

Технічні науки

УДК 32.973.3

Репнікова Наталія Борисівна

*кандидат технічних наук, доцент,
викладач факультету інформатики та обчислювальної техніки
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»*

Репникова Наталия Борисовна

*кандидат технических наук, доцент,
преподаватель факультета информатики и вычислительной техники
Национальный технический университет Украины
«Киевский политехнический институт имени Игоря Сикорского»*

Repnikova Nataliya

*PhD, Associated Professor of the
Faculty of Informatics and Computer Science
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"*

Капченко Поліна Володимирівна

*студент факультету інформатики та обчислювальної техніки
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»*

Капченко Полина Владимировна

*студент факультета информатики и вычислительной техники
Национального технического университета Украины
«Киевский политехнический институт имени Игоря Сикорского»*

Kapchenko Polina

*Student of the Faculty of Informatics and Computer Science of the
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"*

**ВДОСКОНАЛЕННЯ ФУНКЦІОНАЛЬНИХ МОЖЛИВОСТЕЙ
МОДЕЛІ ПРОГНОЗУВАННЯ ЕПІДЕМІЧНИХ ЗАХВОРЮВАНЬ
УСОВЕРШЕНСТВОВАНИЕ ФУНКЦИОНАЛЬНЫХ
ВОЗМОЖНОСТЕЙ МОДЕЛИ ПРОГНОЗИРОВАНИЯ
ЭПИДЕМИЧЕСКИХ ЗАБОЛЕВАНИЙ
IMPROVING THE FUNCTIONAL CAPABILITIES OF THE EPIDEMIC
DISEASE PREDICTION MODEL**

***Анотація.** Розглядається проблема прогнозування розвитку епідемічних захворювань. Пропонується використання базової моделі прогнозування епідемічних захворювань та методу оптимізації з метою наближення прогнозу розвитку епідемії до її реальних параметрів.*

В ході вдосконалення використовуються методи найменших квадратів для покращення точності результатів прогнозування. Приведено результати моделювання моделі з використанням прикладного пакету Matlab на базі статистичних даних захворювання на COVID-19 Сполучених Штатів Америки.

***Ключові слова:** віруси, моделі дослідження, епідемія, метод найменших квадратів, прогнозування.*

***Аннотация.** Рассматривается проблема прогнозирования развития эпидемических заболеваний. Предлагается использование базовой модели прогнозирования эпидемических заболеваний и метода оптимизации с целью приближения прогноза развития эпидемии к реальным параметрам.*

В ходе усовершенствования используются методы наименьших квадратов для улучшения точности результатов прогнозирования. Приведены результаты моделирования модели с использованием прикладного пакета Matlab на базе статистических данных заболевания COVID-19 Соединенных Штатов Америки.

Ключевые слова: вирусы, модели исследования, эпидемия, метод наименьших квадратов, прогнозирование.

Summary. Considered the problem of predicting the epidemic development of diseases. It is proposed to use a basic model for predicting epidemic diseases and an optimization method in order to bring the epidemic development forecast closer to real parameters.

The improvement uses least squares methods to improve the accuracy of the prediction results. The results of modeling are presented in the Matlab application package on the basis of statistical data on the COVID-19 disease in the United States of America are presented.

Key words: viruses, research models, epidemic, least squares method, forecasting.

Вступ. Методи прогнозування інфекційної захворюваності активно розвиваються в наші дні. В останні роки кількість робіт на цю тему стрімко зростає завдяки розгортанню інформаційних систем нагляду та появі великих обсягів статистики, доступної для аналізу. Епідеміологічні прогнози виконуються для різних термінів і в залежності від них служать різним цілям.

Дана тема є досить актуальною в наші дні, враховуючи спалахи епідемій нових штамів грипу, а також зовсім нових хвороб, таких як COVID-19.

Для того, аби як можна швидше попереджувати зараження великої кількості людей використовуються спеціальні заходи зі сторони виконавчих органів країн. Для передбачення того, як швидко ситуація може стати критичною використовують статистичні дані та різні види моделей прогнозування на їхній основі. Існують полігамні моделі в епідеміології, регресійні моделі та багато інших. Недоліком моделей є прогноз на великі

проміжки часу, адже найчастіше, вони показують велике відхилення від реальних.

Основна частина. Найчастіше використовують полігамні моделі (Рис.1) для прогнозування розповсюдження захворювання. Для дослідження цих моделей використовують звичайні диференціальні рівняння (які є детермінованими). В моделі населення поділяється на три групи: люди, які ще не захворіли, інфіковані та видужавші або померлі.

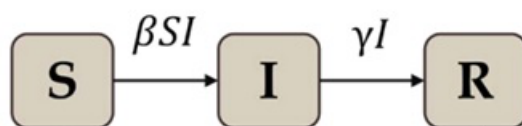


Рис. 1. Базова модель прогнозування розповсюдження захворювання

Модель описується такими рівняннями:

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta IS}{N} \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

де S - люди, які ще не захворіли, I – інфіковані, R – видужавші або померлі

β - інтенсивність зараження, середня кількість контактів на людину протягом періоду часу, γ - інтенсивність одужання .

В даній моделі є припущення, що всі значення є постійними. Це означає, що очікувана тривалість інфекції прямо пропорційна швидкості одужання/смертності.

N – це кількість населення, що є сталим і незмінним числом, і є сумою всіх груп моделі

$$N = S(t) + I(t) + R(t) = Constant$$

Для побудови моделі обрана хвороба COVID-19 та країна Сполучені Штати Америки. З відомих нам параметрів це буде кількість населення станом на 2020 рік і це 333 466 000 осіб.

R_0 – коефіцієнт поширення інфекції. R_0 є безрозмірним числом. Даний параметр для різних країн і різних інфекцій є різним. В залежності від етапу хвороби, також є різним. У період епідемій R_0 завжди більше 1. У випадку з ковідом параметр коливається від 1.4 та 2.8 [1].

Також потрібно враховувати, кількість днів, скільки триває хвороба. Ковід в середньому триває 14 днів, то ж $D=14$. Обираємо $R_0=2$. Побудуємо модель (Рис. 2) за допомогою програмного пакету Matlab на період 365 днів.

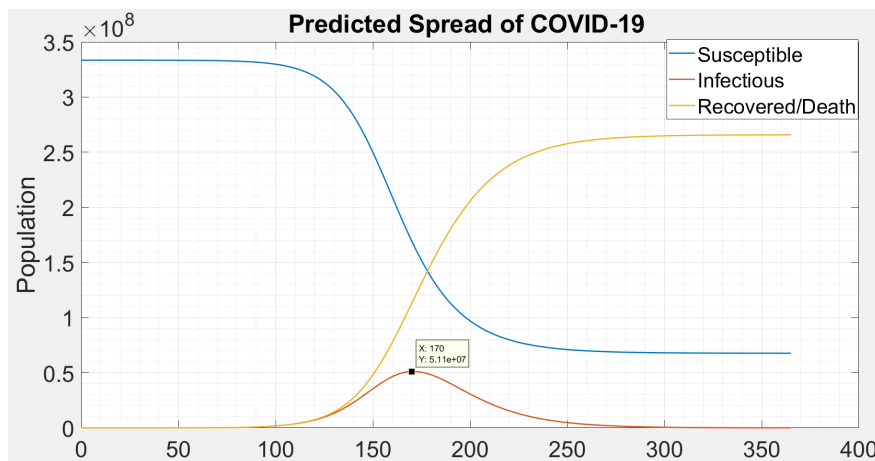


Рис. 2. Модель прогнозування розповсюдження ковіду в США

Джерело: авторська розробка

Для оцінки результатів порівняємо з реальними даними інфікованих від березня 2020 року [2] та побудуємо модель тільки для інфікованих випадків

Після моделювання (Рис. 3) можна зробити висновки, що прогноз має велику похибку, а також не співпадає пік з реальними даними. Проблема полягає в тому, що в базовій моделі не враховується безліч факторів, таких як перебіг хвороби, інкубаційний період, коли особа вже заразна, але ще не відчуває симптомів, а також не враховуються карантинні міри.

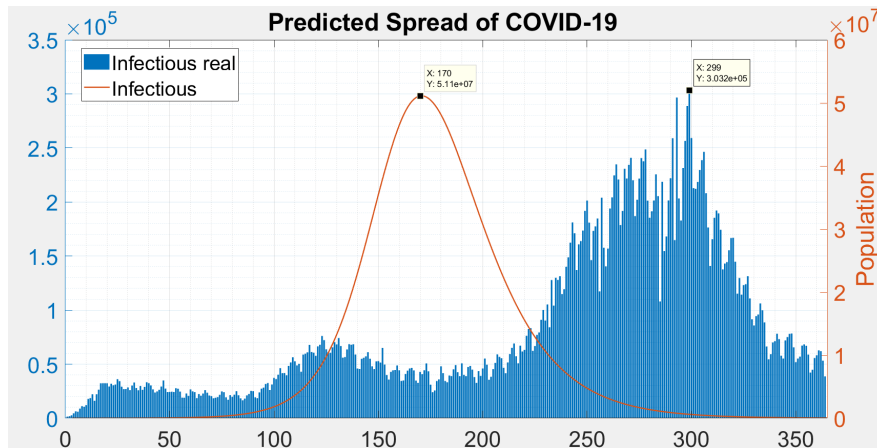


Рис. 3. Модель прогнозування розповсюдження ковіду в США з реальними даними

Джерело: авторська розробка

Існує багато модифікацій полігамної моделі для епідемії, де враховуються деякі з перелічених факторів. Будемо використовувати метод найменших квадратів для покращення результатів прогнозування.

Так як протягом епідемії в моделі кількість населення є сталою, перепишемо рівняння для знаходження кількості інфікованих та одужавших або померлих осіб. Підставляючи N в диференційні рівняння отримаємо :

$$\frac{dI}{dt} = (R_0\beta - \gamma)I - \frac{R_0\beta}{N}I^2 - \frac{R_0\beta}{N}IR$$

$$\frac{dR}{dt} = \gamma I$$

Для оптимізації параметрів R_0 , β та γ використаємо метод найменших квадратів. Для вектор функції $f(x, \alpha)$, оптимізуємо параметр α

$$\alpha = [\gamma \beta R_0]$$

$$F(\alpha) = \sum_{i=1}^m [y_i - f(x_i, \alpha)]^2$$

Задаємо початкові параметри в α та виконуємо оптимізацію, використовуючи програмний пакет Matlab, а також ode45 для вирішення диференційних рівнянь та метод найменших квадратів [3], який реалізовано за допомогою lsqcurvefit.

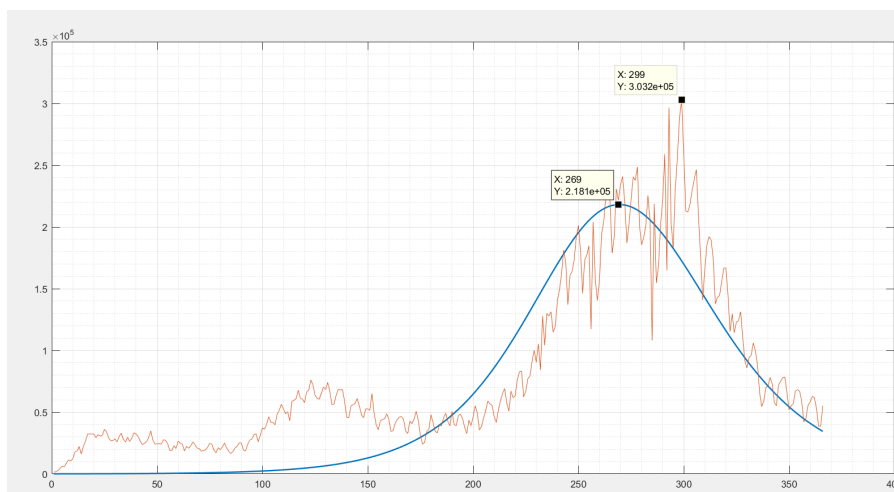


Рис. 4. Моделювання з оптимізованими параметрами

Джерело: авторська розробка

Таблиця 1

Результати моделювання

Метод	Кількість інфікованих	День піку	Відхилення від реальних даних
Реальні дані	303 200	299	-
Базова модель	51 100 000	170	50 796 800
Вдосконалена модель	218 100	269	85 100

Джерело: авторська розробка

На Рис.4 видно, що метод найменших квадратів досить сильно наблизив прогнозування розповсюдження хвороби до реальних даних в порівнянні з вихідною моделлю. Результати порівняння наведені в Таблиці 1.

Хоча вдосконалена модель дає набагато кращі результати, які наближені до реальних, вона не враховує велику кількість параметрів і не передбачуваних ситуацій, наприклад: скупчення людей на мітингах, концерти і таке інше, в один з днів, там де ймовірність зараження є більшою. Пік прогнозований на 269 день є близьким до реальних прогнозів, адже

одразу йде плавний спад інфікованих осіб і є тільки два непередбачуваних спалахи.

Висновки. В даній науковій роботі запропоновано використання методу найменших квадратів для вдосконалення існуючої моделі прогнозування розповсюдження вірусів. Виконано прогнозування розповсюдження епідемічних захворювань та порівняння з статистичними даними реальної епідемії коронавіруса у Сполучених Штатах Америки, що дало високий відсоток співпадання.

Результати отримані в ході дослідження можуть використовуватися надалі у якості бази для подальших модифікацій та покращень полігамних моделей прогнозування розповсюдження епідемічних захворювань.

Література

1. Batista M. Estimation of the final size of the coronavirus epidemic by the SIR model / M.Batista. RegeGate, 2020.
2. WHO Emergency Committee. Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-ncov). 2020. URL: [https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov))
3. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений / Ю.В. Линник. Москва: Государственное издательство физико-математической литературы, 1962. 326 с.