheder/Coronavirus/2020/Maj/~/media/Filer%20-%20dokumenter/01-corona/App/Politisk-aftale-om-smittessporingsappen.pdf

during a 15 minute interval. These ID exchanges are sufficient for infected individuals running the smitte|stop app to notify people whom they have plausible infected.

Part of the strategy of the danish government to stop a second wave relies on the smitte|stop app having a significant reduction in the number of new infections. However, the impact of such an app relies heavily on the adoption of the population. We thus demonstrate the use of UPPAAL to asses the impact of smitte|stop with varying adoption rates.

To model the effect of the smitte|stop application, we extend the agent based model presented in Section 2.3. However, to reduce the complexity, we make the following assumptions.

1. Infected using the app that has tested positive immediately warns other users,
2. people who receive a warning are tested immediately,
3. test results are received after a fixed amount of `testDelay` days, and
4. when a person gets infected that interaction is accurately and specifically captured by smitte|stop.

We believe that assumption 1-3 are quite realistic. However, we can see that assumption 4 is a crude simplification as it captures neither that 1. two persons might infect each-other, but not interacting long enough for smitte|stop to register it, nor 2. two persons who were infected each by separate third parties may by caught by the smitte|stop accidentally. Nonetheless, we expect these effects to be minor. Furthermore, we restrict ourselves to a 1.000-agent simulation due to the required computational effort, as noted in Section 2.3

Each individual in our world is modeled by two different automata: one automata (`Health`) models the health condition of each person while a second automaton (`Test`) models a persons behaviour in regards to testing policies.

Notice that the `Health` automaton (Figure 16) is a modified version of Figure 4: an extra location `Q` for quarantine has been added while the hospitalisation has been merged into the `R` location. We merged the hospitalisation into `R` as the number of hospitalisations is not interesting for this particular scenario. The extra `Q` location captures the (assumed) non-interaction of persons in quaranteen - a health state reached via a synchronisation on the person-specific `quarantine` channel. During location changes `Health` also updates a shared variable `s` to reflect its new health status such that the behaviour of the `Test` template is modified accordingly.

Figure 17 shows the `Test` automaton. Initially a choice is made as to whether this person uses smitte|stop. Afterwards (in `S`) we wait for the `Health` automaton to signal it has been exposed. Upon exposure an existing infected individual is selected (at random) as the source of the exposure- if no-one is infected, the edge guarded by `no_infectious ()` can be taken, in which case the source of infection is assumed to be external. Notice the increment of `frandom` and `fapp`: these are counters of how many infected are found using random testing and how many are found using the app.

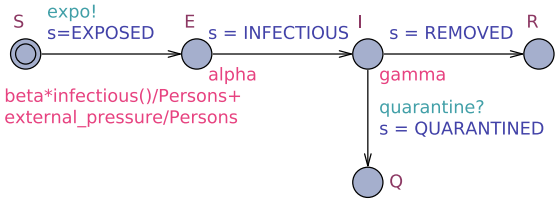


Fig. 16: `Health` of the smitte|stop model. The `external_pressure/Persons` models people might be exposed from outside (e.g. travels)

In the `E` location a person can either be selected for random testing through synchronisation on `test`, or via a warning from smitte|stop, modeled via synchronization on (`positive_c[who]?`). The automaton follows a similar pattern in both cases: we go to a location where it waits for a test result for `testDelay` days. If the test is positive (with probability weight `testInf ()`) then the individual moves to isolation – if the individual is using smitte|stop, the person emits a warning via smitte|stop (`pos`, received on `positive_c[who]` when `who` matches the ID of the transmitting individual). If the test is negative (with probability weight `testNInf()`) then for the random testing path (red rectangle) the individual returns to `E` while the current protocol (blue rectangle) for a smitte|stop case mandates a second test after a two day waiting/incubation period [7].

Remark 1. We mentioned above that people could be selected for random testing by synchronisation on `test`, but did not mention who controls this channel. It is controlled by an additional automaton that continuosly chooses a delay from an exponential distribution. After that delay it selects a person to be tested a uniformly. The exponential distribution has rate T where T is the amount of people tested each day.

To assess the potential effect of using smitte|stop, we estimated the number of people found using the app within 200 days, and the number of people found using random testing within 200 days for different adoption rates of smitte|stop, and for different levels of testing accuracies. In Table 1 we summarise the result of these simulations. Worth noting in Table 1:

1. The amount of infected using smitte|stop (smitte|stop column) does increase with higher adoption rate of the app,
2. the `Random`+smitte|stop column indicates that we do find more infected in total with higher adoption rates,
3. A superficial scan over Table 1 could easily lead to the conclusion that smitte|stop is not useful - with the highest adoption rate (highlighted in Table 1) it only finds 1.14 infected. However, it should be taken into consideration that random testing only find 4.59 infected, and random testing needs to find an infected individual before smitte|stop can alert people. Increasing test capacity should have a positive effect on both the number of

[7] `https://smittestop.dk/spoergsmaal-og-svar`

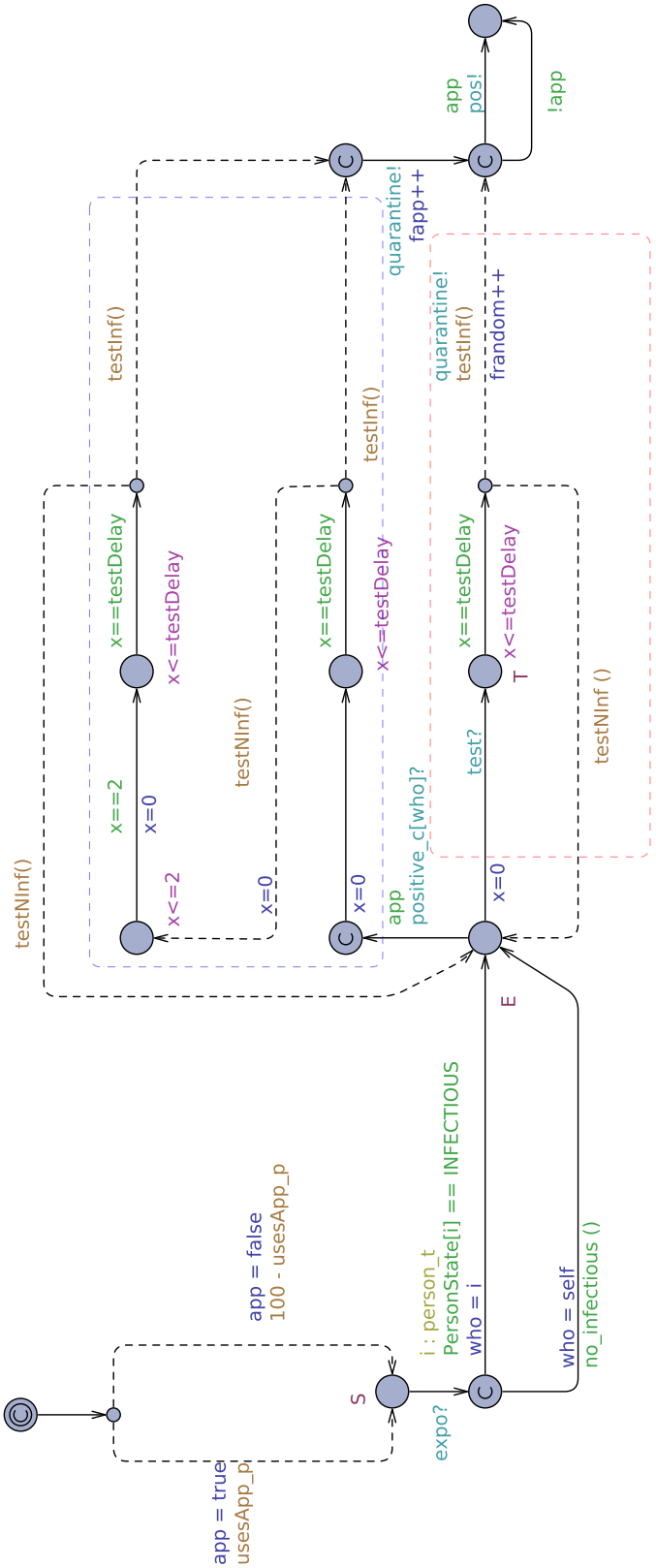


Fig. 17: Excerpt of `Test` of the smitte|stop model. The area in the blue rectangle is the smitte|stop testing procedure and the area red rectangle is the random test procedure.

Persons	TestAcc	Tests/day	%smitte|stop	Random	smitte|stop	Random+smitte|stop
1000	.25	2	.25	1.75	0.02	1.79
1000	.25	2	.50	1.75	0.07	1.82
1000	.25	2	.75	1.70	0.17	1.87
1000	.50	2	.25	3.35	0.05	3.40
1000	.50	2	.50	3.25	0.23	3.48
1000	.50	2	.75	3.22	0.59	3.81
1000	.75	2	.25	4.72	0.11	4.83
1000	.75	2	.50	4.69	0.49	5.18
1000	.75	2	.75	4.59	1.14	5.73

Table 1: Estimation Data for the smitte|stop model: `Persons` is total size of the population, `TestAcc` is the accuracy of the tests, `Tests/day` is the amount of people tested using random testing per day, %smitte|stop is the percentage of people using smitte|stop, `Random` is the number of infected found using random testing while smitte|stop is the amount found using the smitte|stop app. The `Random` and smitte|stop columns are estimated over 10000 runs.

infected found by random testing and the number of infected found using smitte|stop.

7 Conclusion

In this paper we have demonstrated how Uppaal SMC may be used to model the ongoing Covid-19 epidemic in several ways. The span of models allows for a range of analyses to be made. This includes analysis at population level for crucial estimation of the sufficiency of hospital capacity. Also analyses at the level of individuals is possible using fluid models, where consequences of various social behavioural patterns may be predicted.

We are convinced that the graphical and rich modelling formalism of Uppaal SMC has been crucial for the rapid speed by which these models have been constructed and analysed. As for the analyses the current efficiency of Uppaal SMC has proven adequate with respect to the scenarios considered. However, we have a number of ideas for optimizations (e.g. precomputing the solutions to the ODE component of fluid models, sweeping of parameters, exploting cluster computing facilities) that will be needed for scaling to more complex scenarios.

In the NNF project domain specific modeling-notations and ways of visualizing results more suited for end-users (doctors, politicians, etc.) are being developed. We plan to support these notations.

One overall important aspect that we have not considered is the estimation of parameters and initial condition based on real observed measurements. In the sister NNF project a number of approaches for this has been examined.

References

1. R. Alur and D. L. Dill. A theory of timed automata. *Theor. Comput. Sci.*, 126(2):183–235, 1994.
2. L. Bortolussi and J. Hillston. Fluid model checking. In M. Koutny and I. Ulidowski, editors, *CONCUR 2012 - Concurrency Theory - 23rd International Conference, CONCUR 2012, Newcastle upon Tyne, UK, September 4-7, 2012. Proceedings*, volume 7454 of *Lecture Notes in Computer Science*, pages 333–347. Springer, 2012.
3. A. David, K. G. Larsen, A. Legay, M. Mikučionis, and D. B. Poulsen. Uppaal SMC tutorial. *Int. J. Softw. Tools Technol. Transf.*, 17(4):397–415, 2015.
4. Jensen, Jørgensen, Larsen, Mikucionis, and Poulsen. Fluid models in uppaal for covid-19 – full models. `https://github.com/DEIS-Tools/uppaal-models/tree/master/CaseStudies/Covid-19`.