

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 předpověď šíření pandemie COVID-19 na území Itálie a České republiky. Tato práce implementuje již existující model, který předpovídá vývoj pandemie na území Itálie a opakuje provedené experimenty. Tento model byl vytvořen Guiliou Giordanem a dalšími [2]. Smyslem projektu je předpověď šíření nemoci COVID-19 na území České republiky v následujících měsících, aplikování různých opatření proti šíření a zhodnocení, zda byly dosavadní 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 [2]. Při pokusu aplikovat model na prostředí České republiky byly použity data z oficiálních stránek Ministerstva zdravotnictví České republiky¹.

1.2 Ověření validity

Validitu modelu jsme ověřovali zkoumáním převzatého modelu, zejména porovnání 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

Model byl převzat z odborného článku [2] publikovaného v časopise Nature Medicine². Všechny data o šíření viru COVID-19 v České republice byly převzaty z oficiálních stránek Ministerstva zdravotnictví České republiky³.

2.1 Popis použitých postupů

Bylo 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 třeba použít nestandardní knihovny, jelikož se model skládá pouze z řady diferenciálních rovnic a parametrů. Pro vykreslení grafů byl použit nástroj GNUPlot⁴.

¹<https://onemocneni-aktualne.mzcr.cz/covid-19>

²<https://www.nature.com/nm/>

³<https://onemocneni-aktualne.mzcr.cz/covid-19>

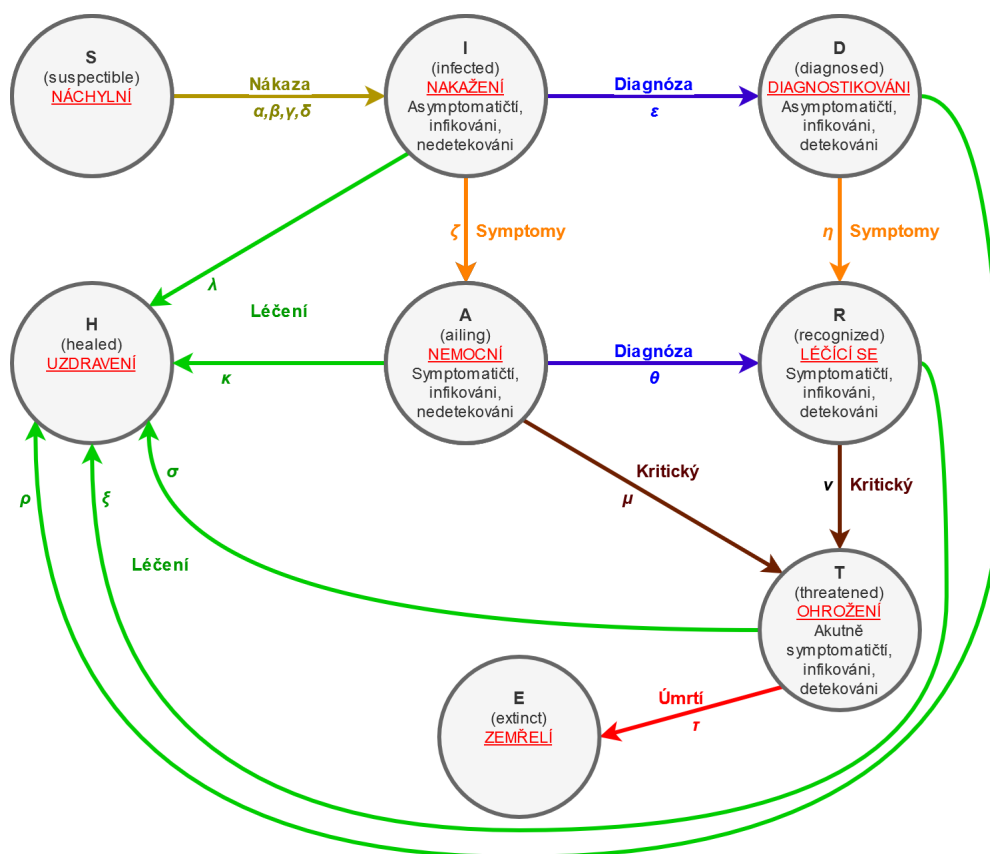
⁴<http://www.gnuplot.info/>

3 Koncepce abstrakčního modelu

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

- S - Náchylní (Susceptible) k onemocnění.
- I - Nakažení (Infected) - asymptomatictí, vir u nich nebyl detekován.
- D - Diagnostikovaní (Diagnosed) - asymptomatictí, vir u nich byl detekován.
- A - Nemocní (Ailing) - symptomatictí, nebyli testováni.
- R - Léčící se (Recognized) - symptomatictí, byl u nich detekován vir.
- T - Životně ohrožení (Threatened) - mají těžký průběh nemoci, vyžadují intenzivní péči.
- H - Vyléčení (Healed).
- E - Mrtví (Extinct).

Obrázek 1: Grafické znázornění modelu, převzato z [2] a přeloženo



Tento model nepočítá s možností, že se jedinec může virem nakazit opakovaně. I když bylo prokázáno, že se jedinec znovu nakazit může [3], 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 sto padesát 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 t .

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

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

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

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

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

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

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

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

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

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

4 Architektura simulačního modelu

V implementovaném programu existují dvě struktury, kde jsou uložena všechna data pro provádění simulací. První struktura v sobě uchovává počty jedinců ve skupinách a další informace o území, na kterém se experimenty provádějí. V této struktuře je také implementováno chování modelu, které za pomoci druhé struktury a aktuálních informací předpovídá vývoj nemoci. To se provádí iterováním výpočtů. Druhá struktura popisuje aktuální chování nemoci, obsahuje tedy všechny parametry pravděpodobnostního charakteru, jež se využívají v simulacích pro určité časové období a pomáhají k lepší definici následků zpřísnění nebo rozvolnění opatření.

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. Proto jsou údaje pro všechny provedené experimenty uloženy přímo ve zdrojovém 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 dostává 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

Cílem experimentování je předpovědět šíření pandemie COVID-19 na území České republiky, zhodnotit zavedená opatření, či zda bylo rozvolnění těchto nebo jiných opatření tím správným přístupem.

5.1 Postup experimentování

Experimentování s modelem je rozděleno do dvou částí. V části 5.2 se pokusíme zopakovat některé experimenty prováděné autory modelu [2]. Ve části 5.3 aplikujeme model na prostředí České republiky, kde budeme zkoumat vývoj pandemie v období od 1. listopadu 2020 do 31. ledna 2021. Pro dosavadní vývoj budeme upravovat parametry modelu tak, aby v jednotlivých dnech co nejvíce odpovídaly reálným datům, které byly získány z oficiálních stránek Ministerstva zdravotnictví České republiky⁵.

5.2 Zopakování experimentů

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

5.2.1 Experiment 1

Experiment lze replikovat spuštěním příkazu `make run1`.

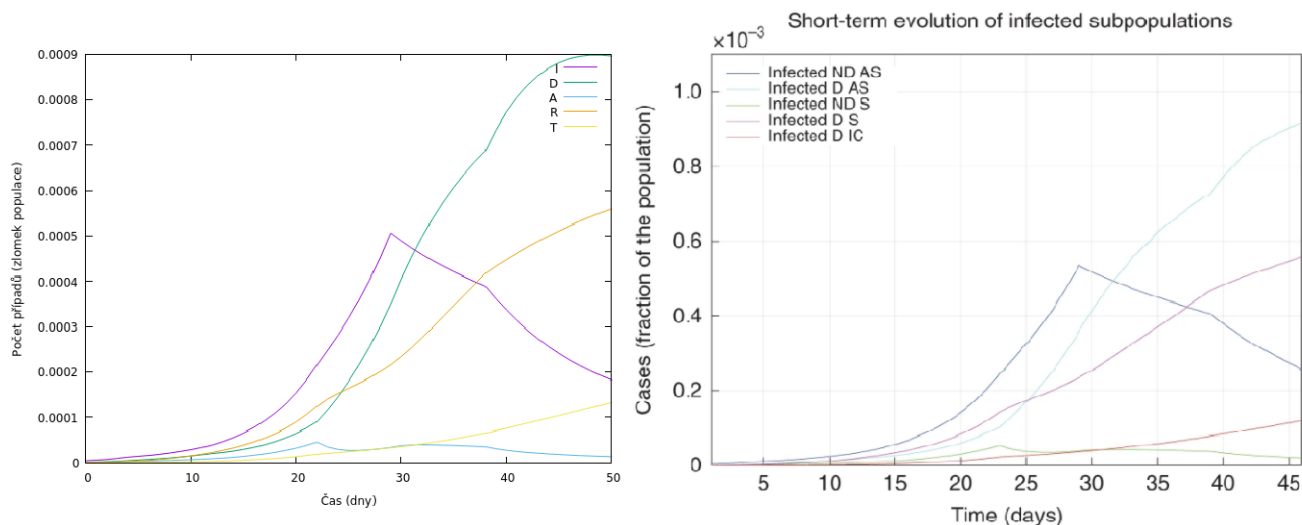
Tento experiment se snaží simulovat šíření nemoci v Itálii od 20. února 2020 do 5. dubna 2020. Celkem se tedy zkoumá čtyřicet šest dní. Během této doby byla v Itálii implementována řada opatření, jejichž následky můžeme sledovat v obrázku 2.

Výstup programu:

```
----- RESULTS -----  
S = 99.6836 % (59810139/60000000)  
I = 0.0182 % (10941/60000000)  
D = 0.0896 % (53739/60000000)  
A = 0.0013 % (793/60000000)  
R = 0.0559 % (33533/60000000)  
T = 0.0133 % (7961/60000000)  
H = 0.1363 % (81802/60000000)  
E = 0.0018 % (1089/60000000)
```

⁵<https://onemocneni-aktualne.mzcr.cz/covid-19>

Obrázek 2: Porovnání výsledků experimentu 5.2.1



Na levém obrázku můžeme vidět graf generovaný z výsledků našeho experimentu. Na pravém obrázku je graf referenční (převzat z [2]).

5.2.2 Experiment 2

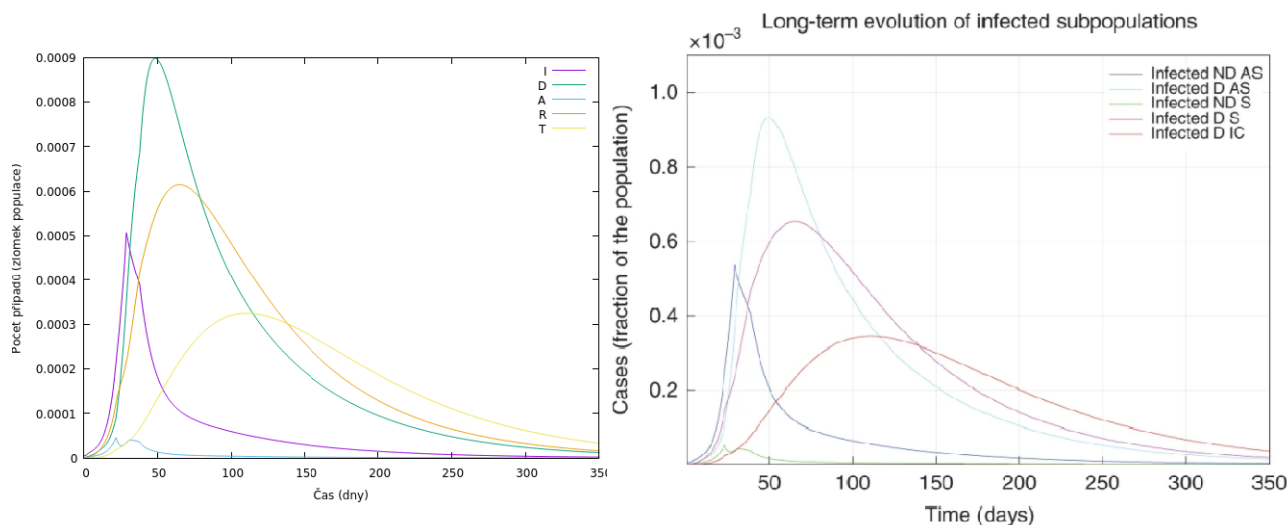
Experiment lze replikovat spuštěním příkazu `make run2`.

Tento experiment navazuje na experiment 5.2.1. Pokouší se simulovat šíření nemoci na větším časovém úseku. Celkem 350 dní od 20. února 2020.

Výstup programu:

```
----- RESULTS -----
S = 99.4270 % (59656223/60000000)
I = 0.0002 % (110/60000000)
D = 0.0012 % (706/60000000)
A = 0.0000 % (7/60000000)
R = 0.0016 % (960/60000000)
T = 0.0033 % (1972/60000000)
H = 0.5113 % (306791/60000000)
E = 0.0554 % (33228/60000000)
```

Obrázek 3: Porovnání výsledků experimentu 5.2.2



Na levém obrázku můžeme vidět graf generovaný z výsledků našeho experimentu. Na pravém obrázku je graf referenční (převzat z [2]).

5.3 Aplikování modelu na Českou republiku

V této kapitole je popsáno, jak jsme model aplikovali na prostředí České republiky. Bylo nutné vybrat vhodné období pro experimentování a zdokumentovat všechna vládní opatření aplikovaná v tomto období nebo případné rozvolňování těchto opatření.

5.3.1 Experiment 1

Tento experiment lze replikovat příkazem `make run`.

Experiment se pokouší předpovědět šíření nemoci v období od 1. listopadu 2020 do 31. ledna 2021. Průběh experimentů je rozdělen do tří menších období dle měnících se opatření.

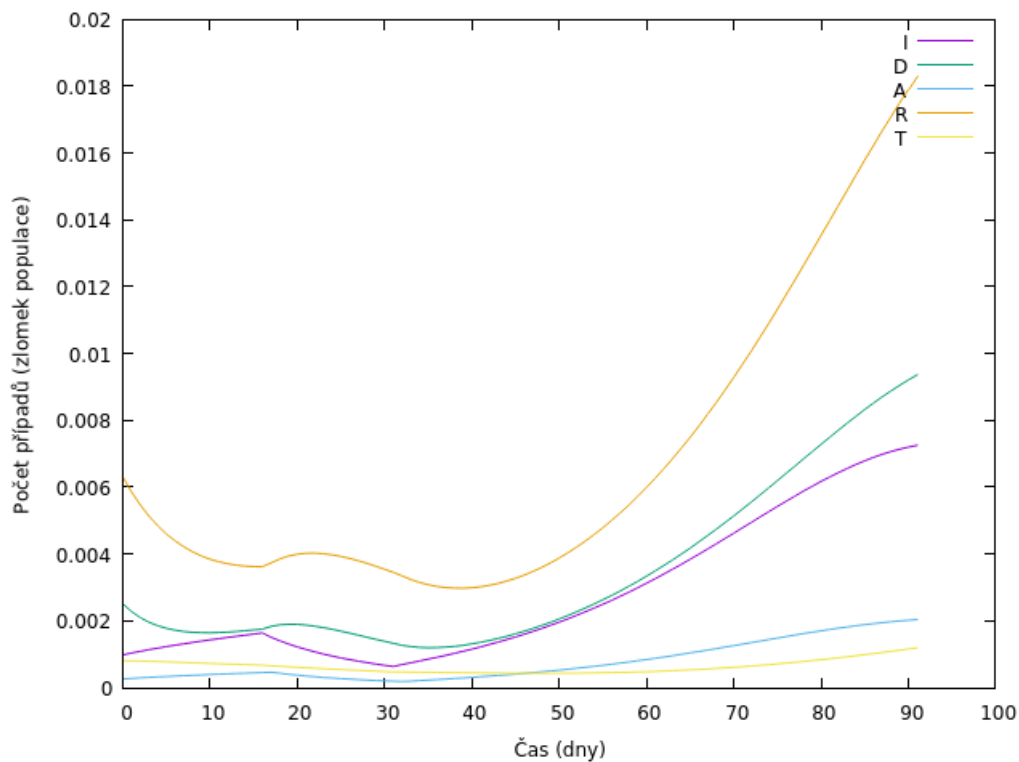
1. 11. 2020 - 17. 11. 2020 - Pro toto období byly parametry nastaveny tak, aby výsledky co nej přesněji kopírovaly reálná data o nakažených.
18. 11. 2020 - 2. 12. 2020 - V tomto období došlo k různým opatřením, největšími faktory jsou však omezení počtu lidí v obchodech a návrat žáků prvních a druhých tříd, návrat závěrečných ročníků středních škol a návrat žáků druhého stupně základních škol.
3. 12. 2020 - 31. 1. 2021 - Zde je samotná předpověď, jak by se nákaza mohla šířit.

Výstup programu:

```
----- RESULTS -----
S = 84.1108 % (8991444/10690000)
I = 0.7247 % (77468/10690000)
D = 0.9355 % (100000/10690000)
A = 0.2032 % (21724/10690000)
R = 1.8282 % (195437/10690000)
T = 0.1190 % (12716/10690000)
```

H = 11.9061 % (1272765/10690000)
 E = 0.1725 % (18442/10690000)

Obrázek 4: Graf šíření nemoci



5.3.2 Experiment 2

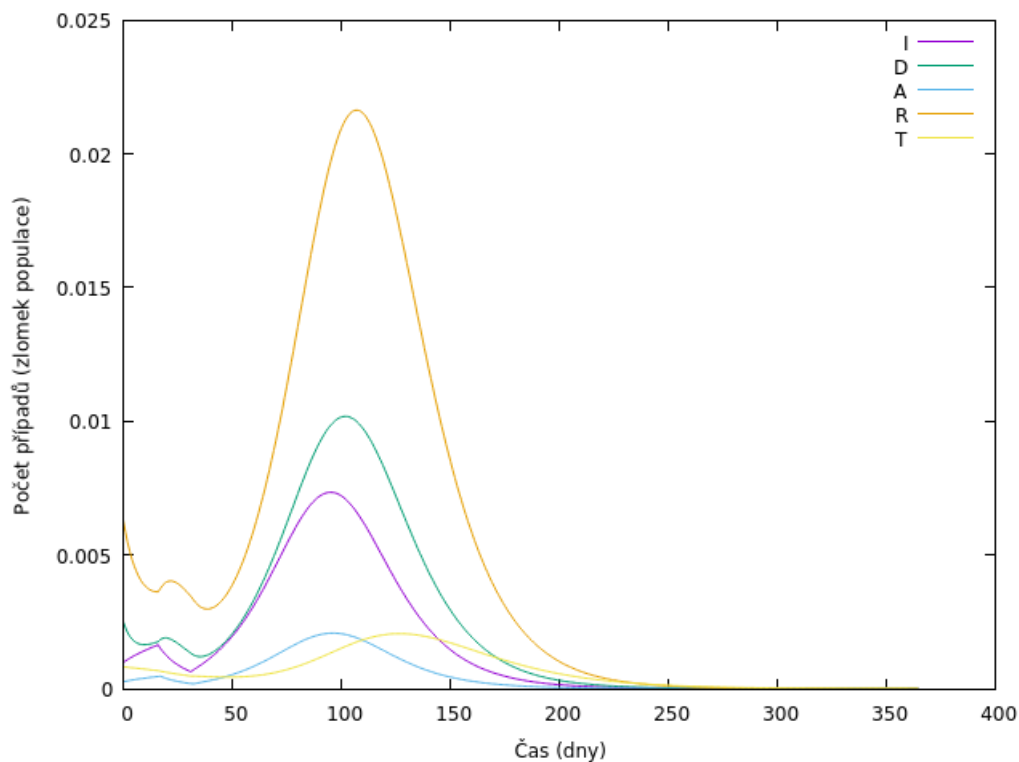
Tento experiment lze replikovat příkazem `make run3`.

Experiment navazuje na experiment 5.3.1, avšak ukazuje, jak by se nemoc mohla šířit po dobu jednoho roku. Mluvíme tedy o období od 1. listopadu 2020 do 31. října 2021.

Výstup programu:

```
----- RESULTS -----
S = 70.2598 % (7510769/10690000)
I = 0.0000 % (0/10690000)
D = 0.0000 % (1/10690000)
A = 0.0000 % (0/10690000)
R = 0.0000 % (5/10690000)
T = 0.0003 % (27/10690000)
H = 29.1180 % (3112718/10690000)
E = 0.6219 % (66477/10690000)
```


Obrázek 5: Graf šíření nemoci



5.3.3 Experiment 3

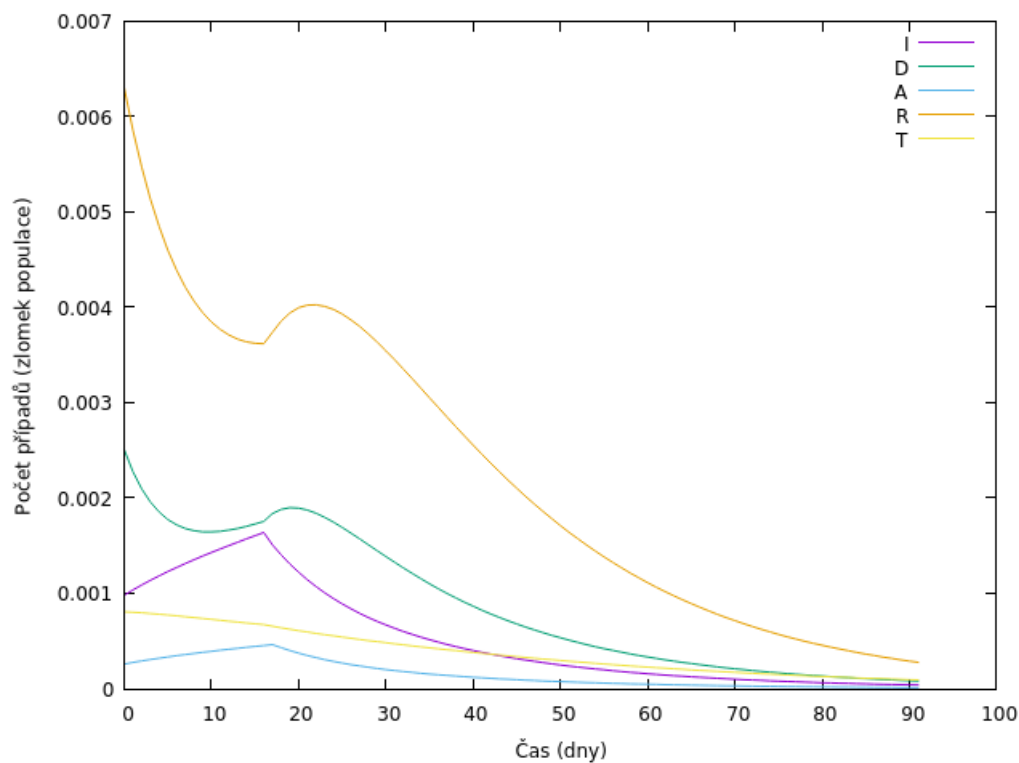
Tento experiment lze replikovat příkazem `make run4`.

Zde se podíváme, jak by šíření vypadalo v období od 1. listopadu 2020 do 31. ledna 2021 v případě, že by vláda nezmírnila opatření otevřením restaurací a služeb dne 3. prosince 2020.

Výstup programu:

```
----- RESULTS -----
S = 94.2215 % (10072274/10690000)
I = 0.0038 % (406/10690000)
D = 0.0080 % (853/10690000)
A = 0.0011 % (120/10690000)
R = 0.0277 % (2956/10690000)
T = 0.0092 % (979/10690000)
H = 5.6100 % (599710/10690000)
E = 0.1188 % (12698/10690000)
```

Obrázek 6: Graf šíření nemoci



5.3.4 Experiment 4

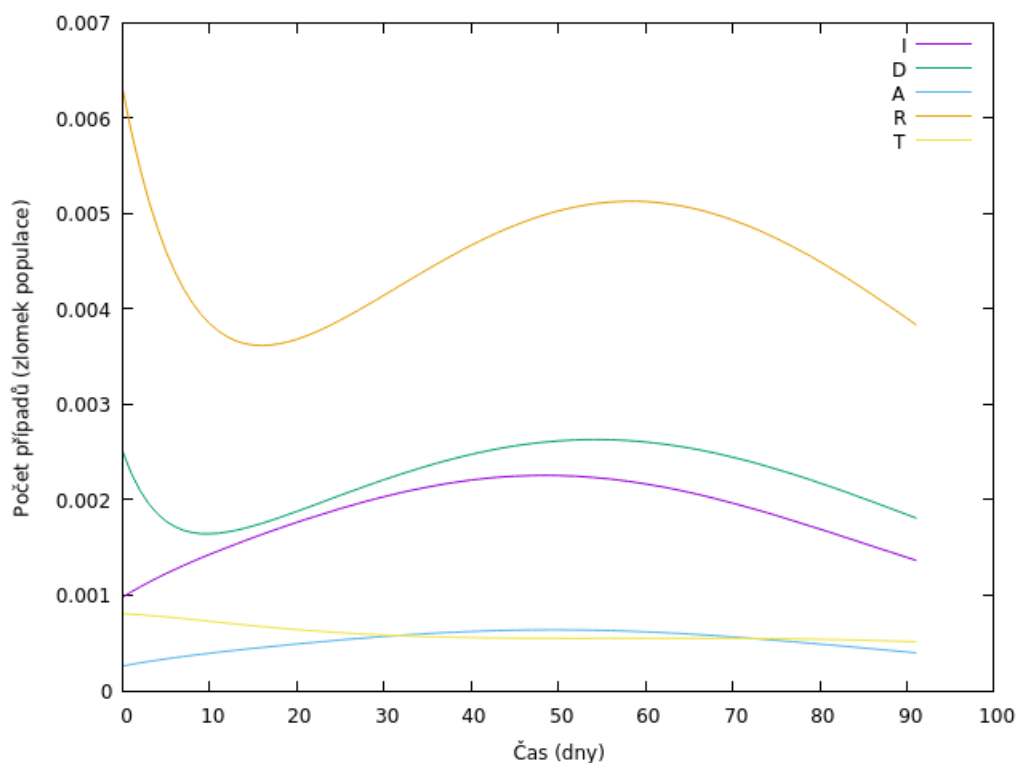
Tento experiment lze replikovat příkazem `make run5`.

Poslední experiment ukazuje, jak by mohly následky vypadat v případě, že by vláda nezmírnila žádná opatření, avšak neomezila maximální počet lidí v obchodech.

Výstup programu:

```
----- RESULTS -----
S = 87.2285 % (9324728/10690000)
I = 0.1366 % (14605/10690000)
D = 0.1810 % (19348/10690000)
A = 0.0396 % (4227/10690000)
R = 0.3835 % (41000/10690000)
T = 0.0512 % (5477/10690000)
H = 11.8343 % (1265089/10690000)
E = 0.1452 % (15521/10690000)
```

Obrázek 7: Graf šíření nemoci



5.4 Závěry experimentů

Bylo provedeno celkem šest experimentů, z nichž první dva ukázaly šíření nemoci v Itálii. Bylo také provedeno porovnání našich výsledků s výsledky autorů modelu [2]. Experimenty v prostředí České republiky nám ukázaly, jak by se mohla pandemie vyvíjet v blízké době a také v době jednoho roku za implementace různých protipatření.

6 Shrnutí simulačních experimentů a závěr

V rámci projektu vznikl nástroj na předpovídání šíření nemoci COVID-19 na území České republiky. Nástroj vychází z modelu publikovaného ve vědeckém článku [2] a rozšiřuje jej o další experimenty. Z experimentu 5.3.1 je patrné, že při současných opatřeních bude pandemie nabírat na síle a bylo by vhodné implementovat další opatření. V experimentu 5.3.2 si všimněme, že i při současných opatřeních pandemie do jednoho roku téměř vymizí, avšak za cenu 66 477 mrtvých. Naopak experiment 5.3.3 nám ukázal, že pokud by vláda 3. prosince 2020 neotevřela restaurace a služby, šíření pandemie by se velmi zpomalilo a nakažených by začalo rapidně ubývat. Můžeme tedy konstatovat, že otevření restaurací nebyl ze strany vlády ten správný krok.

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] 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>
- [3] 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)