

Изследване на паралелната ефективност при решаване на задача от еко-епидемиологията

Силви - Мария Гюрова

Институт по информационни и комуникационни технологии - БАН,
София, България

Конференция по проект EuroCC - България:
"Високопроизводителни пресмятания в полза на
изследователите и обществото"
21 - 22.11.2022

- 1 Математическо моделиране
 - Математическо моделиране в епидемиологията
 - Модели на взаимодействащи популации
- 2 Модел хищник-жертва с епидемия в популацията на жертвата
 - Модел хищник-жертва с епидемия в популацията на жертвата
 - Безразмерна версия на модела
 - Математически твърдения и репродуктивно число
 - Явен метод на Ойлер за решаване на система ОДУ
- 3 Паралелна имплементация
- 4 Използвани НРС системи
- 5 Изследване на паралелната ефективност
- 6 Числени резултати
 - Числени резултати за неустойчив случай
 - Числени резултати за устойчив случай
- 7 Заключение & благодарности

1.1 Математическо моделиране в епидемиологията

Модел	ОДУ	Означения
SIS	$S' = -\beta SI + \gamma I + \mu N - \nu S$ $I' = \beta SI - \gamma I - \nu I$ $N = S + I$	μ - параметър на раждаемост ν - параметър на смъртност
SIR	$S' = -\beta SI + \mu N - \nu S$ $I' = \beta SI - \gamma I - \nu I$ $R' = \gamma I - \nu R$ $N = S + I + R$	β - параметър на инфекция
SIRS	$S' = -\beta SI + \mu N - \nu S + \zeta R$ $I' = \beta SI - \gamma I - \nu I$ $R' = \gamma I - \nu R - \zeta R$ $N = S + I + R$	γ - параметър на възстановяване
SEIR	$S' = -\beta SI + \mu N - \nu S$ $E' = \beta SI - \sigma E - \nu E$ $I' = \sigma E - \gamma I - \nu I$ $R' = \gamma I - \nu R$ $N = S + E + I + R$	σ - параметър, при който латентните индивиди стават инфекциозни
SEIRS	$S' = -\beta SI + \mu N - \nu S + \zeta R$ $E' = \beta SI - \sigma E - \nu E$ $I' = \sigma E - \gamma I - \nu I$ $R' = \gamma I - \nu R - \zeta R$ $N = S + E + I + R$	ζ - параметър, при който възстановените индивиди се връщат към податливите поради загуба на имунитет

Таблица: Видове епидемиологични модели

1.2 Модели на взаимодействиящи популации

Системата

$$\begin{cases} \frac{dN}{dt} = N(a - bP) = aN - bNP \\ \frac{dP}{dt} = P(cN - d) = cNP - dP, \end{cases}$$

е известна като модел на Лотка-Волтера, който предлага прост модел за хищничеството на един вид с друг, където $N(t)$ е популацията на жертвата, а $P(t)$ – популацията на хищника за време t и a, b, c, d са параметри, описващи раждаемостта на жертвата, хищничеството и смъртта на хищника. Моделът може да бъде променян, ако параметрите a, b, c, d се заменят с функции, които зависят от N, P и се добавят нови членове към уравненията, описващи процеси като миграция, смъртност на жертвата и други.

2.1 Модел хищник-жертва с SIERS епидемия в популацията на жертвата

$$\begin{aligned}
 \frac{dS}{dt} &= rS\left(1 - \frac{S+E+I+R}{K}\right) - \beta SI - q_1 SP - d_1 S - m_1 S && \underbrace{+\zeta R}_{\text{възстановените видове стават податливи}} \\
 \frac{dE}{dt} &= \underbrace{\beta SI}_{\text{инфекция}} - \underbrace{q_2 EP}_{\text{хищничество}} - \underbrace{\sigma E}_{\text{изложените видове стават заразни}} - \underbrace{d_2 E}_{\text{смъртност}} - \underbrace{m_2 E}_{\text{миграция}} \\
 \frac{dI}{dt} &= \sigma E - \underbrace{\gamma I}_{\text{заразените видове се възстановяват}} - \underbrace{q_3 IP}_{\text{хищничество}} - \underbrace{d_3 I}_{\text{смъртност}} - \underbrace{m_3 I}_{\text{миграция}} - \underbrace{cI}_{\text{смъртност от инфекцията}} \\
 \frac{dR}{dt} &= \gamma I - \underbrace{q_4 RP}_{\text{хищничество}} - \underbrace{d_4 R}_{\text{смъртност}} - \underbrace{m_4 R}_{\text{миграция}} - \zeta R \\
 \frac{dP}{dt} &= \underbrace{q_1 SP}_{\text{хищничество}} + \underbrace{q_2 EP}_{\text{хищничество}} + \underbrace{q_3 IP}_{\text{хищничество}} + \underbrace{q_4 RP}_{\text{хищничество}} - \underbrace{d_5 P}_{\text{смъртност}} - \underbrace{m_5 P}_{\text{миграция}},
 \end{aligned} \tag{1}$$

$S(0) = S_0 \geq 0$, $E(0) = E_0 \geq 0$, $I(0) = I_0 \geq 0$, $R(0) = R_0 \geq 0$, $P(0) = P_0 \geq 0$,
 където $S(t)$ – податливи жертви, $E(t)$ – изложени на зараза жертви,
 $I(t)$ – инфектирани жертви, $R(t)$ – възстановени жертви и $P(t)$ – хищник.

2.2 Безразмерна версия на модела

Системата (1) може да се представи в безразмерен вид (2), когато са извършени следните полагания:

$$\bar{s} = \frac{S}{K}, \quad \bar{e} = \frac{E}{K}, \quad \bar{i} = \frac{I}{K}, \quad \bar{r} = \frac{R}{K}, \quad \bar{p} = \frac{P}{K}, \quad \bar{t} = rt,$$

$$\bar{q}_i = \frac{q_i K}{r} \quad i = \overline{1,4}, \quad \bar{m}_i = \frac{m_i}{r}, \quad \bar{d}_i = \frac{d_i}{r} \quad i = \overline{1,5},$$

$$\bar{\beta} = \frac{\beta K}{r}, \quad \bar{c} = \frac{c}{r}, \quad \bar{\gamma} = \frac{\gamma}{r}, \quad \bar{\sigma} = \frac{\sigma}{r}, \quad \bar{\zeta} = \frac{\zeta}{r},$$

където K – капацитет на насищане, r – параметър на раждаемост, t – време.

$$\left| \begin{array}{l} \frac{d\bar{s}}{dt} = \bar{s}[1 - (\bar{s} + \bar{e} + \bar{i} + \bar{r})] - \bar{\beta}\bar{s}\bar{i} - \bar{q}_1\bar{s}\bar{p} - \bar{d}_1\bar{s} - \bar{m}_1\bar{s} + \bar{\zeta}\bar{s} \\ \frac{d\bar{e}}{dt} = \bar{\beta}\bar{s}\bar{i} - \bar{q}_2\bar{e}\bar{p} - \bar{\sigma}\bar{e} - \bar{d}_2\bar{e} - \bar{m}_2\bar{e} \\ \frac{d\bar{i}}{dt} = \bar{\sigma}\bar{e} - \bar{\gamma}\bar{i} - \bar{q}_3\bar{i}\bar{p} - \bar{d}_3\bar{i} - \bar{c}\bar{i} - \bar{m}_3\bar{i} \\ \frac{d\bar{r}}{dt} = \bar{\gamma}\bar{i} - \bar{q}_4\bar{r}\bar{p} - \bar{d}_4\bar{r} - \bar{m}_4\bar{r} - \bar{\zeta}\bar{r} \\ \frac{d\bar{p}}{dt} = \bar{q}_1\bar{s}\bar{p} + \bar{q}_2\bar{e}\bar{p} + \bar{q}_3\bar{i}\bar{p} + \bar{q}_4\bar{r}\bar{p} - \bar{d}_5\bar{p} - \bar{m}_5\bar{p}, \end{array} \right. \quad (2)$$

$$\bar{s}(0) \geq 0, \quad \bar{e}(0) \geq 0, \quad \bar{i}(0) \geq 0, \quad \bar{r}(0) \geq 0, \quad \bar{p}(0) \geq 0.$$

2.3 Математически твърдения и репродуктивно число (1/2)

Математически твърдения:

- 1 Решенията на двете системи (1) и (2) са дефинирани в R_+^5 и са ограничени.
- 2 Популацията на податливите жертви (S) е ограничена.

За двете системи (1) и (2), репродуктивното число се представя във вида:

$$R_0 = \frac{\bar{\sigma}\bar{\beta}s^*}{(\bar{q}_2p^* + \bar{\sigma} + \bar{d}_2 + \bar{m}_2)(\bar{\gamma} + \bar{q}_3p^* + \bar{d}_3 + \bar{c} + \bar{m}_3)}, \quad (3)$$

където ако $R_0 > 1$ - инфекцията ще продължи да се разпространява, а когато $R_0 < 1$ - инфекцията изчезва. Като се използва

безразмерните стойности на параметрите и стойностите на координати в **равновесна точка без заболяване** $E(s^*, 0, 0, 0, p^*)$ се пресмята репродуктивното число.

2.3 Математически твърдения и репродуктивно число (2/2)

Направени са следните предположения:

- 1 Зададени са различни числени стойности за параметрите на естествена смъртност за отделните популации (d_i $i = \overline{1,5}$).
- 2 Числени стойности за процеса хищничество са различни, защото когато жертвата е изложена на заболяване или е инфектирана, тя по-лесно може да бъде хваната от хищника, за разлика от податливата или възстановената жертва.
- 3 Числени стойности за процеса миграция m_1, m_4 имат по-високи от тези на m_2, m_3 , тъй като податливата или оздравялата жертва може да изминава по-дълго разстояние сравнение с изложената на зараза или инфектираната жертва.

2.4 Явен метод на Ойлер за решаване на система ОДУ

Методът на Ойлер се използва за решаване на системата, описваща модела хищник-жертва с епидемично заболяване SEIRS в жертва. Този числен подход е един от най-популярните методи за решаване на системи с нелинейни ОДУ.

$$\left\{ \begin{array}{l} \frac{s_{j+1}-s_j}{h} = s_j \cdot [1 - (s_j + e_j + i_j + r_j)] - \beta s_j \cdot i_j - q_1 \cdot s_j \cdot p_j - d_1 \cdot s_j - m_1 \cdot s_j + \zeta \cdot s_j \\ \frac{e_{j+1}-e_j}{h} = \beta \cdot s_j \cdot i_j - q_2 \cdot e_j \cdot p_j - \sigma \cdot e_j - d_2 \cdot e_j - m_2 \cdot e_j \\ \frac{i_{j+1}-i_j}{h} = \sigma \cdot e_j - \gamma \cdot i_j - q_3 \cdot i_j \cdot p_j - d_3 \cdot i_j - c \cdot i_j - m_3 \cdot i_j \\ \frac{r_{j+1}-r_j}{h} = \gamma \cdot i_j - q_4 \cdot r_j \cdot p_j - d_4 \cdot r_j - m_4 \cdot r_j - \zeta \cdot r_j \\ \frac{p_{j+1}-p_j}{h} = q_1 \cdot s_j \cdot p_j + q_2 \cdot e_j \cdot p_j + q_3 \cdot i_j \cdot p_j + q_4 \cdot r_j \cdot p_j - d_5 \cdot p_j - m_5 \cdot p_j, \end{array} \right. \quad (4)$$

и

$$h = \frac{t_{j+1} - t_j}{N}, \quad j = \overline{0, N-1} \quad N - \text{брой стъпки} \quad t - \text{време}$$

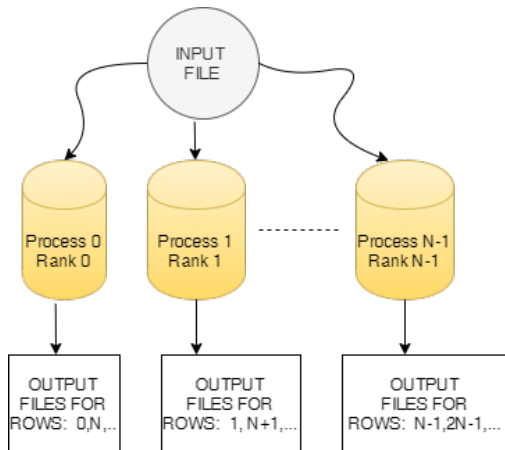
$$s(0) \geq 0, \quad e(0) \geq 0, \quad i(0) \geq 0, \quad r(0) \geq 0, \quad p(0) \geq 0.$$

3. Паралелна имплементация (1/2)

- Създадена е паралелна версия, използваща MPI, написана на C++. Като предварителен процес на производство всичките 21 параметъра са групирани в 4 групи, които са дискретизирани с фиксирана стъпка в даден интервал.
- Получени са повече от 60 000 на брой входящи данни.
- Като се използват тези точки, по метода на Ойлер за различни начални данни, се решава нелинейната ОДУ система. За получаването на добро приближение на решенията, във времевия интервал от 0 до 240 дни, са направени 24 000 итерации.
- Получени са общо 69 230 изходни файла. Те съдържат решение на системата във фиксирана точка, която описва промените в популацията в даден интервал от време.
- Паралелното изпълнение на алгоритъма използва модулна аритметика. Това означава, че всеки ред от входната матрица се чете и данните се изпращат към съответния процес, определен чрез модулната аритметика, за да се изчисли решението на системата ОДУ, като се използва методът на Ойлер.

3. Паралелна имплементация (2/2)

Фигурата показва как се разпределя работата като се използва модулна аритметика, когато имаме N на брой процеса и N на брой ранга.



Фигура: Паралелна схема за разпределение на входната матрица с 69 230 реда и 21 колони

4. Използвани HPC системи (1/2)

1. ARGO HPC system - Department of Informatics and Telecommunications of the University of Athens (Greece)

- Number of Nodes: 12 Dell OptiPlex 7050 MT with Quad-Core Intel-Core i5-6500 @ 3.20 GHz processors offers a total of 96 cores (CPU cores)
- Operating System: CentOS/RHEL
- Compiler: GNU8
- MPI: mpich, mvapich2, openmpi3
- InfiniBand FDR
- Batch system: PBS Workload Manager to distribute jobs
- Storage: Lustre file system to write output files

4. Използвани HPC системи (2/2)

2. AVITOHOL HPC system - ИКТ-BAS (Bulgaria)

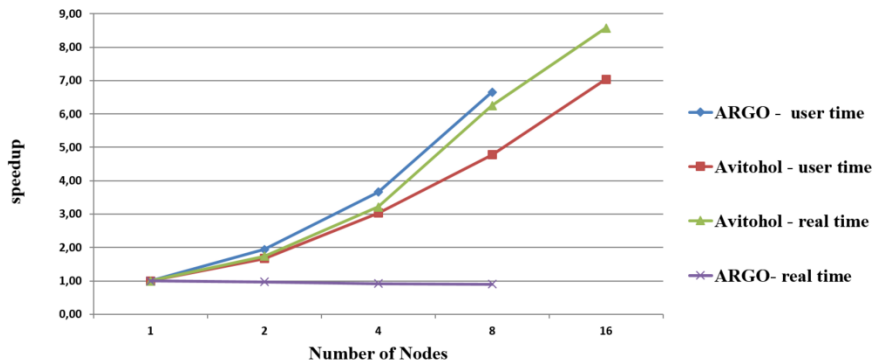
- Number of Nodes: 150 HP Cluster Platform SL250S GEN8 servers each with 2 Intel Xeon E 2650 v2 CPUs and offers a total of 2400 CPU cores with accelerators dual Intel Xeon Phi 7120p
- Operating System: Red Hat Enterprise Linux for HPC
- Compiler: Intel Composer XE 2015
- InfiniBand FDR
- Batch system: PBS Workload Manager to distribute jobs
- SHM (Shared Memory) system for write to output files and Lustre file system

5. Изследване на паралелната ефективност (1/3)

HPC system	nodes	total core	user time (s)	real time (s)
ARGO	1	8	652.70	874,11
	2	16	337.01	902,94
	4	32	178.31	953,16
	8	64	98.16	975,86
AVITOHOL	1	16	38.18	589,81
	2	32	22.81	339,94
	4	64	12,62	183,57
	8	128	7,98	94,41
	16	256	5,42	68,80

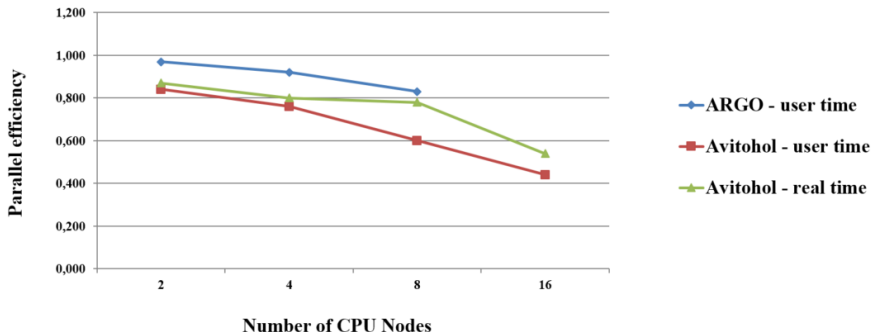
Фигура: Числени резултати въз основа на двете HPC системи

5. Изследване на паралелната ефективност (2/3)



Фигура: Скорост

5. Изследване на паралелната ефективност (3/3)



Фигура: Паралелна ефективност

6.1 Числени резултати за неустойчив случай (1/2)

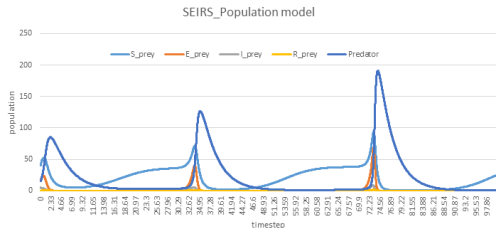
r	K	β	q_1	q_2	q_3	q_4	m_1	m_2	m_3	m_4
0.80	100	0.20	0.01	0.03	0.10	0.01	0.40	0.10	0.03	0.15
m_5	d_1	d_2	d_3	d_4	d_5	c	σ	γ	ζ	
0.25	0.08	0.07	0.04	0.03	0.05	0.16	0.50	0.25	0.25	

Таблица: Числени стойности на параметрите за модела (1)

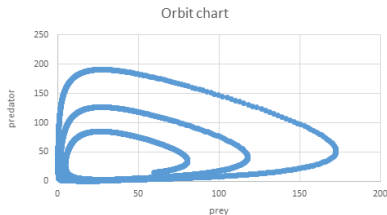
r	K	β	q_1	q_2	q_3	q_4	m_1	m_2	m_3	
1	-	25	1.25	3.75	12.50	1.25	0.50	0.125	0.0625	
m_4	m_5	d_1	d_2	d_3	d_4	d_5	c	σ	γ	ζ
0.1875	0.3125	0.10	0.0875	0.05	0.0375	0.075	0.20	0.625	0.3125	0.025

Таблица: Числени стойности на параметрите за безразмерния модел (2)

6.1 Числени резултати за неустойчив случай (2/2)



Фигура: А) Популяционната динамика на модела (1) при начални данни:
 $S(0) = 40, E(0) = 15, I(0) = 5, R(0) = 0, P(0) = 10$



Фигура: В) Орбитална диаграма на модела хищник-жертва

Забележка

- Няма устойчива равновесна точка.
- $R_0 = 2.86605 > 1$ – има епидемия.

6.2 Числени резултати за стабилен случай (1/2)

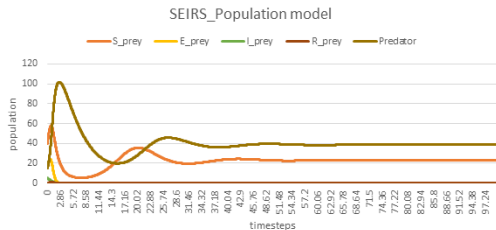
r	K	β	q_1	q_2	q_3	q_4	m_1	m_2	m_3	m_4
0.80	100	0.20	0.01	0.03	0.10	0.01	0.20	0.10	0.05	0.15
m_5	d_1	d_2	d_3	d_4	d_5	c	σ	γ	ζ	
0.20	0.03	0.04	0.05	0.03	0.03	0.15	0.40	0.30	0.25	

Таблица: Числени стойности на параметрите за модела (1)

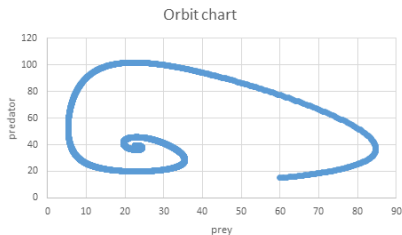
r	K	β	q_1	q_2	q_3	q_4	m_1	m_2	m_3	
1.00	-	0.25	1.25	3.75	12.50	1.250	0.250	0.125	0.0625	
m_4	m_5	d_1	d_2	d_3	d_4	d_5	c	σ	γ	ζ
0.1875	0.250	0.0375	0.05	0.0625	0.0375	0.0375	0.1875	0.50	0.375	0.025

Таблица: Числени стойности на параметрите за безразмерния модел (2)

6.2 Числени резултати за устойчив случай (2/2)



Фигура: А) Популяционна динамика на модела (1) с начални данни:
 $S(0) = 40, E(0) = 15, I(0) = 5, R(0) = 0, P(0) = 10$



Фигура: В) Орбитална диаграма на модела хищник-жертва

Забележка

- Съществува устойчива равновесна точка.
- $R_0 = 0.245721 < 1$ – няма епидемия.

7. Заключение & благодарности (1/2)

- 1 Паралелното внедряване на алгоритъма ни позволява да получим голям брой решения на системата, описваща взаимодействието хищник-жертва със SEIRS епидемично заболяване в жертвата.
- 2 Като се използват тези резултати, може да се изберат точки на равновесие, в които динамиката на популациите да е стабилна и заболяването да се контролира.
- 3 Големият брой решения позволява да се създаде видео, демонстрирайки как коефициентите влияят на популационната динамика на двата животински вида в епидемична среда.
- 4 Резултатите от таблицата за изследване на паралелната ефективност показват, че добрата паралелна ефективност на алгоритъма зависи от архитектурата на HPC системите и от избора на файловата система за отпечатване на изходните файлове.

7. Заключение & благодарности (2/2)

Благодарности:

- 1 ЕуроСС Национален център за компетентност България се финансира от Европейската програма за научни изследвания и иновации – Хоризонт2020 (проект EUROCC) съгласно споразумение за безвъзмездна помощ No 951732 и от Министерството на образованието и науката съгласно споразумение No Д0-49/02.03.2021 г.
- 2 Тази работа е извършена с частичната подкрепа от договор No BG05M2OP001-1.001-0003 "ЦВП по Информатика и ИКТ финансиран чрез Оперативна програма „Наука и образование за интелигентен растеж“ 2014-2020 от Европейския фонд за регионално развитие.