

UNIVERSIDADE NOVE DE JULHO - UNINOVE
PROGRAMA DE PÓS GRADUAÇÃO EM INFORMÁTICA E GESTÃO
DO CONHECIMENTO - PPGI

EMERSON DA SILVA

MODELAGEM E SIMULAÇÃO DE PROPAGAÇÃO DE INFORMAÇÕES
PARA POPULAÇÕES BASEADAS EM AUTÔMATOS CELULARES
PROBABILISTAS CONTÍNUOS ASSÍNCRONOS

São Paulo
2018

EMERSON DA SILVA

**MODELAGEM E SIMULAÇÃO DE PROPAGAÇÃO DE INFORMAÇÕES
PARA POPULAÇÕES BASEADAS EM AUTÔMATOS CELULARES
PROBABILISTAS CONTÍNUOS ASSÍNCRONOS**

Exame de Defesa apresentado à Universidade Nove de Julho - UNINOVE, como parte dos requisitos para a obtenção do título de Mestre em Informática e Gestão do Conhecimento.

Prof. Orientador: Dr. Pedro Henrique Triguis Schimit

**São Paulo
2018**

Silva, Emerson da.

Modelagem e simulação de propagação de informações para populações baseadas em autômatos celulares probabilistas contínuos assíncronos. / Emerson da Silva. 2018.

42 f.

Dissertação (Mestrado) - Universidade Nove de Julho - UNINOVE, São Paulo, 2018.

Orientador: Prof. Dr. Pedro Henrique Triguís Schimit.

1. Autômatos celulares assíncronos.
2. Propagação de informações.
3. Propagação de rumores.
4. Autômatos celulares probabilistas.
5. Modelo SIR.

I. Schimit, Pedro Henrique Triguís.

II. Título.

CDU 004

PARECER – EXAME DE DEFESA

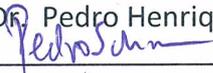
Parecer da Comissão Examinadora designada para o exame de defesa do Programa de Pós-Graduação em Informática e Gestão do Conhecimento a qual se submeteu o aluno **Emerson da Silva**

Tendo examinado o trabalho apresentado para obtenção do título de “Mestre em Informática e Gestão do Conhecimento”, com Dissertação intitulada “**Modelagem e Simulação de Propagação de Informações para Populações Baseadas em Autômatos Celulares Probabilistas Contínuos Assíncronos**”, a Comissão Examinadora considerou o trabalho:

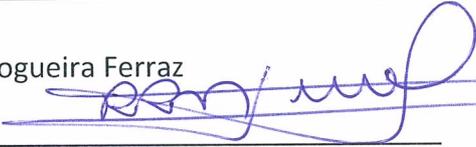
- () Aprovado () Aprovado condicionalmente
() Reprovado com direito a novo exame () Reprovado

EXAMINADORES

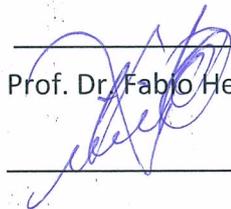
Prof. Dr. Pedro Henrique Triguís Schimit



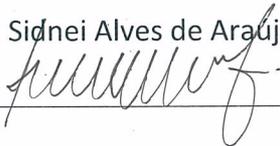
Prof. Dr. Renato Ribeiro Nogueira Ferraz



Prof. Dr. Fabio Henrique Pereira



Prof. Dr. Sidnei Alves de Araújo



São Paulo, 25 de setembro de 2018.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me dado as condições necessárias para ingressar nesse curso e ter ajudado a me manter nele, pois por diversas vezes pensei em desistir, devido às dificuldades que tive que enfrentar, mas com a ajuda de Deus consegui supera-las e hoje estou aqui concluindo meu Mestrado.

Agradeço também ao professor Sidnei Alves de Araujo e ao professor André Felipe Henriques Librantz, por todo incentivo que me deram desde a primeira vez em que perguntei para eles sobre o curso de Mestrado.

Agradeço à professora Cristina Koyama por sempre me admoestar de forma positiva para que eu não desanimasse do curso e também por me ajudar nos momentos em que mais precisei de ajuda.

Agradeço ao meu orientador, o professor Pedro Henrique Triguís Schimit, primeiro por ter escolhido me orientar e segundo por fazer desta orientação um processo de muito aprendizado, com conversas, explicações, conselhos e críticas, sempre muito bem colocadas de maneira muito franca e transparente.

Agradeço a todos os meus companheiros de trabalho no setor de Tecnologia da Informação da Uninove, se não fosse pelo apoio deles eu não teria conseguido completar o curso. Agradeço principalmente ao meu gerente, o professor Angelo Ravanelli, que sempre me apoiou em todos os momentos de dificuldade.

Agradeço à Uninove por ter me proporcionado as oportunidades de crescimento profissional e acadêmico.

Por fim, agradeço à minha mãe, Hideko Siga, que superou diversas dificuldades para criar, sozinha, três filhos pequenos, dando educação e tornando-os pessoas de bem. Agradeço também à minha esposa, Elaine Laudicea dos Santos Silva, por todo apoio e compreensão nos vários momentos em que não consegui lhe dar a devida atenção pois tinha que me dedicar aos trabalhos do Mestrado.

Informações se propagam entre pessoas das mais diversas maneiras. Seja por meio de processos que envolvem informantes e informados, seja por meio de busca de informações por parte de indivíduos. Em comum com a propagação de doenças está o processo de contato entre dois indivíduos que, ao invés de transmitirem um vírus ou uma bactéria, trocam informações. Neste trabalho, apresenta-se dois modelos de propagação de informações: um discreto, baseado num modelo de propagação de doenças, e um contínuo, baseado em modelos educacionais entre dois agentes. Nesse último, assume-se que há a possibilidade de campanhas de marketing a uma parcela da população, sendo mais uma maneira dos indivíduos receberem informações. Com o primeiro modelo, estuda-se um exemplo de propagação de boatos numa empresa, e com o segundo, o impacto das escolhas da campanha de marketing no processo de ganho de informação da população. Os dois modelos são baseados em autômatos celulares probabilistas, sendo o primeiro discreto e síncrono, e o segundo, contínuo e assíncrono. Os resultados sugerem uma boa aplicabilidade dos modelos em cenários reais de propagação de informações, sendo que o modelo discreto, mais simples, mostra que campanhas educacionais em empresas podem ajudar a conter o espalhamento de um boato, e o modelo contínuo, mais facilmente configurável, mostra que não basta um investimento alto em marketing, mas que campanhas direcionadas a uma parcela da população, com conteúdo adequado, retornam melhores resultados.

Palavras-chave: autômatos celulares assíncronos, propagação de informações, propagação de rumores, autômatos celulares probabilistas, modelo SIR.

Information spreads amongst people in different forms. Either by learners and informers agents, or individuals which look for information in media, for instance. There is a similarity with disease propagation, but instead of virus or bacterial infection in the contact between two individuals, they change information. In this work, we present two models for rumor propagation: a discrete, based on a disease spreading model, and a continuous, based on the learning process between two agents. In this case, we assume the possibility of marketing campaigns to a fraction of the population, and this is another way for an individual to obtain information. By using the first model, we study a case of rumor propagation inside a company; and by using the second model, we study the impact of marketing campaigns parameters choice in the population information level. Both models are based in probabilistic cellular automata, the first being discrete and synchronous, and the second, continuous and asynchronous. Results of discrete model show that educational campaigns in companies may save costs to stop a rumor spreading, and the continuous model suggests that marketing campaign must be directed to a parcel of population with adequate content.

Keywords: asynchronous cellular automata, information propagation, rumor propagation, probabilistic cellular automata, SIR model.

Lista de Figuras	6
Lista de Abreviaturas	8
1 Introdução	9
2 Fundamentação Teórica	12
2.1 Autômatos Celulares	12
2.2 Modelo SIR e Propagação de Informações	15
2.3 Trabalhos Correlatos	17
3 Os Modelos Discreto e Contínuo	20
3.1 O Modelo Discreto LPD	20
3.2 O Modelo Contínuo Assíncrono	21
4 Resultados	26
4.1 Modelo Discreto Para Propagação de Informações Falsas em Empresas	26
4.2 Modelo Contínuo para Simulação de Processos de Marketing	29
4.2.1 O modelo Sem Marketing e Sem Esquecimento	30
4.2.2 O Modelo Sem Marketing e Com Esquecimento	31
4.2.3 O Modelo Com Marketing e Com Esquecimento	31
4.2.4 A influência do Marketing no Modelo Proposto	33
5 Conclusões	38
Referências Bibliográficas	40

LISTA DE FIGURAS

2.1	Vizinhança de von Neumann.	14
2.2	Vizinhança de Moore com $r=1$	14
2.3	No Jogo da Vida, o estado de cada célula depende do estado das células vizinhas. A vizinhança é composta por oito células.	14
2.4	Modelo SIR de (SCHIMIT; MONTEIRO, 2009).	15
2.5	Modelo SIHR de (ZHAO et al., 2012).	16
3.1	Diagrama de estados do modelo LPD.	21
3.2	Processos de ganho e perda de informações de cada indivíduo.	22
3.3	Ganho de informação do indivíduo i por influência de um indivíduo j dada por $G_{i,j}^n$, considerando $I_j > I_i$, para $k_i = 1$, $k_i = 2$ e $k_i = 6$	24
3.4	Ganho de informação do indivíduo i por influência de marketing positivo dada por G_i^{mp} , para $k_m = 1$, $k_m = 2$ e $k_m = 3$	24
3.5	Perda de informação do indivíduo i por esquecimento dada por G_i^f , para $k_e = 0,1$, $k_e = 1$ e $k_e = 2$	25
4.1	Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de K_b , com $P_s = 0,6$ e $P_f = 0,3$	28
4.2	Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de P_s , com $K_b = 1$ e $P_f = 0,3$	28
4.3	Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de P_f , com $K_b = 1$ e $P_s = 0,6$	29
4.4	Evolução temporal da informação (\bar{I}), influência da vizinhança ($\overline{G_i^n}$), marketing positivo ($\overline{G_i^m}$), esquecimento ($\overline{G_i^f}$) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0$, $m = 0$ e $I_0 = 0,4$	30
4.5	Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0$, $m = 0$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$	31
4.6	Evolução temporal da informação (\bar{I}), influência da vizinhança ($\overline{G_i^n}$), marketing positivo ($\overline{G_i^m}$) e esquecimento ($\overline{G_i^f}$) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0,25$, $m = 0$ e $I_0 = 0,4$	32

4.7	Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0,25$, $m = 0$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$	32
4.8	Evolução temporal da informação (\bar{I}), influência da vizinhança ($\overline{G_i^n}$), marketing positivo ($\overline{G_i^m}$), esquecimento ($\overline{G_i^f}$) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 10$, $k_e = 0,5$, $m = 0,1$ e $I_0 = 0,4$	33
4.9	Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 10$, $k_e = 0,5$, $m = 0,1$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$	33
4.10	Ganho com marketing ($\overline{G^m}$) em função de m e k_m para $k_i = 2, 6, 10$ e $k_e = 0,1, 1, 2$ usando $I_0 = 0,2$, $\mu = 0,5$	35
4.11	Média da informação (\bar{I}) em função de m e k_m para $k_i = 2, 6, 10$ e $k_e = 0,1, 1, 2$ usando $I_0 = 0,2$	36
4.12	Média da informação (\bar{I}) [a, b] e ganho com marketing ($\overline{G^m}$) [c, d] em função de μ e m com $k_m = 1$ [a, c] em função de μ e k_m com $m = 0,3$ [b, d] usando $I_0 = 0,4$	37

LISTA DE ABREVIATURAS

AC	Autômato Celular
ACP	Autômato Celular Probabilista
LPD	Leigo Propagador Desinteressado
SCIR	Suscetível Contatado Infectado Refratário
SI	Suscetível Infectado
SIHR	Suscetível Infectado Hibernando Recuperado
SIR	Suscetível Infectado Recuperado
SIS	Suscetível Infectado Suscetível

Informação, rumor, boato. Na literatura da área de modelagem de propagação de notícias ou conhecimentos, esses são os três termos mais usados. Rumor se refere a algum assunto mal explicado, baseado em alguma informação corrente não confirmada ou reavivada dentro de um determinado contexto que envolva o interesse comum das pessoas. Já o boato é normalmente utilizado com o mesmo significado de rumor, porém, o boato não tem base em informações verdadeiras sendo considerado sempre uma informação falsa, passada de boca em boca, sem comprovação. O conteúdo dos rumores pode variar desde fofocas simples até propaganda e material de marketing. Rumores com conteúdos diversos podem ser propagados para pessoas em diferentes níveis sociais e podem levar a diferentes tipos de impactos e interpretações, desde a pequena preocupação pública até o grande pânico social (REULE, 2008). Em determinadas ocasiões as informações veiculadas são insuficientes para esclarecer os fatos, a falta de detalhes e a veiculação de informações distorcidas acabam criando ambiente propício para a geração de rumores e boatos.

A informação circula com maior ou menor força de acordo com o valor intrínseco que agrega para o grupo social no qual circula. Quanto mais informação se tem a respeito de um assunto, mais fácil se torna a tomada de decisão referente àquele assunto, ao mesmo tempo que, quanto menos informação se tem, mais difícil se torna a decisão e mais versões da verdade circulam de forma aleatória (ZHAO et al., 2013). Um exemplo dessa situação ocorreu no Japão, em 2011. Um terremoto de grandes proporções atingiu a região de Tohoku, e uma usina nuclear foi gravemente afetada, com vazamento de radiação que acabou sendo espalhada pelo vento para locais distantes. Surgiu um rumor de que para se proteger da radiação, as pessoas deveriam ingerir alimentos preparados com sal iodado, o que levou a uma corrida desenfreada de moradores para comprar grandes quantidades de sal iodado. Apesar de verdadeira a relação do iodo como proteção para a radiação, os estoques feitos por diversas famílias mostrou-se desproporcional ao evento, num caso típico de histeria coletiva (ZHAO et al., 2013). Em inglês, rumor não tem a mesma conotação que boato como na língua portuguesa, os artigos que falam de rumores são mais relacionados com informações do que com boatos, portanto, neste trabalho usa-se o termo informação. Qualificaremos a informação quando necessário para definir quando se trata de uma informação desejada ou uma informação indesejada.

Nos primórdios da civilização as informações eram propagadas de boca em boca, também foram utilizados desenhos em paredes nas cavernas, tempos depois desenvolveu-se a escrita e a imprensa. Hoje, a internet e as mídias sociais facilitam a velocidade de transmissão das notícias. Neste contexto ainda temos o problema das notícias falsas que circulam livremente fazendo com que em instantes todos saibam dos assuntos em circulação. Essa velocidade de qualquer tipo de notícia é útil, mas pode gerar grande pânico

social (REULE, 2008). Além disso, pode ser interessante possibilitar que uma certa informação chegue a todos os interessados da maneira mais rápida possível, para que decisões sejam tomadas da melhor maneira possível.

Para entender, analisar e prever comportamentos do espalhamento de informações, muitos estudos começaram a usar modelos de propagação de doenças como base. Desde a década de 1960, há um grande aproveitamento de modelos variados a partir do modelo SIR para uso no campo de propagação de informações (XIONG et al., 2012), (ZHAO et al., 2012). Nesse modelo, os estados que um indivíduo pode assumir são: *Suscetível*, quando não possui a doença e pode contrair por contato com um indivíduo *Infectado*, que quando curado, se torna *Recuperado*, ganhando imunidade permanente, como descrito no modelo SIR original de Kermack-McKendric, de 1927 (KERMACK; MCKENDRICK, 1927).

O processo de propagação de informações é: Um indivíduo Suscetível não conhece a informação. Ao entrar em contato com Infectados (conhecedores da informação), um indivíduo Suscetível, pode conhecer a informação e também começar a propagá-la, e após propagar por algum tempo se tornar um indivíduo Recuperado, não passando essa informação para mais ninguém. Esse paralelo foi usado na modelagem de um mercado consumidor (SCHIMIT; BARROS JUNIOR; ARAÚJO, 2017), para caracterizar a propagação de uma informação pelo *microblog Twitter* (XIONG et al., 2012; SU; HUANG; ZHAO, 2015) em redes sociais (NEKOVEE et al., 2007; ZHAO et al., 2012), e para avaliar quais os nós que mais propagavam informações numa rede social (LIN et al., 2014; ZENG; ZHANG, 2013). Outros modelos epidemiológicos também são traduzidos para o mundo das informações, como o modelo SIS (OLINKY; STONE, 2004), em que a cura não confere imunidade, ou o modelo SI, em que o indivíduo morre com a doença (XIONG et al., 2012).

Assim, alguns modelos de propagação de informações começaram a mostrar uma possibilidade de analisar sistemas complexos e de difícil abordagem. A maneira como a informação se propaga pela internet (ZHAO et al., 2013) e por redes sociais (CHIERICHETTI; LATTANZI; PANCONESI, 2011; XIONG et al., 2012; ZHAO et al., 2012) tem ganhado bastante atenção da comunidade acadêmica, assim como o efeito de notícias falsas nessas redes (CARVALHO; KLAGGE; MOENCH, 2011). Além disso, os efeitos de marketing também são modelados para identificar os processos chaves para que uma informação seja mais conhecida numa rede social, ou na população em geral (CHOU; TIAN; YIN, 2015; GOLDENBERG; LIBAI; MULLER, 2001; KOSFELD, 2005; SCHIMIT; BARROS JUNIOR; ARAÚJO, 2017).

Os processos que fazem um agente (que pode ser um indivíduo, conjunto de indivíduos, perfis em redes sociais) ganhar informação são: aprendizado com outro indivíduo ou imprensa (ZHOU; LIU; LI, 2007; ZHAO et al., 2012; ROSNOW; ESPOSITO; GIBNEY, 1988), e campanhas de marketing (CHOU; TIAN; YIN, 2015; GOLDENBERG; LIBAI; MULLER, 2001). Por outro lado, esse agente perde informações por esquecimento (ROS-

NOW; ESPOSITO; GIBNEY, 1988; XIONG et al., 2012; ZHOU; LIU; LI, 2007), ou por efeitos negativos do marketing (KOSFELD, 2005).

Portanto, essas são as variáveis que os modelos deste trabalho usam para analisar e controlar os impactos de uma informação. Seja um rumor com efeitos negativos sobre uma população (DIFONZO; BORDIA; ROSNOW, 1994), uma notícia que vazou sobre o fechamento de uma cadeia de lojas em Hong-Kong que gerou uma histeria coletiva (LAI; WONG, 2002), ou um problema com fake news, como o caso de um artigo falso sobre problemas financeiros da United Airlines (que era uma informação antiga publicada novamente), que levou a uma queda de 11,2% em suas ações em 2008 (CARVALHO; KLAGGE; MOENCH, 2011), o modelo pode contemplar todas as características desses processos.

Assim como na epidemiologia matemática, a modelagem de propagação de informações encontrou no autômato celular (WOLFRAM, 1994) uma ferramenta de fácil implementação, baixo custo computacional e eficiente para simular diversos cenários. Além do mercado consumidor já citado (SCHIMIT; BARROS JUNIOR; ARAÚJO, 2017), a ferramenta também já foi usada para modelar a persistência e oscilação de rumores em populações (KAWACHI et al., 2008).

Todos os trabalhos apresentados até aqui abordam alguma ferramenta multi-agente que considera estados discretos para os indivíduos da população, ou seja, o indivíduo sabe ou não sabe sobre determinada informação. Assim o objetivo deste trabalho é propor um modelo baseado em autômato celular probabilista contínuo assíncrono, em que os indivíduos de uma população não possuem conhecimento binário sobre uma informação, mas um nível contínuo de conhecimento sobre um fato, estando entre 0 e 1. Com esse sistema, mostrou-se que modelar uma campanha de marketing fica mais facilmente configurável. Antes de propor esse modelo contínuo, foram feitas adaptações em outro modelo, mais simples, de propagação de informações, baseado em autômato celular discreto, cujas simulações indicaram que o modelo é limitado, justificando a proposta do modelo contínuo.

Este trabalho foi organizado da seguinte forma: o próximo capítulo contém a fundamentação teórica para o trabalho, mostrando as ferramentas matemáticas utilizadas e os trabalhos correlatos à área aqui estudada. Os modelos são apresentados no capítulo 3, com resultados no capítulo 4. Por fim, uma discussão final é feita no capítulo 5.

FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados alguns fundamentos sobre autômatos celulares e propagação de doenças, com um paralelo para propagação de informações.

2.1 AUTÔMATOS CELULARES

Autômatos Celulares foram propostos por John von Neumann e Stanislaw Marcin Ulam para modelar organismos auto-reprodutores e crescimento celular, na década de 1940. De maneira geral, um Autômato Celular (AC) é uma ferramenta matemática computacional de fácil implementação usada para representar sistemas físicos compostos por elementos discretos com interações locais. Tais interações podem mudar os estados das células, e o autômato evolui de acordo com o tempo. Um AC pode ser usado para representar sistemas evolutivos a partir de uma situação inicial aleatória, fazendo com que cada componente do sistema tenha sua evolução baseada em um vizinho mais próximo com um conjunto de regras iguais para todos os componentes (MELOTTI, 2009). As células evoluem respeitando regras que dependem do estado em que elas se encontram e de como suas vizinhas se comportam, a vizinhança se relaciona com cada célula de forma local e uniforme (WEIMAR, 1997). As definições clássicas de um AC são (WOLFRAM, 1994):

1. *Formato*: estrutura das células, com uma, duas ou três dimensões. O AC elementar é uni-dimensional, populações geralmente são modeladas pelo reticulado bidimensional. As condições de contorno geralmente são periódicas (MURRAY, 2008);
2. *Estados das células*: As células de um AC possuem um número finito de estados, sendo dois para o AC elementar, e três quando se aplica o modelo SIR;
3. *Regras de transição*: Determinam as regras para evolução temporal das células, ou seja, o estado futuro de uma célula é função de algumas características de um passo de tempo anterior;
4. *Vizinhança*: É o conjunto de células que se conecta com cada célula do reticulado. No caso do AC elementar, cada célula tem duas células vizinhas. A vizinhança de Moore para o reticulado bidimensional tem 8 células vizinhas para cada célula.

Outras características que também definem um AC, e surgiram em trabalhos mais recentes são:

1. *Estados das células*: Além do caso *discreto*, discutido anteriormente, é possível que as células tenham um estado *contínuo* (CASTRO; CASTRO, 2008), sendo que ao

invés de um número finito de estados, cada célula pode ter qualquer valor dentro de um intervalo;

2. *Atualização dos estados das células*: Em muitos modelos epidemiológicos, os estados futuros das células são atualizados de maneira *síncrona*, geralmente ao final de cada passo tempo (MONTEIRO; CHIMARA; BERLINCK, 2006; SCHIMIT; MONTEIRO, 2009), porém, quando as células são atualizadas em instantes diferentes, ou no momento em que se avalia sua vizinhança e condições para definir o próximo estado, tem-se uma atualização *assíncrona*;
3. *Regras de transição*: Que podem ser *determinísticas* ou *probabilistas* (WEIMAR, 1997).

A cada passo de tempo, cada célula está em um determinado estado, que pode ser alterado no passo de tempo seguinte em função do seu próprio estado e do estado das células que formam sua vizinhança. As regras de transição de estados podem ser deterministas (modelo discreto: quando dada uma configuração no tempo t o estado seguinte é sempre um estado conhecido em $t + 1$), ou probabilistas (modelo contínuo: quando o estado seguinte é definido com base numa probabilidade) (CASTRO; CASTRO, 2008), (WEIMAR, 1997).

Muitos pesquisadores utilizam modelos de AC em aplicações de biologia evolutiva, dinâmica de reações químicas, sistemas dinâmicos da física, comportamento de mercados de consumidores e etc. Normalmente, nos modelos epidemiológicos baseados em AC, os indivíduos formam uma matriz de $n \times n$ células com condições de contorno periódicas (MURRAY, 2008; MONTEIRO; CHIMARA; BERLINCK, 2006).

Neste trabalho, dois modelos de AC são usados: um modelo probabilista discreto síncrono, baseado no modelo epidemiológico SIR, e um modelo probabilista contínuo assíncrono.

Três autores se destacaram no começo do estudo de AC: John von Neumann, John Holton Conway e Stephen Wolfram. O AC de Neumann, foi criado em 1966 e tinha espaço bidimensional com células que se ligavam de forma ortogonal, uma perpendicular à outra. Cada célula possuía 4 vizinhas, sendo duas na vertical e duas na horizontal, conforme Figura 2.1. O objetivo do estudo de Neumann era criar um autômato que se auto reproduzisse, assim, com a ajuda de seu colega Stanislaw Ulam que trabalhava no Laboratório Nacional de Los Alamos, Novo México, conseguiu criar um modelo matemático abstrato baseado numa grelha bidimensional infinita onde cada célula podia assumir 29 estados diferentes (CASTRO; CASTRO, 2008).

Outra vizinhança bastante usada nos estudos com AC foi a vizinhança de Moore, em que uma célula se conecta com as 8 vizinhas mais próximas, como mostrado na Figura 2.2. De maneira genérica, uma vizinhança de Moore de raio r é composta por células que

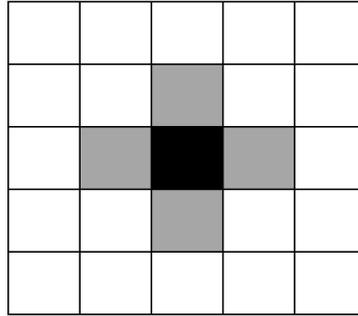


Figura 2.1: *Vizinhança de von Neumann.*

fazem parte de um quadrado de lado $2r + 1$ centralizado na célula em questão (CASTRO; CASTRO, 2008).

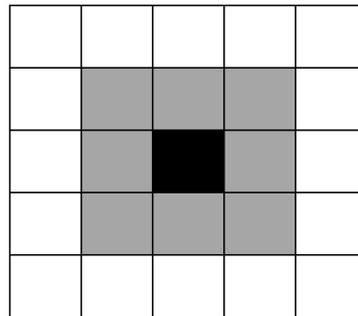


Figura 2.2: *Vizinhança de Moore com $r=1$.*

John Holton Conway criou o jogo da vida, na década de 1970. A ideia era simular alterações em populações de seres vivos com regras locais determinísticas e simples. Usando uma vizinhança de Moore de raio $r = 1$, células morrem ou nascem, de acordo com a presença de células vivas em suas vizinhanças, como na Figura 2.3.

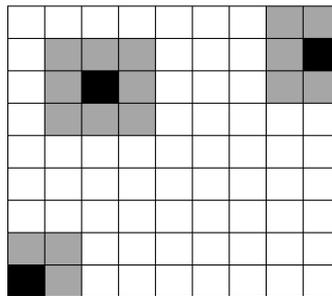


Figura 2.3: *No Jogo da Vida, o estado de cada célula depende do estado das células vizinhas. A vizinhança é composta por oito células.*

Por fim, Stephen Wolfram estudou a complexidade de sistemas evolutivos baseando-se em idealizações matemáticas de sistemas físicos, nos quais o espaço e o tempo são discretos, e os atributos assumem um conjunto de valores também discretos (WOLFRAM, 1994). Tendo por base todos os trabalhos que desenvolveu, Wolfram criou o Princípio da

Equivalência Computacional onde afirma que os ACs podem ser usados para representar todos os tipos de sistemas evolutivos (CASTRO; CASTRO, 2008).

2.2 MODELO SIR E PROPAGAÇÃO DE INFORMAÇÕES

Modelos de propagação de doenças epidemiológicas como o modelo clássico SIR (KERMACK; MCKENDRICK, 1927) considera uma população fixa com três estados possíveis: S para Suscetíveis, I para Infectados e R para Recuperados. Neste modelo os indivíduos Suscetíveis ao serem infectados pela doença se tornam em indivíduos “I”, mas ao serem curados passam para o estado R. Um exemplo de modelo SIR é o estudado em (SCHIMIT; MONTEIRO, 2009). Nesse trabalho, a população é modelada por um autômato celular bidimensional probabilista discreto, em que cada célula representa um indivíduo que pode estar em um dos três estados: Suscetível, Infectado ou Recuperado. O tamanho do reticulado é $n \times n = N$, e a vizinhança é baseada na vizinhança de Moore: cada indivíduo se conecta a outro C (vizinho) dentro de um raio r , sendo $r = 1$ uma vizinhança de Moore clássica composta apenas pela primeira camada que circunda o indivíduo. A probabilidade do indivíduo se conectar com outro da camada i é dada por $q_i = 2(r + 1 - i)/r(r + 1)$, fazendo com que haja maior probabilidade de conexão com indivíduos mais próximos. Considere $r = 3$, a probabilidade do indivíduo se conectar a um indivíduo da camada 1 é de $1/2$, camada 2 é de $1/3$ e camada 3 de $1/6$. As transições de estado do modelo são:

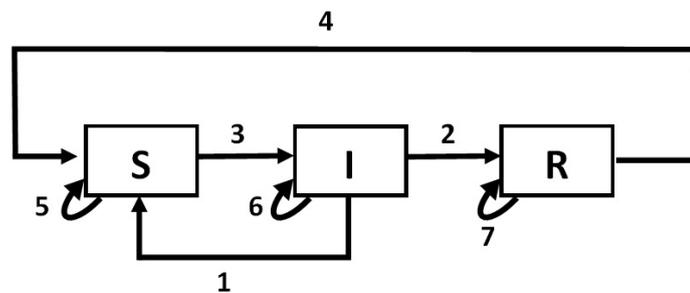


Figura 2.4: Modelo SIR de (SCHIMIT; MONTEIRO, 2009).

1. $I \rightarrow S$: Morte de infectado causado pela doença
2. $I \rightarrow R$: Cura de infectado
3. $S \rightarrow I$: Infecção de suscetível
4. $R \rightarrow S$: Morte de recuperado não causada pela doença
5. $S \rightarrow S$: Suscetível permanece suscetível
6. $I \rightarrow I$: Infectado permanece infectado

7. $R \rightarrow R$: Recuperado permanece recuperado

Quando Infectados morrem pela doença, ou Recuperados morrem por outras causas, nasce um Suscetível em seu lugar, mantendo a população constante.

O trabalho de (ZHAO et al., 2012) usa ideias de um modelo epidemiológico similar ao apresentado em (SCHIMIT; MONTEIRO, 2009), porém, adaptado para a propagação de informações, neste modelo, em comparação com o modelo SIR do trabalho de (SCHIMIT; MONTEIRO, 2009), além dos estados representados pelas letras S, I e R, adiciona-se o estado H (Hibernador). Com vistas à epidemiologia, é um indivíduo infectado que para de transmitir a doença ao se tornar hibernador por um certo tempo, e volta a transmitir a doença quando volta ao estado infectado.

Apesar de usar redes complexas e equações diferenciais ordinárias, as transições de estado consideradas são parecidas, sendo que Suscetível é chamado de Ignorante, Infectado de Propagador e Removido de Removido. Além das transições da figura 2.5, todos podem permanecer no mesmo estado ao final do passo de tempo.

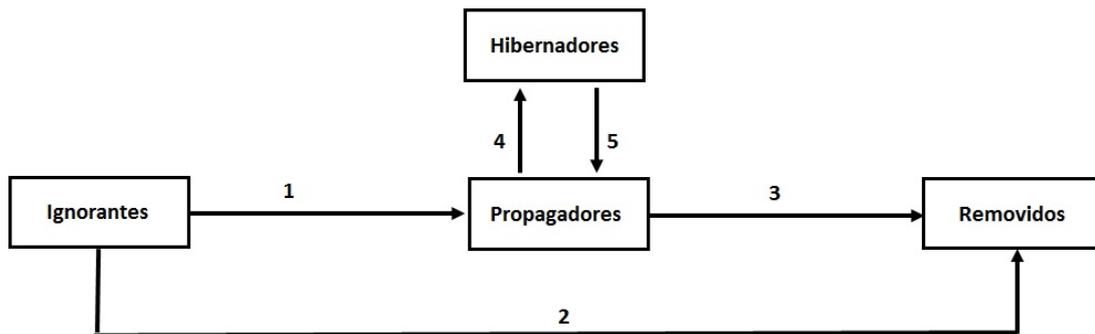


Figura 2.5: Modelo SIHR de (ZHAO et al., 2012).

1. $I \rightarrow P$: Ignorante se torna Propagador
2. $I \rightarrow R$: Ignorante se torna Removido
3. $P \rightarrow R$: Propagador se torna Removido
4. $P \rightarrow H$: Propagador se torna Hibernador
5. $H \rightarrow P$: Hibernador se torna Propagador

Note que um Ignorante pode passar para o estado Removido sem saber da informação. Essa etapa de esquecer e lembrar uma informação ($P \rightarrow H$ e $H \rightarrow P$) se mostrou importante para entender as características de propagação da informação no estudo, sendo que essa interrupção temporária na propagação do rumor por parte de alguns indivíduos ajudou a controlar a máxima influência do rumor na população.

2.3 TRABALHOS CORRELATOS

Informação, rumor e boato muitas vezes são separados por uma linha tênue antes que se verifique todos os fatos envolvidos numa situação. Em situações de crise, uma informação não verificada pode piorar o controle da situação em até 65% dos casos, de acordo com o Relatório da Comissão Consultiva Nacional sobre Desordem Civil norte-americano (tradução livre de *Report of the National Advisory Commission on Civil Disorder*) (ROSNOW; ESPOSITO; GIBNEY, 1988). Uma informação muito relevante fornecida de maneira incompleta para a população pode ser o estopim para uma avalanche de rumores e boatos sobre o tema.

Quatro fatores contribuem para a propagação de rumores. O primeiro é a incerteza da informação fornecida; o segundo é a importância do assunto que gerou o rumor; o terceiro é a ansiedade do informado, que pode ficar emocionalmente envolvido com a notícia e fazer uma má interpretação das notícias disponíveis; por fim, o quarto fato é a vontade de acreditar no rumor, pois se um ouvinte quiser acreditar no rumor que ouviu, ele será um propagador, independente da veracidade da informação, conforme relatado no trabalho de (PEZZO; BECKSTEAD, 2006). Nesse trabalho, buscou-se analisar a relação entre ansiedade e crença em situações de propagação de rumores em duas situações reais: a morte de um aluno norte-americano por meningite, e um tiroteio em Washington D.C. Os resultados mostraram uma correlação forte entre propagação dos rumores e crença, com pouca interferência da ansiedade, sugerindo que quanto maior a ansiedade, maior o nível de crença no rumor.

Ansiedade e crença também são confrontados em (BORDIA et al., 2006). O trabalho mostra uma pesquisa entre funcionários de um importante hospital metropolitano que passava por uma grande reestruturação organizacional, envolvendo mudança de instalações, realocação de serviços e privatização parcial dos serviços oferecidos, ou seja uma situação propícia para funcionários ansiosos e terreno fértil para propagação de rumores. De fato, o trabalho conclui que ansiedade e crença se relacionam na propagação de um rumor, sendo que indivíduos muito ansiosos tendem a propagar mais o rumor, e os menos ansiosos, a propagar menos.

Exatamente para entender esses processos, há a necessidade de modelar e simular a propagação de rumores numa sociedade. O objetivo geralmente são dois: entender o processo e propor estratégias para diminuir os impactos negativos da situação. Conforme dito anteriormente, muitos modelos da área de epidemiologia matemática foram aproveitados para a propagação de rumor, pois os pesquisadores começaram a ver similaridades entre a maneira como uma doença e um rumor se espalham. Um exemplo interessante desse uso é o trabalho de (SU; HUANG; ZHAO, 2015), que criou um modelo com base no SIR clássico para testar a eficácia de espalhamento e influência das informações em sites de Microblog. Um Microblog, como o Twitter, por exemplo, é uma forma de blog em que usuários trocam

informações curtas sobre os mais variados temas, permitindo que milhares de mensagens sejam postadas a cada minuto.

O trabalho de (XIONG et al., 2012) tem a figura do indivíduo Refratário, que não propaga mais uma informação (comparável ao Removido, que não propaga mais uma doença). O modelo analisa o comportamento em microblogs, tendo como referência a dinâmica da troca de informações entre os usuários. Além do Refratário, tem-se também o estado Contatado, formando a sigla SCIR. Os contatados são os indivíduos que receberam a informação, mas por algum motivo não a propagam (XIONG et al., 2012).

Em (KOSFELD, 2005) buscou-se avaliar o impacto da propagação de rumores na economia, utilizando-se de um modelo de influência entre vizinhos. No modelo, o rumor é passado de boca em boca pelos vizinhos, e a probabilidade de um indivíduo acreditar em um rumor está ligado à quantidade de vezes que ele ouve esse rumor, assim como a probabilidade de esquecê-lo. Se o indivíduo ouvir o rumor de apenas um vizinho, este poderá acreditar ou esquecer o rumor, mas se o indivíduo ouvir o rumor mais de uma vez, cresce a crença no rumor e esse indivíduo pode passar a propagar também.

Redes complexas também são usadas como modelos para populações no estudo de propagação de informação. Em (ZHOU; LIU; LI, 2007), alguns modelos de redes complexas foram testadas, em que os estados dos indivíduos são baseados no modelo SIR. O estudo concluiu por simulações numéricas que a estrutura das redes têm influencia direta na propagação das informações. Em (NEKOVEE et al., 2007), utilizou-se de um modelo estocástico para estudar a propagação de rumores em redes complexas. Um modelo estocástico não oferece uma solução única, mas sim uma distribuição de resultados expressos em termos de probabilidades. Por fim, (BARGMANN; JORDAN, 2011) faz a classificação dos nós da rede por decomposição, que ajuda na criação de estratégias para impedir a propagação de informações indesejadas ou contribuir para o espalhamento de informações úteis.

A relação de proximidade entre vizinhos na área de marketing de relacionamento é testada em (CHIERICHETTI; LATTANZI; PANCONESI, 2011), com o uso de um modelo *Push-Pull*. Nesse modelo, cada nó ou elemento que conhece a informação seleciona um vizinho aleatoriamente e envia esta informação (Push), ou seja, uma informação foi propagada. No mesmo instante cada nó ou elemento da matriz que ainda não tem a informação, seleciona um vizinho aleatoriamente e solicita a informação (Pull), que será transmitida caso o vizinho selecionado a possua. Nesse modelo, objetiva-se a propagação da informação por todos os indivíduos.

Em (GOLDENBERG; LIBAI; MULLER, 2001), um autômato celular estocástico é usado para generalizar um modelo de difusão de informação de marketing. Estuda-se como as conexões entre indivíduos influenciam a propagação de uma informação fornecida por marketing e absorvida de maneira diferente para cada indivíduo. Os processos de marketing aqui apresentados foram baseados de (ELIASHBERG; LILIEN, 1993), sendo

que a modelagem para a saturação de marketing foi baseada em (BERGER; NASR, 1998).

No próximo capítulo, os modelos são apresentados.

OS MODELOS DISCRETO E CONTÍNUO

3.1 O MODELO DISCRETO LPD

Nesta seção, apresenta-se um modelo SIR modificado, que batizamos de LPD. O LPD nada mais é que um modelo de propagação de informações adaptado a partir do modelo utilizado em (SCHIMIT; MONTEIRO, 2009) também motivado pelos estados que os indivíduos podem assumir (SILVA et al., 2016). Nesse modelo, tem-se o indivíduo *Leigo*, que desconhece a informação, e pode tomar conhecimento dela a partir de um indivíduo *Propagador*, também se tornando Propagador. Após algum tempo propagando a informação, o Propagador pode se tornar *Desinteressado*, parando de falar sobre o assunto com seus vizinhos. Também se considera as transições de estados de Propagador e Desinteressado para Leigo, para prever a troca/renovação da população. Como a atualização dos estados dos indivíduos é simultaneamente feita ao final de cada passo de tempo, o modelo é classificado como síncrono.

Assim como no modelo epidemiológico, a população é representada por um Autômato Celular Probabilista Discreto Síncrono, em que cada célula representa um indivíduo. O autômato é bidimensional, de dimensão $n \times n = N$, em que a primeira e última colunas se conectam, assim como as linhas superior e inferior, formando uma superfície toroidal. A vizinhança é a mesma usada em (SCHIMIT; MONTEIRO, 2009), em que cada indivíduo se conecta com C indivíduo dentro de uma vizinhança de Moore de raio r , sendo $r = 1$ uma vizinhança de Moore clássica composta apenas pela primeira camada que circunda o indivíduo. A vizinhança é definida conforme apresentado na seção 2.2.

Esse modelo é apenas uma interpretação do modelo proposto em (SCHIMIT; MONTEIRO, 2009), e o diagrama de estados se encontra na Figura 3.1. Leigo se torna propagador com probabilidade $P_p(v) = 1 - e^{-kv}$, sendo v a quantidade de indivíduos com informação com quem o Leigo teve contato, e k um parâmetro específico do tipo de informação. Um Propagador se torna Desinteressado com probabilidade P_s , e se torna Leigo com probabilidade P_f . Um Desinteressado se torna Leigo com probabilidade P_q . Além disso, os indivíduos podem não mudar de estado num passo de tempo.

A tabela 3.1 mostra uma comparação entre o modelo apresentado em (SCHIMIT; MONTEIRO, 2009) e o modelo LPD.

Vale ressaltar que alguns trabalhos usam essas comparações entre modelos em seus estudos. Em (SCHIMIT; BARROS JUNIOR; ARAÚJO, 2017), por exemplo, o modelo SIR foi adaptado para contemplar a informação que os indivíduos tinham a respeito de um produto. Em (ZHAO et al., 2013), um modelo SIR foi modificado para se prever uma taxa de esquecimento da informação (que aqui pode ser comparada com a probabilidade P_q).

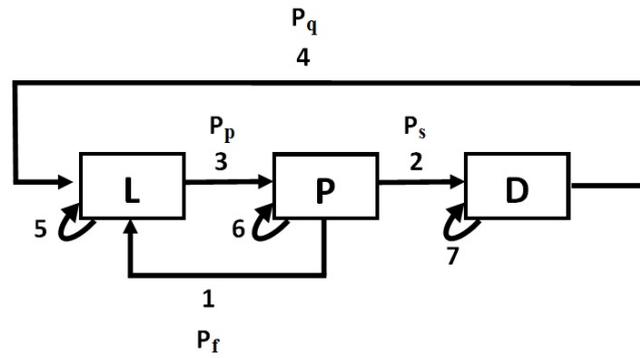


Figura 3.1: Diagrama de estados do modelo LPD.

Tabela 3.1: Comparação entre as variáveis dos dois modelos.

Modelo epidemiológico		Modelo de propagação de informação	
Variável	Interpretação	Variável análoga	Interpretação
S	Suscetível	L	Leigo
I	Infectado	P	Propagador
R	Recuperado	D	Desinteressado
P_i	Probabilidade de infecção	P_p	Probabilidade de conhecer a informação
P_c	Probabilidade de cura	P_s	Probabilidade do Propagador parar de propagar a informação
P_d	Probabilidade de morte pela doença	P_f	Probabilidade de demissão do Propagador
P_n	Probabilidade de morte natural	P_q	Probabilidade do Desinteressado morrer ou se renovar
K	Parâmetro relacionado à infectividade da doença	K_b	Parâmetro relacionado à força da informação
N	Tamanho da população	N	Quantidade de funcionários

Os resultados serão mostrados no próximo capítulo, mas é interessante notar que nesse modelo discreto baseado no modelo SIR, não há espaço para configurar campanhas de marketing com variações de comportamento, essa limitação fica evidente quando se observa que o modelo discreto possui apenas uma variável (P_s) para trabalhar as opções de propagação das informações. Para contornar esse ponto, outro modelo será proposto: o modelo contínuo assíncrono.

3.2 O MODELO CONTÍNUO ASSÍNCRONO

O modelo contínuo é baseado num Autômato Celular Probabilista Contínuo Assíncrono. Diferente do modelo discreto LPD mostrado anteriormente, aqui, o estado do indivíduo é baseado numa variável contínua, que é a quantidade de informação que o indivíduo

possui num determinado passo de tempo t . Seja essa variável denotada por $I_i(t)$ para um indivíduo i , tem-se que a informação desse indivíduo está no intervalo $0 \leq I_i(t) \leq 1$. Ou seja, um indivíduo com pouca informação, ou que fala pouco sobre determinado assunto tem um $I_i(t)$ baixo. Já um indivíduo com bastante informação sobre um tema, e que propaga essa informação com seus próximos tem um $I_i(t)$ alto. Esse valor contínuo como estado do indivíduo classifica o Autômato Celular como contínuo. Como as atualizações da informação de cada indivíduo são feitas assim que se ganha ou se perde informação (e não somente ao final de cada passo de tempo, como no modelo discreto), chama-se o modelo de assíncrono.

Note que nesse caso a quantidade de informação pode ter diferentes interpretações: pode ser a quantidade efetiva de informação que um indivíduo possui sobre um tema, pode ser a probabilidade do indivíduo falar sobre o tema e passar para seus próximos, ou pode ser uma combinação desses dois fatores.

A população é baseada num reticulado bidimensional com $n \times n = N$ células, com um indivíduo por célula. Cada indivíduo interage com seus vizinhos conforme apresentado na seção 2.2.

Os processos que fazem variar a informação de um indivíduo são quatro: influência de vizinhos, marketing positivo, marketing negativo e esquecimento. Um resumo desses processos é mostrado na figura 3.2.

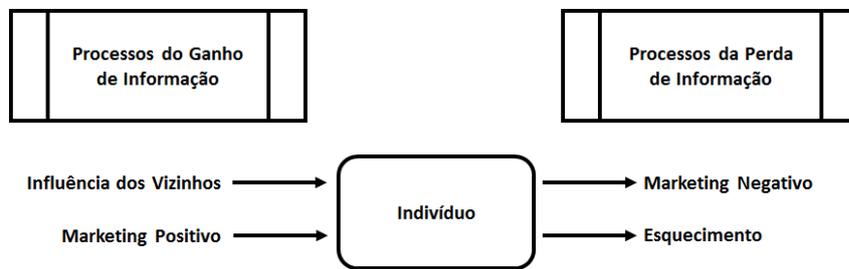


Figura 3.2: *Processos de ganho e perda de informações de cada indivíduo.*

Portanto, considerando um indivíduo i , os processos de ganho de informação são dados por:

- **Influência de cada vizinho j :** quando dois indivíduos têm diferentes valores de informação, aquele com maior informação “ensina” uma fração de informação para o outro. Um dos objetivos do modelo é a flexibilidade de se trabalhar com diferentes tipos de informação, como aquelas fáceis de se explicar e de fazer a população ganhar informação (uma campanha de vacinação, por exemplo), e aquelas mais complicadas, com menor passagem de informação (como tentar explicar a estrutura tributária brasileira). Para isso, considerando que $I_j(t) > I_i(t)$ sem perda de generalidade, o ganho de informação é dado por $G_{i,j}^n(t) = I_j(t) \cdot (I_j(t) - I_i(t)) \cdot e^{-k_i \cdot (I_j(t) - I_i(t))}$, e a informação de j não é modificada. Assim, a cada passo de tempo, o indivíduo i tem

um ganho total de $\sum_{j=1}^C \max(G_{i,j}^n(t), 0)$, considerando os C vizinhos. O tipo ou a intensidade de espalhamento da informação pode ser ajustada pelo parâmetro k_i ;

- Marketing positivo: a cada passo de tempo, uma parcela m da população recebe uma quantidade de informação em função da informação que possui. Há uma flexibilidade do modelo para que as informações sejam absorvidas pelos indivíduos de maneira diferente, de acordo com sua informação, e com o tipo de informação sendo veiculada. O ganho é dado por: $G_i^{mp}(t) = (1 - I_i(t)) \cdot e^{-k_m/(1-I_i(t))}$, sendo que k_m pode ser configurável de acordo com a campanha.

Por outro lado, os indivíduos perdem informação por conta dos seguintes processos:

- Marketing negativo: Caso um indivíduo receba uma campanha de marketing por dois passos de tempo seguidos, ele perde parte da informação recebida no passo anterior de acordo com $G_i^{mn}(t) = -\mu \cdot I_i(t) \cdot G_i^{mp}(t-1)$, sendo μ o peso do grau de rejeição do marketing exaustivo ($0 \leq \mu \leq 1$);
- Esquecimento: A cada passo de tempo, todos os indivíduos esquecem parte da informação que possuem. Portanto, o esquecimento é dado por $G_i^f(t) = I_i(t) \cdot (1 - e^{-k_e \cdot I_i(t)})$, sendo k_e uma variável que controla a quantidade esquecida, e pode ser ajustável para o tipo de informação circulando;

Ou seja, k_i , k_m e k_e são parâmetros que podem ser ajustados de acordo com a informação a ser analisada.

Para entender como esses parâmetros influenciam o ganho de informação, as Figuras 3.3, 3.4 e 3.5 contêm os ganhos em função da informação do indivíduo, ou da diferença para outro indivíduo, para diferentes valores de k_i , k_m e k_e . Por exemplo, para $k_i = 1$, tem-se uma informação que é bastante transferida entre indivíduos caso a diferença de informação deles seja grande, o que não ocorre para $k_i = 6$. Caso se deseje trabalhar com informações que são muito esquecidas, pode-se adotar $k_e = 2$.

No começo de cada simulação, alguns indivíduos começam com uma certa quantidade de informação, no intervalo $0 \leq I_i(0) \leq I_0$. Ou seja, a informação máxima de um indivíduo no instante inicial é I_0 . Por exemplo, considere que todos os 100 indivíduos de uma população comece com um valor aleatório uniformemente distribuído entre 0 e I_0 no instante inicial. Nesse cenário, tem-se $\overline{I_i(0)} = I_0/2$.

A tabela 3.2 contém um resumo das variáveis, definição e descrição.

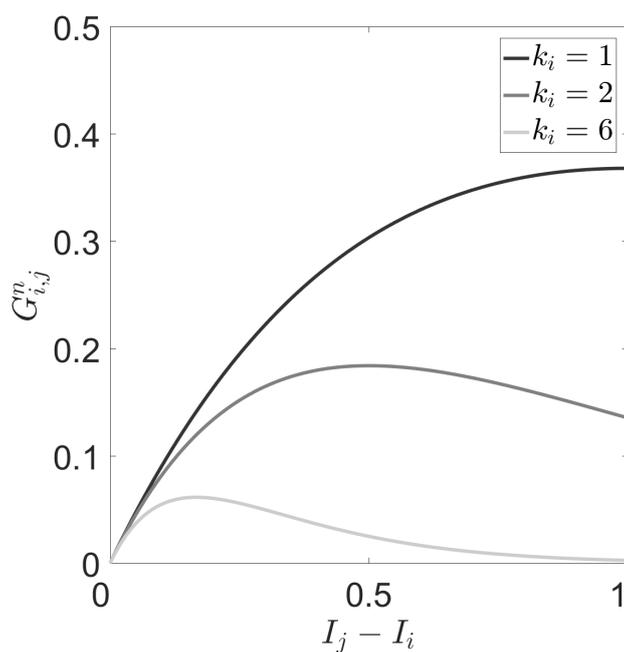


Figura 3.3: Ganho de informação do indivíduo i por influência de um indivíduo j dada por $G_{i,j}^n$, considerando $I_j > I_i$, para $k_i = 1$, $k_i = 2$ e $k_i = 6$.

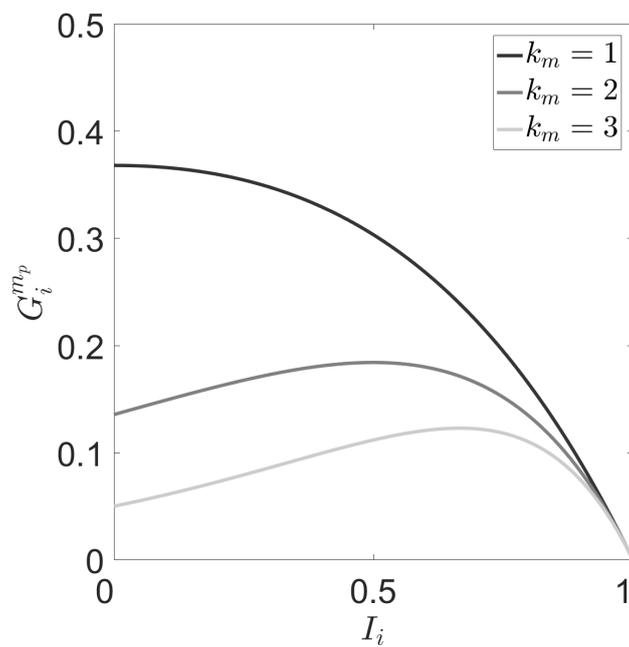


Figura 3.4: Ganho de informação do indivíduo i por influência de marketing positivo dada por G_i^{mp} , para $k_m = 1$, $k_m = 2$ e $k_m = 3$.

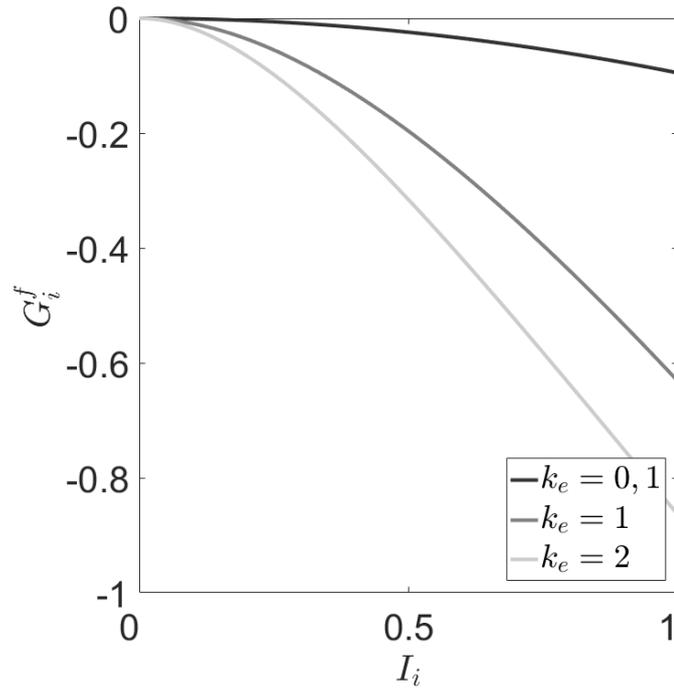


Figura 3.5: Perda de informação do indivíduo i por esquecimento dada por G_i^f , para $k_e = 0,1$, $k_e = 1$ e $k_e = 2$.

Tabela de Notação

Notação	Definição	Descrição
$G_{i,j}^m$	$\in [0,1]$	Processo do ganho de informação por influência dos vizinhos.
G_i^{mp}	$\in [0,1]$	Processo do ganho de informação por influência do processo de marketing positivo.
G_i^f	$\in [0,1]$	Processo de perda de informação devido ao esquecimento.
G_i^{mn}	$\in [0,1]$	Processo de perda de informação devido ao marketing negativo.
k_i	$\in [0,\infty]$	Parâmetro relacionado à G_i^m .
k_m	$\in [0,\infty]$	Parâmetro relacionado à G_i^{mp} .
k_e	$\in [0,\infty]$	Parâmetro relacionado à G_i^f .
m	$\in [0,1]$	Percentual de indivíduos que sofrem influência do marketing positivo.
μ	$\in [0,1]$	Rejeição das campanhas de marketing, relacionado ao processo de marketing negativo.
$\overline{I(0)}$	$\in [0,1]$	Média de informação da população como condição inicial.
I_0	$\in [0,1]$	Quantidade máxima de informação de um indivíduo no instante inicial.

Tabela 3.2: Notação usada neste trabalho.

Este capítulo contém os resultados das simulações realizadas com os modelos apresentados. O modelo discreto foi usado para a avaliação do impacto da propagação de informações falsas dentro de uma empresa fictícia com aproximadamente 50.000 funcionários. Já o modelo contínuo foi usado para avaliar o impacto de campanhas de marketing em diferentes situações.

4.1 MODELO DISCRETO PARA PROPAGAÇÃO DE INFORMAÇÕES FALSAS EM EMPRESAS

O modelo LPD é usado com um ACP bidimensional de tamanho $n = 223$, com $N = n^2$, com $N = 49.429$ indivíduos. Cada indivíduo representa um funcionário numa empresa fictícia, e espera-se avaliar ações para diminuir a propagação de informações falsas dentro da empresa. Assim, em todas as simulações tem-se a seguinte condição inicial: $L(0) = 99,5\%$, $P(0) = 0,05\%$, $D(0) = 0\%$ com todos os indivíduos homogeneamente espalhados pelo reticulado, com as probabilidades $P_s = 0,6$, $P_f = 0,3$, $P_q = 0,1$ e $K_b = 1$. As variáveis do modelo LPD são interpretadas conforme Tabela 4.1. Todas as simulações feitas nessa seção rodam por $t_s = 100$ passos de tempo, sendo que verifica-se que o sistema atinge o regime permanente após o passo de tempo $t = 80$ para todos os casos. Dessa maneira, as médias de todas as variáveis são tomadas pelos valores dos últimos 20 passos de tempo das simulações. Além disso, a vizinhança é configurada com $C = 8$ e $r = 1$.

Nesse ambiente de uma empresa, cada indivíduo pode permanecer em seu estado ao final de cada passo de tempo, ou mudar para outro estado, de acordo com as transições a seguir:

1. $P \rightarrow L$: Indivíduo que propagava a informação falsa e é demitido da empresa;
2. $P \rightarrow D$: Indivíduo para de propagar a informação falsa;
3. $L \rightarrow P$: Indivíduo que acredita na informação falsa e passa a propagá-la;
4. $D \rightarrow L$: Indivíduo que se desliga da empresa sem relação direta com a informação falsa.

Para entender a influência de parâmetros relacionados com um possível combate à propagação de uma informação na empresa, foram realizados os seguintes experimentos:

1. Fixa-se $P_f = 0,3$ e $P_s = 0,6$, e varia-se K_b para os valores $K_b = 0, 0,1, 0,2, \dots, 1$;
2. Fixa-se $K_b = 1$ e $P_f = 0,3$, e varia-se P_s para os valores $P_s = 0, 0,1, 0,2, \dots, 1$;

Tabela 4.1: *O Modelo LPD e o Espalhamento de Informações Falsas Numa Empresa.*

Modelo de Propagação de Informações		Detalhamento do Modelo
Variável	Interpretação	Interpretação
L	Leigo	Indivíduo que desconhece informação
P	Propagador	Indivíduo que propaga a informação
D	Desinteressado	Indivíduo que ignora a informação
P_p	Probabilidade de conhecer a informação	Indivíduo que pode se interessar pela informação
P_s	Probabilidade do Propagador parar de propagar a informação	Devido à empresa aplicar medidas para esclarecer os fatos com minimizando o impacto das informações falsas
P_f	Probabilidade de demissão do propagador	Empresa demite o funcionário que propaga informações falsas
P_q	Probabilidade do Desinteressado se desligar da empresa por outras causas	Indivíduo que se demite ou é demitido, mas sem relação com a informação em circulação
K_b	Parâmetro relacionado à força da informação	Força da informação se espalha
N	Quantidade de indivíduos	Quantidade de funcionários

3. Fixa-se $K_b = 1$ e $P_s = 0,6$, e varia-se P_f para os valores $P_f = 0, 0,1, 0,2, \dots, 1$;

Essas simulações servem para determinar qual a melhor situação para a empresa:

1. investir em campanhas educativas para os funcionários (associado a P_s), para que não acreditem e nem espalhem informações falsas;
2. demitir todo e qualquer funcionário que disseminar informações falsas (associado a P_f), sendo uma forma de punição por comportamento reprovável;
3. Agir para alterar a força de propagação da informação falsa, trabalhando para mostrar que é falsa (associada a K_b);

Há de se levar em conta os custos de demissão de um funcionário e o custo de uma campanha contínua contra informações falsas na empresa para avaliar qual a melhor medida a ser adotada (LI et al., 2015). As campanhas educativas em empresas podem ser realizadas na forma de cartazes, envio de e-mails e palestras educativas.

As Figuras 4.1, 4.2 e 4.3 mostram a quantidade média de propagadores da informação falsa e a quantidade média de novos indivíduos que passam a conhecer a informação falsa (transição $L \rightarrow P$) em função de, respectivamente, K_b , P_s e P_f .

Note que para $K_b < 0,2$ na Figura 4.1, a informação praticamente não se propaga na população. Acima desse valor, há um crescimento conforme se aumenta K_b , porém a variação é cada vez menor, indicando um equilíbrio para valores grandes de K_b . Uma

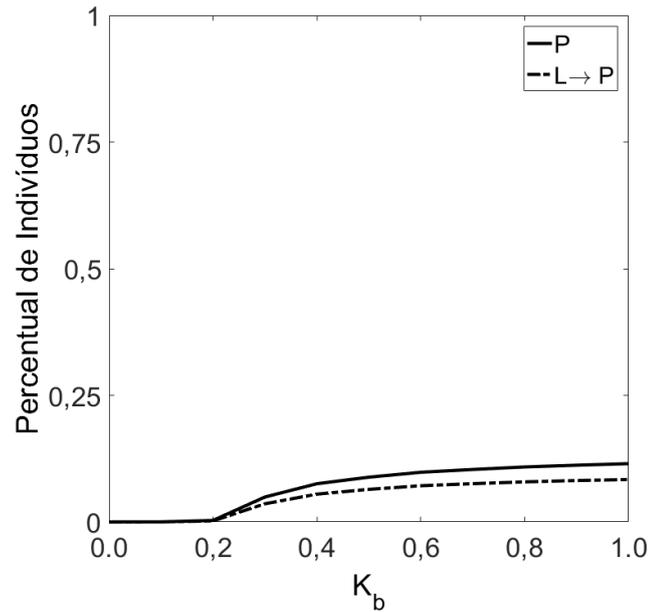


Figura 4.1: Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de K_b , com $P_s = 0,6$ e $P_f = 0,3$.

empresa teria dificuldades de mudar o valor desse parâmetro, a não ser que, entre os indivíduos propagadores, mostre-se informações que tirem o crédito da informação.

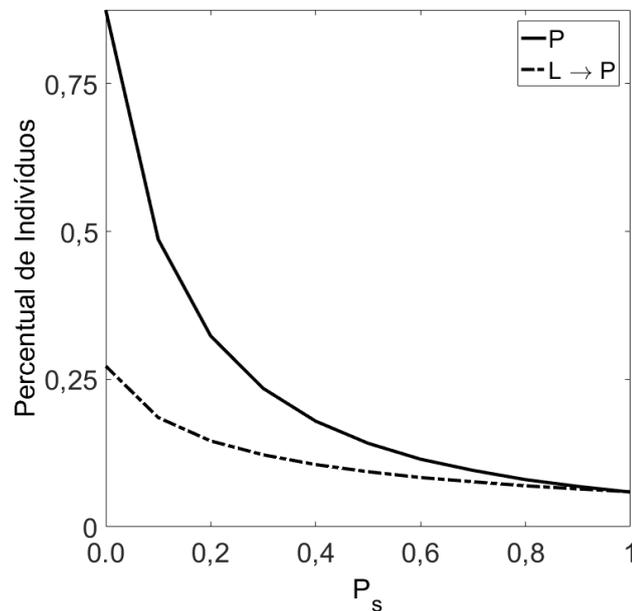


Figura 4.2: Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de P_s , com $K_b = 1$ e $P_f = 0,3$.

Na Figura 4.2, percebe-se que se a probabilidade do Propagador parar de falar sobre a informação for pequena, uma grande parcela da população fala sobre isso, com muitos funcionários sendo propagadores e uma alta taxa de transição de Leigos para Propagadores. Um pequeno aumento em P_s (de $P_s = 0,2$ para $P_s = 0,4$, por exemplo) pode-se reduzir à metade a quantidade de Propagadores na empresa. Isso pode ser feito com campanhas

educativas, mostrando aos funcionários que aquela informação falsa não procede.

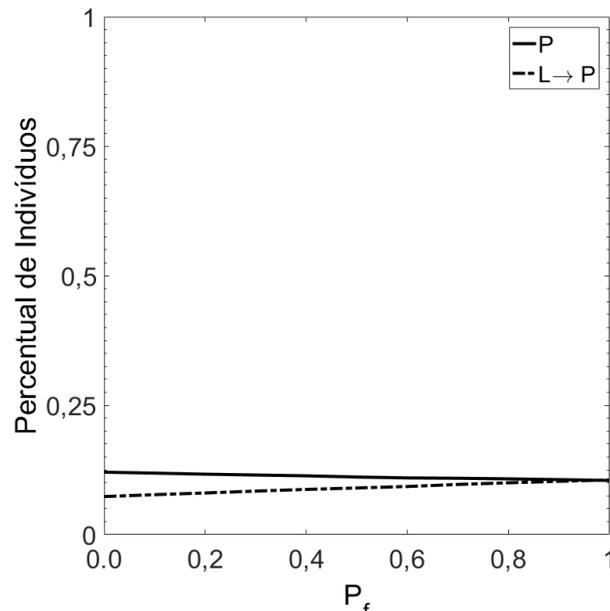


Figura 4.3: Quantidade de Propagadores (P) e transições de Propagadores para Leigos ($L \rightarrow P$) em função de P_f , com $K_b = 1$ e $P_s = 0,6$.

A demissão de funcionários, aqui associada à variável P_f , parece ser pouco efetiva, conforme mostra a Figura 4.3. Uma redução grande nessa probabilidade não leva a uma redução de Propagadores na empresa, e diminui muito pouco a taxa de transição de Leigos para Propagadores.

Frente aos resultados mostrados, o parâmetro a ser considerado para mudar o cenário de propagação de uma informação na empresa é o P_s , associado à campanhas educativas. Os outros parâmetros, além de estarem relacionados à ações mais custosas e de difícil aplicação, ainda são pouco efetivas.

Em termos de ações de marketing, esse modelo se mostra limitado. Para contornar isso, foi necessário propor o modelo contínuo, e com ele, criar variáveis adequadas para levar a informação da melhor maneira à população considerada.

4.2 MODELO CONTÍNUO PARA SIMULAÇÃO DE PROCESSOS DE MARKETING

Para entender a dinâmica do modelo contínuo, apresenta-se simulações com diferentes processos de ganho e perda de informação. Em seguida, diferentes configurações das variáveis relacionadas ao marketing são testadas para entender sua influência na população. Vale lembrar que uma das vantagens do modelo contínuo é a facilidade de adaptação do modelo para outros similares da área, apenas escolhendo seus parâmetros.

Em todas as simulações feitas nessa seção foi usado um reticulado de tamanho $n = 100$, $N = n^2$, $N = 10.000$, com a vizinhança definida por $C = 4$ e $r = 4$. Como uma informação geralmente começa a ser espalhada a partir de alguns agentes dentro de uma

população, os indivíduos i com $I_i(0) > 0$ serão somente os indivíduos presentes num quadrado centralizado no reticulado de tamanho 10×10 .

4.2.1 O MODELO SEM MARKETING E SEM ESQUECIMENTO

Considerando que a população passa pelo efeito da propagação de informações apenas pelo contato entre indivíduos e sem perda de informação, tem-se o caso da propagação de uma informação falsa em um certo período de tempo. Não há declarações por parte da mídia, por ser um intervalo curto de tempo pode-se se considerar que as pessoas não esquecem as informações. Portanto, configura-se uma situação sem marketing e sem esquecimento, com parâmetros $k_m = 0$, $k_e = 0$ e $m = 0$. Além disso, escolheu-se $k_i = 10$, $I_0 = 0,4$ $\overline{I(0)} = 0,002$.

A figura 4.4 mostra a evolução temporal das variáveis de média de informação \overline{I} , média de ganho de informação por vizinhos $\overline{G_{i,j}^n}$, média com ganho de marketing $\overline{G_i^{m_p}}$ e média de esquecimento $\overline{G_i^f}$. A distribuição inicial de indivíduos com informação é centralizada, com $I_0 = 0,4$, e como esses 100 indivíduos correspondem a 1% da população e somente eles têm informação, a média de população no início da simulação é baixa ($\overline{I(0)} = 0,002$). Além disso, todos os ganhos de outros processos de passagem de informações ficam zerados ao longo da simulação, devido aos parâmetros escolhidos.

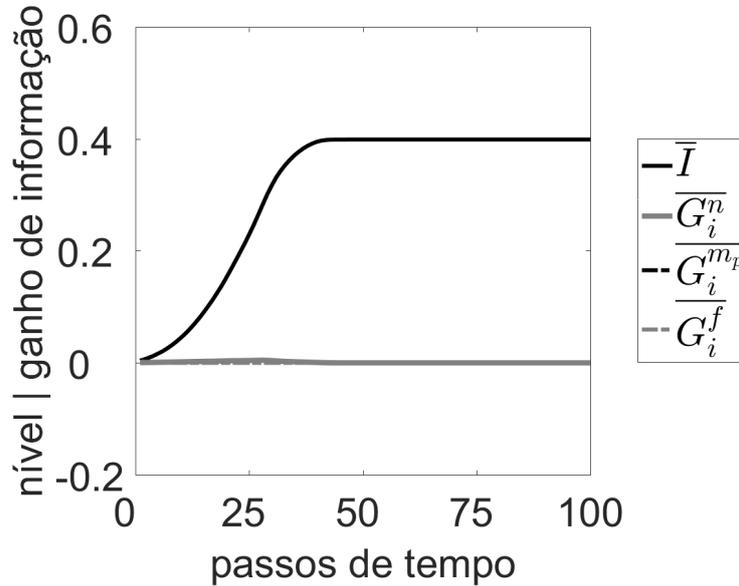


Figura 4.4: Evolução temporal da informação (\overline{I}), influência da vizinhança ($\overline{G_i^n}$), marketing positivo ($\overline{G_i^{m_p}}$), esquecimento ($\overline{G_i^f}$) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0$, $m = 0$ e $I_0 = 0,4$.

Para verificar o comportamento espacial da propagação da informação, a figura 4.5 contém imagens do reticulado nos instantes de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$ e

$t = 100$. Não há ganho de informação na população, somente seu espalhamento, já que todos os indivíduos da população tem $I_i(100) \approx 0,4$ ao final da simulação, que é um valor próximo do valor máximo de um indivíduo na condição inicial. Ou seja, ao final, todos sabem, no máximo, o que o indivíduo com mais informação na população possui de informação.

Além disso, o centro do círculo de espalhamento vai descendo conforme a simulação é efetuada. Isso se deve ao tratamento das células do reticulado começando na parte superior esquerda, fazendo todas as linhas da esquerda para a direita, até a última linha. Soma-se o fato de se ter uma atualização assíncrona das informações, e uma superfície toroidal, visto que no passo de tempo $t = 25$, tem-se a onda de informação passando de baixo para cima no reticulado.



Figura 4.5: *Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0$, $m = 0$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$.*

4.2.2 O MODELO SEM MARKETING E COM ESQUECIMENTO

Considere o mesmo conjunto de parâmetros da seção anterior, porém considerando que os indivíduos esquecem parte da informação a cada passo de tempo. Para isso, toma-se $k_e = 0,25$ e a evolução temporal é apresentada na Figura 4.6. Note que o nível de informação da população é baixo, e como o esquecimento é parte dessa informação baixa, também se mantém próximo de zero.

A difusão espacial desse caso está na Figura 4.7. A informação se espalha pelo reticulado, porém vai perdendo intensidade dado o esquecimento. Esse caso pode representar o comportamento de uma notícia de média relevância que tem um pico de espalhamento, muitas pessoas passam a conhecer a informação que depois é naturalmente esquecida.

4.2.3 O MODELO COM MARKETING E COM ESQUECIMENTO

Agora considere que além dos processos de passagem de informação por vizinhos e esquecimento, há também a renovação da informação por campanhas de marketing. Nesse caso, 10% ($m = 0,1$) dos indivíduos recebem uma campanha de marketing, com a força

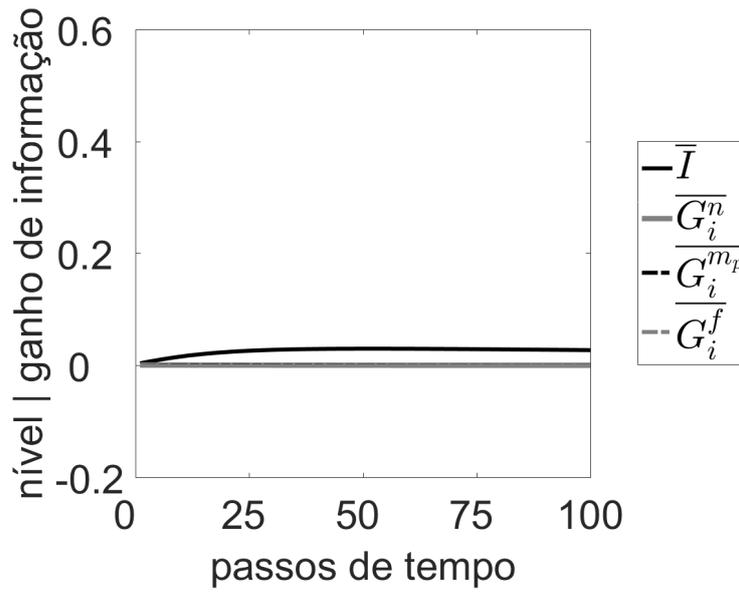


Figura 4.6: Evolução temporal da informação (\bar{I}), influência da vizinhança ($\overline{G_i^n}$), marketing positivo ($\overline{G_i^{m_p}}$) e esquecimento ($\overline{G_i^f}$) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0,25$, $m = 0$ e $I_0 = 0,4$.



Figura 4.7: Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 0$, $k_e = 0,25$, $m = 0$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$.

de influência do marketing com $k_m = 10$. A evolução temporal para esse caso está na Figura 4.8. Há bastante perda de informação por esquecimento, um pequeno ganho por marketing e por influência de vizinhos.

Esse caso pode ser interpretado como, por exemplo, o caso de doenças que podem ser evitadas por vacinação. Há campanhas para informar à população sobre a vacinação, porém o que leva as pessoas a voltarem a falar sobre o assunto são indivíduos que contraem a doença e geram informação.

A evolução espacial desse caso não tem um espalhamento a partir do centro, e em pouco tempo, a informação está totalmente espalhada pelo reticulado, já que as campanhas de marketing acontecem em todo o reticulado.

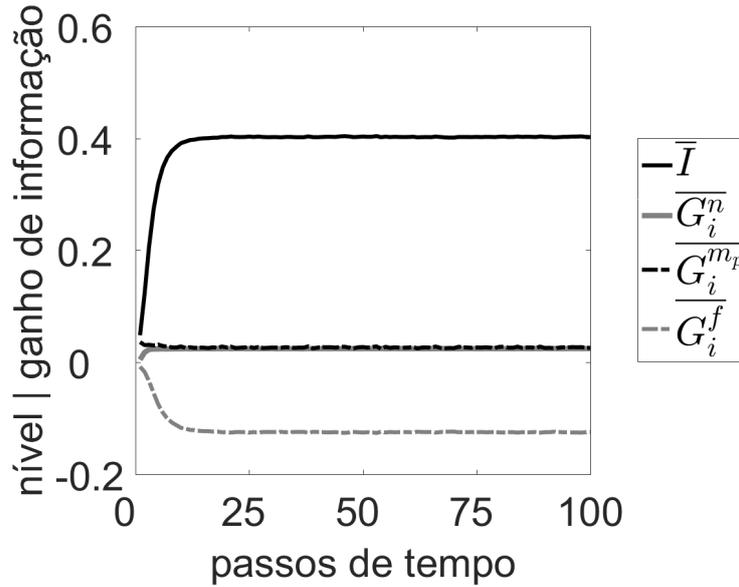


Figura 4.8: *Evolução temporal da informação (\bar{I}), influência da vizinhança (\overline{G}_i^n), marketing positivo ($\overline{G}_i^{m_p}$), esquecimento (\overline{G}_i^f) usando os parâmetros $C = 4$, $r = 4$, $k_i = 10$, $k_m = 10$, $k_e = 0,5$, $m = 0,1$ e $I_0 = 0,4$.*

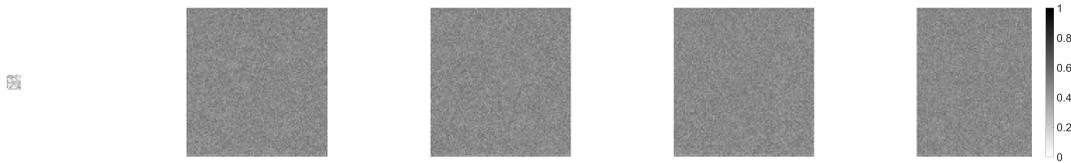


Figura 4.9: *Propagação da informação quando a condição inicial consiste de um quadrado 10×10 centralizado no reticulado com informação dos indivíduos aleatoriamente inicializada entre 0 e 0,4. Os parâmetros do modelo são: $C = 4$, $r = 4$, $k_i = 10$, $k_m = 10$, $k_e = 0,5$, $m = 0,1$ e $I_0 = 0,4$. Representação do reticulado para os passos de tempo $t = 0$, $t = 25$, $t = 50$, $t = 75$, $t = 100$.*

4.2.4 A INFLUÊNCIA DO MARKETING NO MODELO PROPOSTO

O modelo de marketing apresentado contém três variáveis: a porcentagem de pessoas que receberão a campanha de marketing m , o peso da campanha para esses indivíduos k_m e a rejeição à exposições seguidas de marketing μ . Apesar de ser difícil de analisar as variações e cenários possíveis dessas três variáveis, tal modelagem é necessária para que seja possível usar esses modelos em outros diversos modelos da área de modelagem de propagação de informação.

Além disso, simulações com essas variáveis ajudam a responder perguntas do tipo: “É mais interessante uma campanha com informações mais detalhadas para uma pequena quantidade de indivíduos, ou informações mais básicas a uma grande quantidade de indivíduos?”, “Qual possui o melhor custo-benefício para aumentar o nível de informação na

população?”, ou “Como a rejeição a uma certa campanha de marketing pode ajudar na escolha de melhores configurações de campanha?”.

Para responder a esses pontos, diversas simulações foram feitas variando-se m , k_m e μ , e analisando como essas variáveis interferem no ganho médio de informação por marketing ($\overline{G}_i^m = \overline{G}_i^{m_p} + \overline{G}_i^{m_n}$) e na média de informação na população (\bar{I}). Portanto, a Figura 4.10 contém \overline{G}_i^m em função de m e k_m para $k_i = 2, 6, 10$ e $k_e = 0, 1, 2$ usando $I_0 = 0, 2$, $\mu = 0, 5$. As simulações foram feitas para todas as combinações de $m = 0, 0, 02, 0, 04, \dots, 1$ e $k_m = 0, 0, 02, 0, 04, \dots, 3$.

É possível avaliar que quando há pouco esquecimento na população (Figura 4.10.a, 4.10.d, 4.10.g), é melhor configurar a campanha de marketing para atingir cerca de 30% da população, passando bastante informação para os indivíduos. Conforme o esquecimento começa a crescer, os melhores resultados de ganho de informação na população por marketing se dão quando a campanha é feita a cerca de 50% da população, com maior investimento na quantidade de informações por indivíduos atingidos na campanha.

Quando se olha para a quantidade média de informação na população para esses cenários, deve-se levar em consideração que é interessante para os elaboradores de campanhas de marketing que a informação fique forte na população. Para isso, a Figura 4.11 contém a quantidade média de informação para as mesmas variações de m e k_m na população. O pouco esquecimento satura a quantidade de informação na população, sendo relevante tomar cuidado ao elaborar a campanha de marketing quando a passagem de informações entre indivíduos é baixa (Figura 4.11.g).

É relevante ressaltar que apesar das Figuras 4.10.c, 4.10.f, 4.10.i, indicarem que o ganho de marketing é máximo quando a campanha é feita a 50% da população com o máximo de informações por indivíduo, as Figuras 4.11.c, 4.11.f, 4.11.i, mostram que o nível de informação é mais alto quando os mesmos 50% recebem uma informação média. Isso pode ocorrer pois o esquecimento e o marketing negativo acabam tirando mais informação proporcionalmente de populações com maior nível de informação.

Por fim, para entender a rejeição a uma campanha de marketing a Figura 4.12 mostra a quantidade média de informação na população e o ganho de marketing em função de m , k_m e μ , variando-se $m = 0, 0, 02, 0, 04, \dots, 1$, $k_m = 0, 0, 02, 0, 04, \dots, 3$ e $\mu = 0, 0, 02, 0, 04, \dots, 1$. Em cada figura, fixa-se m ou k_m .

Mesmo variando a rejeição a uma campanha de marketing, é possível apostar de maneira segura em algumas configurações para manter a informação alta na população. Por exemplo, ao escolher $m = 0, 3$ e $k_m = 0$, nota-se que levam a regiões nas figuras com maior nível de informação e de ganho de marketing.

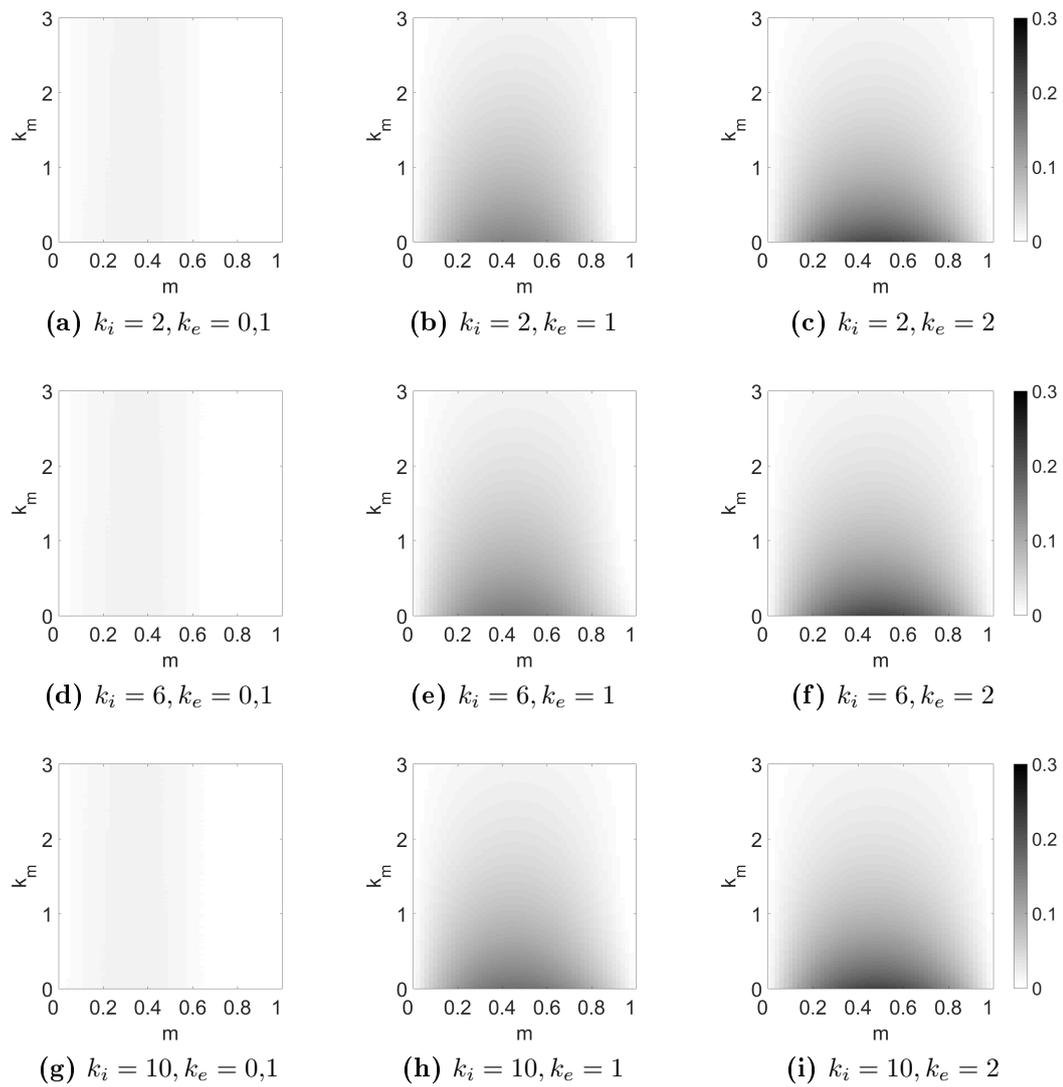


Figura 4.10: *Ganho com marketing ($\overline{G^m}$) em função de m e k_m para $k_i = 2, 6, 10$ e $k_e = 0,1, 1, 2$ usando $I_0 = 0,2$, $\mu = 0,5$.*

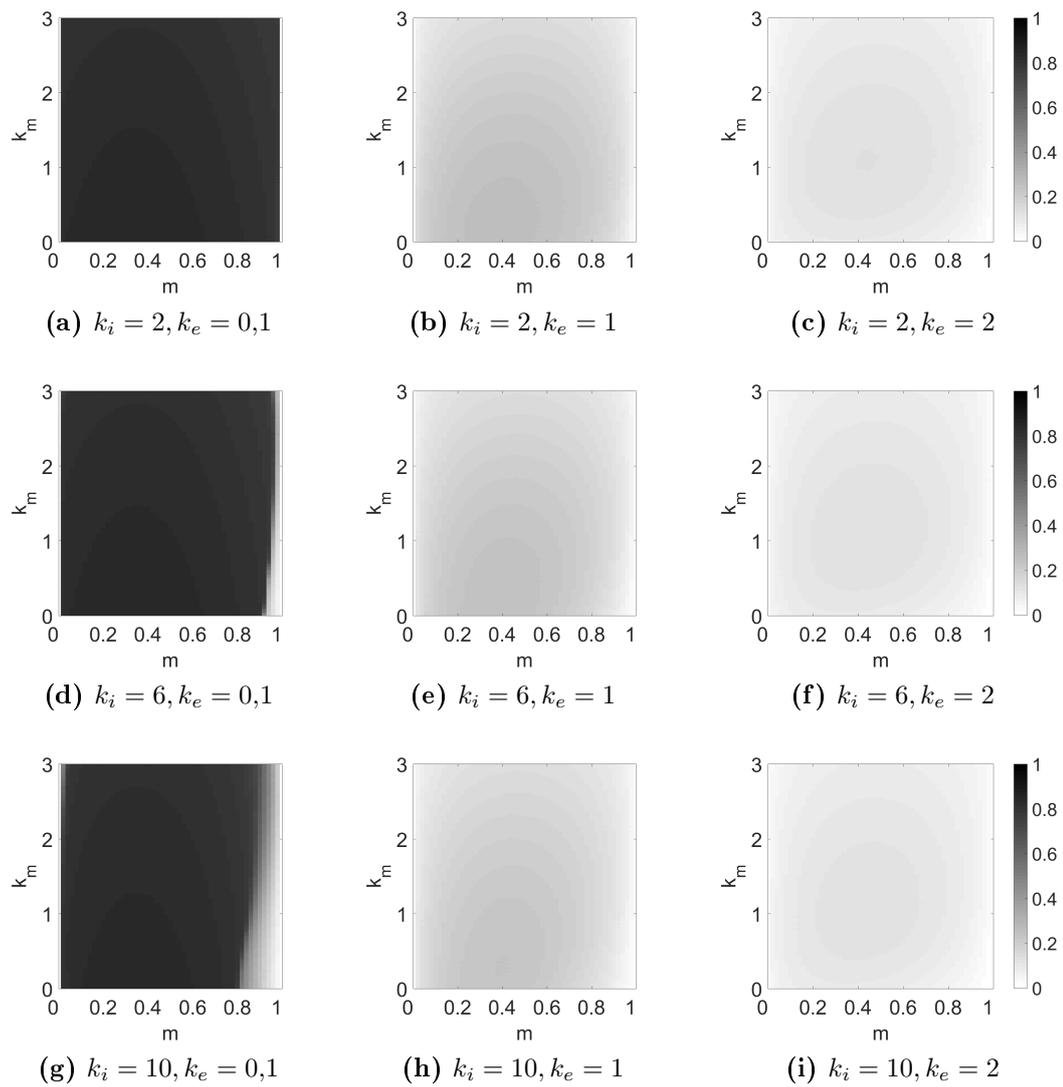


Figura 4.11: Média da informação (\bar{I}) em função de m e k_m para $k_i = 2, 6, 10$ e $k_e = 0,1, 1, 2$ usando $I_0 = 0,2$.

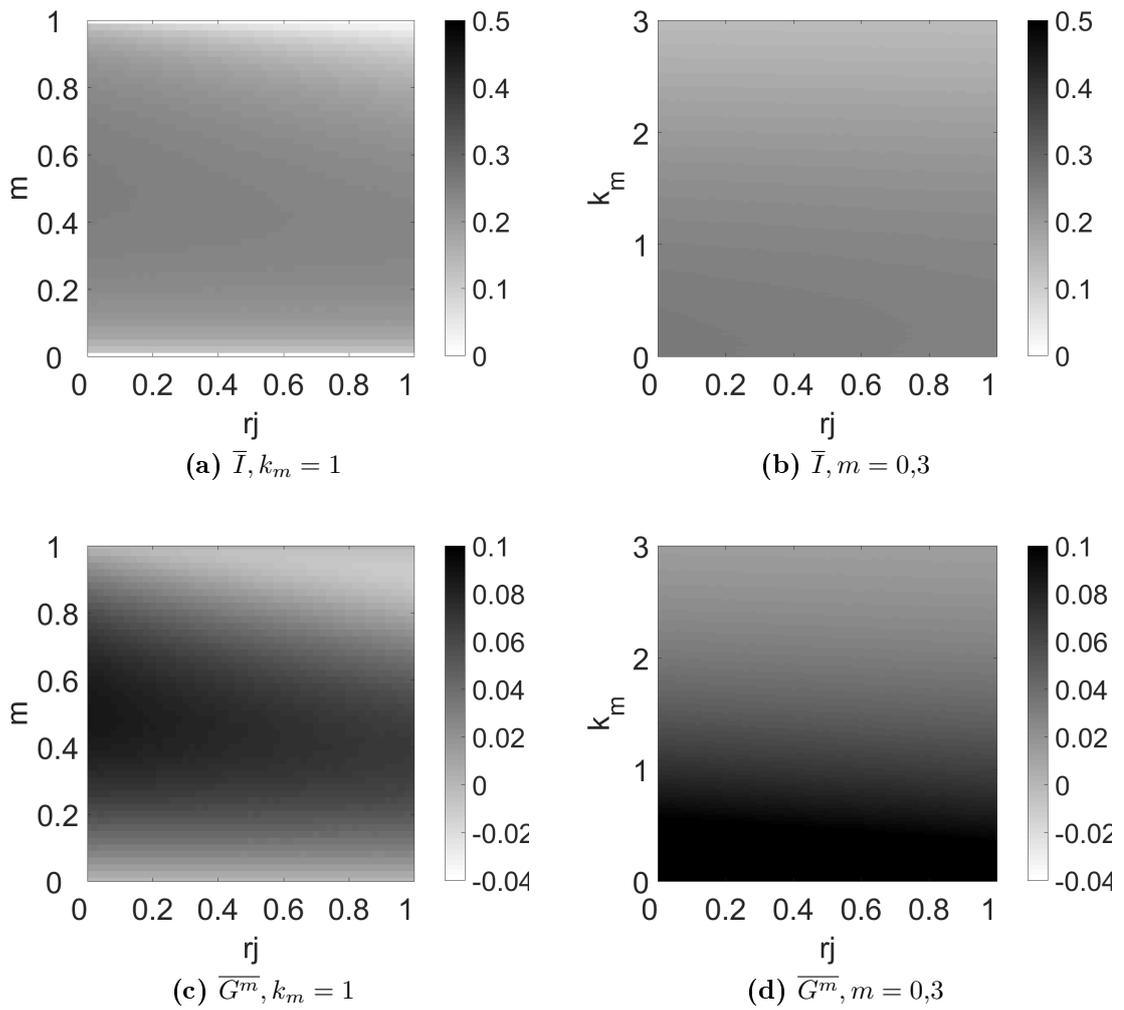


Figura 4.12: Média da informação (\bar{I}) [a, b] e ganho com marketing (\bar{G}^m) [c, d] em função de μ e m com $k_m = 1$ [a, c] em função de μ e k_m com $m = 0,3$ [b, d] usando $I_0 = 0,4$.

CONCLUSÕES

A propagação de uma informação é um processo complexo e pode-se representá-la através de diversos modelos criados para demonstrar o fluxo destas informações. No presente trabalho demonstrou-se alguns cenários possíveis onde o espalhamento de informações falsas pôde ser analisado tendo como base a teoria do espalhamento de uma doença epidemiológica. Isso foi feito com dois modelos: um discreto e síncrono, próximo a um modelo SIR, e outro contínuo e assíncrono, melhor adaptado para trabalhar com outras questões relacionadas à propagação de uma informação, como processos de marketing, por exemplo.

O modelo discreto foi usado para o caso de propagação de informações falsas em empresas. A importância de se resolver rapidamente um surto de uma informação incerta é algo bastante estudado na literatura, como por exemplo os trabalhos de (BORDIA et al., 2006; PEZZO; BECKSTEAD, 2006). Aqui, as simulações sugeriram que o melhor caminho é o investimento em campanhas educativas para mostrar aos indivíduos que a informação não é (totalmente) verdadeira, acalmando os ânimos. O modelo se mostrou limitado para ampliar os métodos da campanha educativa, e uma melhor abordagem de marketing em populações foi abordada no modelo contínuo.

Nesse modelo contínuo, alguns cenários de propagação de informação foram mostrados, ressaltando a influência de cada parâmetro do modelo. Três parâmetros foram mais explorados para dar ferramentas para uso do modelo com campanhas de marketing: a rejeição às campanhas sucessivas de marketing, a porcentagem de indivíduos que recebem essas campanhas, e a intensidade da informação para os indivíduos que recebem a campanha de marketing. Nesse caso, uma opção segura para a variedade de cenários simulados é fazer uma campanha para 30% da população com uma quantidade média de informação para os indivíduos que recebem a campanha. Uma das vantagens desse modelo é a flexibilidade para adaptar métodos de outros modelos e comparar resultados. Mesmo trabalhos baseados em modelos discretos, como (XIONG et al., 2012; ZHAO et al., 2012; ZHOU; LIU; LI, 2007) podem ser adaptados para o modelo contínuo, trocando os estados por uma certa quantidade de informações nos indivíduos.

No modelo apresentado em (BERGER; NASR, 1998), mostrou-se que o ciclo de vida de compra dos consumidores tende a um patamar máximo quando expostos a longas campanhas de marketing. Esse foi o princípio usado no trabalho para que a informação na população não ficasse subindo indefinidamente no modelo, e chegasse a um máximo, que também depende de outras variáveis consideradas. Inclusive, o modelo sugere que essas outras variáveis não podem ser desconsideradas, pois os melhores resultados são atingidos quando se conhece as características da informação e da população.

Com esses resultados, os próximos passos desse projeto podem ser nas seguintes direções:

- Configurar o modelo contínuo para o uso no caso de propagação de boato numa empresa apresentado na Seção 3.1;
- Uso do modelo apresentado para outras situações apresentadas em outros trabalhos da área para comparação de resultados;
- Uso de dados reais da propagação de informação numa população para ajustar os parâmetros k_i , k_e e k_m para testar diferentes cenários de campanhas de marketing;
- Uso do modelo em conjunto com outro problema, por exemplo, a influência da (falta de) informação para vacinação (SCHIMIT; MONTEIRO, 2011), ou para a compra de determinado produto (SCHIMIT; BARROS JUNIOR; ARAÚJO, 2017);
- Uso do modelo com outra modelagem populacional, como por exemplo, grafos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARGMANN, S.; JORDAN, P. M. A second-sound based, hyperbolic SIR model for high-diffusivity spread. *Physics Letters, Section A: General, Atomic and Solid State Physics*, Elsevier B.V., v. 375, n. 5, p. 898–907, 2011. Citado na pág. 18.
- BERGER, P. D.; NASR, N. I. Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing*, v. 12, n. 1, p. 17–30, 1998. Citado na pág. 19, 38.
- BORDIA, P.; JONES, E.; GALLOIS, C.; CALLAN, V. J.; DIFONZO, N. Management are aliens!: Rumors and stress during organizational change. *Group and Organization Management*, v. 31, n. 5, p. 601–621, 2006. Citado na pág. 17, 38.
- CARVALHO, C.; KLAGGE, N.; MOENCH, E. The persistent effects of a false news shock. *Journal of Empirical Finance*, Elsevier B.V., v. 18, n. 4, p. 597–615, 2011. Citado na pág. 10, 11.
- CASTRO, M. L. A.; CASTRO, R. D. O. Autômatos celulares: implementações de von Neumann, Conway e Wolfram. *Revista de Ciências Exatas e Tecnologia*, v. 3, n. 3, p. 89–106, 2008. Citado na pág. 12, 13, 14, 15.
- CHIERICHETTI, F.; LATTANZI, S.; PANCONESI, A. Rumor spreading in social networks. *Theoretical Computer Science*, Elsevier B.V., v. 412, n. 24, p. 2602–2610, 2011. Citado na pág. 10, 18.
- CHOU, H. I.; TIAN, G. Y.; YIN, X. Takeover rumors: Returns and pricing of rumored targets. *International Review of Financial Analysis*, Elsevier Inc., v. 41, p. 13–27, 2015. Citado na pág. 10.
- DIFONZO, N.; BORDIA, P.; ROSNOW, R. L. Reining in rumors. *Organizational Dynamics*, v. 23, n. 1, p. 47 – 62, 1994. Citado na pág. 11.
- ELIASHBERG, J.; LILIEN, G. L. Mathematical marketing models: Some historical perspectives and future projections. *Handbooks in Operations Research and Management Science*, v. 5, n. C, p. 3–23, 1993. Citado na pág. 18.
- GOLDENBERG, J.; LIBAI, B.; MULLER, E. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Academy of Marketing Science Review*, v. 9, n. 3, p. 1–18, 2001. Citado na pág. 10, 18.
- KAWACHI, K.; SEKI, M.; YOSHIDA, H.; OTAKE, Y.; WARASHINA, K.; UEDA, H. A rumor transmission model with various contact interactions. *Journal of Theoretical Biology*, v. 253, n. 1, p. 55–60, 2008. Citado na pág. 11.
- KERMACK, W. O.; MCKENDRICK, A. G. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 115, n. 772, p. 700–721, 1927. Citado na pág. 10, 15.
- KOSFELD, M. Rumours and markets. *Journal of Mathematical Economics*, v. 41, n. 6, p. 646–664, 2005. Citado na pág. 10, 11, 18.

- LAI, G.; WONG, O. The tie effect on information dissemination: The spread of a commercial rumor in Hong Kong. *Social Networks*, v. 24, n. 1, p. 49–75, 2002. Citado na pág. 11.
- LI, D.; MA, J.; TIAN, Z.; ZHU, H. An evolutionary game for the diffusion of rumor in complex networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 433, p. 51–58, 2015. Citado na pág. 27.
- LIN, J. H.; GUO, Q.; DONG, W. Z.; TANG, L. Y.; LIU, L. G. Identifying the node spreading influence with largest k-core values. *Physics Letters, Section A: General, Atomic and Solid State Physics*, Elsevier B.V., v. 378, n. 45, p. 3279–3284, 2014. Citado na pág. 10.
- MELOTTI, G. Aplicação de Autômatos Celulares em Sistemas Complexos : Um Estudo de Caso em Espalhamento de Epidemias. p. 116, 2009. Citado na pág. 12.
- MONTEIRO, L.; CHIMARA, H.; BERLINCK, J. Big cities: Shelters for contagious diseases. *Ecological Modelling*, v. 197, n. 1, p. 258 – 262, 2006. ISSN 0304-3800. Citado na pág. 13.
- MURRAY, J. D. *Mathematical Biology II - Spatial Models and Biomedical Applications*. [S.l.: s.n.], 2008. 814 p. Citado na pág. 12, 13.
- NEKOVEE, M.; MORENO, Y.; BIANCONI, G.; MARSILI, M. Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, v. 374, n. 1, p. 457–470, 2007. ISSN 03784371. Citado na pág. 10, 18.
- OLINKY, R.; STONE, L. Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission. *Phys. Rev. E*, American Physical Society, v. 70, p. 030902, Sep 2004. Citado na pág. 10.
- PEZZO, M. V.; BECKSTEAD, J. W. A multilevel analysis of rumor transmission: Effects of anxiety and belief in two field experiments. *Basic and Applied Social Psychology*, v. 28, n. 1, p. 91–100, 2006. Citado na pág. 17, 38.
- REULE, D. A dinâmica dos rumores na rede: a web como espaço de propagação de boatos virtuais. 2008. Citado na pág. 9, 10.
- ROSNOW, R. L.; ESPOSITO, J. L.; GIBNEY, L. Factors influencing rumor spreading: replication and extension. *Language & Communication*, v. 8, n. 1, p. 29–42, 1988. Citado na pág. 10, 11, 17.
- SCHIMIT, P. H. T.; BARROS JUNIOR, D. F.; ARAÚJO, S. A. On estimating the impact of information spreading in a consumer market modeled by probabilistic cellular automata and ordinary differential equations. *Simulation*, v. 93, n. 12, p. 1063–1072, 2017. Citado na pág. 10, 11, 20, 39.
- SCHIMIT, P. H. T.; MONTEIRO, L. H. A. On the basic reproduction number and the topological properties of the contact network: An epidemiological study in mainly locally connected cellular automata. *Ecological Modelling*, v. 220, n. 7, p. 1034–1042, 2009. Citado na pág. 6, 13, 15, 16, 20.

- SCHIMIT, P. H. T.; MONTEIRO, L. H. A. A vaccination game based on public health actions and personal decisions. *Ecological Modelling*, Elsevier B.V., v. 222, n. 9, p. 1651–1655, 2011. Citado na pág. 39.
- SILVA, E.; PRADO, K.; CARDOSO, M. V.; SCHIMIT, P. H. T. Propagação de informações a partir de um modelo epidemiológico baseado em autômatos celulares. *Anais do Seminário em Tecnologia da Informação Inteligente - SeTII 2016*, v. 1, p. 282–288, 2016. Citado na pág. 20.
- SU, Q.; HUANG, J.; ZHAO, X. An information propagation model considering incomplete reading behavior in microblog. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 419, p. 55–63, 2015. Citado na pág. 10, 17.
- WEIMAR, J. R. Cellular automata for reaction-diffusion systems. *Parallel Computing*, v. 23, n. 11, p. 1699 – 1715, 1997. Cellular automata. Citado na pág. 12, 13.
- WOLFRAM, S. *Cellular Automata and Complexity: Collected Papers*. [S.l.]: Westview Press, New York, 1994. Citado na pág. 11, 12, 14.
- XIONG, F.; LIU, Y.; ZHANG, Z. J.; ZHU, J.; ZHANG, Y. An information diffusion model based on retweeting mechanism for online social media. *Physics Letters, Section A: General, Atomic and Solid State Physics*, Elsevier B.V., v. 376, n. 30-31, p. 2103–2108, 2012. Citado na pág. 10, 11, 18, 38.
- ZENG, A.; ZHANG, C. J. Ranking spreaders by decomposing complex networks. *Physics Letters, Section A: General, Atomic and Solid State Physics*, Elsevier B.V., v. 377, n. 14, p. 1031–1035, 2013. Citado na pág. 10.
- ZHAO, L.; CUI, H.; QIU, X.; WANG, X.; WANG, J. SIR rumor spreading model in the new media age. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 392, n. 4, p. 995–1003, 2013. Citado na pág. 9, 10, 20.
- ZHAO, L.; WANG, J.; CHEN, Y.; WANG, Q.; CHENG, J.; CUI, H. SIHR rumor spreading model in social networks. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 391, n. 7, p. 2444–2453, 2012. Citado na pág. 6, 10, 16, 38.
- ZHOU, J.; LIU, Z.; LI, B. Influence of network structure on rumor propagation. *Physics Letters, Section A: General, Atomic and Solid State Physics*, v. 368, n. 6, p. 458–463, 2007. Citado na pág. 10, 11, 18, 38.