

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

Modelování a simulace
Predikce epidemie COVID-19

7. prosince 2020

Radek Švec (xsvecr01)
Martin Kostelník (xkoste12)

1 Úvod

Tématem práce jsou epidemiologické modely na makroúrovni, konkrétně předpovídání šíření pandemie COVID-19 na území Itálie a České republiky. Tato práce implementuje již existující model vytvořený Giulio Giordanem a dalšími [3], který předpovídá vývoj pandemie na území Itálie a opakuje provedené experimenty. Smyslem projektu bude předpovědět šíření nemoci COVID-19 na území České republiky v následujících měsících, vzít v úvahu různá opatření proti šíření a zhodnotit, zda byly dosavadně zavedené opatření dostačující, případně zda bylo rozvolnění těchto opatření správným rozhodnutím.

1.1 Autoři a zdroje informací

Autory tohoto projektu jsou:

- Radek Švec, xsvecr01@stud.fit.vutbr.cz
- Martin Kostelník, xkoste12@stud.fit.vutbr.cz

Jak bylo řečeno v úvodu, samotný model byl převzat z odborné literatury [3]. Při pokusu aplikovat model na prostředí České republiky byly použity data z oficiálních stránek ministerstva zdravotnictví České republiky [2].

1.2 Ověření validity

Validitu modelu jsme ověřovali zkoumáním převzatého modelu, zejména porovnáním výsledků prováděných simulací s reálným vývojem pandemie v Itálii. Validita modelu v prostředí České republiky je opět ověřována porovnáním simulačních dat s reálnými.

2 Rozbor tématu a použitých metod/technologií

2.1 Popis použitých postupů

Model byl převzat z odborného článku [3] publikovaného v časopise Nature Medicine¹. Bylo tedy nutné model detailně nastudovat a zkontrolovat dosažené výsledky. Tento konkrétní model nám přišel na modelování pandemie vhodný, jelikož rozděluje společnost do většího množství skupin a počítá s různě těžkými projevy nemoci.

2.2 Popis použitých technologií

Pro provedení experimentů bylo nutné vytvořit simulační model, který jsme implementovali v jazyce C++. Nebylo nutné použít žádné nestandardní knihovny, jelikož se model skládá pouze z řady diferenciálních rovnic. Pro generování grafů byl použit nástroj GNUPlot².

3 Koncepce abstraktního modelu

Model zvaný SIDARTHE je rozšířením jednoduššího SIR modelu, který rozděluje populaci na tři skupiny lidí (náchylné, infikované a vyléčené) o dalších pět skupin. Vzniká nám tedy následujících osm skupin.

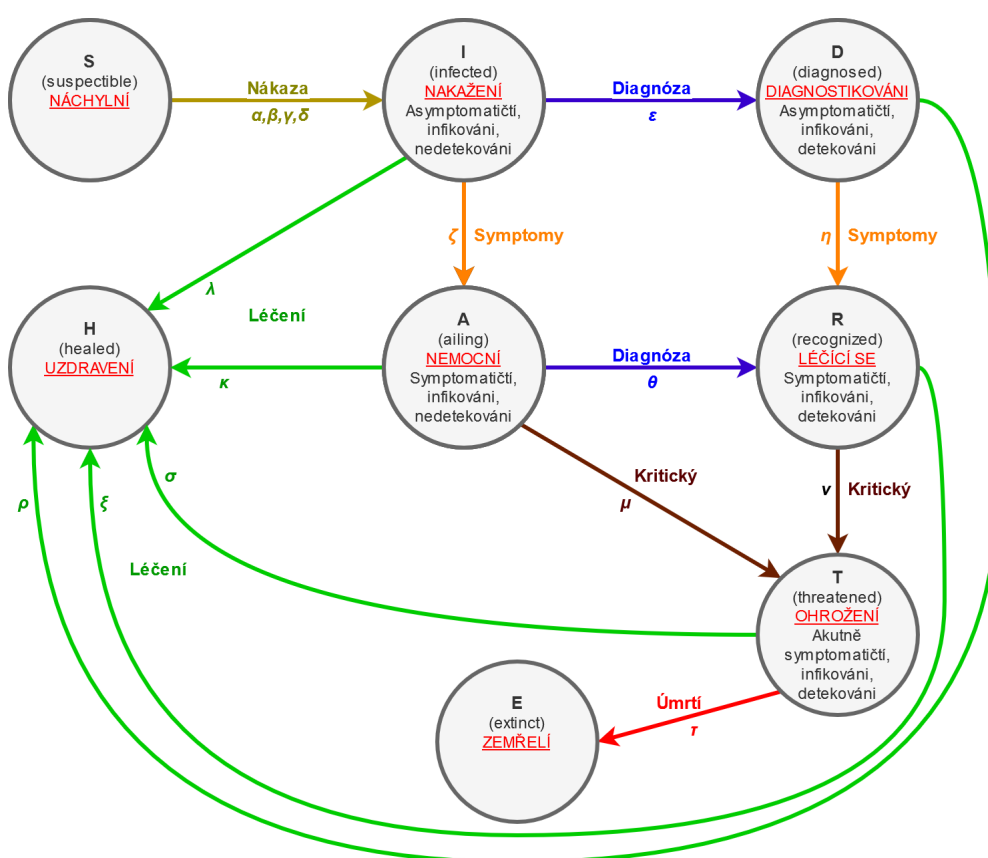
- S – Náchylní (Susceptible) k onemocnění.
- I – Nakažení (Infected), kteří jsou asymptotičtí, ale vir nebyl detekován.

¹<https://www.nature.com/nm/>

²<http://www.gnuplot.info/>

- D – Diagnostikovaní (Diagnosed), kteří jsou asymptomaticí, avšak vir u nich byl detekován.
- A – Nemocní (Ailing), kteří jsou symptomaticí, ale nebyli testováni.
- R – Léčící se (Recognized), kteří mají symptomy i u nich byl vir detekován.
- T – Životně ohrožení (Threatened), kteří mají těžký průběh nemoci a je u nich vyžadována intenzivní péče.
- H – Vyléčení (Healed).
- E – Mrtví (Extinct).

Obrázek 1: Grafické znázornění modelu, převzato z [3] a přeloženo



Tento model nepočítá s možností, že se jedinec může virem nakazit opakovaně. I když bylo prokázáno, že se jedinec znovu nakazit může [4], není stále přesně známo, jakou imunitu si jedinec vybuduje. Takto nakažených je navíc velice malé množství, například ve Švédsku v polovině října to bylo pouhých 150 případů [1], tudíž je v modelu zanedbáváme.

3.1 Matematické vyjádření modelu

Model můžeme matematicky popsat celkem osmi diferenciálními rovnicemi, které popisují vývoj jedinců v každé skupině jako zlomek celkové populace státu v čase.

$$S'(t) = -S(t)(\alpha I(t) + \beta D(t) + \gamma A(t) + \delta R(t)) \quad (1)$$

$$I'(t) = S(t)(\alpha I(t) + \beta D(t) + \gamma A(t) + \delta R(t)) - (\epsilon + \zeta + \lambda)I(t) \quad (2)$$

$$D'(t) = \epsilon I(t) - (\eta + \rho)D(t) \quad (3)$$

$$A'(t) = \zeta I(t) - (\theta + \mu + \kappa)A(t) \quad (4)$$

$$R'(t) = \eta D(t) + \theta A(t) - (\nu + \xi)R(t) \quad (5)$$

$$T'(t) = \mu A(t) + \nu R(t) - (\sigma + \tau)T(t) \quad (6)$$

$$H'(t) = \lambda I(t) + \rho D(t) + \kappa A(t) + \xi R(t) + \sigma T(t) \quad (7)$$

$$E'(t) = \tau T(t) \quad (8)$$

Kromě množství jedinců ve skupinách, se v rovnicích vyskytují následující parametry:

- $\alpha, \beta, \gamma, \delta$ představují pravděpodobnost přenosu viru mezi náchylným jedincem a jedincem infikovaným, diagnostikovaným, nemocným a léčícím se v tomto pořadí, vynásobenou průměrným počtem kontaktů za den.
- ϵ, θ představují pravděpodobnost detekce viru u nakažených a nemocných.
- ζ, η – představují pravděpodobnost, že se u nakaženého a diagnostikovaného jedince vyskytnou symptomy nemoci.
- μ, ν představují pravděpodobnost, že se z nemocného a léčícího se jedince stane kriticky ohrožený jedinec.
- τ představuje úmrtnost.
- $\lambda, \kappa, \xi, \rho$ a σ představují pravděpodobnost, že se nakažený, nemocný, léčící se, diagnostikovaný, nebo ohrožený jedinec z nemoci vyléčí.

4 Architektura simulačního modelu

Simulační model je implementován v programovacím jazyce C++. Pro jeho vytvoření nebylo nutné použít jakékoliv nestandardní knihovny, jelikož samotné jádro simulace je pouhé iterování výpočtů.

V programu existují dvě struktury, ve kterých jsou uloženy všechny data pro provádění simulací. První struktura v sobě uchovává počty jedinců ve skupinách a další informace o území, na kterém se pokusy provádí. V této strukture je také implementováno chování modelu, které za pomoci druhé struktury a aktuálních informací předpovídá vývoj nemoci. Druhá struktura popisuje aktuální chování nemoci, uchovává totiž v sobě všechny parametry pravděpodobnostního charakteru, které se využívají v simulacích pro určité časové období a pomáhají nám lépe definovat následky opatření, nebo jejich rozvolňování.

U tohoto modelu je problematické zadání všech parametrů uživatelem. Uživatel by musel zadat velké množství hodnot a tento přístup by byl velice náchylný k chybám. Rozhodli jsme se proto všechny provedené experimenty uložit přímo do zdrojového kódu programu. Jednotlivé experimenty je možné spouštět různými cíly definovanými v souboru `Makefile`. Konkrétní příkazy pro spuštění experimentů jsou v kapitole 5.

Po provedení simulace program vypíše výsledné množství jedinců v každé ze skupin na standardní výstup a zároveň data z každého dne zapisuje do souboru `data.txt` pro možnost generování grafů.

4.1 Mapování abstraktního modelu do modelu simulačního

Diferenciální rovnice popsané v kapitole 3.1 jsou implementovány ve funkci `predict()`, která je uložena ve struktuře s názvem `SIDARTHE`. Tato funkce bere na vstupu tři parametry. Prvním je struktura `Disease` popisující nemoc, druhým a třetím parametrem je počáteční a koncový den simulace. Funkce poté iterativně počítá přírůstky jedinců do každé ze skupin a každá iterace představuje jeden den.

5 Podstata simulačních experimentů a jejich průběh

Experimentování s modelem je rozděleno do dvou částí. V první části se pokusíme zopakovat některé experimenty prováděny autory modelu [3]. V druhé části aplikujeme model na prostředí České republiky, kde budeme zkoumat vývoj pandemie v období od 1. listopadu 2020 do 31. ledna 2021 a pokusíme se zhodnotit zavedené opatření a případně doporučit zavedení dalších opatření v blízké době.

6 Zopakování experimentů

V této kapitole se pokusíme zopakovat některé z již provedených experimentů a porovnat naše výsledky s referenčními.

6.1 Experiment 1

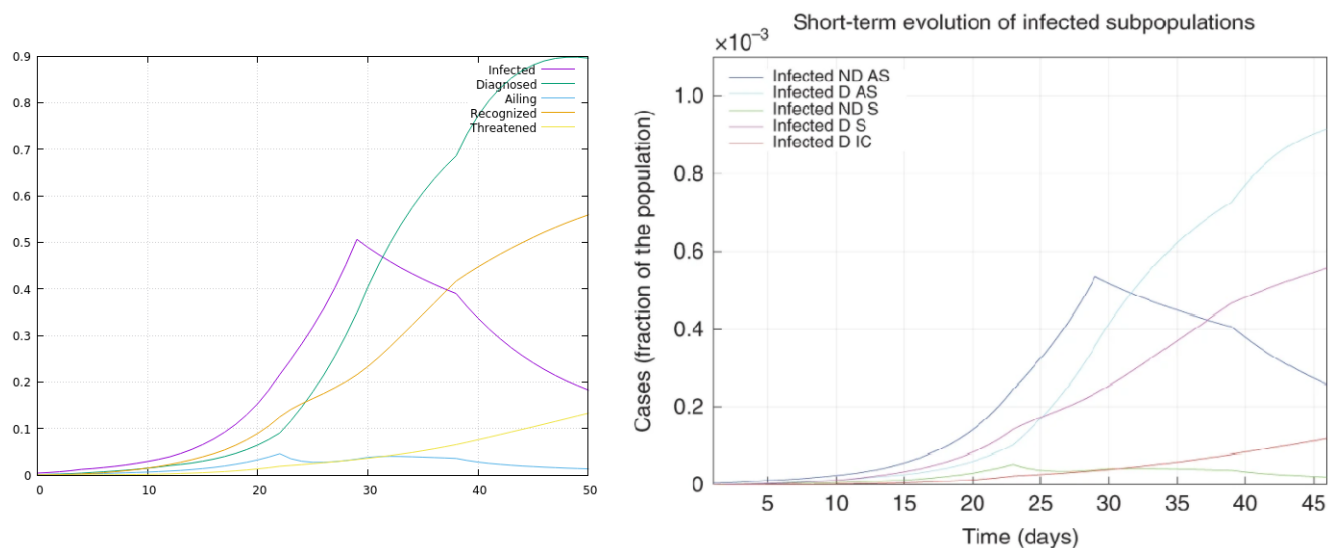
Experiment lze replikovat spuštěním příkazu `make run1`.

Tento experiment se snaží simulovat šíření nemoci v Itálii od 20. února 2020 do 5. dubna 2020. Celkem se tedy zkoumá čtyřicet šest dní. Během této doby byla v Itálii implementována řada opatření, jejichž následky můžeme sledovat v obrázku 2.

Výstup programu:

TODO: VYSTUP PROGRAMU PRO EXPERIMENT 1

Obrázek 2: Porovnání výsledků



Na levém obrázku můžeme vidět graf generovaný z výsledků našeho experimentu. Na pravém obrázku je graf referenční.

6.2 Experiment 2

Tento experiment navazuje na experiment 6.1. Pokouší se simulovat šíření nemoci na větším časovém úseku.

7 Aplikování modelu na Českou republiku

8 Shrnutí simulačních experimentů a závěr

asdfsdfasdfsdfasdfsdfasdfsdfasdfsdfasdfsdfasdfsdf asdfsdfasdfsdfasdfsdfasdfsdf

Literatura

- [1] Akpan, N.: Why coronavirus reinfections are happening. 2020.
Dostupné z: <https://www.nationalgeographic.com/science/2020/12/why-coronavirus-reinfections-are-happening/>
- [2] ČERNÝ, M.: Znakové sady v typografických systémech. 1999.
- [3] Giordano, G.; aj.: Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. , č. 26, 2020: s. 855–860.
Dostupné z: <https://doi.org/10.1038/s41591-020-0883-7>
- [4] Kerwin, H.; aj.: Genomic evidence for reinfection with SARS-CoV-2: a case study. 2020.
Dostupné z: [https://doi.org/10.1016/S1473-3099\(20\)30764-7](https://doi.org/10.1016/S1473-3099(20)30764-7)