

Modelowanie Systemów Dyskretnych

**CovSim - dyskretny model
epidemii COVID-19**

Kacper Piastowicz, Michał Jasiński, Wojciech Ciężobka



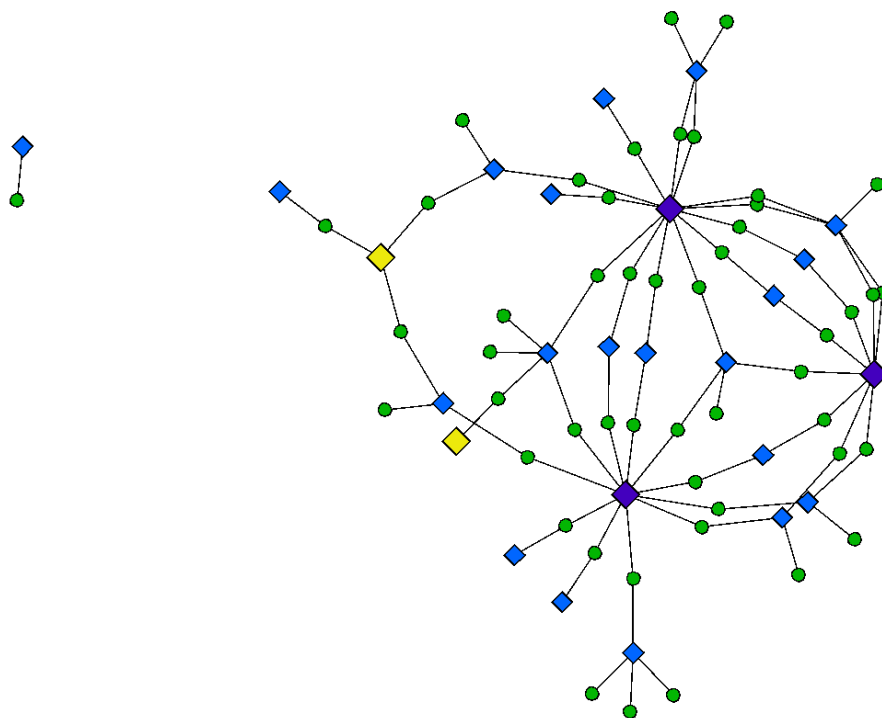
AGH

I. Zarys Modelu i Cele

Tematem projektu jest zamodelowanie epidemii wirusa SARS-CoV-2, która przetoczyła się przez cały świat w 2020 roku. Zjawisko oparliśmy na modelu SEIR czyli klasycznym modelu epidemii, rozszerzonym o stan **Exposed**, a do rozprzestrzeniania się wirusa wykorzystaliśmy siatkę kontaktów (szczegóły w punkcie II). Celem projektu jest reprezentacja małych społeczności w konfrontacji z wirusem. Przez małe społeczeństwa rozumiemy populację $N \sim 10^4$. Budując model zapewniliśmy regulację takich parametrów jak stosunek osób noszących maski oraz ich skuteczność, odpowiedzialność socjalną, rozumianą jako stosunek osób izolujących się od społeczeństwa po wystąpieniu symptomów.

II. Opis Działania Symulacji i Algorytmu

Spółeczeństwo reprezentujemy grafem (rys. 1), w którym wierzchołkami są mieszkańcy, domy, szkoły oraz miejsca pracy. Krawędź grafu to powiązanie mieszkańca z każdą z tych instytucji. Każdy z mieszkańców ma przypisany swój dom. Dodatkowo pracujący mieszkańcy (to jest w wieku produkcyjnym, którzy nie są bezrobotni) są powiązani z konkretną pracą lub szkołą w zależności od wieku.



Rys. 1: Przykładowy graf reprezentujący społeczność, $N=50$, liczba domostw - 20 (małe niebieskie romby), liczba szkół - 2 (żółte romby), liczba miejsc pracy - 3 (fioletowe romby)

W każdym dniu (dzień jest jedną iteracją algorytmu) chorzy mieszkańcy, czyli zarówno w stanie **Exposed** jak i **Infectious** rozprzestrzeniają wirusa w swoich domach i

miejscach pracy. Rozprzestrzenianie w konkretnym budynku jest realizowane przez przejście po każdej parze aktualnie w nim się znajdujących, i jeżeli choć jeden z pary zaraża, to drugi z pary zostaje zarażony lub nie, w zależności od parametrów. Realizowane jest to odpowiednio przez metody:

```
def spreadInfection(buildingRate: Double, buildingNode: String) {
  for (i <- inhabitants) {
    if (i.isInfesting()) for (j <- inhabitants) {
      if (i ne j) j.getInfectedOrNot(i, buildingRate, buildingNode)
    }
  }
}

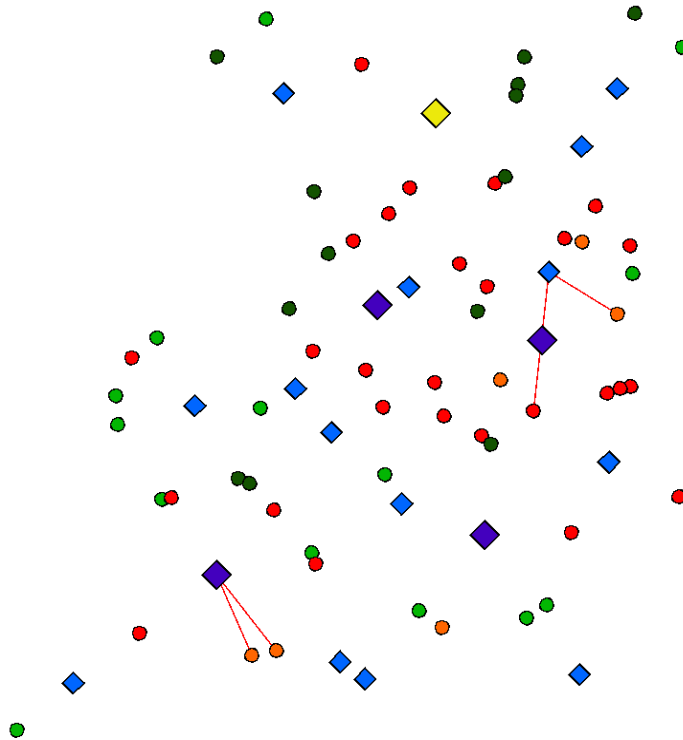
def getInfectedOrNot(from: Citizen, contagionModerator: Double, where: String): Unit = {
  if (
    currentState == Susceptible &&
    (from.currentState == Exposed || (from.currentState == Infectious && !from.responsible))
  ) &&
    RNG.nextDouble() < calcProbability(
      RAW_CONTAGION_RATE,
      contagionModerator,
      if (wearsMask) MASK_EFFICIENCY else 1.0
    )
  ) {
    infected = true
    infectionData = Some(("p${from.id}", where))
  }
}
```

Snippet 1: Metoda spreadInfection() klasy Building, oraz getInfectedOrNot() klasy Citizen

Infekcja wirusem przebiega następująco:

- Początkowo mieszkaniec jest zdrowy, w stanie **Susceptible**.
- Jeśli został zainfekowany, będąc w stanie **Susceptible**, przechodzi w stan **Exposed**.
- Mieszkaniec pozostaje w stanie bezobjawowym - **Exposed**, przez ilość czasu która jest wyznaczana na podstawie rozkładu czasu inkubacji wirusa. Po tym czasie przechodzi on w stan **Infectious**.
- Analogicznie, mieszkaniec pozostaje w stanie **Infectious**, przez ilość czasu wyznaczoną zgodnie z odpowiednim rozkładem.
- Następnym, finalnym stanem jest wejście w stan **Recovered** lub alternatywnie w stan **Dead**.

Z każdą iteracją graf jest odświeżany, a aktualny stan epidemii zapisywany do pliku, z którego następnie generujemy wykresy.



Rys. 2: Przykładowa iteracja, krawędzie obrazujące relacje zostały usunięte dla czytelności, czerwone krawędzie reprezentują aktualnie zarażane pary. Kolory wierzchołków - mieszkańców zmieniają się w zależności od stanu, od jasnozielonego, przez pomarańczowy, czerwony, aż do ciemno zielonego lub czarnego krzyżyka

III. Parametry modelu

- **FATALITY_RATE:**
 - Współczynnik śmiertelności choroby. Dla COVID-19 może wynosić od 0.7% aż do 14% na początku pandemii.
- **RAW_CONTAGION_RATE**
 - Podstawowy współczynnik zaraźliwości, lub inaczej, szansa że ktoś zarazi się po kontakcie z osobą zarażoną. Przy determinacji zarażenia jest on mnożony przez inne czynniki.
- **MASK_WEAR_RATE**
 - Współczynnik populacji, która nosi maseczki.
- **MASK_EFFICIENCY**
 - Współczynnik skuteczności maseczki.
- **SOCIAL_RESPONSIBILITY**
 - Współczynnik zarażonych, którzy unikają kontaktu, czyli przebywają w kwarantannie/izolują się.
- **INFECTION_TIME**

- Czas (w epokach), przez jaki trwa stan jawnego zakażenia, czyli osoba jest zainfekowana i może zarażać innych, mając symptomy choroby.
- **EXPOSURE_TIME**
 - Czas (w epokach), przez jaki trwa stan ekspozycji, czyli osoba jest zainfekowana, lecz nie ma symptomów.
- **UNEMPLOYMENT_RATE**
 - Współczynnik bezrobocia populacji. Wyższa wartość będzie skutkować mniejszą ilością połączeń między domami a miejscami pracy.

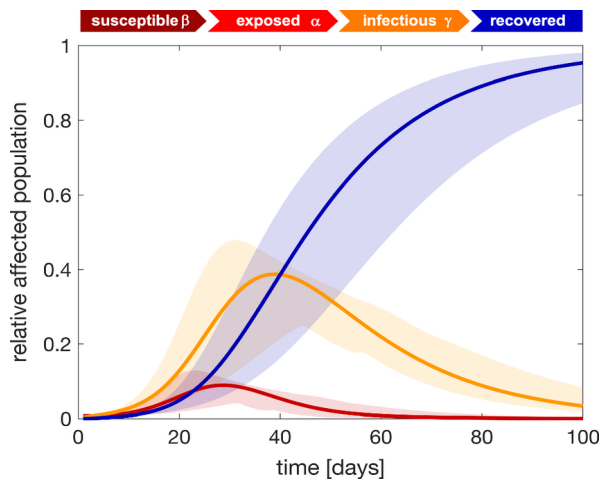
IV. Walidacja modelu

Aby wyznaczyć parametr **RAW_CONTAGION_RATE**, dobraliśmy jego wartość analitycznie, tak, aby współczynnik R_0 (*basic reproductive number*) symulacji był możliwie najbliższy rzeczywistemu R_0 wirusa. R_0 w modelu liczyliśmy ze wzoru:

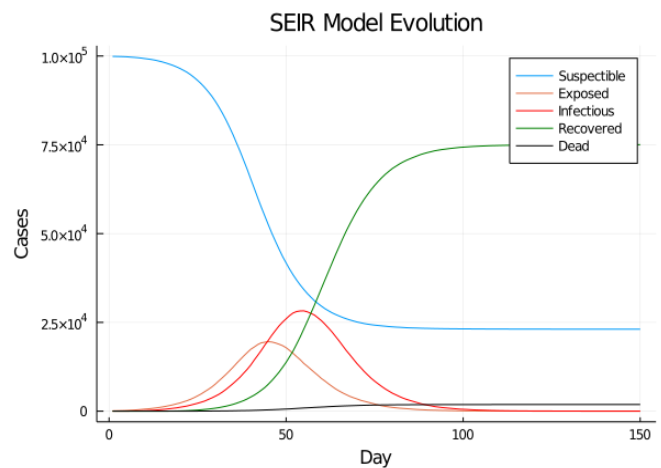
$$R_0 = c \times p \times d$$

gdzie c to liczba kontaktów osoby chorej ze zdrowymi w danej jednostce czasu, p to prawdopodobieństwo zakażenia w czasie kontaktu, a d to czas trwania okresu zaraźliwości. Choć R_0 jest zmienne w czasie, skorzystaliśmy z estymacji $R_0=2.9$.

Ponieważ czas ekspozycji i zainfekowania nie jest dla wszystkich taki sam, posłużyliśmy się odpowiednimi rozkładami aby je zróżnicować. Dane takie jak bezrobocie, czy rozkład wieku zostały pobrane z danych statystycznych.

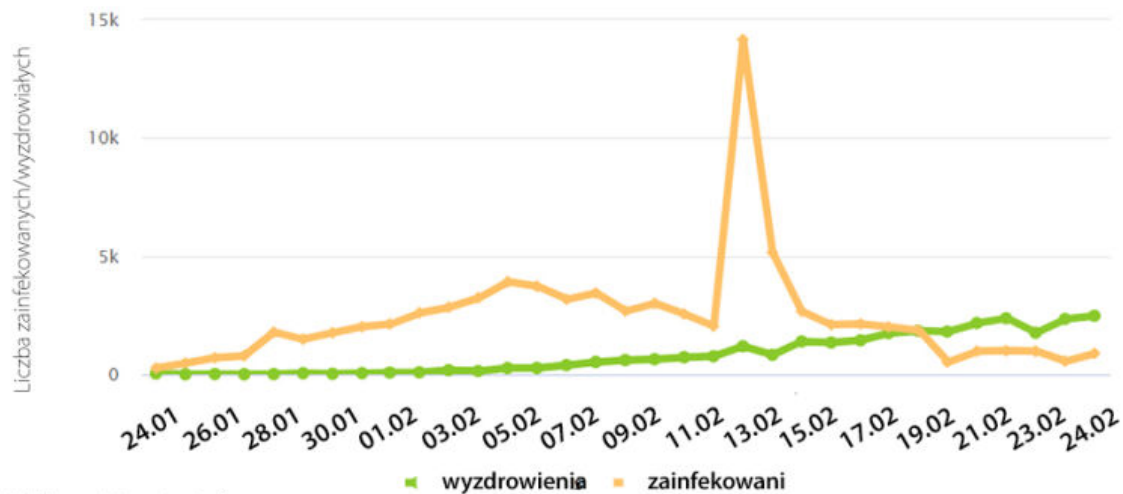


Symulacja SEIR rozprzestrzeniania się wirusa w USA (Ogólnokrajowa symulacja, osobne parametry każdego stanu)



Wynik naszej symulacji dla 100 tysięcznej populacji

Nowe przypadki oraz wyzdrowienia



źródło: worldometers.info

Zestawienie liczby zainfekowanych do czasu, wykres obrazuje nieregularność przebiegu epidemii w rzeczywistym środowisku.

Trudno jest zestawiać nasze wyniki z rzeczywistymi statystykami, ponieważ:

- Ze względu na charakter testów na COVID-19, wśród zarażonych nie rozróżnia się osób z symptomami i bez.
- Ludzkość dynamicznie reaguje na postępującą pandemię, to znaczy reguluje obostrzenia względem aktualnego stanu epidemii. W naszym modelu wszystkie parametry ustalane są na początku.
- Nie sposób zarejestrować wszystkich przypadków. Oficjalne statystyki reprezentują tylko część właściwych zakażeń.
- Występują false positives i false negatives.

Mając na uwadze powyższe, jesteśmy w stanie zwalidować nasz model jedynie w pewnym ograniczonym stopniu.

V. Wnioski

Przeprowadzając symulację dla licznych populacji, graf służący wizualizacji przestaje być przejrzysty ze względu na rosnącą liczbę nachodzących na siebie krawędzi. Można natomiast zupełnie wyłączyć wizualizację w czasie rzeczywistym i korzystać jedynie z wygenerowanych danych. Algorytm rozprzestrzeniania się wirusa działa dobrze dla dużych N ze względu na lokalność zarażeń - uniknęliśmy w ten sposób złożoności: $O(N^2)$, która mogłaby okazać się problematyczna przy bardziej licznych populacjach.

Model dobrze symuluje pandemię wirusa. Generacja społeczeństwa z rozkładów skutkuje drobnymi niedoskonałościami w małej skali, jednak zwiększając N przestają być one znaczące..

Powyższe wnioski przypominają nam, że model jest jedynie **uproszczeniem rzeczywistości** i nie będzie idealnie jej imitował. Może jednak być z powodzeniem wykorzystany do analizowania scenariuszy pandemicznych.

Praca nad tym projektem pokazała pewną słabość wizualizacji na podstawie grafu. Już dla populacji wielkości około 1000 graf przestaje być użyteczny jako metoda wizualizacji przebiegu pandemii.

VI. Future Works

Ciekawym rozwinięciem może być analiza indywidualna, lub zbieranie szczegółowych statystyk dotyczących wieku i zachowania. Skutkowałyby to uzasadnieniem stwierdzeń typu “noszenie maseczki zmniejsza ryzyko śmierci o x%”. Umożliwiłoby to również wyszczególnienie w analizie danych grup wiekowych oraz porównanie ze sobą ich statystyk. Za miejsce na udoskonalenie można uznać relacje w rodzinie/domostwie skutkujące wspólną izolacją lub naśladowaniem rodziców w zachowaniu/odpowiedzialności. Z uwagi na dużą customizację można stworzyć bazę parametrów różnych wirusów/chorób. Może ona posłużyć do dalszej kalibracji samego modelu.

Kolejnym możliwym rozwiązaniem jest rozbudowa projektu o kolejne możliwe miejsca, w których ludzie mogą się zarażać np. transport publiczny, sklepy, wydarzenia szerokiej skali. Byłoby to jednak szerokie przedsięwzięcie wymagające dogłębnego zapoznania się z częstotliwością z jaką przeciętni ludzie przemieszczają się po mieście i uczęszczają w aktywnościach z innymi członkami społeczeństwa; co być może wymagałoby kolejnych modeli.

VII. Bibliografia

- Institute for Disease Modelling: <https://www.idmod.org/>
- Ciekawy artykuł o rozprzestrzenianiu się Covida w związku z powrotem do szkół:
https://iazipvnewgrp01.blob.core.windows.net/source/2021-03/reports/pdf/Stepping_Back_to_School_2.pdf
- Covid Symulator od IDM:
<https://docs.idmod.org/projects/covasim/en/latest/>
- Modele SIR:
https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology
- Covid Symulator od IDM:
<https://docs.idmod.org/projects/covasim/en/latest/>
- Rola masek w rozprzestrzenianiu się covida:
<https://www.medrxiv.org/content/10.1101/2020.09.02.20186742v1>
- Kolejny model wpływu noszenia masek:
<https://www.medrxiv.org/content/10.1101/2020.09.28.20202937v2.full.pdf+html>
- Face masks: what the data say:
<https://www.nature.com/articles/d41586-020-02801-8>
- Ryzyko związane z powrotem do szkół:
<https://www.medrxiv.org/content/10.1101/2020.09.08.20190942v1.full.pdf+html>
- Incubation Period:
<https://www.worldometers.info/coronavirus/coronavirus-incubation-period/>
- Mask wearing ratio:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7561164/>
- Model pandemic w USA i Chinach
https://www.researchgate.net/publication/340958551_Outbreak_dynamics_of_COVID-19_in_China_and_the_United_States
- Biblioteka GraphStream
<https://graphstream-project.org/>