

Unioeste - Universidade Estadual do Oeste do Paraná
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
Colegiado de Ciência da Computação
Curso de Bacharelado em Ciência da Computação

Um Modelo Multiagente *Bitstring* em CUDA para Simular a Propagação de Hipotéticas Doenças Baseadas em Modelagem Compartmental Tipo SEIRS

Wesley Luciano Kaizer

CASCABEL
2016

WESLEY LUCIANO KAIZER

**UM MODELO MULTIAGENTE *BITSTRING* EM CUDA PARA
SIMULAR A PROPAGAÇÃO DE HIPOTÉTICAS DOENÇAS BASEADAS
EM MODELAGEM COMPARTIMENTAL TIPO SEIRS**

Monografia apresentada como requisito parcial
para obtenção do grau de Bacharel em Ciência da
Computação, do Centro de Ciências Exatas e Tec-
nológicas da Universidade Estadual do Oeste do
Paraná - Campus de Cascavel

Orientador: Prof. Dr. Rogério Luís Rizzi

CASCABEL
2016

WESLEY LUCIANO KAIZER

**UM MODELO MULTIAGENTE *BITSTRING* EM CUDA PARA
SIMULAR A PROPAGAÇÃO DE HIPOTÉTICAS DOENÇAS BASEADAS
EM MODELAGEM COMPARTIMENTAL TIPO SEIRS**

Monografia apresentada como requisito parcial para obtenção do Título de Bacharel em
Ciência da Computação, pela Universidade Estadual do Oeste do Paraná, Campus de Cascavel,
aprovada pela Comissão formada pelos professores:

Prof. Dr. Rogério Luís Rizzi (Orientador)
Colegiado de Matemática, UNIOESTE

Profa. Dra. Claudia Brandelero Rizzi
Colegiado de Ciência da Computação,
UNIOESTE

Prof. Dr. Guilherme Galante
Colegiado de Ciência da Computação,
UNIOESTE

Cascavel, 4 de julho de 2016

DEDICATÓRIA

AGRADECIMENTOS

Lista de Figuras

2.1	Fluxo de transição de estados dos indivíduos no modelo SEIRS	6
2.2	Representação em palavra de 32 bits ao modelo SEIRS baseado em indivíduos .	15

Lista de Tabelas

Lista de Abreviaturas e Siglas

SIS	Modelo Suscetível-Infectado-Suscetível
SIR	Modelo Suscetível-Infectado-Recuperado
SEIR	Modelo Suscetível-Exposto-Infectado-Recuperado
SEIRS	Modelo Suscetível-Exposto-Infectado-Recuperado-Suscetível
API	<i>Application Programming Interface</i>
CUDA	<i>Compute Unified Device Architecture</i>
CPU	<i>Central Processing Unit</i>
GPU	<i>Graphics Processing Unit</i>
GPGPU	<i>General Purpose Graphics Processing Unit</i>
EDO	Equações Diferenciais Ordinárias
PVI	Problema de Valor Inicial
PVIC	Problema de Valor Inicial e de Contorno
AC	Autômato Celular
LGCA	<i>Lattice Gas Cellular Automata</i>

Lista de Símbolos

α	Taxa de infecção
γ	Período de exposição
α	Período de infectância
δ	Período de recuperação

Sumário

Lista de Figuras	vi
Lista de Tabelas	vii
Lista de Abreviaturas e Siglas	viii
Lista de Símbolos	ix
Sumário	x
Resumo	xii
1 Introdução	1
1.1 Objetivos	2
1.2 Motivação e Justificativas	3
1.3 Organização do Trabalho	3
2 Fundamentos	5
2.1 Epidemiologia Computacional	5
2.2 Modelos compartimentais: Tipos e Classificações	5
2.3 Agentes e Multiagentes	10
2.4 Modelagem em Operadores	11
2.5 <i>Bitstring</i>	14
2.6 Refinamento do Modelo	18
3 Metodologias Computacionais	19
3.1 Introdução	19
3.2 SIMULA	19
3.3 Estruturas de Dados, Linguagens e Estratégias de Implementação	19
3.4 CUDA e OpenMP	19

4	Soluções	20
4.1	Introdução	20
4.2	Implementação Padrão com CUDA e OpenMP	20
4.3	Implementação <i>Bitstring</i> com CUDA e OpenMP	20
4.4	Discussões Qualitativas, Quantitativas, Eficiência e Acurácia	20
5	Resultados e Discussões	21
5.1	Introdução	21
5.2	Cases: Discutir as Simulações no Espaço Geográfico Escolhido.	21
	Glossário	22

Resumo

A aplicação de modelagem compartimental na epidemiologia é amplamente estudada, como pode-se observar na extensa literatura disponível. A simulação de dinâmicas epidemiológicas é de particular interesse no estudo, prevenção e controle de doenças transmissíveis. Com base nestas premissas, este trabalho busca abordar o problema de simulação de hipotéticas doenças baseadas em modelagem compartimental SEIRS, por meio de sistema multiagente. Para obter uma adequada formulação à simulação em GPGPU na plataforma CUDA, ao modelo é empregada a metodologia de *bitstring*. Ao ambiente computacional são utilizadas técnicas e software especificamente desenvolvido para manipular dados georreferenciados à especificação e composição de um *lattice* apropriado à simulação computacional. Como resultado pretende-se apresentar uma aplicação funcional e adequada para simular eventos epidemiológicos em uma região da cidade de Cascavel/PR. A avaliação do modelo proposto é executada através da realização de experimentos numérico-computacionais, que são comparados com os obtidos na literatura, sobretudo nos aspectos quantitativos às implementações realizadas.

Palavras-chave: Epidemiologia computacional, Sistema multiagente, Modelos compartimentais, Modelagem *bitstring*, Plataforma Computacional CUDA.

Capítulo 1

Introdução

Epidemiologia pode ser definida como o estudo da frequência, da distribuição e dos estados ou eventos relacionados com o espalhamento de doenças transmissíveis e ocorrências de doenças não transmissíveis em populações específicas, e a aplicação dos resultados desses estudos na prevenção e controle dos problemas decorrentes e relacionados com a saúde pública. Modelos computacionais baseados em indivíduos vêm sendo empregados na epidemiologia para estudar a propagação e transmissão de doenças, que são processos centrais na dinâmica de doenças infecto-contagiosas. O uso destes modelos viabiliza a modelagem de fenômenos de natureza probabilística e da heterogeneidade nas relações entre os indivíduos e o meio, conferindo mais realidade ao modelo estudado. Modelos compartimentais podem ser utilizados à definição de modelos mais complexos, tendo como base a subdivisão da população em categorias, em que os indivíduos fluem entre elas de acordo com determinadas taxas e cenários, podendo respeitar as características particulares de uma doença de interesse.

À implementação de modelos baseados em indivíduos em uma linguagem computacional pode ser relevante utilizar abordagens mais eficientes à codificação do sistema e definir um *lattice* apropriado à execução dos experimentos computacionais. Dependendo da dimensão do *lattice*, da quantidade de indivíduos e da complexidade das dinâmicas modeladas, é desejável otimizar o tempo de execução dos experimentos, utilizando os recursos computacionais de processamento e armazenamento disponíveis da forma mais eficiente possível.

1.1 Objetivos

O objetivo principal deste trabalho é propor, desenvolver, implementar e avaliar um modelo epidemiológico computacional multiagente, com formulação em *bitstring*, para simular computacionalmente a propagação de doenças que possam ser modeladas por modelos compartimentais tipo Suscetível, Exposto, Infectado, Recuperado e Suscetível, SEIRS, utilizando como *lattice* um mapeamento de uma região da cidade de Cascavel/PR. A solução computacional do modelo proposto contempla uma implementação utilizando *Compute Unified Device Architecture*, CUDA, para obter máxima eficiência computacional por meio de paralelismo de dados. Para alcançar este objetivo é necessária a realização de objetivos mais específicos que contemplam:

1. Revisão bibliográfica nas temáticas pertinentes ao trabalho, incluindo temas como epidemiologia computacional, modelagem compartimental, sistema multiagente, formulação *bitstring*, estruturas de dados e plataforma computacional paralela CUDA.
2. Desenvolvimento e implementação de um modelo multiagente em *bitstring*, baseado em formulação compartimental, tendo como *lattice* um mapeamento de uma quadra da cidade de Cascavel.
3. Paralelização em nível de dados do sistema multiagente em *bitstring*, utilizando a plataforma computacional paralela CUDA.
4. O emprego e o aperfeiçoamento de uma ferramenta computacional para viabilizar e otimizar as fases de pré-processamento, processamento e pós-processamento da simulação, como as etapas de configuração e visualização dos resultados obtidos, utilizando o *lattice* gerado por tal software.
5. Realização de experimentos numérico-computacionais, visando avaliar a acurácia da solução implementada, bem como sua eficiência computacional, e demais aspectos computacionais.

1.2 Motivação e Justificativas

Em conformidade com os objetivos estabelecidos, este trabalho é motivado e se justifica pelos seguintes motivos:

- O estudo e aplicação de modelos compartimentais em epidemiologia computacional é relevante, pois mostram-se adequados e flexíveis à modelagem de hipotéticas doenças, bem como são amplamente utilizados, como apresentado na literatura técnica.
- O uso de agentes computacionais em simulações permite a modelagem mais realística dos fenômenos epidemiológicos de interesse, pois os indivíduos têm rica e emergente dinâmica, independentemente uns dos outros, além de viabilizar o emprego de metodologia à sua paralelização.
- A abordagem em multiagente com *bitstring* é relativamente nova e relevante, viabilizando a modelagem de agentes computacionais de forma concisa e eficiente, evitando desperdícios de memória, simplificando o processo de captura e configuração de atributos dos indivíduos e facilitando a implementação na plataforma CUDA por reduzir significativamente a quantidade de dados nas transferências entre CPU e GPU, além de diminuir a complexidade das estruturas de dados utilizadas.
- O uso da plataforma CUDA é atrativo por possibilitar a paralelização massiva do sistema implementado, esperando-se ganhos de desempenho desejáveis nos experimentos computacionais que serão realizados.
- Por fim, o trabalho realizado é relevante à formação técnica do futuro profissional em Ciência da Computação.

1.3 Organização do Trabalho

Este trabalho está organizado da seguinte forma: no Capítulo 2 é apresentada e discutida a fundamentação teórica utilizada para a realização do trabalho de conclusão de curso, envolvendo temáticas como a epidemiologia computacional, modelagem compartimental, sistema multiagente e modelagem *bitstring*. No Capítulo 3 são apresentados os métodos utilizados à

modelagem, implementação e testes do sistema multiagente, como estruturas de dados, linguagens, APIs e demais softwares de apoio. No Capítulo 4 são discutidas as implementações realizadas, fazendo-se comparações entre elas e discutindo demais itens pertinentes. Por fim, o Capítulo 5 discute resultados obtidos por meio da execução de testes utilizando as implementações realizadas.

Capítulo 2

Fundamentos

2.1 Epidemiologia Computacional

Epidemiologia pode ser definida como o estudo da frequência, da distribuição e dos determinantes dos estados ou eventos relacionados com a distribuição de doenças transmissíveis e não transmissíveis em populações específicas, e a aplicação dos resultados desses estudos na prevenção e controle dos problemas de saúde. Doenças transmissíveis têm fases definidas à latência, incubação, infecciosidade e recuperação da infecção. Essa caracterização é determinante na construção de um modelo representativo do cenário endêmico ou epidêmico.

É importante determinar se o modelo será do tipo determinístico ou estocástico, discreto ou contínuo, com população homogênea ou heterogênea, entre outros aspectos. No processo de modelagem, ao abstrair um problema complexo a um outro mais tratável, inevitavelmente simplifica-se características importantes do problema original. É pertinente o estudo da acurácia do modelo através da execução de testes que são usados na sua calibração e validação.

2.2 Modelos compartimentais: Tipos e Classificações

A abordagem usualmente empregada na modelagem de doenças infecto-contagiosas é a de diagrama de blocos ou em compartimentos. Doenças infecciosas admitem uma modelagem em estrutura compartimental, que é obtida por meio da subdivisão da população em categorias, entre as quais os indivíduos fluem com taxas que dependem das características da dinâmica espaço-temporal da doença. Supondo que $N(t)$ designa a quantidade total de indivíduos num nível de tempo t e tomando uma modelagem compartimental, os indivíduos que compõem a população podem ser divididos em classes do tipo suscetíveis, $S(t)$; latentes ou expostos, $E(t)$;

infectados, $I(t)$, e recuperados ou removidos, $R(t)$, de modo que $S(t) + E(t) + I(t) + R(t) = N(t)$, observando-se a conservatividade do número de indivíduos na população.

Cada indivíduo pode pertencer a somente um dos estados em um dado período de tempo: o estado suscetível, S , indica que o indivíduo ainda não contraiu a doença e está apto a adquiri-la. O estado exposto, E , indica que o indivíduo contraiu a doença, mas não pode transmiti-la. O estado infectante, I , indica que o indivíduo está infectado com a doença, sendo capaz de transmiti-la para outros que estejam no estado suscetível, e o estado recuperado, R , indica que o indivíduo já foi infectado com a doença e não pode se contaminar novamente nem transmiti-la para outros em um determinado período.

Considerando-se os estados relacionados com o processo infeccioso, usuais modelos são o Suscetível-Infectado-Suscetível, SIS, empregado em situações em que a doença não confere imunidade, passando o indivíduo de infectado para suscetível novamente, o modelo Suscetível-Infectado-Recuperado, SIR, para casos de doenças em que os indivíduos infecciosos podem recuperar-se e adquirir imunidade permanente, o modelo Suscetível-Exposto-Infectado-Recuperado, SEIR, para doenças com período de latência e imunidade permanente e o modelo Suscetível-Exposto-Infectado-Recuperado-Suscetível, SEIRS, para doenças com período de latência e sem imunidade permanente. Neste trabalho será adotado particularmente o modelo SEIRS, visto que pode ser usado para modelar uma ampla gama de doenças. O fluxo de transição de estado dos indivíduos para o modelo SEIRS é mostrado na Figura 2.1.

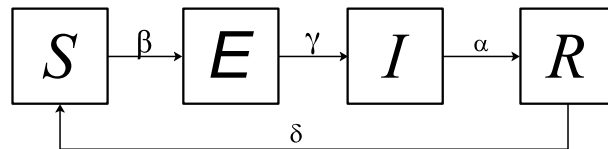


Figura 2.1: Fluxo de transição de estados dos indivíduos no modelo SEIRS

Na Figura 2.1, β designa a taxa de transmissão da doença. No encontro de indivíduos saudáveis com infectantes a doença se propaga a tal taxa. O indivíduo fica um período, γ , no estado exposto. O parâmetro α designa a taxa de recuperação, em que indivíduos doentes se recuperam estando imunes àquela doença por outro período. Este período é caracterizado pelo parâmetro δ que modelará qual é a taxa de perda de imunidade de indivíduos que estejam recuperados, passando-os novamente para o estado suscetível.

As principais modelagens empregadas na simulação das dinâmicas de doenças infecto-

contagiosas são:

- **Equações Diferenciais Ordinárias:** Considerando-se uma população fixa no tempo, isto é, desprezando por simplicidade as taxas vitais de nascimento e de morte, pode-se descrever matematicamente o fluxo de transição dos possíveis estados dos indivíduos no modelo SEIRS através de um sistema de equações diferenciais ordinárias não lineares, EDO. Se $N = S(t) + E(t) + I(t) + R(t)$ designa o tamanho fixo da população de indivíduos, onde suas quantidades nos diferentes estados variam no tempo, de modo que suas quantidades são funções do tempo, $S = S(t)$, $E = E(t)$, $I = I(t)$ e $R = R(t)$, respectivamente, obtém-se os problemas de valor inicial, PVI, para o caso contínuo e discreto, como apresentados em 2.1.

$$\left\{ \begin{array}{l} \frac{d(S(t))}{dt} = -\frac{\beta}{N} I(t) S(t) + \delta R(t) \\ \frac{d(E(t))}{dt} = +\frac{\beta}{N} I(t) S(t) - \gamma E(t) \\ \frac{d(I(t))}{dt} = +\gamma E(t) - \alpha I(t) \\ \frac{d(R(t))}{dt} = +\alpha I(t) - \delta R(t) \end{array} \right. \quad ; \begin{array}{l} S(0) = S_0 \\ E(0) = E_0 \\ I(0) = I_0 \\ R(0) = R_0 \end{array} \quad \left\{ \begin{array}{l} S^n = S^{n-1} - \frac{\beta}{N} I^{n-1} S^{n-1} + \delta R^{n-1} \\ E^n = E^{n-1} + \frac{\beta}{N} I^{n-1} S^{n-1} - \gamma E^{n-1} \\ I^n = I^{n-1} + \gamma E^{n-1} - \alpha I^{n-1} \\ R^n = R^{n-1} + \alpha I^{n-1} - \delta R^{n-1} \end{array} \right. \quad ; \begin{array}{l} S^0 = S_0 \\ E^0 = E_0 \\ I^0 = I_0 \\ R^0 = R_0 \end{array} \quad (2.1)$$

Nas formulações apresentadas em 2.1, S_0 , E_0 , I_0 e R_0 representam, respectivamente, as distribuições iniciais das populações de suscetíveis, expostos, infectantes e recuperados. A solução analítica do PVI contínuo não é conhecida ou não é possível de ser obtida. Resta, pois, obter numericamente sua solução resolvendo a formulação discreta, em que n e $n - 1$ denotam os diferentes níveis de tempo que são avaliados numericamente às quantidades envolvidas no modelo.

- **Autômatos Celulares:** Autômatos celulares, AC, podem ser classificados como sendo modelos de sistemas dinâmicos discretos no espaço e no tempo, que operam sobre um *lattice* contendo uma quantidade finita ou infinita de células. São caracterizados por iterações locais em que cada célula está associada a um estado ou fase por meio de conjunto discreto de valores. A atualização do estado é realizada a partir dos estados das células vizinhas de acordo com regras locais. Autômatos celulares possuem um domínio especificado pelo ambiente ou espaço celular em que estes são operados; um espaço de estados

das células em que cada célula do *lattice* possui determinado estado; a vizinhança de cada célula; e a regra de transição local que atua sobre a célula de modo que seu estado pode variar ao ocorrer uma transição.

O autômato celular é especificado sobre um *lattice* finito com adequado tipo de vizinhança, sendo sua discretização determinada pelo conjunto $\mathcal{L} = \{(i, j); 1 \leq i \leq n, 1 \leq j \leq m; n, m \in \mathbb{N}^*\}$. Para vizinhança tipo Moore cada célula (i, j) tem uma vizinhança de células $\mathcal{V}_{i,j} = \{(i + \alpha_k, j + \beta_k); 1 \leq k \leq 8\}$. Localmente a vizinhança de Moore é $\mathcal{V} = \{(0, 0), (-1, 0), (-1, 1), (0, 1), (1, 1), (1, 0), (1, -1), (0, -1), (-1, -1)\}$, sendo a vizinhança sem a célula central indicada por $\mathcal{V}^* = \mathcal{V} - \{(0, 0)\}$.

A regra de transição local, σ , é especificada dos estados $e, E \in \mathcal{Q}$, de modo que, dados distintos, discretos e consecutivos níveis de tempo t e $t+1$ têm-se $\sigma(e_{(i+\alpha_k, j+\beta_k)}^t) = E_{i,j}^{t+1}$, sendo $e_{(i+\alpha_k, j+\beta_k)}^t$ a designação para o estado de uma célula vizinha da célula (i, j) no nível de tempo t . Para um espaço celular finito deve-se especificar apropriadas condições de contorno, e algumas delas empregadas em autômato celular são a periódica, tipo reflexiva e a tipo Dirichlet, e um caso particular dessa última, a condição homogênea nula. À especificação da condição de contorno periódica faz-se $E_{1,j} = E_{n,j}$ e $E_{i,1} = E_{i,m}$ e à condição de contorno homogênea faz-se $E_{1,j} = E_{n,j} = 0$ e $E_{i,1} = E_{i,m} = 0$, para tais células em $\partial\mathcal{L}$, o conjunto das células de contorno do domínio.

Autômatos celulares são formalmente caracterizados pela especificação da quádrupla $\langle \mathcal{L}, \mathcal{Q}, \mathcal{V}, \sigma \rangle$ que atenda as características de homogeneidade espaço-temporal na sua dinâmica, e da especificação de regras de transição não ambíguas. Como as simulações de eventos epidemiológicos requerem uma distribuição dos estados no tempo inicial, t_0 , é necessário especificar tal condição. Escrevendo o AC como um problema de valor inicial e de contorno discreto, PVIC, a regra de transição é definida em $(i, j) \times t \in \mathcal{L} \times (0, T)$, sendo as condições iniciais dadas $(i, j) \in \mathcal{L}$ para $t_0 = 0$ e as condições de contorno especificadas em $(i, j) \times t \in \partial\mathcal{L} \times [0, T]$, sendo que T denota um determinado tempo. Considerando-se as expressões para o domínio e a regra de transição, e tomando a condi-

ção de contorno periódica, o AC é especificado como (2.2):

$$\begin{cases} s_{i,j}^{t+1} = s_{i,j}^t - \beta s_{i,j}^t i_{i,j}^t - s_{i,j}^t \left(\sum_{(\alpha,\beta) \in \mathcal{V}^*} \mu_{\alpha,\beta}^{(i,j)} \frac{N_{i+\alpha,j+\beta}}{N_{i,j}} i_{i+\alpha,j+\beta}^t \right) + \delta r_{i,j}^t \\ e_{i,j}^{t+1} = e_{i,j}^t + \beta s_{i,j}^t i_{i,j}^t + s_{i,j}^t \left(\sum_{(\alpha,\beta) \in \mathcal{V}^*} \mu_{\alpha,\beta}^{(i,j)} \frac{N_{i+\alpha,j+\beta}}{N_{i,j}} i_{i+\alpha,j+\beta}^t \right) - \gamma e_{i,j}^t \\ i_{i,j}^{t+1} = i_{i,j}^t + \gamma e_{i,j}^t - \sigma i_{i,j}^t \\ r_{i,j}^{t+1} = r_{i,j}^t + \sigma i_{i,j}^t - \delta r_{i,j}^t \\ s_{i,j}^0 > 0, \quad e_{i,j}^0 \geq 0, \quad i_{i,j}^0 > 0, \quad r_{i,j}^0 \geq 0, \end{cases} \quad (2.2)$$

A solução de (2.2) é obtida por meio da evolução espaço-temporal dada pela regra de transição que fornece os estados finais a partir do cálculo dos estados iniciais. A transição se dá num nível de tempo t para o consecutivo nível de tempo $t + \Delta t$, sendo Δt o passo de tempo.

- **Lattice Gas Cellular Automata:** *Lattice Gas Cellular Automata*, LGCA, são um particular tipo de autômato celular que são utilizados principalmente na simulação de fluxos de fluídos. Para a construção de um LGCA deve-se especificar uma quádrupla $\langle \mathcal{L}, \mathcal{Q}, \mathcal{V}, \sigma \rangle$ que atenda as características de AC. Um modelo LGCA é como apresentado em (2.3).

No modelo (2.3) τ indica um estado qualquer, ou seja, $\tau \in \{S, I, R\}$, e o operador $\tilde{\sigma}$ é um algoritmo do operador σ . Ou seja, $\tilde{\sigma} = \tilde{P} \circ \tilde{D} \circ (1 - (1 - p_{i,j}))^{N_I}$, sendo \tilde{P} e \tilde{D} procedimentos ou algoritmos para implementar as fases de propagação e de randomização.

$$\begin{cases} \tilde{\sigma} \left(\tau_{(i+\alpha_k, j+\beta_k)}^t \right) = \tau_{i,j}^{t+1}; & (i, j) \in \mathcal{V}_{i,j} \subset \mathcal{L}, t > 0 \\ \tau_{1,j} = \tau_{n,j}, \quad \tau_{i,1} = \tau_{i,m}; & (i, j) \times t \in \partial \mathcal{L} \times [0, T] \\ \tau_{(i,j)}^0 = \tau(0); & (i, j) \in \mathcal{L} \end{cases} \quad (2.3)$$

- **Sistemas multiagentes:** Agentes baseados em modelos são definidos espaço-temporalmente especificando-se como ocorre a transição do seu estado num intervalo de tempo e seu movimento no ambiente, de uma posição para outra no passo de tempo, em um ciclo de transição. A especificação formal do agente é realizada por meio de um operador de evolução que define o estado atual do agente, quando interagindo com o ambiente, decorre da composição do operador de transição temporal, que realiza a transição do estado do agente de acordo com sua interação com outros agentes e com o ambiente, e o operador de transição espacial, que movimenta o agente de sua posição, considerando sua conectividade e mobilidade.

2.3 Agentes e Multiagentes

Agentes baseados em modelos são definidos espaço-temporalmente especificando-se como ocorre a transição do seu estado num intervalo de tempo e seu movimento no ambiente, de uma posição para outra no passo de tempo, em um ciclo de transição. A especificação formal do agente é realizada por meio de um operador de evolução que define o estado atual do agente, quando interagindo com o ambiente, decorre da composição do operador de transição temporal, que realiza a transição do estado do agente de acordo com sua interação com outros agentes e com o ambiente, e o operador de transição espacial, que movimenta o agente de sua posição, considerando sua conectividade e mobilidade.

À dinâmica são consideradas as operações de movimentação, contato e de transição de estados. Na movimentação, os indivíduos transitam no ambiente matricial por meio de suas vizinhanças de Moore às posições escolhidas aleatoriamente, respeitando os limites espaciais do ambiente. Nas operações de contato ocorre, probabilisticamente, a transmissão da doença por meio dos indivíduos infectados para os indivíduos suscetíveis que ocupam mesma posição no ambiente. Nas transições de estados ocorrem a passagem de indivíduos suscetíveis para expostos, e de expostos para infectantes, e de infectantes para removidos, nessa sequência e uma vez a cada ciclo.

Um agente $\chi(t)$ é definido como $\chi(t) = (P, C, E)$, em que P é a sua posição no ambiente, E designa o seu estado, e C um registro para o contador de ciclos que controla as transições de estados. Cada agente implementa a operação de evolução que atualiza seu estado atual quando interagindo com o ambiente, sendo ele definido pela composição entre os operadores μ e σ . O operador de conecto-mobilidade, μ , movimenta o agente de sua posição considerando-se os atributos de conectividade e mobilidade, e o operador σ realiza a transição do estado do agente por meio de sua interação com outros agentes e o ambiente.

O operador espaço-temporal $\lambda(\chi(t))$ realiza as operações do agente $\chi(t)$ movimentando-o da posição (i, j) para uma posição (ξ, η) no ciclo de tempo, t , para o ciclo de tempo, $t + \Delta t$. Formalmente o estado do agente $\chi(t)$ é representado como em 2.4, por um problema de valor inicial e de contorno:

$$\begin{cases} \lambda(\chi(t)) = \sigma(\mu(\chi(t))) \\ \lambda_{1,j} = \lambda_{n,j}, \lambda_{i,1} = \lambda_{i,m}; (i,j) \times t \in \partial\mathcal{M} \times [0, T] \\ \lambda_{(i,j)}^0 = \lambda(0); \quad (i,j) \in \mathcal{M} \end{cases} \quad (2.4)$$

sendo que $\partial\mathcal{M}$ designa o contorno do domínio e \mathcal{M} o domínio matricial dos indivíduos, o ambiente em que ocorre a dinâmica multiagente.

2.4 Modelagem em Operadores

Os modelos apresentados em (1) serviram à fundamentação do modelo baseado em agentes, que foi desenvolvido e utilizado nas simulações computacionais. A modelagem empregada para simular o espalhamento de hipotética doença de transmissão direta em indivíduos considera agentes baseados em modelos que são definidos espaço-temporalmente especificando-se como ocorre a transição do seu estado num intervalo de tempo e seu movimento no ambiente, de uma posição para outra no passo de tempo. Um passo de tempo é especificado como um ciclo de transição.

Um agente $\chi(t)$ é definido espaço-temporalmente especificando-se como ocorre a transição do seu estado num intervalo de tempo t e seu movimento no espaço. O estado do agente $\chi(t)$ é especificado através do seguinte conjunto de atributos:

$$\chi(t) \equiv (L, X, Y, C, E) \quad (2.5)$$

Cujos significados dos identificadores dos atributos do estado interno do agente $\chi(t)$ são como:

- Lote (L): Identificador do lote que o agente $\chi(t)$ se encontra.
- Posição em x (X): Coordenada x da posição do agente $\chi(t)$ no lote.
- Posição em y (Y): Coordenada y da posição do agente $\chi(t)$ no lote.
- Contador de ciclos (C): Contador de ciclos que controla os períodos de transição entre os estados do agente.
- Estado (E): Identificador do estado atual do agente $\chi(t)$.

De maneira que um agente é designado por $\chi(t) \equiv (L, X, Y, C, E)$.

A especificação formal de um agente é realizada através de um operador de evolução que define o estado atual do agente, quando interagindo com o ambiente. Esse operador decorre da composição entre os operadores de transição temporal, que realiza uma transição do estado interno do agente considerando-se sua interação com outros agentes e com o ambiente, e o operador de transição espacial, que movimenta o agente de sua posição para outra, considerando-se os atributos de conectividade e mobilidade.

À dinâmica populacional são considerados três tipos de operações, as de movimentação, contato e de transição de estados. Nas operações de movimentação, os indivíduos são movimentados dentro de um ambiente virtual com topologia matricial através de suas vizinhanças de Moore às posições escolhidas aleatoriamente, respeitando os limites do ambiente. Nas operações de contato ocorre, probabilisticamente, a transmissão da doença através dos indivíduos infectados para os indivíduos suscetíveis que ocupam uma mesma posição no ambiente. Nas transições de estados, ocorre a passagem de estados dos indivíduos para expostos, e depois de expostos para infectantes e então de infectantes para recuperados. Por fim de recuperados para suscetíveis. Tais operações são realizadas na sequência em que foram apresentadas e uma vez a cada ciclo, que consiste na aplicação dos operadores sobre a população de indivíduos e geração de arquivos de saída. Uma simulação é composta por vários ciclos.

Cada agente implementa uma operação de evolução λ que atualiza o estado atual do agente quando interagindo com o ambiente, definido como $\lambda(\chi(t)) = \sigma(\mu(\chi(t)))$, que decorre da composição entre os operadores μ e σ . O operador de conecto-mobilidade, μ , movimenta o agente de sua posição considerando-se os atributos de conectividade e mobilidade, e o operador σ realiza a transição do estado interno do agente considerando-se sua interação com outros agentes e o ambiente.

O operador espaço-temporal $\lambda(\chi(t))$ realiza as operações do agente $\chi(t)$ movimentando-o da posição (i, j) para uma posição (ξ, η) no ciclo de tempo atual, t , para o ciclo de tempo, $t + 1$. Formalmente o estado resultante das operações no agente $\chi(t)$ é representado como em (2):

$$\lambda(\chi(t)^{(i,j)}) \equiv \sigma(\mu(\chi(t+1))^{(\xi,\eta)}) \quad (2.6)$$

Para especificar algoritmicamente tal operador é necessário determinar configurações e parâmetros que estabelecem quais as condições do ambiente e do estado interno do agente. O ambiente é composto por diversas estruturas matriciais, que representam lotes heterogêneos com $\#L$ linhas e $\#C$ colunas, onde se considera uma estrutura de vizinhança tipo Moore para efeitos de conecto-mobilidade de um agente que está numa posição (i, j) à uma posição (ξ, η) vizinha. Quando da distribuição inicial dos agentes ou, mesmo, da sua movimentação em qualquer ciclo da simulação, não há restrições à quantidade máxima de agentes que ocupam uma mesma posição.

Os parâmetros pertinentes à simulação, como as quantidades iniciais de agentes e as taxas são definidos em faixas, o que significa que cada parâmetro possui um valor mínimo e máximo. Nas operações pertinentes que dependem das taxas, ou na movimentação, é randomizado um valor dentro da sua faixa ou da posição a ser movimentada para estabelecer o carácter probabilístico do modelo. A tabela 1 apresenta os parâmetros utilizados no presente modelo.

P01	Quantidade de simulações que compõe uma simulação tipo Monte Carlo	$MC = 1$
P02	Quantidade de ciclos por simulação	$t_f = 1000$
P03	Quantidade de linhas do lote	$\#L$
P04	Quantidade de colunas do lote	$\#C$
P05	Quantidade inicial de agentes suscetíveis	$S_0 = 990$
P06	Quantidade inicial de agentes expostos	$E_0 = 0$
P07	Quantidade inicial de agentes infectados	$I_0 = 10$
P08	Quantidade inicial de agentes recuperados	$R_0 = 0$
P09	Taxa de infecção	$0.90 \leq \beta \leq 0.95$
P10	Quantidade de ciclos de exposição	$15 \leq \gamma \leq 20$
P11	Quantidade de ciclos de infecção	$30 \leq \alpha \leq 35$
P12	Quantidade de ciclos de recuperação	$40 \leq \delta \leq 45$

Observe-se na tabela 1 que os parâmetros da distribuição inicial, da especificação do domínio de simulação e das taxas ou períodos às operações definem as quantidades ou faixas que são empregadas num experimento computacional. Os agentes transitam probabilisticamente nos estados em faixas que variam entre os valores $0.90 = \beta_{min} \leq \beta \leq \beta_{max} = 0.95$, $15 = \gamma_{min} \leq \gamma \leq \gamma_{max} = 20$, $30 = \alpha_{min} \leq \alpha \leq \alpha_{max} = 35$ e $40 = \delta_{min} \leq \delta \leq \delta_{max} = 45$. Mas a escolha é arbitrária e pode ser modificada para faixas adequadas às doenças em que tais parâmetros sejam conhecidos.

Para a manutenção da densidade dos diferentes estados dos agentes no ambiente de simu-

lação, deve-se considerar que sua parametrização, especificada por $\#L \times \#C$, depende das quantidades dos agentes nos estados suscetível, exposto, infectante e recuperados, $\#S$, $\#E$, $\#I$, $\#R$, respectivamente.

A metodologia tipo Monte Carlo adotada neste trabalho consiste na execução de determinada quantidade de simulações com mesma configuração inicial às quantidades de agentes e à dimensão do ambiente. Porém, pela natureza probabilística do modelo, as simulações têm, eventualmente, diferentes taxas ou períodos de infecção, recuperação e imunização. O valor médio, VM , de cada variável ou estado em cada passo de tempo, em uma simulação Monte Carlo é especificada pela razão VM/MC , onde MC designa a quantidade de simulações executadas. Mas esse valor pode ser configurado para explorar a distribuição dos parâmetros e a eficiência computacional, entre outros fatores da simulação.

2.5 Bitstring

A modelagem bitstring aqui realizada é baseada na manipulação de bits em uma palavra computacional, que caracteriza sem ambigüidade a especificação do agente $\chi(t) = (L, X, Y, C, E)$, com identificador do lote, L , coordenada x da posição, X , coordenada y da posição, Y , contador de controle, C , e estado, E . É uma metodologia frequentemente empregada para agrupar informações de modo otimizado em várias aplicações.

Para o emprego de modelo em bitstring é necessário utilizar uma linguagem de programação que dê suporte apropriado às operações diretas com bits. A implementação realizada à obtenção dos resultados apresentados e discutidos neste trabalho utilizou a linguagem C , que provê suporte aos propósitos de modelagem. O tipo de dado inteiro em C , *int*, contém 32 bits, sendo suficiente à especificação da formulação para o agente $\chi(t)$ em termos de bits. O uso do tipo de dado *unsigned int* descarta a necessidade de preservação de sinal para deslocamentos aritméticos, evitando erros na manipulação dos bits mais significativos quando da modelagem do agente.

À modelagem da posição X, Y de um agente se considera que as quantidades de linhas, $\#L$, e de colunas, $\#C$, do ambiente, são limitadas por $\max(\#L \times \#C) = (512 \times 512)$, onde \max indica o máximo possível de linhas e de colunas que podem ser alocadas à essa escolha. Dimensões maiores demandam outros tipos de dados em C . À modelagem do controle, C ,

de um agente se considera que os ciclos em que o agente fica no estado exposto, infectado ou recuperado, ou que outras operações que eventualmente demandem um controle de tempo, não ultrapassem 64 ciclos de simulação, conforme discutido. À modelagem do estado de um agente se considera que o modelo compartimental é do tipo SEIRS, tendo somente quatro distintos estados. À identificação do lote atual do agente considera-se que o identificador não ultrapasse 64 valores.

Observe-se que $512_{10} = 2^9$, $64_{10} = 2^6$ e $4_{10} = 2^2$, sendo suficiente uma palavra de tamanho 32 bits para armazenar a especificação de $\chi(t) = (L, X, Y, C, E)$ em bits, como ilustrado na figura 2.



Figura 2.2: Representação em palavra de 32 bits ao modelo SEIRS baseado em indivíduos

Na figura 2, os elementos do identificador do lote estão nas posições 31 a 26. Os elementos das coordenadas x e y do lote estão nas posições 25 a 17 e 16 a 8, respectivamente. Os elementos do contador de controle e do estado estão nas posições 7 a 2 e 1 a 0, respectivamente. Considerando as escolhas realizadas à sua modelagem, um agente qualquer do modelo pode ser representado em bits como em (3).

$$\chi(L, X, Y, C, E) \equiv (l_{31}, \dots, l_{26}; x_{25}, \dots, x_{17}; y_{16}, \dots, y_8; c_7, \dots, c_2; e_1, \dots, e_0) \quad (2.7)$$

sendo que os elementos l_i , x_i , y_i , c_i e e_i são ou o bit 0 ou o bit 1, onde seus índices subscritos indicam que cada bit pode variar nos respectivos intervalos que estão especificados.

Para identificar o lote e as coordenadas da posição do agente $\chi(L, X, Y, C, E) = \chi(t)$ no ambiente, assim como seu estado interno e o registro de controle de ciclos às transições de estados, essas características são definidas como em (4), para os ciclos t e $t + 1$, respectivamente como:

$$\begin{cases} L(t+1) = \left(\left(\chi(L, X, Y, C, E) \right) \ll_{arth} (Pl) \right) \gg_{arth} (C\#l) \\ X(t+1) = \left(\left(\chi(L, X, Y, C, E) \right) \ll_{arth} (Px) \right) \gg_{arth} (C\#x) \\ Y(t+1) = \left(\left(\chi(L, X, Y, C, E) \right) \ll_{arth} (Py) \right) \gg_{arth} (C\#y) \\ C(t+1) = \left(\left(\chi(L, X, Y, C, E) \right) \ll_{arth} (Pc) \right) \gg_{arth} (C\#c) \\ E(t+1) = \left(\left(\chi(L, X, Y, C, E) \right) \ll_{arth} (Pe) \right) \gg_{arth} (C\#e) \end{cases} \quad (2.8)$$

em que Pl , Px , Py , Pc e Pe designam, respectivamente, as quantidades de bits posteriores à cada campo e $C\#l$, $C\#x$, $C\#y$, $C\#c$ e $C\#e$ designam, respectivamente, as quantidades de bits complementares à cada campo. As operações \ll_{arth} e \gg_{arth} designam, respectivamente, deslocamento de modo aritmético (*arithmetic shift*) à esquerda e à direita na palavra, na quantidade de bits indicado.

Na modelagem realizada, as especificações às quantidades são $Pl = 0$, $Px = 6$, $Py = 15$, $Pc = 24$, $Pe = 30$, $C\#l = 26$, $C\#x = 23$, $C\#y = 23$, $C\#c = 26$ e $C\#e = 30$, de modo que (4) é então reescrita para os ciclos t e $t+1$ designados como *old* e *new*, respectivamente, como em (5).

$$\begin{cases} L_{new} := \left(\chi(L_{old}, X, Y, C, E) \ll_{arth} (0) \right) \gg_{arth} (26) \\ X_{new} := \left(\chi(L, X_{old}, Y, C, E) \ll_{arth} (6) \right) \gg_{arth} (23) \\ Y_{new} := \left(\chi(L, X, Y_{old}, C, E) \ll_{arth} (15) \right) \gg_{arth} (23) \\ C_{new} := \left(\chi(L, X, Y, C_{old}, E) \ll_{arth} (24) \right) \gg_{arth} (26) \\ E_{new} := \left(\chi(L, X, Y, C, E_{old}) \ll_{arth} (30) \right) \gg_{arth} (30) \end{cases} \quad (2.9)$$

As operações definidas em (5) e (6) capturam informações armazenadas nos agentes, distribuídos e parametrizados consoantes às informações sistematizadas na Tabela 1.

Também são necessárias outras operações para manipular bits, após aquelas definidas em (5). Elas setam no agente as informações necessárias às atualizações de seu lote, posição, estado e contador. São definidas como em (7), sendo que \ll_{circ} e \gg_{circ} designam, respectivamente, deslocamento de modo circular (*circular shift*) à esquerda e à direita na palavra, na quantidade de bits indicado, designando que os bits que são expelidos em um extremo são reinseridos no extremo oposto da palavra.

$$\begin{cases} \ll_{circ} := \left(z(t+1) \ll_{arth} \text{shift} \right) \mid \left(z(t+1) \gg_{arth} (N_BITS - \text{shift}) \right) \\ \gg_{circ} := \left(z(t+1) \gg_{arth} \text{shift} \right) \mid \left(z(t+1) \gg_{arth} (N_BITS - \text{shift}) \right) \end{cases} \quad (2.10)$$

sendo que $z(t+1)$ denota ou $x(t+1)$ ou $y(t+1)$, N_BITS denota a quantidade de bits totais e \mid denota a operação lógica "ou inclusivo" que é realizada bit a bit à palavra. Assim, a especificação às quantidades adotadas na modelagem, (7) é reescrita como em (8).

$$\begin{cases} \ll_{circ} := \left(z(t+1) \ll_{arth} \text{shift} \right) \mid \left(z(t+1) \gg_{arth} (32 - \text{shift}) \right) \\ \gg_{circ} := \left(z(t+1) \gg_{arth} \text{shift} \right) \mid \left(z(t+1) \gg_{arth} (32 - \text{shift}) \right) \end{cases} \quad (2.11)$$

Subsequentemente são utilizadas as operações definidas em (3) e em (5), ou concretamente, as definidas em (5) e (7), para implementar o operador espaço-temporal $\lambda(\chi(t))$ que movimenta o agente da posição (i, j) para uma posição (ξ, η) no ciclo de tempo atual, t (*old*), para o ciclo de tempo, $t+1$ (*new*). A combinação dessas operações resulta que a identificação do lote, as coordenadas x e y da posição, o controle e estado do agente são atualizadas, respectivamente, para $\lambda(\chi(t))$ por (9).

$$\begin{cases} \chi(L(t+1), X(t), T(t), C(t), E(t)) := \left[\left(\left(\left(\chi(X(t), Y(t), C(t), E(t)) \gg_{circ} (Al) \right) \gg_{arth} (\#l) \right) \ll_{arth} (\#l) \right) \mid L(t+1) \right] \ll_{circ} (Al) \\ \chi(L, X(t+1), Y(t), C(t), E(t)) := \left[\left(\left(\left(\chi(X(t), Y(t), C(t), E(t)) \gg_{circ} (Ax) \right) \gg_{arth} (\#x) \right) \ll_{arth} (\#x) \right) \mid X(t+1) \right] \ll_{circ} (Ax) \\ \chi(L, X(t), Y(t+1), C(t), E(t)) := \left[\left(\left(\left(\chi(X(t), Y(t), C(t), E(t)) \gg_{circ} (Ay) \right) \gg_{arth} (\#y) \right) \ll_{arth} (\#y) \right) \mid Y(t+1) \right] \ll_{circ} (Ay) \\ \chi(L, X(t), Y(t), C(t+1), E(t)) := \left[\left(\left(\left(\chi(X(t), Y(t), C(t), E(t)) \gg_{circ} (Ac) \right) \gg_{arth} (\#c) \right) \ll_{arth} (\#c) \right) \mid C(t+1) \right] \ll_{circ} (Ac) \\ \chi(L, X(t), Y(t), C(t), E(t+1)) := \left[\left(\left(\left(\chi(X(t), Y(t), C(t), E(t)) \gg_{circ} (Ae) \right) \gg_{arth} (\#e) \right) \ll_{arth} (\#e) \right) \mid E(t+1) \right] \ll_{circ} (Ae) \end{cases} \quad (2.12)$$

em que Al , Ax , Ay , Ac e Ae designam, respectivamente, as quantidades de bits anteriores à cada campo e $\#l$, $\#x$, $\#y$, $\#c$ e $\#e$ designam, respectivamente, as quantidades de bits de cada campo.

Como $Al = 26$, $Ax = 17$, $Ay = 8$, $Ac = 2$, $Ae = 0$, $\#l = 6$, $\#x = 9$, $\#y = 9$, $\#c = 6$ e $\#e = 2$, a especificação às quantidades adotadas na modelagem, (9) é reescrita como em (10).

$$\begin{cases}
\chi(L_{new}, X_{old}, Y_{old}, C_{old}, E_{old}) := \left[\left(\left(\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{old}) \gg_{circ} (26) \right) \gg_{arth} (6) \right) \ll_{arth} (6) \right] | L_{new} \ll_{circ} (26) \\
\chi(L_{old}, X_{new}, Y_{old}, C_{old}, E_{old}) := \left[\left(\left(\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{old}) \gg_{circ} (17) \right) \gg_{arth} (9) \right) \ll_{arth} (9) \right] | X_{new} \ll_{circ} (17) \\
\chi(L_{old}, X_{old}, Y_{new}, C_{old}, E_{old}) := \left[\left(\left(\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{old}) \gg_{circ} (8) \right) \gg_{arth} (9) \right) \ll_{arth} (9) \right] | Y_{new} \ll_{circ} (8) \\
\chi(L_{old}, X_{old}, Y_{old}, C_{new}, E_{old}) := \left[\left(\left(\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{old}) \gg_{circ} (2) \right) \gg_{arth} (6) \right) \ll_{arth} (6) \right] | C_{new} \ll_{circ} (2) \\
\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{new}) := \left[\left(\left(\chi(L_{old}, X_{old}, Y_{old}, C_{old}, E_{old}) \gg_{circ} (0) \right) \gg_{arth} (2) \right) \ll_{arth} (2) \right] | E_{new} \ll_{circ} (0)
\end{cases} \quad (2.13)$$

Para completar a modelagem é suficiente detalhar e relacionar os identificadores dos lotes e dimensões dos lotes, a faixa de variação do controle, e os possíveis tipos de estados do agente na base 10, com tais características na base 2, para determinar sem ambigüidade os elementos de (2).

Os identificadores dos lotes têm variação de $(0)_{10} = (000\ 000)_2$ até $(63)_{10} = (111\ 111)_2$, totalizando 64 possíveis valores. Os lotes têm dimensão máxima de $512_{10} = 2^9$ linhas por $512_{10} = 2^9$ colunas, sendo possível variarem numa representação matricial da posição $(0;0)_{10} = (000\ 000\ 000; 000\ 000\ 000)_2$ até a posição $(511;511)_{10} = (111\ 111\ 111; 111\ 111\ 111)_2$, totalizando os 512 possíveis valores. Semelhantemente, faz-se uma representação matricial ao controle para que sua variação ocorra de $(0)_{10} = (000\ 000)_2$ até $(63)_{10} = (111\ 111)_2$, totalizando os 64 possíveis valores. Os estados do agente são setados como $(0,0)_2$ para o suscetível, S , $(0,1)_2$ para o exposto, E , $(1,0)_2$ para o infectante, I , e $(1,1)_2$ para o recuperado, R .

2.6 Refinamento do Modelo

Capítulo 3

Metodologias Computacionais

3.1 Introdução

3.2 SIMULA

3.3 Estruturas de Dados, Linguagens e Estratégias de Implementação

3.4 CUDA e OpenMP

Capítulo 4

Soluções

4.1 Introdução

4.2 Implementação Padrão com CUDA e OpenMP

4.3 Implementação *Bitstring* com CUDA e OpenMP

4.4 Discussões Qualitativas, Quantitativas, Eficiência e Acurácia

Capítulo 5

Resultados e Discussões

5.1 Introdução

5.2 Cases: Discutir as Simulações no Espaço Geográfico Escolhido.

Glossário