

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛІАНСЬКА АКАДЕМІЯ»

Кафедра мультимедійних систем факультету інформатики



**Моделювання складних процесів за допомогою клітинних
автоматів**

**Текстова частина до курсової роботи
за спеціальністю „Інженерія програмного забезпечення” 121**

Керівник курсової роботи
доцент, к.н., Жежерун О.П.
(прізвище та ініціали)

(підпис)

“ _____ ” _____ 2021 р.

Виконала _____ студентка
Черепина Є.О.

(прізвище та ініціали)

“ _____ ” _____ 2021 р.

Київ 2021

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА
АКАДЕМІЯ»

Кафедра мультимедійних систем факультету інформатики

ЗАТВЕРДЖУЮ

Зав. Кафедри мультимедійних систем,
доцент, к.н.,

_____ Жежерун О.П.

(підпис)

“ ____ ” _____ 2021 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ
на курсову роботу

Студентці Черепиній Єлизаветі Олексіївні факультету інформатики 1
курсу магістратури

ТЕМА: Моделювання складних процесів за допомогою клітинних
автоматів

Вихідні дані:

Система для моделювання епідемії за допомогою клітинних
автоматів

Зміст ГЧ до курсової роботи:

Вступ

Частина 1: Аналіз предметної області

Частина 2: Теоретичні відомості

Частина 3: Опис реалізації програмного продукту

Висновки

Список використаної літератури

Додатки

Дата видачі „ ____ ” _____ 2020 р. Керівник _____
(підпис)

Завдання отримала _____
(підпис)

Зміст

| | |
|---|----|
| Анотація | 4 |
| Вступ | 5 |
| РОЗДІЛ 1: Аналіз предметної області. Постановка завдання курсової роботи | 6 |
| 1.1. Аналіз сучасного стану питання та обґрунтування теми | 6 |
| 1.2. Огляд існуючих аналогів розробки..... | 8 |
| 1.3. Постановка задачі | 11 |
| РОЗДІЛ 2: Теоретичні відомості | 13 |
| РОЗДІЛ 3: Опис реалізації програмного продукту | 17 |
| 3.1. Аналіз технічного завдання..... | 17 |
| 3.2. Обґрунтування алгоритму й структури програми | 17 |
| 3.3. Обґрунтування вибору засобів розробки | 17 |
| 3.4. Опис розробки програми | 17 |
| 3.5. Створення об'єктів і розробка головної програми | 17 |
| 3.6. Опис файлів даних та інтерфейсу програми | 17 |
| Висновки | 18 |
| Список використаної літератури | 19 |
| Додаток А. Програмний код | 20 |

Анотація

Дана робота спрямована на розробку системи для моделювання розповсюдження інфекційних хвороб за допомогою клітинних автоматів.

Для реалізації було використано мову програмування Python та pygame — набір крос-платформових модулів для Python, призначених для створення відеоігор.

Вступ

Розповсюдження інфекцій є дуже актуальною темою сьогодні. Багато людей постраждало внаслідок пандемії від COVID-19. Всі ми на власному досвіді відчули наслідки цієї хвороби. Але для підтримання обізнаності серед людей, важливо давати візуальну картину можливих наслідків порушення режиму карантину. Створення системи, яка б давала візуальне прогнозування епідемій хвороб мало б велику значимість.

Мета цієї роботи – дослідити, що за моделі для прогнозування епідемій існують.

Об'єктом дослідження є системи, які вже мають вбудовані моделі прогнозування епідемій.

Предмет дослідження – прогнозування різних природних процесів, критеріїв їхньої оцінки та

Враховуючи результати, які будуть отримані можна буде використати як основу одну з моделей прогнозування епідемій, за потреби її модифікувати та реалізувати систему для моделювання розповсюдження хвороби.

Робота складається зі вступу, трьох розділів, висновків, списку використаних джерел та додатків.

РОЗДІЛ 1: Аналіз предметної області. Постановка завдання курсової роботи

1.1. Аналіз сучасного стану питання та обґрунтування теми

Людство за весь час існування потерпало від незліченної кількості епідемій. Різні хвороби постійно з'являються та впливають на якість життя людей. Оскільки такі проблеми потрібно контролювати задля забезпечення безпеки населення, науковці навчилися прогнозувати розповсюдження захворювань на основі моделювання.

Математичне моделювання є невід'ємною частиною сучасної медицини, оскільки остання включає в себе вивчення складних процесів, де неможливо проводити реальні експерименти. Прикладом таких є епідеміологічні процеси. Математичні моделі використовують для аналізу поточних станів епідемій, прогнозування їх розширення, впливу на смертність та появу імунітету до певної інфекції вже давно використовують. Таким чином, вчені можуть прорахувати ризики та наслідки епідемій. Перевагою математичного прогнозування розповсюдження хвороб різного типу є можливість деталізації, точності та завчасного реагування на положення. Всі основні вхідні дані можуть бути статистично прораховані та передбачені можливі зміни у реакції інфекції на організми, мутації.

Стрімка глобалізація, великий потік людей, що контактують один з одним у транспорті, закладах освіти, закладах харчування – основний спосіб неконтрольованого розповсюдження інфекцій. З 2019 року весь світ потерпає від жахливої пандемії коронавірусу COVID-19. Для всього людства поява нового виду вірусу стало великим потрясінням, появою переживань щодо власної безпеки та кардинальною зміною образу життя. Однак, хоча це і була нова хвороба, що мала на початок 2020 року невідомі наслідки для людства, вчені змогли вже напочатку спрогнозувати

перебіг хвороби та її розповсюдження з урахуванням певних критеріїв, а саме швидкість розповсюдження, дотримання соціального дистанціювання, відсоток смертності, а в подальшому – процес появи імунітету у людей, що перехворіли.

Дані дослідження є важливими не тільки для запровадження заходів для контролю епідеміологічної ситуації, але й для інформування населення щодо можливого перебігу епідемії та необхідності дотримання оголошених правил для урегулювання ситуації.

Предметна область даного дослідження – існуючі моделі для прогнозування епідемій та можливі реалізації цих моделей для побудови візуального представлення розповсюдження хвороб. Наразі існує досить багато перевірених часом схожих моделей, але способи їхніх реалізацій відрізняються.

Одним із способів реалізації моделювання природних процесів, а саме епідемій, є використання клітинних автоматів. Доведено, що обчислення за допомогою клітинних автоматів, запропонованих вперше фон Нейманом, є корисним для складних систем та процесів в різних галузях науки. Саме тому їх давно використовують для дослідження поширення інфекційних хвороб.

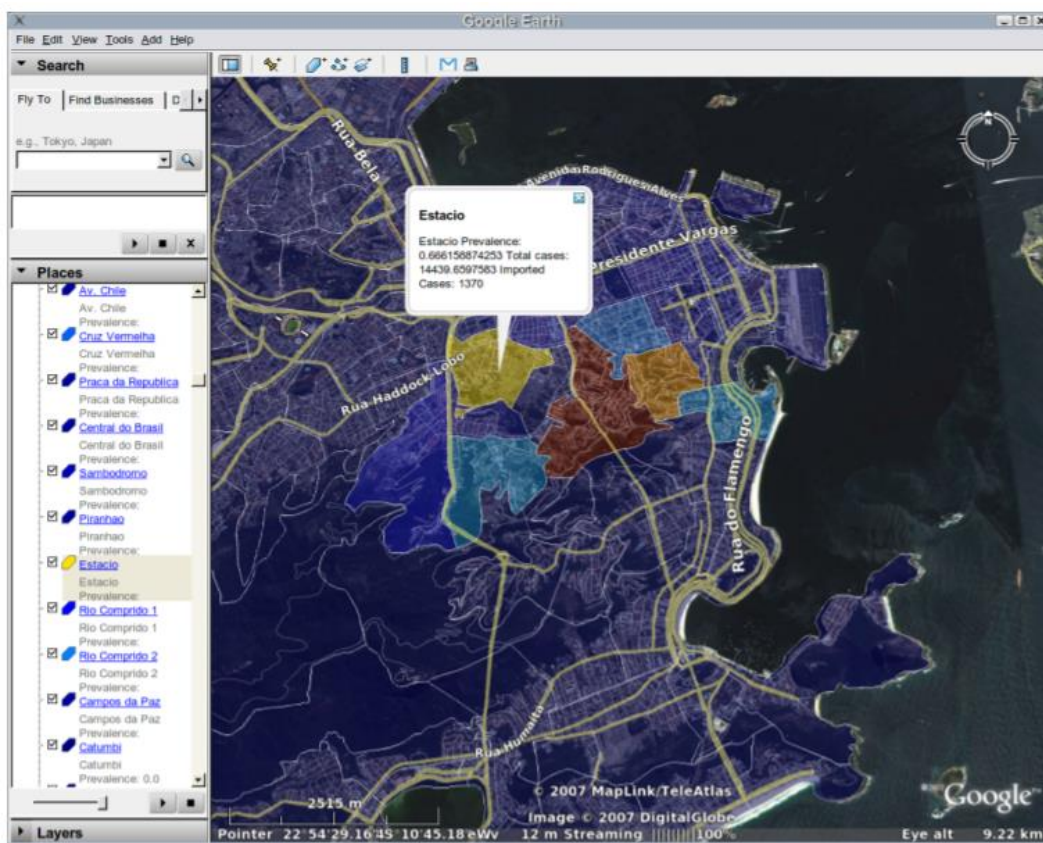
З огляду на отримані дані дослідження, розробка системи моделювання поточної ситуації розповсюдження епідемії може бути корисною для населення, оскільки матиме візуальне відображення ситуації. Це може стати поштовхом для усвідомлення важливості дотримання правил для запобігання великих масштабів пандемії. Саме тому, можна зробити припущення, що дана тема є актуальною для сьогодення.

1.2. Огляд існуючих аналогів розробки

Оскільки задача прогнозування епідемій постала перед людством давно, на сьогоднішній день існує досить багато готових систем, які моделюють різні природні процеси, в тому числі – епідемії. Розглянемо основні з них.

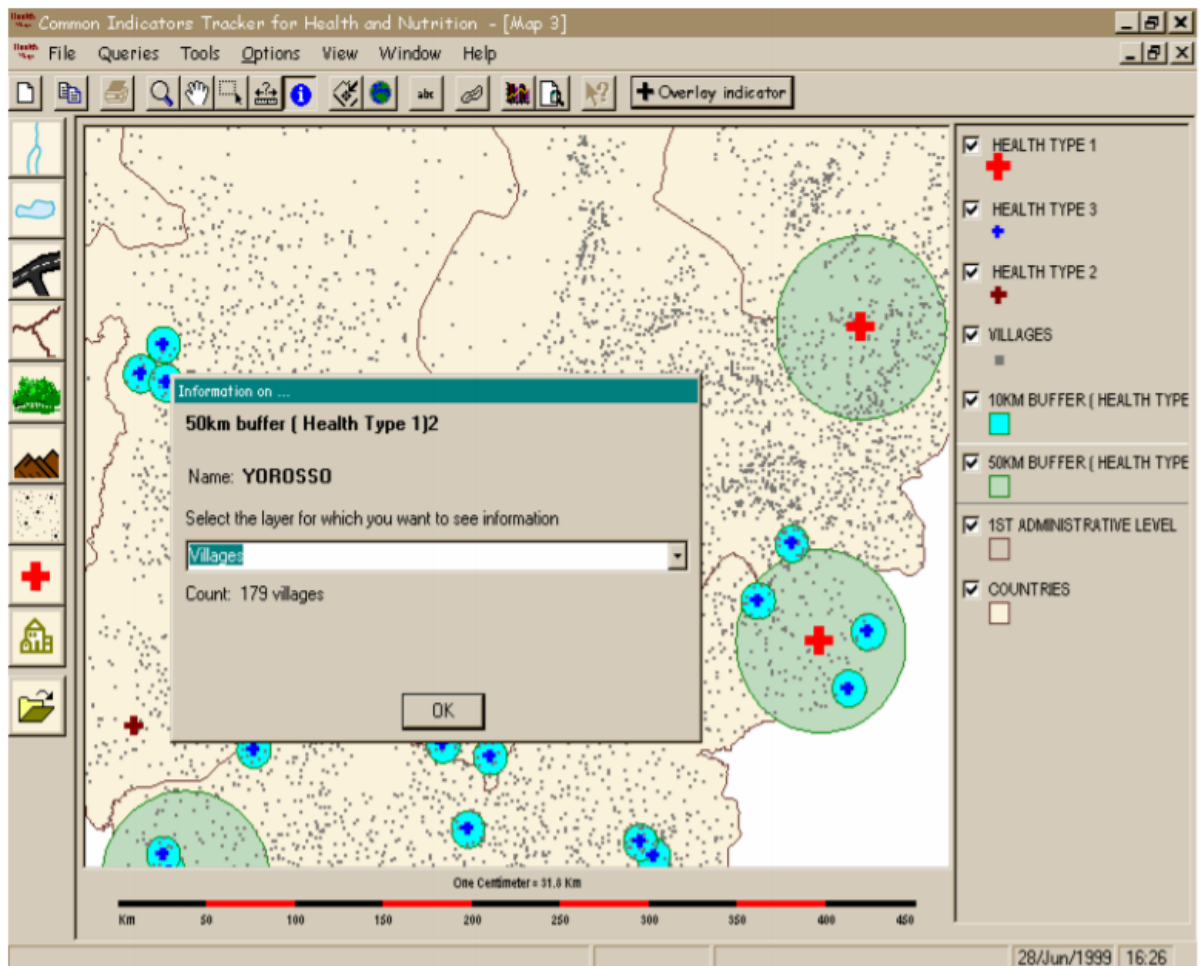
Одним із яскравих прикладів є платформа EpiGrass для побудови та моделювання та аналізу епідеміологічних процесів. Дане програмне забезпечення дає змогу здійснювати просторово-часові моделювання та містить готовий пакет моделей епідемій.

Географічні мережі, в яких відбуваються епідеміологічні процеси, можуть бути дуже прямо представлені в об'єктно-орієнтованій системі. У такій системі вузли та краї географічних мереж є об'єктами зі своїми атрибутами та методами. Після того, як вузлові та реберні об'єкти визначені відповідними атрибутами та методами, будується кодове представлення реальної системи, де міста (або інші географічні населені пункти) та транспортні шляхи є екземплярами відповідно вузлових та реберних об'єктів. Вся мережа - це також об'єкт із цілою колекцією атрибутів та методів. Цей фреймворк веде до компактної та ієрархічної обчислювальної моделі, що складається з мережевого об'єкта, що містить змінну кількість вузлових та реберних об'єктів. Для кінцевих споживачів дана система є зрозумілою, оскільки імітує природну структуру реальної системи. Навіть після того, як модель перетворена на об'єкт коду, усі її компоненти залишаються доступними для моделей користувача.



Ще один приклад – Health Mapper. Це програмне забезпечення, розроблене Всесвітньою організацією здоров'я, що є потужним інструментом для спостереження, візуалізації і розуміння глобальних даних про стан здоров'я. Health Mapper дає доступ до показників здоров'я на рівні країни які можуть бути відображені у вигляді карти, таблиці, гістограми, тощо.

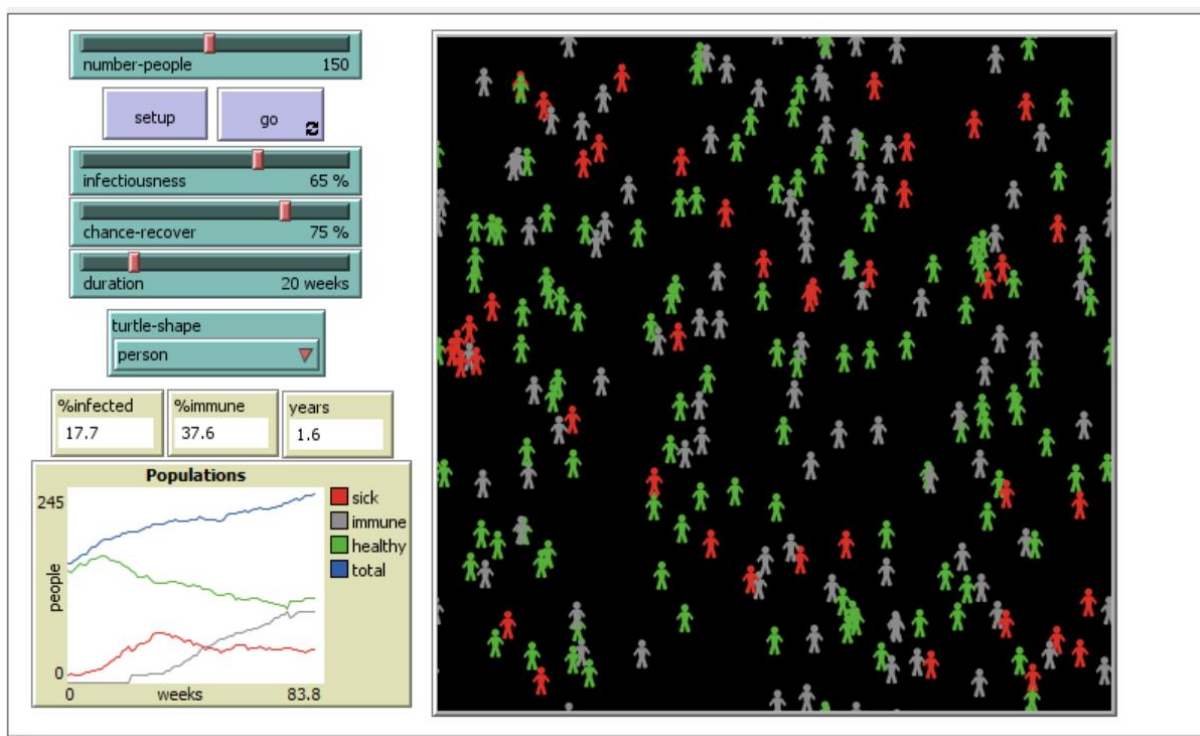
Система містить базу даних щодо географічних, демографічних та медико-санітарних даних, зокрема розташування громади, охорони здоров'я та освіти, доступності доріг, доступ до безпечної води, тощо. Дана система не є прямим аналогом для моделювання епідемій, однак є дуже потужним інструментом для візуалізації схожих процесів.



Розглянемо також середовище для програмування NetLogo. Воно використовується для побудови моделей ситуацій, що відбуваються у світі, на базі мультиагентних систем. Дана система має велику кількість різних готових моделей, що описують біологічні, хімічні, соціальні процеси, тощо.

Розробник має можливість задати всі параметри та вказівки для агентів, які в свою чергу діють незалежно один від одного та мають можливість хаотично пересуватися у межах побудованої моделі.

Модель розповсюдження вірусів також присутня у середовищі, однак, оскільки NetLogo призначене для більш широкого спектру задач, епідеміологічні прогнози не деталізовані на достатньому рівні.



З огляду на проаналізовані аналоги системи, можна стверджувати, що всі вище зазначені системи є популярними серед вчених та дослідників, однак не серед звичайного населення.

1.3. Постановка задачі

Проаналізувавши дані дослідження та беручи до уваги важливість моделювання розповсюдження епідемій, було сформовано задачу щодо розробки відповідного продукту.

Ключова задача – розробити зрозумілу та просту у використанні систему з використаною моделлю прогнозування епідемій за допомогою клітинних автоматів, з можливістю кастомізувати систему під потреби користувача.

Реалізація даної задачі повинна включати в себе наступні дії:

- Аналіз існуючих моделей розповсюдження інфекцій
- Оцінка необхідності модифікації цих моделей
- Визначити додаткові фактори, що впливають на модель

- Розробити дизайн системи
- Розробити систему для прогнозування епідемій
- Протестувати програму

РОЗДІЛ 2: Теоретичні відомості

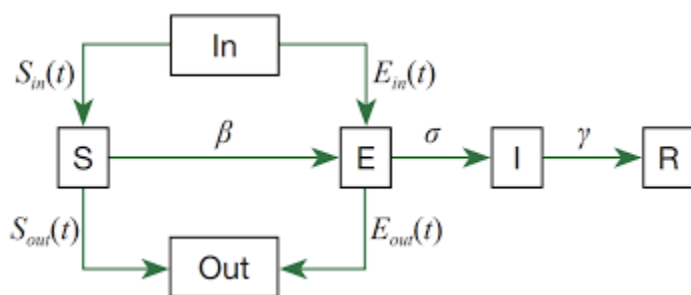
Для реалізації системи, що прогнозує розповсюдження епідемії потрібно дослідити, по-перше, існуючі моделі інфікування людей, а також доречність використання клітинних автоматів для цієї розробки.

Однією з найпоширеніших епідеміологічних моделей є модель Кермака-МакКендріка - SIR. Вона складається з 3 компонентів, що відповідають літерам абрєвіатури:

- S – Susceptible (кількість вразливих організмів)
- I – Infected (кількість заражених організмів)
- R – Recovered (організми, що одужали)

Таким чином, за цією моделлю населення ділиться на 3 групи, що зазначені вище.

Дана модель доволі проста, оскільки передбачає, що розмір популяції є константним (не враховується життєвий цикл населення), немає інкубаційного періоду, тощо. Але цією моделі достатньо для побудування поширення багатьох інфекцій (кір, краснуха, тощо). Існує також розширена версія цієї моделі – модель SEIR. Дана модель передбачає, що люди, які мали контакт з інфікованими носіями, можуть бути зараженими, а отже і термін передачі хвороби стає довшим, що більше відповідає характеристикам епідемії.



SEIR модель складається з 4 компонентів, які включають себе компоненти SIR моделі та додатковий компонент E – Exposed (незахищений, тобто той, хто контактував з інфікованою особою, але може мати, а може і не мати симптоми).

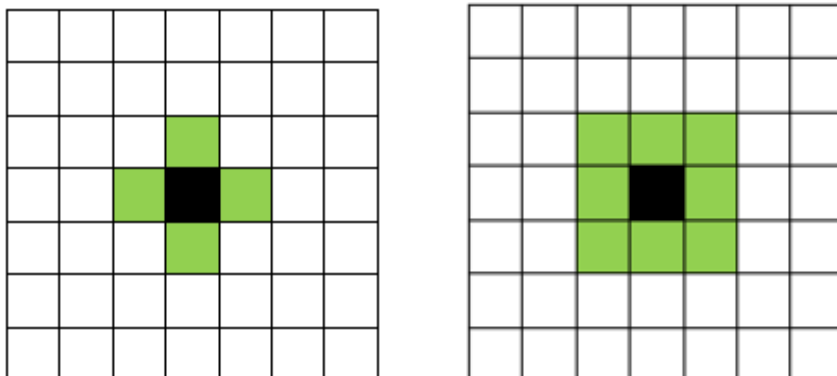
Клітинні автомати – це дискретні динамічні системи, що складаються з однакових клітин, кожна з яких є кінцевим автоматом. Вони об'єднані між собою та їх сукупність утворює решітку клітинного автомату. Для кожної клітини можна визначити наступні характеристики:

- Вона належить решітці
- Вона має набір станів (в простих клітинних автоматах має 2 стани – «жива» чи «мертва»)
- Кожна клітина має сусідів – вони можуть бути визначені різними способами, але це зазвичай клітини, що знаходяться в околі основної клітини

Час у системі клітинних автоматів вимірюється дискретно, одиниця виміру – такт. Майбутній стан клітини (в наступний такт) визначається як функція від власного стану та станів сусідів в поточний момент часу.

Розглянемо основні 2 види клітинних автоматів – одновимірні та двовимірні. В найпростішому одновимірному автоматі у нас є одновимірний масив, клітини якого можуть приймати один із станів – жива або мертва клітина. Цей стан визначається станами двох сусідніх клітинок за певним правилом. Але одновимірного автомату нам не достатньо для моделювання епідеміологічної ситуації.

Проаналізуємо двовимірний автомат. Для даного виду властиві також набір станів, що визначаються сусідніми клітинами. В даному випадку використовується 2 найпоширеніших окола – окіл фон Неймана та окіл Мура.



Окіл фон Неймана – це сукупність чотирьох клітин , що мають спільну сторону з даною клітиною, а окіл Мура клітини – це сукупність восьми клітин, що оточують спільну.

Для побудови моделі надалі будемо використовувати двовимірні клітинні автомати та модель SIR із додатковими параметрами . Для написання правил переходу та перелік станів кожної клітини, зазначимо параметри, які впливають на розповсюдження хвороби та визначимо їх правила обчислення.

Передбачаємо, що переважна більшість клітин є вразливими. Кожна з них має певну ймовірність захворіти. Для цього використовується наступна формула :

$$P_{\text{infect}} = 1 - (1 - p)^R ,$$

де p – ймовірність, що інфікована клітина передасть хворобу здоровій клітині.

Окрім ймовірності захворіти, у кожної клітини також є ймовірність одужання від хвороби.

$$P_{\text{recover}} = q ,$$

де значення P та q лежить у межах від 1 до 0.

Обидві ймовірності – рандомні числа, значення яких, впливають на стан клітини, а саме зробить її хворою, здоровою, або залишить в тому ж самому стані.

РОЗДІЛ 3: Опис реалізації програмного продукту

3.1. Аналіз технічного завдання

У даній роботі передбачається, що буде розроблена певна система, яка буде мати візуальне відображення моделі розповсюдження епідемії. Ми будемо працювати з даними, які є статистичними а знаходяться у вільному доступі.

Серед основних можливостей, повинні бути впровадженні додаткові параметри, які впливають на розповсюдження епідемії, а саме соціальне дистанціювання, носіння маски, тощо. Цей функціонал повинен бути виділений окремою формою, або іншим способом для того, щоб розпізнати його. Клітини повинні мати можливість заразитися, вилікуватися, вмерти через хворобу. У системі повинні задаватися всі вхідні значення, як ймовірності, кількість клітинок, кількість хворих.

Система повинна мати можливість побудувати модель та за вказаний період часу вивести статистичні дані результатів моделювання.

3.2. Обґрунтування алгоритму й структури програми

3.3. Обґрунтування вибору засобів розробки

3.4. Опис розробки програми

3.5. Створення об'єктів і розробка головної програми

3.6. Опис файлів даних та інтерфейсу програми

Як джерело даних

Висновки

Дана робота включає в себе реалізовані пункти поставленого завдання.

У цій роботі я проаналізувала різні існуючі моделі для прогнозування розповсюдження хвороб, що допомогло сформуванню загального бачення системи.

Було визначено основні критерії, що впливають на процес розповсюдження інфекцій, а саме соціальне дистанціювання, носіння маски, тощо. Було визначено 3 основних стани клітин – вразливі, хворі, та ті, що одужали.

Зважаючи на актуальність теми розповсюдження хвороб, дана розробка буде корисною для користувачів з різними намірами. А це ще раз доводить актуальність даної роботи.

Список використаної літератури

1. <https://www.who.int/hac/techguidance/training/predeployment/Health%20Mapper.pdf>
2. <https://www.anylogic.ru/>
3. <https://www.researchgate.net/post/What-is-the-difference-between-agent-based-modelling-and-cellular-automata-approach-to-modelling>
4. <http://science.lpnu.ua/sites/default/files/journal-paper/2019/apr/15992/vis673ism-154-162.pdf>
5. Cellular Automata Model for Epidemics, Sharon Chang
6. https://uk.wikipedia.org/wiki/%D0%9F%D0%BE%D0%BB%D1%96%D0%B3%D0%B0%D0%BC%D0%BD%D1%96_%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D1%96_%D0%B2_%D0%B5%D0%BF%D1%96%D0%B4%D0%B5%D0%BC%D1%96%D0%BE%D0%BB%D0%BE%D0%B3%D1%96%D1%97
7. http://sedici.unlp.edu.ar/bitstream/handle/10915/56001/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y
8. The Foundations for a New Kind of Science, Chapter 1 -2> Stephen Wolfram
9. Forecasting the 2019-ncov Epidemic in Wuhan by SEIR and Cellular Automata Model. SiYou Wang¹, Hui Fang¹, ZhiHong Ma² and XiuFen Wang¹
10. A Cellular Automata SIR Model for Landscape Epidemiology . Erin L Landguth

Додаток А. Програмний код