

# Aplicação de Teoria de Informação na Modelagem de Disseminação de Opiniões em Redes Complexas

**Patrick Oliveira de Paula**



Universidade Federal do ABC

**Título:** Aplicação de Teoria de Informação na Modelagem de Disseminação de Opiniões em Redes Complexas

**Autor:** Patrick Oliveira de Paula

**Orientador:** Prof. Dr. Alejandra Rada

**Banca Examinadora:**

**Catalina Rúa Álvarez**

Universidad de Nariño - Colombia

**Daniel Miranda Machado**

Universidade Federal do ABC

Santo André, Agosto, 2022.

2.1	Diagrama de um canal de transmissão com ruído. Compreende-se que a distorção da informação ocorre bit por bit. . . . .	13
4.1	Evolução temporal da entropia média e tempo de execução em função do número de vértices e tamanho da memória. Foram utilizados os parâmetros $\kappa = 0$ e $\gamma = 0$ para a obtenção das curvas. . . . .	24
4.2	Evolução temporal da entropia média e proximidade ideológica média em função dos parâmetros $\kappa$ e $\gamma$ para um grafo completo. . . . .	25
4.3	Evolução temporal da entropia média para um grafo regular em função dos parâmetros $\kappa$ , $\gamma$ e $\delta$ . . . . .	26
4.4	Evolução temporal da proximidade ideológica média para um grafo regular em função dos parâmetros $\kappa$ , $\gamma$ e $\delta$ . . . . .	26
4.5	Relações entre o número médio de informações efetivamente transmitidas e grau do vértice. . . . .	27
4.6	Relações entre o número médio de informações aceitas e grau do vértice. . . . .	28
4.7	Evolução temporal da entropia média para uma rede de Barabasi-Albert em função dos parâmetros $\kappa$ e $\gamma$ . . . . .	29
4.8	Gráfico de calor da evolução temporal da distribuição de probabilidade das informações no sistema. . . . .	30
4.9	Curvas da evolução temporal da distribuição de probabilidade das informações no sistema. . . . .	31
4.10	Evolução temporal da entropia média para uma rede de Barabasi-Albert em função dos parâmetros $\kappa$ e $\gamma$ , enfatizando diferentes intervalos de tempo de simulação. . . . .	32
4.11	Evolução temporal da proximidade média para um grafo de Barabasi-Albert em função dos parâmetros $\kappa$ e $\beta$ . . . . .	33

<b>1</b>	<b>Introdução</b>	<b>7</b>
1.1	Organização do Texto . . . . .	9
<b>2</b>	<b>Fundamentos Teóricos</b>	<b>10</b>
2.1	Redes Aleatórias e Livre de Escala . . . . .	10
2.2	Teoria da informação . . . . .	12
2.2.1	Entropia . . . . .	13
2.2.2	Medidas de Divergência de Informação . . . . .	16
<b>3</b>	<b>Modelo Proposto</b>	<b>19</b>
3.1	Definições . . . . .	19
3.2	Algoritmo . . . . .	22
<b>4</b>	<b>Resultados</b>	<b>23</b>
4.1	Análise das Simulações . . . . .	24
4.1.1	Grafos Completos e Regulares . . . . .	25
4.1.2	Rede de Barabasi-Albert . . . . .	27
4.1.3	Comparação com a Literatura . . . . .	34
4.2	Conclusão . . . . .	35
4.2.1	Trabalhos Futuros . . . . .	38
<b>5</b>	<b>Apêndice</b>	<b>39</b>
5.1	Simulação de Variáveis Aleatórias . . . . .	39
5.2	Algoritmos . . . . .	40
	Referências . . . . .	42

Em memória à Giovanna S., cuja inspiração deu início ao que vivi e realizei até então, acompanhado sempre das boas lembranças, que mantiveram acesos o meu ímpeto e entusiasmo ao longo deste percurso.

À minha família, em especial à minha mãe e avó, pelo suporte e cuidados por toda a minha vida, e pelo privilégio inestimável de poder dedicar-me integralmente aos meus estudos.

Às minhas orientadoras, Alejandra Rada e Catalina Álvarez, pela paciência, atenção e incentivo, pelas ótimas sugestões, correções e críticas durante a execução deste trabalho.

Aos excelentes professores que tive durante a minha graduação, especialmente os professores Marcus Marrocos, Ana Boero e Daniel Miranda, pelos cursos mais formativos que tive a oportunidade de assistir.

O compartilhamento de opiniões e a troca de informações são aspectos essenciais da interação social humana. A disseminação de informações e rumores pode ter efeitos significativos na sociedade e nas suas estruturas político-econômicas. A internet tornou possível o surgimento de redes sociais digitais massivas, aumentando o alcance e a influência da propagação de informações, introduzindo, para além dos benefícios evidentes da facilidade de troca de ideias, desafios como o fenômeno das *fake news*. Dessa forma, é de grande interesse teórico e prático o desenvolvimento de modelos mais realistas de disseminação de informações em rede com respeito aos fatores cognitivos atuantes na troca de opiniões. Partindo das primeiras propostas baseadas em modelos epidemiológicos, os modelos mais modernos buscam introduzir progressos de diversas áreas como as ciências cognitivas, a economia comportamental e a teoria dos jogos para a modelagem da dinâmica de interações sociais e propagação de informações. Nesse sentido, este trabalho propõe um modelo que utiliza o conceito de Entropia de Informação e a medida de divergência de Jensen-Shannon para simular computacionalmente a disseminação de opiniões levando em conta a memória individual, a proximidade ideológica entre indivíduos e efeitos de distorção potencial de informações devido à recuperação imperfeita de informações da memória.

**Palavras Chaves:** Entropia, Redes Complexas, Propagação, Rumores, Teoria de Informação

Information exchange and opinion sharing are essential aspects of human social interaction. The dissemination of informations and rumours can have significant effects on society and its political and economical structures. The internet made possible the emergence of massive digital social networks, increasing the reach and influence of propagated informations, introducing, in addition to the evident benefits of facilitated exchange of ideas, challenges such as the *fake news* phenomenon. In this sense, it is of great theoretical and practical interest the development of models for information dissemination in networks which seeks more realism regarding the cognitive factors relevant to opinion exchanges. Starting from the first proposals based on epidemiological models, modern models seek to introduce the progress made in several areas such as the cognitive sciences, behavioural economics and game theory for modelling of the dynamics of social interactions and information propagation. In this sense, this work proposes a model that makes use of the Information Entropy concept and the Jensen-Shannon information divergence function to simulate computationally the dissemination of opinions taking into account individual memory, the ideological proximity between individuals and effects of potential distortion due to imperfect retrieval of information from memory.

**Keywords:** Entropy, Complex Networks, Rumour, Propagation, Information Theory

A disseminação de rumores e opiniões é um elemento importante na dinâmica social de seres humanos, resultando, em alguns casos, em impactos significativos na sociedade. Esse fluxo de informações pode influenciar a opinião pública sobre conjunturas políticas (Galam, 2003), impactar mercados financeiros (Kosfeld, 2005; Kimmel, 2004) e criar tensões sociais e ansiedade generalizada desnecessária (Grein et al., 2000; Thomas, 2007) sobre questões públicas possivelmente falsas. Mais recentemente, a disseminação de falsos rumores tem dificultado o desafio de solucionar a pandemia de Covid-19, aumentando a hesitação pública diante das novas vacinas (Carrieri et al., 2019) e favorecendo a adoção de intervenções sem respaldo científico quanto sua eficácia para o tratamento da Covid-19 (Touret & de Lamballerie, 2020; Ferner & Aronson, 2020). O conteúdo de informações disseminadas pode variar de simples fofoca à esquemas complexos de marketing viral (Richardson & Domingos, 2002).

O surgimento da internet e das redes sociais facilitou a troca de informações e permitiu que qualquer indivíduo tivesse a oportunidade de compartilhar as suas opiniões para um número considerável de pessoas, o que até então era um poder limitado aos grandes veículos de mídia. Os benefícios sociais são evidentes, entretanto a facilidade de disseminação de informações fragiliza a sociedade diante de boatos espúrios e estratégias deliberadas de desinformação que têm, como consequência, a polarização contraproducente do debate público (Lazer et al., 2018). Um modelo de dinâmica de opiniões mais realista com respeito aos fatores cognitivos que afetam a troca de informações tem, portanto, relevância teórica e prática, por exemplo, contribuindo para a formulação de estratégias de controle de disseminação de boatos em redes sociais, de divulgação eficiente de informações públicas relevantes, ou a compreensão de aspectos mais gerais acerca da dinâmica social de disseminação de informações.

É possível fazer a analogia de opiniões como “infecções da mente”, apresentando semelhanças com a dinâmica de epidemias (Nekovee et al., 2007). De fato, esta similaridade motivou os primeiros modelos de propagação de opiniões, baseados em modelos epidemiológicos clássicos: o modelo SIS (*Susceptible and Infected*) e o modelo SIR (*Susceptible, Infected and Recovered*). O Modelo DK, introduzido em (Daley & Kendall, 1965), juntamente com sua variação, o Modelo MT (Maki-Thompson, 1973), fundamentaram uma série de estudos de propagação de rumores inspirados por modelos epidemiológicos (Lefevre & Picard, 1994; Sudbury, 1985; Zhang & Zhang, 2009; Noymer, 2001). No modelo DK, uma popula-

ção fechada e homogênea é dividida em três grupos: um grupo ignorante de um dado rumor, aqueles que ouviram e ativamente disseminaram, e aqueles que ouviram mas optaram por não disseminá-lo. Estes grupos são chamados, respectivamente, de *ignorantes*, *disseminadores* e *abafadores*. O rumor é propagado em contatos par-a-par entre disseminadores e o restante da população. No caso em que um indivíduo ignorante é “contaminado”, este se torna um disseminador; caso o contato se dê com outro disseminador ou um abafador, há a chance de que os envolvidos se tornem (ou se mantenham) abafadores.

Um limite desta classe de modelos é o fato de que ao assumir que a distribuição de graus dos vértices é homogênea, a topologia subjacente da rede social não é levada em consideração, ou é utilizada alguma topologia simplificada, como um grafo regular (Sudbury, 1985; Zhang & Zhang, 2009). Isso faz com que esta classe de modelos seja inadequada para modelar a disseminação de opiniões em redes massivas de interação social, como a Internet, que envolve dezenas de milhares ou mesmo milhões de vértices. Estas redes apresentam complexos padrões de conectividade, envolvendo, por exemplo, a presença de *hubs*, i.e. indivíduos com um grande número de conexões sociais (Csányi & Szendrői, 2004). O primeiro modelo a considerar uma topologia não-trivial foi introduzido por Zanette (Zanette, 2002), estudando um modelo de disseminação epidêmica de opiniões em redes de Watts-Strogatz (Watts & Strogatz, 1998). Em seguida, surgiram versões do modelo MT aplicado em redes livre-de-escala (Moreno et al., 2004) e o estudo da dinâmica de disseminação enfatizando aspectos estruturais como a influência de hubs (Boccaletti et al., 2006).

A evolução da opinião pública de uma população também é um fator de interesse no estudo de propagação de rumores e troca de informações. No modelo apresentado em (Hegselmann et al., 2002), intervalos de confiança são utilizados para determinar as opiniões que os indivíduos levam em consideração; as opiniões fora destes intervalos são descartadas. Um modelo de decisão por maioria foi proposto por Chen *et al.* (Chen & Redner, 2005), onde a probabilidade de aceitação de uma opinião por um indivíduo é proporcional ao número de indivíduos em sua vizinhança que aderem àquela opinião. A integração de mecanismos de persuasão baseados em Teoria dos Jogos, proposto por Mare *et al.* (Di Mare & Latora, 2007), faz parte do esforço importante de incluir *insights* acerca do comportamento social na construção destes modelos. Ademais, é importante considerar o progresso das ciências cognitivas no entendimento dos mecanismos envolvidos na absorção e troca de informações por seres humanos. Sabe-se que efeitos de distorção no processo de difusão de opiniões e a disseminação proposital de informações distorcidas são aspectos influentes na transmissão de rumores. Quando uma informação é disseminada, o estado psicológico do indivíduo, aliado à falhas na recuperação de memórias, entre outros fatores, podem acarretar na distorção da informação transmitida. Nessa linha, Wang *et al.* (Wang et al., 2017) introduz um modelo que inclui efeitos de memória e propensão de distorção de informação, utilizando para isso o conceito de Entropia de Informação.

Propomos, partindo do desenvolvimento descrito, um modelo de propagação de opiniões que considera a capacidade de memória dos indivíduos. As informações são modeladas

como códigos binários de tamanho especificado. A presença de uma memória de informações em cada indivíduo permite a aplicação de Entropia de Informação como uma medida de heterogeneidade de opiniões, a partir da qual pode-se incluir os efeitos de distorção decorrentes de recuperação imperfeita de informações da memória. Além disso, utilizamos uma medida de divergência de informação, em particular a medida de Jensen-Shannon, para quantificar a proximidade ideológica entre indivíduos, identificada como o peso das conexões da rede. Essa proximidade passa a ser considerada no cálculo da probabilidade de aceitação de informações.

## 1.1 Organização do Texto

No que segue, será feita uma breve introdução teórica de alguns tópicos fundamentais para a descrição do modelo proposto. A fim de tornar esta introdução concisa, apenas os principais conceitos serão apresentados, omitindo, por exemplo, conceitos básicos de Teoria dos Grafos e de Probabilidade.

Em seguida, será apresentada a proposta do modelo. Os algoritmos principais para o seu funcionamento, considerando também um método para simulação de variáveis aleatórias, estarão disponíveis no Apêndice.

Por fim, será feita uma análise dos resultados obtidos com a simulação do modelo através principalmente do estudo das curvas de entropia média do sistema e da proximidade média entre os indivíduos.

### 2.1 Redes Aleatórias e Livre de Escala

Uma *rede aleatória* é um grafo  $G$  simples e conexo em que os pares de vértices são conectados aleatoriamente com alguma probabilidade pré-definida. Ou seja, inicia-se com uma coleção de  $n$  vértices e para cada uma das  $\binom{n}{2}$  arestas possíveis, adicionamos uma aresta  $\langle u, v \rangle$  com probabilidade  $p_{uv}$ .

Os conceitos fundamentais de grafos aleatórios foram inicialmente elaborados por Paul Erdős e Alfréd Rényi (Erdős, 1959) que introduziram o modelo de rede aleatória Erdős-Rényi, ou ER.

**Definição 2.1** Um grafo aleatório de Erdős-Rényi com  $n$  vértices, também definido como um grafo aleatório ER, é um grafo não-direcionado  $G_{n,p}$  em que cada par de vértices distintos é conectado por uma aresta com probabilidade  $p$ .

A notação  $ER(n, p)$  é utilizada para denotar um conjunto de grafos construídos tal como a definição, tendo em vista que pode haver dois grafos  $G_{n,p}^1$  e  $G_{n,p}^2$  com parâmetros iguais, tais que uma aresta  $e = (u, v) \in E(G_{n,p}^1)$  mas  $e \notin E(G_{n,p}^2)$ .

O grau de um vértice pode ser tratado como uma variável aleatória  $\delta$ , e assim podemos falar sobre a probabilidade de que um dado vértice tenha grau igual a  $k$ , escrevendo  $P(\delta = k)$ . Para grafos  $ER(n, p)$ , tomando um vértice  $u$  qualquer, seja  $P(\delta(u) = k)$  a probabilidade de que o grau de  $u$  seja  $k$ . Existem  $\binom{n-1}{k}$  possibilidades de escolha de  $k$  diferentes vizinhos para  $u$  e, sabendo que a probabilidade de que  $u$  seja adjacente a exatamente  $k$  vértices é  $p^k \cdot (1-p)^{n-1-k}$ , temos que

$$P(\delta(u) = k) = \binom{n-1}{k} p^k (1-p)^{n-1-k},$$

então, pela generalidade de  $u$ , a distribuição de graus em um grafo  $ER(n, p)$  é binomial, com valor esperado dado por

$$\bar{\delta} := E[\delta] := \sum_{k=1}^{n-1} k \cdot P(\delta = k) = p(n-1).$$

A construção de redes aleatórias com diferentes propriedades estruturais pode ser feita pela modificação no mecanismo de atribuição de conexões entre os vértices. Uma topolo-

gia alternativa são as denominadas *redes livre de escala*, popularizadas pelos trabalhos de Albert-László Barabási e seu então estudante de doutorado Réka Albert, caracterizadas por possuírem em sua estrutura um pequeno conjunto de vértices com elevado grau de conexões, chamados de *hubs*, uma propriedade comum às estruturas de redes empíricas.

Dizemos que uma rede é livre de escala se o grau dos vértices é uma variável aleatória  $K$  cuja distribuição segue uma *Lei de Potência*. Grosso modo, isso significa que a probabilidade de que um nodo tenha grau  $k$  é tal que  $P(\delta = k) = P(k) \propto k^{-\alpha}$ , em que  $\alpha > 1$  é um *coeficiente de escala*. Para compreender a razão do título *livre de escala*, considere que uma função  $f(x)$  é dita livre de escala se, dado um parâmetro  $b$ , ela satisfaz a propriedade  $f(bx) = C(b) \cdot f(x)$ , em que  $C(b)$  é uma constante dependente apenas de  $b$ . Entende-se que o comportamento geral da função  $f(x)$  não se altera quando tomamos valores de  $x$  diferentes apenas por um fator  $b$ . Nota-se que distribuições que seguem uma lei de potência obedece essa propriedade, e para muitas redes empíricas, tem-se que  $2 < \alpha < 3$  (Van Steen, 2010).

Redes livres de escala são distintas no sentido de que podem ser construídas apenas por um processo de crescimento com preferências de vinculação (Dorogovtsev & Mendes, 2002). Observa-se, pois, a forma como novos vértices tendem a se conectar com os vértices pré-existentes. Barabási e Albert foram os primeiros a elaborar um procedimento de construção de redes livres de escala; no que segue, considere um grafo aleatório de Erdős-Rényi  $G_0$  com um conjunto de vértices  $V_0$  de tamanho  $n_0$  e um conjunto de arestas  $E_0$ .

---

**Algoritmo 1:** Algoritmo de construção de um grafo livre de escala
 

---

**Dados:**  $n > 0, m > 0$

$s \leftarrow 1$ ;

**enquanto**  $s < n$  **faça**

    Seja  $v_s$  um novo vértice;

    Faça  $V_s \leftarrow V_{s-1} \cup \{v_s\}$ ;

**para**  $u \in V_{s-1}$  **faça**

$$P(u) = \frac{\delta(u)}{\sum_{w \in V_{s-1}} \delta(w)};$$

**fim**

    Faça  $E_s \leftarrow V_{s-1} \cup \{\langle v_s, u_1 \rangle, \langle v_s, u_2 \rangle, \dots, \langle v_s, u_m \rangle\}$ , onde cada vértice  $u_i \in V_{s-1}$  é escolhido com probabilidade  $P(u_i)$ ;

    Faça  $s \leftarrow s + 1$ ;

**fim**

---

O grafo resultante é chamado de *Grafo Aleatório de Barabási-Albert*, ou *Grafo BA*. Denotamos o conjunto de grafos BA de parâmetros  $n, n_0, m$  por  $BA(n, n_0, m)$ . Observe que em cada etapa da construção, a probabilidade de escolha de um vértice  $u$  para a criação de uma nova conexão é proporcional ao grau de  $u$  naquele momento. Pode-se mostrar que a probabilidade  $P(k)$  de que um vértice  $v$  qualquer em um grafo BA tenha grau  $k$  é proporcional a  $k^{-3}$

(Barabási & Albert, 1999). Mais precisamente, temos (Van Steen, 2010)

**Teorema 2.1** *Para qualquer grafo  $G \in BA(n, n_0, m)$ , a probabilidade de que um vértice  $v \in V(G)$  tenha grau  $k \geq m$  é dada por*

$$\mathbb{P}[k] = \frac{2m(m+1)}{k(k+1)(k+2)} \propto \frac{1}{k^3}.$$

## 2.2 Teoria da informação

A perspectiva deste trabalho enfatiza a evolução das opiniões, segundo algumas métricas, conforme os indivíduos interagem entre si a partir de regras definidas. Desta perspectiva, pode-se pensar que a rede estabelece um sistema de comunicação com ruído, tornando conveniente a aplicação de conceitos e técnicas advindas da Teoria da Informação para a elaboração e análise do modelo.

A Teoria da Informação foi introduzida por Claude Shannon no seu trabalho *A Mathematical Theory of Communication* (Shannon, 1948), motivado primeiramente pelo problema de compressão e transmissão de informações em canais de comunicação imperfeitos, estudando as condições que permitiriam preservar as informações transmitidas e reduzir tanto quanto o possível o custo de transmissão (i.e. compressões eficientes). Em um canal de comunicação imperfeito, considera-se uma fonte de informação transmitindo códigos digitais, i.e. códigos binários, para um receptor. Ao longo da transmissão, o código pode ser contaminado por uma fonte de ruído, de tal modo que a informação recebida difere do código original. O modelo mais simples de transmissão com ruído de informação entre dois entes é denominado *canal binário simétrico*. Indicando um bit emitido por  $x$  e o bit efetivamente recebido por  $y$ , sendo  $f$  a probabilidade de interferência por ruído, à informação recebida estão associados as probabilidades

$$P(y = 0|x = 0) = 1 - f;$$

$$P(y = 1|x = 0) = f;$$

$$P(y = 0|x = 1) = f;$$

$$P(y = 1|x = 1) = 1 - f.$$

O principal conceito introduzido juntamente com a Teoria da Informação é o conceito de *entropia de informação*, sendo uma forma de quantificar o conteúdo de informação de um dado código, medido em *bits*<sup>1</sup>. Em torno deste conceito são obtidos resultados relativos à, por exemplo, capacidade de transmissão de um canal, limites de compressão sem perda de informação, e outros mecanismos de medição baseados em entropia, como a Informação

<sup>1</sup>Não confundir com a variável *bit* que assume os valores 0 ou 1.

Mútua, que já possuem um caráter mais abstrato e que podem ser utilizadas para estudar e comparar distribuições de probabilidade de forma geral.

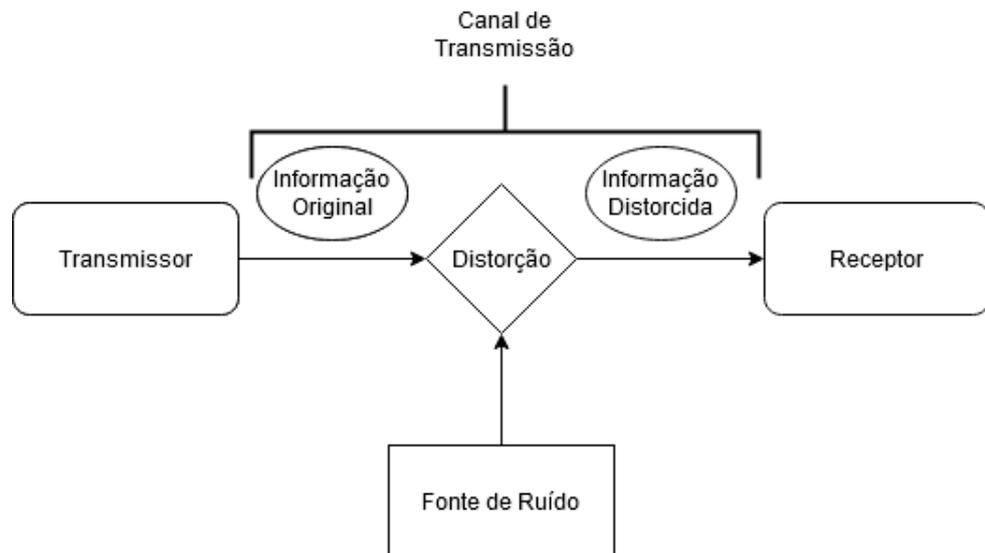


Figura 2.1: Diagrama de um canal de transmissão com ruído. Compreende-se que a distorção da informação ocorre bit por bit.

No que segue, as definições e medidas pertinentes para a apresentação do modelo serão introduzidas. Será dada ênfase nas definições das funções de entropia e algumas medidas de divergência.

### 2.2.1 Entropia

Considere um *alfabeto*  $\mathcal{A}$  como conjunto de todas as sequências de códigos binários de tamanho  $m$ . Associamos a este alfabeto a uma distribuição de probabilidade  $P(X)$  onde  $X$  é uma variável aleatória definida sobre as observações tomadas de  $\mathcal{A}$ . Entende-se que  $X$  representa as informações emitidas tendo como repositório o alfabeto  $\mathcal{A}$ . No que segue, faremos a simplificação  $P(X = x) = P(x)$ .

Se  $x \in \mathcal{A}$  é uma informação qualquer, a *quantidade de informação* - segundo Shannon - de  $x$ , medida em *bits*, pode ser calculada por

$$h(x) = \log_2 \frac{1}{P(x)}.$$

Trataremos apenas de códigos binários, portando emitiremos a base do logaritmo, ainda que seja possível definir esta quantidade de informação tomando valores distintos na base. A *entropia de Shannon*, ou simplesmente a entropia de um emissor, tem a seguinte definição:

**Definição 2.2** A entropia de Shannon  $H(X)$  de uma variável aleatória discreta  $X$  é definida como

$$H(X) = - \sum_{x \in \mathcal{A}} P(x) \log P(x) = \sum_{x \in \mathcal{A}} P(x) \log \frac{1}{P(x)},$$

onde dado que

$$\lim_{\theta \rightarrow 0^+} \theta \log \frac{1}{\theta} = 0.$$

convencionamos que para  $P(x) = 0$  vale que  $P(x) \times \log \frac{1}{P(x)} = 0$ .

Ademais,  $H(X)$  admite as propriedades:

1.  $H(X) \geq 0$  com igualdade se e somente se  $P(x) = 1$  para algum  $x \in \mathcal{A}$ , ou  $P(x) = 0$  pela convenção.
2. A entropia é máxima quando a distribuição  $P(X)$  é uniforme, i.e. vale que  $H(X) \leq \log(|\mathcal{A}|)$ , com igualdade se e somente se  $P(x) = \frac{1}{|\mathcal{A}|}$  para todo  $x \in \mathcal{A}$ .

Em relação à primeira propriedade, note que, para qualquer  $x \in \mathcal{A}$ , temos que  $P(x) \in [0, 1]$  e, portanto,  $\log 1/P(x) \geq 0$ , dessa forma teremos  $H(X) = 0$  se, e somente se,  $P(x) \log 1/P(x) = 0$  para todo  $x \in \mathcal{A}$ , o que ocorre somente quando  $P(x) = 1$  para algum  $x \in \mathcal{A}$  e  $P(y) = 0$  para qualquer outro  $y \neq x$ .

Para provar a segunda propriedade faremos uso da Desigualdade de Jensen (Rudin, 1987),

**Teorema 2.2** Seja  $\mu$  uma medida positiva definida em uma  $\sigma$ -álgebra  $\mathcal{F}$  sobre um conjunto  $\Omega$ , tal que  $\mu(\Omega) = 1$ . Se  $f$  é uma função real em  $L^1(\mu)$ , e  $a < f(x) < b$  para todo  $x \in \omega$ , e se  $\phi$  é uma função convexa em  $(a, b)$ , então

$$\phi\left(\int_{\Omega} f d\mu\right) \leq \int_{\Omega} (\phi \circ f) d\mu.$$

Se  $\Omega$  é um conjunto finito de pontos  $x_1, \dots, x_n$ , e colocamos  $\mu(\{x_i\}) = 1/n$ , então obtemos a desigualdade

$$\phi\left(\frac{1}{n}(f(x_1) + \dots + f(x_n))\right) \leq \frac{1}{n}(\phi(f(x_1)) + \dots + \phi(f(x_n))).$$

No caso em que  $\phi$  é uma função côncava, invertemos a desigualdade. Podemos considerar a função  $\phi(y) = y \log 1/y$ , côncava no intervalo  $(0, 1)$ , e para cada  $x_i \in \mathcal{A}$  a função  $f(x_i) = P(x_i)$ , o que nos permite escrever

$$\frac{1}{|\mathcal{A}|} \log |\mathcal{A}| \geq \frac{1}{|\mathcal{A}|} \sum_{x \in \mathcal{A}} P(x) \log \frac{1}{P(x)},$$

e tomando a distribuição uniforme  $P(x) = 1/|\mathcal{A}|$  obtemos a igualdade.

Se  $g(X) : \mathbb{R} \rightarrow \mathbb{R}$  é uma função qualquer de uma variável aleatória  $X$ , então o valor esperado  $E[g(X)]$  é dado por

$$E[g(X)] = \sum_x p(x)g(x),$$

o que nos permite escrever a definição de  $H(X)$  simplesmente por

$$H(X) = -E[\log p(X)] = E[h(X)].$$

Seguindo essa notação, definimos a entropia *conjunta* dada de forma análoga à definição básica e que pode ser estendida para várias variáveis. Ela pode ser interpretada no contexto em que se observa certas informações em sequência, e busca-se entender o quanto a incerteza de uma variável é reduzida pela observação de uma outra, isto é, sua dependência.

**Definição 2.3** A entropia conjunta  $H(X, Y)$  de um par de variáveis aleatórias  $X$  e  $Y$  é definida como

$$H(X, Y) = E \left[ \log \frac{1}{p(X, Y)} \right].$$

Se  $\mathcal{A}$  é um alfabeto sobre o qual são definidas duas distribuições de probabilidade  $P$  e  $Q$ , relativas às variáveis aleatórias  $X$  e  $Y$ , pode nos interessar quantificar a informação compartilhada e medir suas diferenças. A entropia  $H(X)$  pode ser escrita por  $H(P)$  sem risco de confusão, podendo ser mais pertinente quando se quer focar a atenção nas distribuições de probabilidade.

Introduzimos a definição da medida de *Informação Mútua*.

**Definição 2.4** Para duas variáveis aleatórias  $X$  e  $Y$ , a informação mútua entre  $X$  e  $Y$  é definida como

$$I(X; Y) = \sum_{x,y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)} = E \left[ \log \frac{P(X, Y)}{P(X)P(Y)} \right].$$

A proposição a seguir (Yeung, 2008) permite relacionar a medida de informação mútua e as definições de entropia:

**Proposição 2.1** *Se as seguintes entropias são finitas, vale a igualdade:*

$$I(X; Y) = H(X) + H(Y) - H(X, Y).$$

Observa-se que  $I(X; Y)$  é simétrica,  $I(X; Y) \geq 0$ , e que  $I(X; X) = H(X)$ , de modo que pode-se referir a  $H(X)$  como *auto-informação* de  $X$ .  $I(X; Y)$  é uma medida de informação compartilhada entre  $X$  e  $Y$ , referindo-se à interceptação dos alfabetos atrelados a  $X$  e  $Y$ .

## 2.2.2 Medidas de Divergência de Informação

Apesar da utilidade da medida de informação mútua como meio de comparação de similaridade entre distribuições - definidas sobre certos conjuntos -, sua definição depende da definição da distribuição conjunta (ou condicional), o que não é imediato a depender do problema modelado, ou ainda depende que as variáveis tratadas não sejam independentes, caso contrário tem-se que  $I(X; Y) = 0$ . Mesmo em um caso de independência pode ser necessário introduzir uma medida de distância entre distribuições de probabilidade definidas sobre um mesmo conjunto.

Introduzimos a *divergência informacional* ou *informação relativa*.

**Definição 2.5** A divergência informacional entre duas distribuições de probabilidade  $P$  e  $Q$  definidas sobre um alfabeto comum  $\mathcal{A}$  é definida como

$$D(P; Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} = E_P \left[ \log \frac{P(X)}{Q(X)} \right],$$

onde  $E_P$  denota o valor esperado relativo à distribuição  $P$ . Por convenção, consideramos que  $c \log \frac{c}{0} = \infty$  se  $c > 0$ . Essa propriedade que convencionamos permite afirmar que  $S_Q \subset S_P$ , em que  $S_Q$  é o suporte de  $Q$  (análogo para  $P$ ), o que significa dizer que para todo  $x$  tal que  $Q(x) = 0$  temos também que  $P(x) = 0$ .

A medida  $D(P; Q)$ , também chamada de *Distância de Kullback-Leibler* (Kullback & Leibler, 1951), não é uma distância propriamente dita ao passo que não satisfaz a desigualdade triangular e não é simétrica em  $P$  e  $Q$ . Apesar disso, vale o teorema (Yeung, 2008).

**Teorema 2.3** Para quaisquer duas distribuições de probabilidade  $P$  e  $Q$  em um alfabeto comum  $\mathcal{A}$ ,

$$D(P; Q) \geq 0,$$

com igualdade se e somente se  $P = Q$

A propriedade acima valida o uso da medida  $D(P; Q)$  como mecanismo comparativo. A assimetria pode ser trabalhada definindo uma segunda medida, a *Divergência J*, a partir de  $D(\cdot; \cdot)$ .

**Definição 2.6** Dadas duas distribuições de probabilidade  $P$  e  $Q$  definidas em um alfabeto comum  $\mathcal{A}$ , a Divergência  $J$  é definida como

$$\begin{aligned} J(P; Q) &= D(P; Q) + D(Q; P) \\ &= \sum_x (P(x) - Q(x)) \log \frac{P(x)}{Q(x)}, \end{aligned}$$

Naturalmente,  $J(\cdot; \cdot)$  e  $D(\cdot; \cdot)$  compartilham muitas de suas propriedades, e como consequência de que a definição de  $D(\cdot; \cdot)$  exige que  $P$  seja absolutamente contínua em relação a  $Q$ , aqui temos que  $P$  e  $Q$  devem ser absolutamente contínuas em relação uma a outra a fim de que  $J(\cdot; \cdot)$  esteja bem definida - em termos discretos, seus conjuntos de suporte devem ser idênticos -, o que introduz alguma dificuldade para o seu uso.

Os problemas anteriormente descritos podem ser em certa medida transpostos pela introdução de novas medidas, como as medidas  $K$  e  $L$  dadas por, respectivamente,

$$K(P; Q) = \sum_x P(x) \log \frac{P(x)}{\frac{1}{2}P(x) + \frac{1}{2}Q(x)},$$

e

$$\begin{aligned} L(P; Q) &= K(P; Q) + K(Q; P) \\ &= H\left(\frac{P+Q}{2}\right) - \frac{H(P) + H(Q)}{2}, \end{aligned}$$

ao passo que  $K(\cdot; \cdot)$  não é uma medida simétrica. A simetria de uma medida de divergência é uma propriedade relevante para o propósito deste trabalho tendo em vista que queremos medir a similaridade entre dois indivíduos  $u$  e  $v$ , não importando, portanto, a ordem da comparação. Por análise pode-se verificar que  $K$  e  $L$  são medidas bem comportadas ao longo do seu domínio (Lin, 1991) (i.e.  $K(P; Q) < +\infty$  e  $L(P; Q) < +\infty$  enquanto que  $D(\cdot; \cdot)$  e  $J(\cdot; \cdot)$  divergem para  $+\infty$  conforme alguma probabilidade tende a zero. Mais especificamente, é possível demonstrar (Lin, 1991) que

$$K(P; Q) \leq 1;$$

$$L(P; Q) \leq 2.$$

Na realidade, a medida  $L$  assim definida é um caso particular de uma medida de divergência mais geral, a *divergência generalizada de Jensen-Shannon*, e passível de aplicação em conjuntos finitos arbitrários de distribuições de probabilidade sobre um mesmo alfabeto.

**Definição 2.7** Sejam  $P_1, P_2, \dots, P_n$  distribuições de probabilidades, e  $\pi_1, \pi_2, \dots, \pi_n$  pesos escolhidos satisfazendo  $\sum_i \pi_i = 1$ . A divergência generalizada de Jensen-Shannon pode ser definida como

$$JSD_{\pi}(P_1, P_2, \dots, P_n) = H\left(\sum_{i=1}^n \pi_i P_i\right) - \sum_{i=1}^n \pi_i H(P_i).$$

Assim, obtêm-se a medida  $L$  fazendo  $n = 2$  e  $\pi_1 = \pi_2 = 0.5$ . No caso geral, valem as desigualdades (Lin, 1991)

$$0 \leq JSD(P_1, P_2, \dots, P_n) \leq \log_2 n,$$

de modo que para  $n = 2$  temos que  $0 \leq JSD_\pi(P; Q) \leq 1$ . Estas propriedades tornam a medida de divergência de Jensen-Shannon como uma função ideal para a comparação de distribuições de probabilidade.

## 3.1 Definições

Os indivíduos serão modelados como vértices em um grafo não-orientado livre de escala de Barabási-Albert de tamanho  $N$ , com pesos nas arestas. Cada vértice  $u$  é caracterizado por uma tripla  $(X_u, L_u, P_u)$ , em que  $L_u$  é uma memória de tamanho  $\mu$  de códigos binários de tamanho  $m$ , sendo possível, assim, representar  $2^m$  diferentes tipos de informações - denotaremos o conjunto de informações possíveis - alfabeto - como  $\mathcal{A}$ . Calculando as frequências empíricas das informações na memória  $L_u$  pode-se definir uma distribuição de probabilidades  $P_u$  sobre o alfabeto, associada à variável aleatória  $X_u$  que modela a seleção de uma informação  $x \in \mathcal{A}$  por  $u$  para disseminação, i.e.

$$P_u(x) = P_u(X_u = x) = f_x.$$

A frequência  $f_x$  pode ser vista não somente como uma probabilidade de seleção da informação  $x$  mas também como uma *medida de saliência*. De posse de uma medida de saliência das informações na memória dos indivíduos, torna-se possível analisar efeitos de conformidade e realizar comparações qualitativas entre indivíduos, formação de comunidades em função da informação mais saliente, etc. Além disso, para cada indivíduo  $u$  calculamos empiricamente a entropia de Shannon,

$$H(P_u) = - \sum_{x \in L_u} f_x \log_2 f_x,$$

em que consideramos o logaritmo na base 2. Por conveniência, algumas vezes utilizaremos a notação  $H(P_u) = H_u$ . A entropia cresce conforme a distribuição de diferentes tipos de informação se torna mais uniforme (i.e.  $f_x \rightarrow \frac{1}{|\mathcal{A}|} = \frac{1}{2^m}$ ) na memória do indivíduo, e diminui conforme um tipo de informação passa a prevalecer ( $f_x \rightarrow 1$  para algum  $x$ ).

Segue, naturalmente, uma medida geral de entropia no grafo, a *entropia média da população* (a qual pode ser definida para um subgrafo  $Q \subset G$ ),  $\bar{H}$ ,

$$\bar{H} = \frac{1}{|V(G)|} \sum_{u \in V(G)} H_u.$$

A entropia oferece uma medida de ruído com a qual o indivíduo precisa lidar ao interagir com diferentes tipos de informação, e nesse sentido é possível incluir uma tendência de

distorção da informação por transmissão imperfeita. Quanto maior a entropia, mais imperfeita é a memória do indivíduo, incorrendo em maior frequência de erros de transmissão. A *probabilidade de distorção* é definida como

$$\delta_u = \frac{1}{\exp \left[ \left( \frac{H_{\max} - H_u}{H_{\max}} \right) \cdot \kappa \right] + 1},$$

onde  $\kappa \geq 0$  é um *fator de conservação*, sendo um parâmetro de controle da tendência geral de distorção de informação e essa tendência será maior para valores grandes de  $\kappa$ , e vice-versa. O valor  $H_{\max}$  é a entropia máxima, para  $f_x = 1/2^m$ . Conforme  $H_u \rightarrow H_{\max}$ , a probabilidade de distorção atinge um valor máximo de  $1/2$ , enquanto que conforme  $H_u \rightarrow 0$ , a probabilidade tende ao valor mínimo  $1/(\exp \kappa + 1)$ .

Neste modelo, entendemos que a transmissão de informações entre vértices define um canal de transmissão com ruído, ou seja, cada código selecionado é transmitido bit por bit, incorrendo uma probabilidade de distorção na transmissão de cada bit. A distorção do canal pode ser construída a fim de levar em consideração diferentes fatores de interação entre indivíduos que modificaria a informação transmitida. Neste caso, em que os efeitos de memória imperfeita são simulados, a probabilidade  $\delta_u$  introduz a chance de uma mutação aleatória em cada bit transmitido por  $u$ . Dessa forma, a transmissão de informações por um indivíduo neutro configura um *canal binário simétrico*. Para um indivíduo  $u$  as probabilidades de distorção são dadas por

$$P(y = 0|x = 0) = 1 - \delta_u;$$

$$P(y = 1|x = 0) = \delta_u;$$

$$P(y = 0|x = 1) = \delta_u;$$

$$P(y = 1|x = 1) = 1 - \delta_u.$$

A atribuição dos pesos nas arestas buscará modelar a proximidade de pensamento entre dois indivíduos, servindo de base para a parametrização da aceitação de informações entre sujeitos. Partindo do modo como a memória de um indivíduo foi modelada, torna-se inviável a princípio definir uma distribuição de probabilidade conjunta ou condicional entre a informação observada entre dois indivíduos, ao passo que as distribuições de probabilidade obtidas pela frequência empírica sobre as memórias  $L_u$  e  $L_v$  são independentes. Dessa forma, medidas tais como *Informação Mútua* não podem ser utilizadas. Assim, aplicamos a medida de divergência de Jensen-Shannon (JSD) - *raio de informação* - sobre as distribuições  $P_u$  e  $P_v$  de dois vizinhos  $u$  e  $v$  para o cálculo da sua similaridade:

$$\begin{aligned} JSD(P_u; P_v) &= \frac{1}{2}D(P_u; M) + \frac{1}{2}D(P_v; M) \\ &= H(M) - \frac{H(P_u) + H(P_v)}{2}, \end{aligned}$$

onde  $M = \frac{1}{2}(P_u + P_v)$ . Com isso, definimos a *proximidade entre  $u$  e  $v$*  como  $\mathcal{S}(u, v) = 1 - JSD(P_u, P_v)$ .

A aceitação das informações de um indivíduo  $v$ , por parte de um indivíduo  $u$  depende de uma probabilidade que considerará a proximidade de  $u$  e  $v$  e a percepção de  $u$  sobre  $v$ . Essa percepção pode levar em consideração a influência, medido pelo número de conexão dos vértices, ponderado pela proximidade medida por  $\mathcal{S}(u, v)$  entre um vértice e seus vizinhos. Considere primeiramente a distribuição sobre os vértices  $u \in Q$ , onde  $Q \subset V(G)$ ,

$$\sigma(u) = \sum_{v \in N(u)} \mathcal{S}(u, v) \leq k_u,$$

onde  $k_u$  é o grau do vértice  $u$ , e  $N(u)$  é o conjunto de vizinhos de  $u$ . A ponderação da proximidade entre  $u$  e seus vizinhos pode ser interpretada como um peso do suporte da comunidade adjacente à  $u$ , tendo em vista que um vértice  $u$  com muitas conexões ideologicamente distantes não poderia ser entendido como *popular* - o esperado é que um indivíduo inserido em uma comunidade ideologicamente distinta não teria suas opiniões facilmente aceitas.

A *probabilidade de aceitação*  $\eta_{u \rightarrow v}$  de uma informação de  $v$  por  $u$  é definida como

$$\eta_{u \rightarrow v} = \frac{2}{\mathcal{S}(u, v)^{-1} + \left( \frac{\sigma(u)^\gamma}{\max_{w \in N(v) \cup \{v\}} \sigma(w)^\gamma} \right)^{-1}}, \quad (3.1)$$

em que  $\gamma$  é um parâmetro denominado *fator de confiança*. Se  $\gamma > 0$ , um indivíduo  $v$  terá maior tendência de aceitar a informação de vizinhos mais influentes, e vice-versa. Para  $\gamma = 0$ ,  $v$  tenderá a confiar em todos os vizinhos igualmente. A aplicação da média harmônica faz com que a probabilidade seja sensível a variações, isto é, a separação elevada entre  $u$  e  $v$  ou a impopularidade (ou popularidade, a depender de  $\gamma$ ) do emissor tem um peso negativo considerável. Esta probabilidade de aceitação é nova dentro da literatura estudada e que serviu de inspiração para este trabalho (Wang et al., 2017; Wang, Koh, Cheong, & Xie, 2019), e acreditamos que esta probabilidade pode refletir uma dinâmica de troca de informações mais verossímil por medir a proximidade ideológica dos indivíduos.

### 3.2 Algoritmo

O grafo será inicializado com a memória de todos os indivíduos totalmente preenchida aleatoriamente, a fim de que seja possível calcular os valores iniciais para os pesos das arestas. A memória funcionará como uma estrutura de fila do tipo *FIFO*, *first-in-first-out*, onde a inserção de uma nova informação na fila desloca as informações mais antigas, descartando a última da fila. Se definirmos  $\varphi(\cdot)$  como um operador que converte números inteiros em suas representações binárias, então o estado inicial da memória de um vértice  $u$ ,  $L_u(0)$ , é dado por

$$L_u(0) = \{x_i \mid i = 1, \dots, \mu; x_i = \varphi(X), X \sim B(2^m, 0.5)\},$$

onde  $B(2^m, 0.5)$  é uma distribuição binomial com parâmetros  $(n, p) = (2^m, 0.5)$ , aproximando uma distribuição normal com parâmetros  $(\mu, \sigma^2) = (2^{m-1}, 2^{m-2})$ . A motivação para iniciarmos as memórias com esta distribuição consiste no fato dela ser simétrica quando ordenamos as informações segundo ao seu valor decimal, tornando mais fácil visualizar o movimento da distribuição do sistema, partindo de uma faixa central de informações disponíveis. É cabível observar que os parâmetros definidos dependem da memória dos indivíduos em cada instante, que é alterada ao longo da simulação. Há, portanto, uma dependência temporal desses parâmetros. Por simplicidade omitiremos na notação a consideração de que as definições e processos descritos nesta seção dizem respeito a um dado instante  $t$ .

Em uma primeira fase da simulação, admitindo que exista uma representação do grafo por uma lista de arestas  $E$ , o algoritmo percorrerá todos os elementos  $\langle u, v \rangle \in E$  realizando trocas de informações entre  $u$  e  $v$ , i.e. escolhe-se ao acaso uma informação  $x \in X_u$  e uma informação  $y \in X_v$  partindo das distribuições  $P_u$  e  $P_v$ , e faz-se a transmissão de  $x$  para  $v$  e  $y$  para  $u$  levando em consideração as probabilidades de distorção definidas anteriormente para os indivíduos de cada grupo. Portanto, a informação recebida por um indivíduo não corresponde necessariamente à informação emitida. Nesse processo, também é garantido que todo vértice transmitirá alguma informação (não necessariamente a mesma) para cada um dos seus vizinhos.

Após a seleção se verifica a aceitação da informação emitida, segundo a probabilidade  $\eta_{u \rightarrow v}$  dada pela Equação 3.1. As informações rejeitadas são descartadas, enquanto que as informações aceitas são armazenadas em uma lista temporária até a conclusão da disseminação na rede, em que os dados da rede são atualizados.

Ao término de um momento  $t$  da simulação, cada vértice  $u \in V(G)$  passa a ter uma nova memória  $L_u(t+1)$  com a adição das informações recebidas, e o grafo é atualizado com o cálculo dos novos pesos  $JSD(u, v)$  e dos parâmetros que dependem deles.

Uma série de simulações foi executada para investigar o modelo apresentado. Definiu-se o parâmetro  $s = 5$ , o que estabelece uma quantidade de informações (i.e. códigos binários) distintas igual a  $s^5 = 32$ . Um primeiro conjunto de simulações foi executado utilizando grafos completos com  $N = 75$  vértices, e regulares com  $N = 100$  vértices com grau variando em  $\delta \in \{5, 20, 50\}$ . Em um primeiro momento, o uso de topologias simples possibilita validar a implementação computacional do modelo por podermos elaborar expectativas concretas sobre como a dinâmica do sistema deve transcorrer em tais condições. Ademais, torna-se possível analisar mais facilmente os resultados obtidos com simulações em topologias mais complexas por termos como base de comparação as diferenças estruturais entre as redes, que conhecemos claramente (e.g. distribuição de graus dos vértices).

Por ser um modelo teórico mais fiel às redes empíricas, a análise principal será feita sobre a rede de Barabasi-Albert (BA). A rede será inicializada com um número  $m_0 = 5$  de nodos, e de acordo com o algoritmo de construção de redes BA, novos nodos foram adicionados individualmente e iterativamente, onde em cada iteração o novo nodo foi conectado a  $m = 2$  nodos existentes com uma probabilidade proporcional ao número total de arestas existentes de acordo com o algoritmo Alg. 1. A evolução temporal do modelo será analisada a partir de um conjunto de métricas: a entropia global  $\overline{H}$ , a proximidade média entre vértices e as distribuições de probabilidade de cada código binário dentro do conjunto de informações de todo o sistema.

Para estudar a dinâmica do modelo em função da variação dos parâmetros principais, quais sejam,  $\kappa$  e  $\gamma$ , é necessário fixar os parâmetros da simulação e da estrutura do modelo, sendo eles o número de simulações para que seja extraída a média das estatísticas calculadas, o número de iterações por simulação, o tamanho da rede e da memória. Todos estes parâmetros influem no tempo de processamento necessário para executar todas as simulações, portanto é preciso escolher uma configuração de parâmetros de simulação que não implique em um tempo de execução excessivo e que garanta a robustez dos resultados.

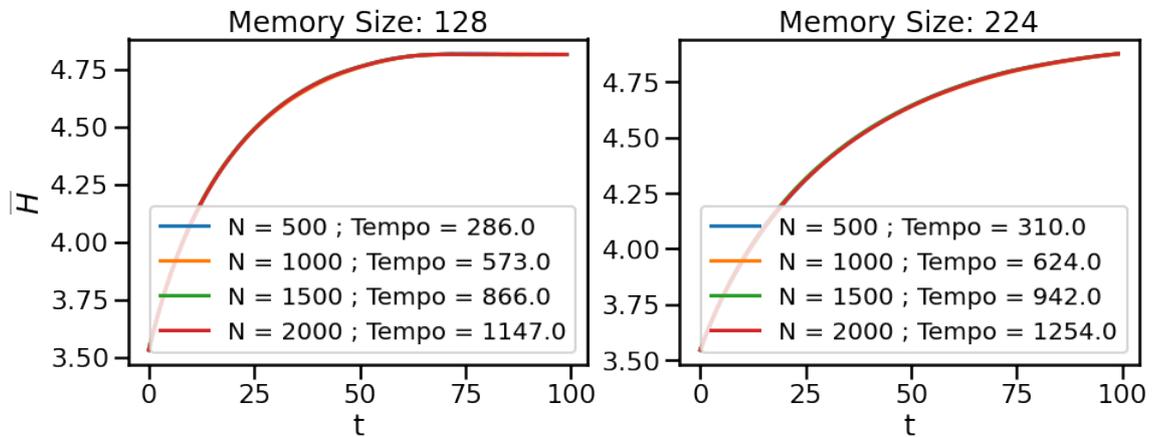


Figura 4.1: Evolução temporal da entropia média e tempo de execução em função do número de vértices e tamanho da memória. Foram utilizados os parâmetros  $\kappa = 0$  e  $\gamma = 0$  para a obtenção das curvas.

Foi realizada então uma série de simulações variando o tamanho da rede BA e da memória a fim de averiguar o tempo de execução e eventuais diferenças na dinâmica do sistema<sup>1</sup>. Considerou-se  $N$  em  $\{500, 1000, 1500, 2000\}$  e  $\mu$  em  $\{32 \cdot n | n = 1, \dots, 7\}$ ,  $\kappa$  em  $\{0, 1, 5\}$ ,  $\gamma$  em  $\{-1, 0, 1\}$ . Alguns resultados podem ser vistos na Figura 4.1. Em primeiro lugar, os aspectos qualitativos da dinâmica do sistema não se alteram consideravelmente com a variação destes parâmetros, notando-se apenas maior estabilidade da curva de entropia média para redes com um maior número de vértices e tamanho da memória. Uma curva mais suave também pode ser obtida com um maior número de simulações, o que é preferível por aumentarmos os números de amostras utilizadas no cálculo das estatísticas. O tempo de execução cresce quase linearmente com o número de vértices na rede. Para as simulações escolhemos, portanto,  $N = 500$  e  $\mu = 256$  dado que os resultados são similares porém com um tempo de execução consideravelmente menor.

## 4.1 Análise das Simulações

Buscaremos analisar os efeitos que a distorção natural decorrente de ruído tem na dinâmica de disseminação de informações da rede.

Nas seções que seguem, serão discutidas as curvas de entropia média obtidas sobre grafos completos (Figura 4.2) e regulares (Figura 4.3), e em seguida os resultados obtidos sobre a rede BA (Figura 4.7). De forma geral, nota-se que na ausência de qualquer preferência por parte dos indivíduos, a tendência dos sistemas é o de aumento de entropia com o aparecimento das variações possíveis de códigos binários até o ponto em que estejam todos

<sup>1</sup>O modelo foi simulado em um computador com as seguintes configurações: processador I5 7300HQ, 8GB de RAM DDR4, SSD WD Black Nvme de 256gb.

presentes nas memórias. A velocidade com que isso ocorre dependerá, naturalmente, da configuração dos parâmetros.

#### 4.1.1 Grafos Completos e Regulares

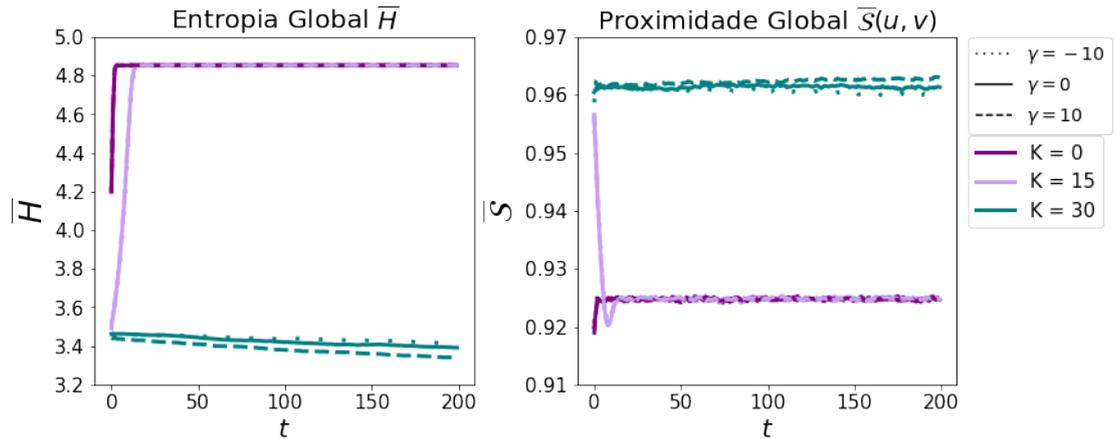


Figura 4.2: Evolução temporal da entropia média e proximidade ideológica média em função dos parâmetros  $\kappa$  e  $\gamma$  para um grafo completo.

O grafo completo oferece a dinâmica mais simples para a disseminação de informações. Na Figura 4.2 são apresentadas as curvas da entropia média  $\bar{H}$  e da proximidade média  $\bar{S}$  utilizando o grafo completo e com os valores de  $\kappa$  e  $\gamma$  nos conjuntos  $\{0, 10, 30\}$  e  $\{-10, 0, 10\}$ , respectivamente. Com exceção do caso em que  $\kappa = 30$ , em que a entropia média permanece próxima ao seu valor inicial, a entropia do sistema sobe rapidamente para um valor máximo. Valores maiores de  $\kappa$  atrasam a convergência para o máximo mas de forma moderada, e o parâmetro  $\gamma$  não gera efeito algum sobre o sistema, como nota-se pela sobreposição das curvas para diferentes valores de  $\gamma$ . A convergência da entropia corresponde à convergência da medida de proximidade para um ponto ligeiramente inferior ao seu valor inicial - observa-se que com  $\kappa = 15$  ocorre uma queda inicial de proximidade em decorrência do aparecimento de novas informações por comunicação imperfeita, mas a recuperação é quase imediata pela facilidade de disseminação dessas variantes.

O comportamento observado está de acordo com o esperado para esta topologia da rede; posto que cada vértice está conectado a todos os outros, não é possível que haja preferências em função da percepção de popularidade e a aceitabilidade de novas informações dependerá somente da proximidade ideológica dos indivíduos, que permanece alta em razão de termos inicializado as memórias de todos os indivíduos com a mesma distribuição. Dessa forma, as distorções de informação geradas por um indivíduo se propagam para toda a rede, saturando-a rapidamente. Valores suficientemente altos de  $\kappa$  garantem apenas o controle sobre o processo de distorção, mantendo a comunicação entre os indivíduos em torno so-

mente das informações já disponíveis no início da simulação.

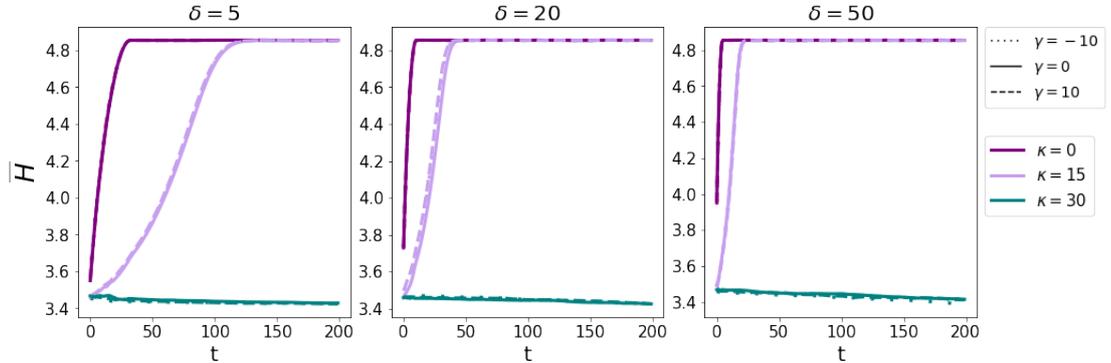


Figura 4.3: Evolução temporal da entropia média para um grafo regular em função dos parâmetros  $\kappa$ ,  $\gamma$  e  $\delta$ .

Quanto ao grafo regular, é preciso levar em conta a variação do parâmetro do grau dos vértices. São mantidos os conjuntos de valores de  $\kappa$  e  $\gamma$ , e consideramos os valores de  $\delta \in \{5, 20, 50\}$  para o grau. Novamente, o parâmetro  $\gamma$  não altera o comportamento do modelo dado que a distribuição de graus é constante. A variação de  $\delta$  tem um efeito sobre a velocidade do aumento da entropia na rede (Figura 4.3), sendo mais lento para um número menor de conexões, e viabilizando um efeito mais agudo sobre o atraso na convergência da entropia em decorrência do aumento do parâmetro  $\kappa$ . Nesta situação, todos os indivíduos estão submetidos às mesmas limitações de compartilhamento de informação, portanto a dificuldade de propagação de distorções de informação é comum à toda rede.

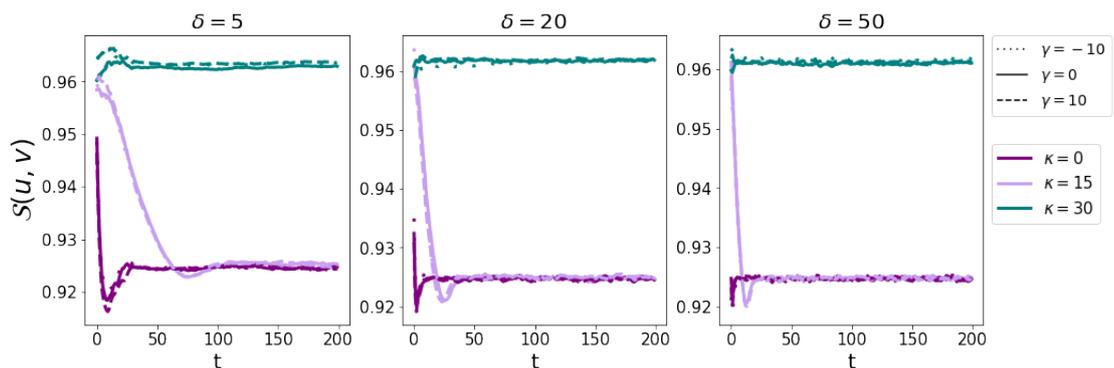


Figura 4.4: Evolução temporal da proximidade ideológica média para um grafo regular em função dos parâmetros  $\kappa$ ,  $\gamma$  e  $\delta$ .

Diferentemente do que foi visto no grafo completo, para valores menores de  $\delta$  observa-se claramente a queda e recuperação da curva de proximidade (Figura 4.4) também para  $\kappa = 0$  e  $\kappa = 15$ . Dada a dificuldade de propagação de distorções pelo número menor de conexões, há um intervalo até que surjam na rede todas as variantes de informações mesmo quando não

há controle algum sobre as distorções, e a recuperação a partir deste ponto, em que a distribuição de probabilidade das informações nas memórias passa a tender para um estado mais homogêneo, é também mais lenta. No caso em que a propagação de distorções é suficientemente dificultada ( $\kappa = 15$ ) essa oscilação é suavizada ao ponto de se tornar negligível, posto que as distorções não aparecem com velocidade suficiente para afastar ideologicamente os indivíduos.

O comportamento aqui descrito está também de acordo com o que esperaríamos para o grafo regular por conta da homogeneidade dos graus dos vértices. Estes resultados corroboram a confiabilidade da implementação do modelo e, por conseguinte, dos resultados que obteremos com topologias de maior complexidade.

#### 4.1.2 Rede de Barabasi-Albert

Podemos resumir o comportamento do modelo em três dinâmicas possíveis. Um caso consiste na manutenção da distribuição inicial de informações da população, atrelada à baixa variação da entropia média. Do outro lado, teríamos a descaracterização dessa distribuição inicial através da maximização da entropia do sistema, onde a distribuição de informações caminharia para próximo da Uniforme. O terceiro caso se daria pela convergência do sistema para uma configuração específica de informações, resultando na queda da entropia.

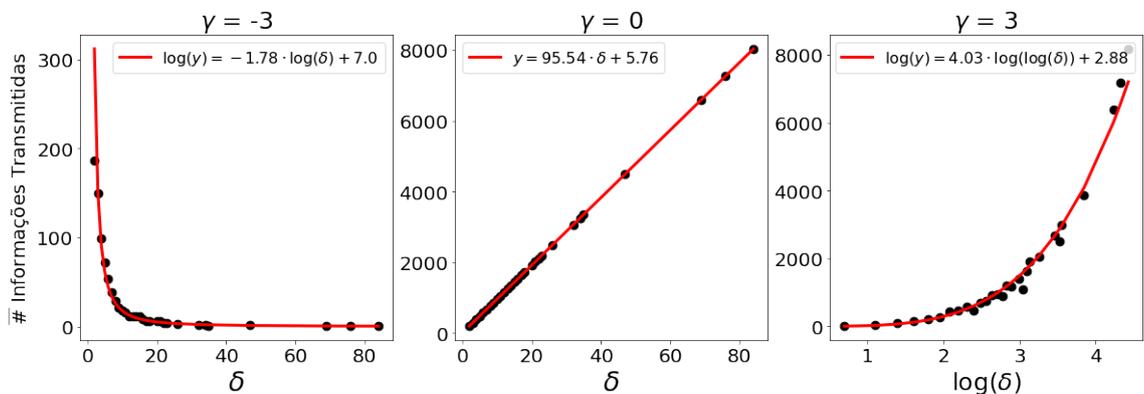


Figura 4.5: Relações entre o número médio de informações efetivamente transmitidas e grau do vértice.

A manutenção da distribuição inicial pode ser realizada pelo aumento do parâmetro  $\kappa$ , reduzindo o efeito de distorção e tornando-o negligível para valores muito grandes, ou variando  $\gamma$  para valores diferentes de zero, atribuindo aos indivíduos da rede uma preferência ou rejeição relativa às informações transmitidas por sujeitos mais populares. Neste contexto em que as diferenças no número de conexões dos indivíduos afetam a confiabilidade do discurso, nos interessa estudar de que forma o grau de um vértice pode impactar a sua interação com seus vizinhos ao longo da simulação. Para isso, podemos extrair do modelo o número total de transmissões e adesões bem sucedidas (i.e. em que os indivíduos envolvi-

dos efetivamente aceitaram a informação transmitida) de informações para cada vértice da rede, e então apresentar os valores em função do grau  $\delta$ .

O gráfico de pontos da média do número total de informações transmitidas com sucesso ( $y$ ) em função do grau do vértice emissor para diferentes valores  $\gamma$  é apresentado na Figura 4.5. Na geração desse gráfico observou-se que a distribuição dos pontos não se alterava significativamente pela variação do  $\kappa$ , então a fim de obter um retrato global da distribuição dos pontos, tirou-se também a média ao longo do conjunto de valores de  $\kappa$ . Nota-se que o  $\gamma$  tem um efeito notável na distribuição de pontos. Para  $\gamma = 0$ , os pontos formam uma distribuição evidentemente linear. Para  $\gamma = 3$ , é mantida uma correlação positiva entre o grau do vértice e o total de transmissões, e em particular, podemos modelar esta relação como uma lei de potência  $y = a(\log \delta)^b$ , em que  $a, b > 0$  são parâmetros obtidos pelo ajuste da curva. Isso significa que as transmissões caem consideravelmente para vértices com menor grau, havendo indivíduos que realizaram com sucesso poucas ou nenhuma transmissão ( $\delta \leq 5$ ), enquanto os indivíduos mais conectados ( $\delta > 60$ ) realizaram mais de 6000 transmissões, dominando a difusão de informações na rede.

Para  $\gamma = -3$  nota-se o contrário, como esperado: os indivíduos mais conectados realizaram poucas ou nenhuma transmissão, enquanto que os indivíduos com menos conexões realizaram mais transmissões. Entretanto, neste caso o número total de transmissões (251) é bastante reduzido em comparação aos casos para  $\gamma = 0$  (8186) e  $\gamma = -3$  (8264). A relação entre  $y$  e  $\delta$  pode ser modelada também como uma lei de potência  $y = a\delta^b$ , com  $a > 0$  e  $b < 0$ .

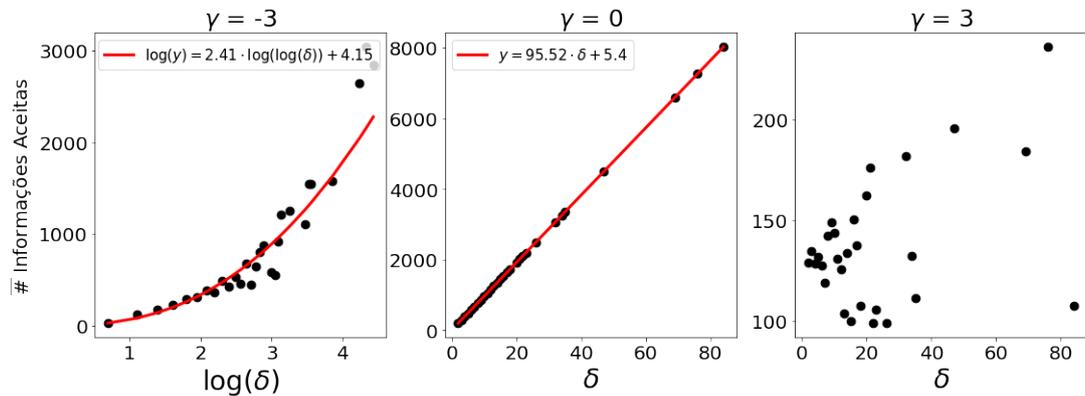


Figura 4.6: Relações entre o número médio de informações aceitas e grau do vértice.

Na Figura 4.6 é apresentado o gráfico de pontos da média do número total de informações aceitas em função do grau do vértice receptor para diferentes valores de  $\gamma$ . Para  $\gamma = 0$ , a distribuição é essencialmente a mesma (linear) obtida no gráfico anterior. Para  $\gamma = -3$ , obtemos novamente uma relação do tipo  $y = a(\log \delta)^b$ , com  $a, b > 0$ . Nesta situação, os indivíduo mais conectados são mais receptíveis às informações que recebem, aceitando-as em um volume muito maior (em decorrência ao número de conexões) que os seus vizinhos. Para  $\gamma = 3$  não é possível determinar uma tendência característica para os pontos, porém cabe notar que as

maiores quantidades de adesões podem ser encontradas nos vértices de menor grau.

Do que foi exposto, podemos concluir que o parâmetro  $\gamma$  determina padrões de dominância na disseminação de informações da rede. No caso em que  $\gamma = 0$ , não há nenhuma preferência e portanto a quantidade de informações transmitidas e recebidas por um vértice será de fato igual ou próximo ao seu número de conexões, variando apenas em função da proximidade ideológica com seus vizinhos. Todavia, para  $\gamma = 3$ , passa a haver uma preferência maior por indivíduos populares, estes dominam a transmissão de informação no sentido de que os indivíduos da rede passam a aceitar informações somente deste pequeno grupo de maior popularidade, enquanto que os indivíduos com pouquíssimas conexões tornam-se receptores passivos, posto que a aceitabilidade das suas informações torna-se exageradamente baixa. Do outro lado, para  $\gamma = -3$  os indivíduos com maior conexão tornam-se os receptores passivos, ao passo que há agora uma desconfiança generalizada para com estes sujeitos, e a aceitação das informações oriundas dos indivíduos menos populares, agora o grupo que domina a transmissão de informações na rede, é maior.

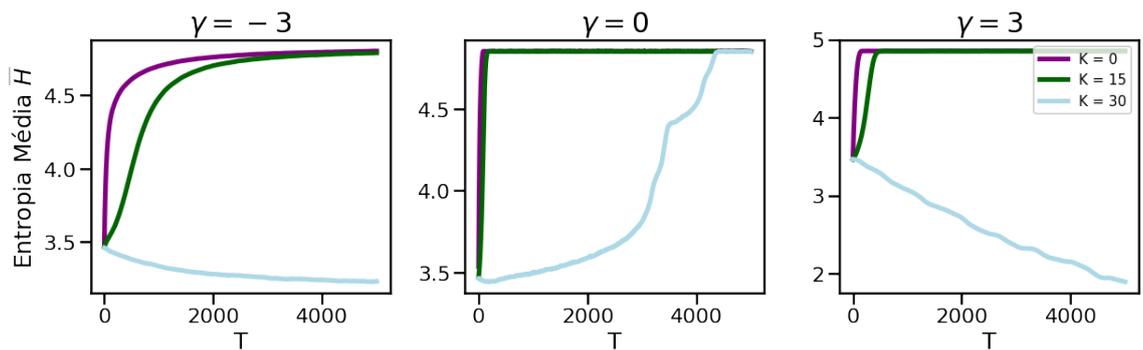


Figura 4.7: Evolução temporal da entropia média para uma rede de Barabasi-Albert em função dos parâmetros  $\kappa$  e  $\gamma$ .

Retornando à análise da dinâmica do modelo, na Figura 4.7 encontram-se as curvas de evolução da entropia média para vários valores de  $\kappa$  e  $\gamma$ . Tanto para valores positivos quanto negativos de  $\gamma$  nota-se que a entropia converge mais lentamente para o seu valor máximo ( $\kappa < 30$ ), entretanto o efeito é mais acentuado para valores negativos. Isso se dá precisamente pelo efeito de dominância discutido anteriormente. Com  $\gamma \neq 0$  há sempre um conjunto de indivíduos com menor probabilidade de transmitir com sucesso as suas informações. Em particular, com  $\gamma$  negativo e maior confiança sobre indivíduos menos populares, há uma grande quantidade de indivíduos transmissores, i.e. com maior probabilidade de aceitação das suas informações, dada a distribuição de  $\delta$  no modelo de Barabasi-Albert, porém cada um deles é capaz de transmitir informações para um número muito limitado de vizinhos, reduzindo a velocidade de propagação de informações e, por consequência, da disseminação de distorções pela rede. Essa menor velocidade se reflete na evolução da curva da entropia média do sistema que deve apresentar um crescimento consideravelmente mais lento se

comparado às curvas obtidas com  $\gamma \geq 0$ . Já para  $\gamma$  positivo, há um grupo menor de indivíduos dominando a disseminação de informações mas cada um deles transmite para um número muito grande de vizinhos, de modo que qualquer variante obtida por distorção é disseminada muito rapidamente na rede, ainda que menos que no caso  $\gamma = 0$  em virtude da menor probabilidade de um indivíduo conseguir transmitir uma informação distorcida.

Fixado um  $\gamma$ , a variação do  $\kappa$  desacelera a convergência da curva de entropia. Podemos notar mais facilmente que para  $\gamma$  positivo suficientemente grande, os valores de  $\kappa$  dividem-se em dois grupos, um para os quais a entropia média cresce e outro para os quais ela decresce. A entropia passa a diminuir a partir do ponto em que os efeitos de distorção tornam-se negligíveis ( $\kappa$  grande) e a variação na composição de informações na rede se torna dependente sobretudo das preferências sociais dos indivíduos, ou seja, da similaridade ideológica e da percepção de popularidade. Neste ponto, há essencialmente uma seleção das informações disponíveis e torna-se possível que um grupo de sujeitos domine a disseminação de informações na rede. A escolha de valores positivos de  $\gamma$  aumenta a probabilidade de aceitação das informações dos indivíduos que possuem também o maior alcance de disseminação, ao mesmo tempo que reduz a probabilidade de um indivíduo bem conectado aceitar informações dos seus vizinhos, a não ser que a similaridade ideológica entre eles aumente. Por conta disso, torna-se fácil que a distribuição final do sistema seja composta por um conjunto fixo de informações oriundas dos indivíduos mais bem conectados da rede. Este resultado pode ser visto na Figura 4.8, em que a evolução temporal da distribuição de probabilidade de informações no sistema é apresentada como um mapa de calor. É evidente pelo gráfico correspondente ao caso descrito ( $\kappa = 30$ ,  $\gamma = 10$ ) que as informações inicialmente mais salientes possuem maior probabilidade de permanecerem no sistema, mas pode ocorrer de informações salientes perderem espaço para outras até então menos frequentes.

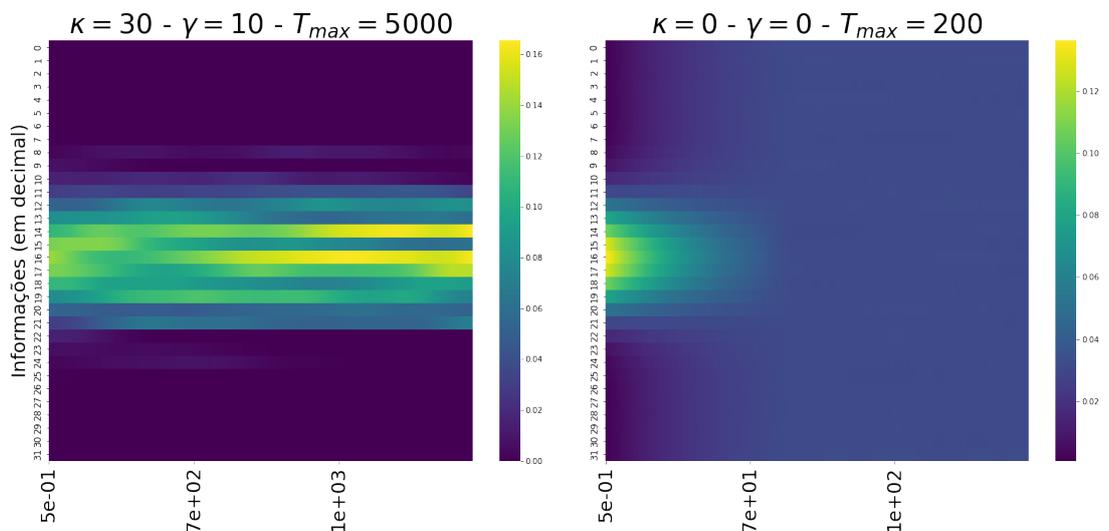


Figura 4.8: Gráfico de calor da evolução temporal da distribuição de probabilidade das informações no sistema.

Na ausência de qualquer fator que atribua aos indivíduos uma preferência por tipos particulares de informação, o modelo tenderá, como comentado, a introduzir no sistema todas as variedades possíveis de informações, descaracterizando a distribuição inicial, eventualmente convergindo para a distribuição uniforme (entropia máxima) ou alguma distribuição próxima, como vemos na Figura 4.8 para  $\kappa = 0$  e  $\gamma = 0$ , e na Figura 4.9, onde é possível ver mais claramente que as probabilidades de todos os códigos convergem rapidamente para o ponto  $0.0312 \pm 0.0002$ , mantendo um equilíbrio dinâmico.

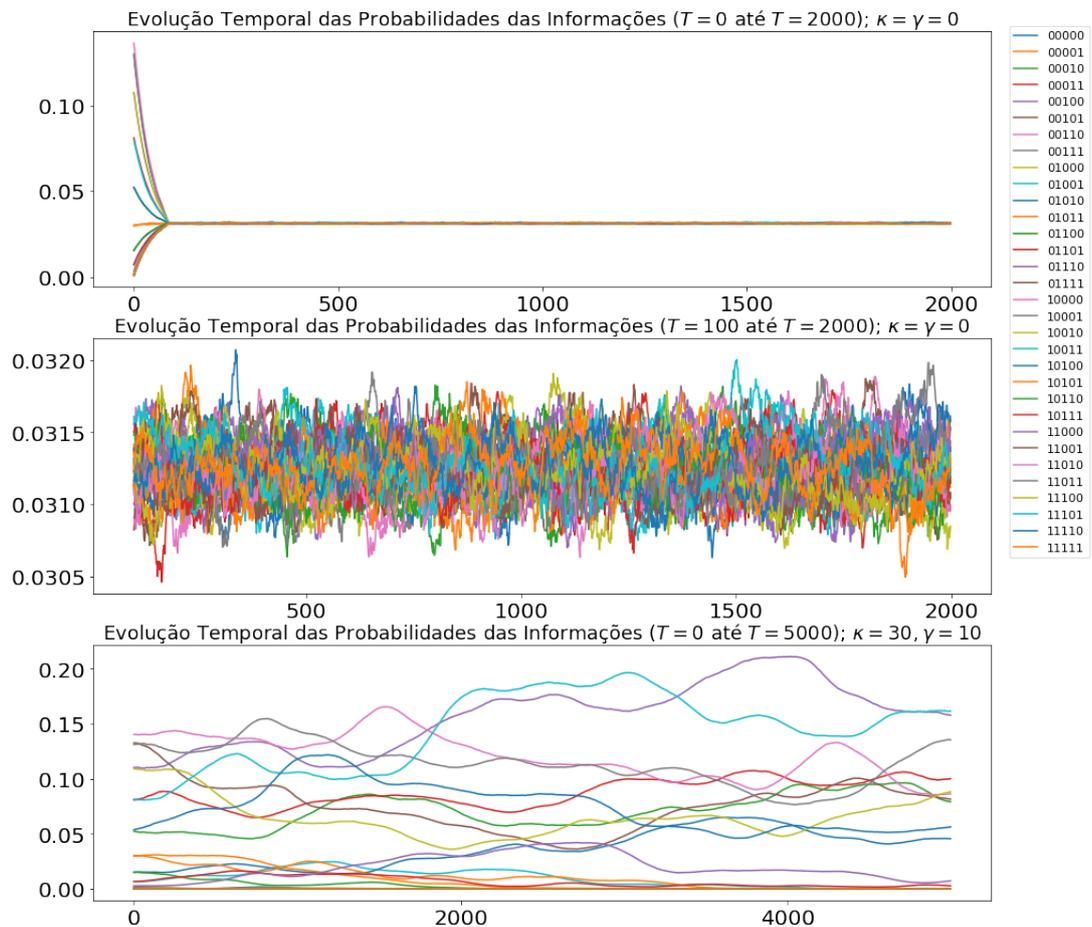


Figura 4.9: Curvas da evolução temporal da distribuição de probabilidade das informações no sistema.

É importante salientar o comportamento não-linear observado sobre a variação da entropia média do sistema ao utilizarmos a rede de Barabasi-Albert com os parâmetros  $\kappa = 30$  e  $\gamma = 0$ . Pela curva apresentada em 4.7, podemos concluir que é necessário simular o modelo por um número suficientemente alto de iterações para que certas possibilidades de dinâmicas sejam analisadas. Se plotarmos as mesmas curvas com um número menor de simulações, como apresentado na Figura 4.10, vemos que seria impossível prever o comportamento do modelo nestas condições ao observarmos a sua evolução ao longo de 5000 iterações. Ainda

assim, para os casos  $\gamma = \pm 3$ , nota-se que 5000 iterações não é ainda suficiente para observar a convergência da entropia  $\bar{H}$  para todos os valores de  $\kappa$ . Isso insere um desafio para a investigação do modelo, posto que o custo computacional de simulá-lo por um número elevado de iterações e repetições (a fim de extrair a média) é consideravelmente elevado, sendo inviável para o escopo do presente projeto.

Esse comportamento não-linear pode ser explicado da seguinte forma: dado que  $\kappa$  é suficientemente alto para tornar os efeitos de distorção insignificantes, a distribuição de informações na rede aparenta não sofrer alteração nenhuma nos instantes iniciais da simulação, o que se reflete na curva de entropia aparentemente estável. A entropia não cai como observado para  $\gamma = \pm 3$  por conta da ausência de grupos dominantes, uma vez que não há preferências sociais com  $\gamma = 0$ , portanto qualquer variante de informação que surja no sistema poderá ser mais facilmente propagado. Nesse sentido, passa a ocorrer um lento acúmulo de novas informações no sistema, que não são removidas do sistema também devido à baixa taxa de distorção de informações. A partir de um certo instante, parece haver um número suficientemente alto de informações variadas no sistema para que seja mais provável, em termos de distorções de bits, que estas aumentem em quantidade e que novas informações surjam no sistema, e então observamos um crescimento mais acelerado da entropia do sistema até o ponto de convergência para a entropia máxima.

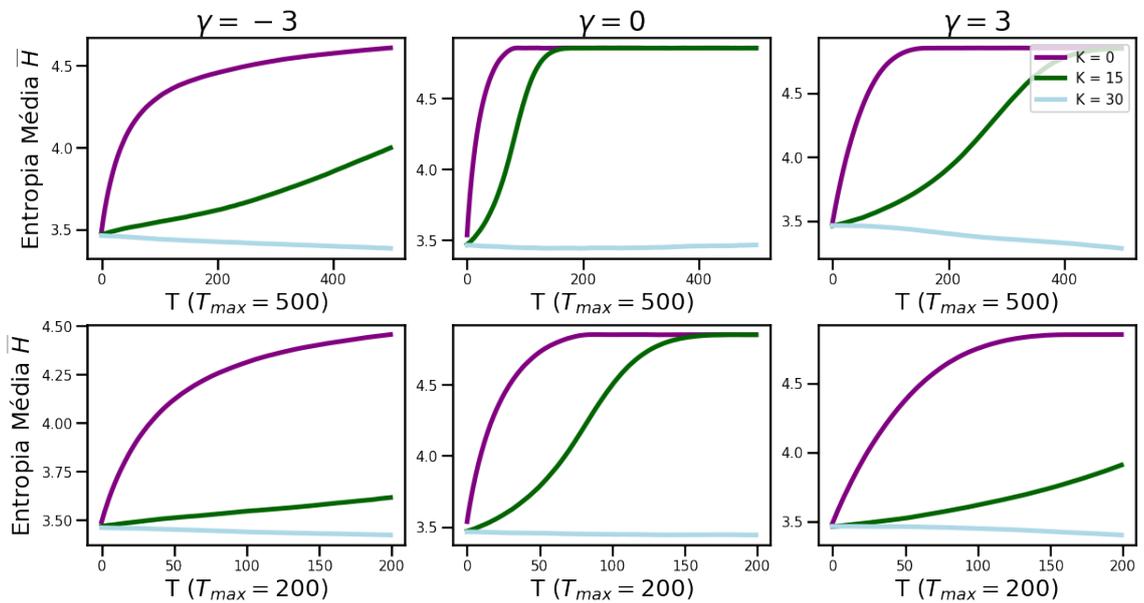


Figura 4.10: Evolução temporal da entropia média para uma rede de Barabasi-Albert em função dos parâmetros  $\kappa$  e  $\gamma$ , enfatizando diferentes intervalos de tempo de simulação.

Uma outra forma de estudar o comportamento do sistema é através da evolução temporal da proximidade ideológica entre indivíduos, i.e. os pesos das arestas. Os gráficos para múltiplos valores de  $\kappa$  e  $\gamma$  são apresentados na Figura 4.11. Na maioria dos casos nota-se

uma queda acelerada, tão acentuada quanto menor o valor de  $\kappa$ , seguida de um crescimento da proximidade até convergir para um ponto próximo a 0.925. A variação dos parâmetros não parece alterar o ponto de equilíbrio, entretanto, como observado pelos gráficos, atribuir ao  $\gamma$  valores diferentes de zero, em particular negativos, torna a convergência mais tardia, além de reduzir o valor mínimo de proximidade. Notando que a queda abrupta de proximidade nos instantes iniciais da simulação é mais acentuada para valores menores de  $\kappa$ , pode-se dizer que ela tem relação com o maior grau de distorções ocorrendo nas informações propagadas na rede. A partir da distribuição inicial, devido ao mecanismo de distorção potencialmente qualquer outra informação pode ser propagada, ainda que não esteja, a princípio, presente no sistema. Entretanto, cada indivíduo pode receber informações distintas, de modo que inicialmente o conteúdo da sua memória vai divergir das informações guardadas por seus vizinhos, o que se reflete na queda observada na curva de proximidade. Essa queda ocorre até o ponto em que todas as informações estão presentes no sistema, e posto que não há uma preferência dos indivíduos, não há mais como os indivíduos divergirem ideologicamente e a tendência geral do sistema é o de homogeneizar a distribuição de informação. Nessa etapa a curva de proximidade passa a crescer até convergir para um ponto entre seu valor inicial e o ponto de mínimo.

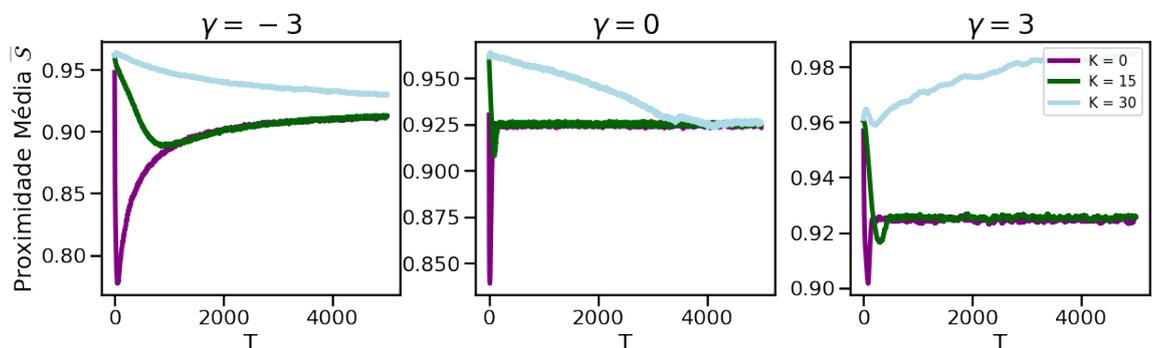


Figura 4.11: Evolução temporal da proximidade média para um grafo de Barabasi-Albert em função dos parâmetros  $\kappa$  e  $\beta$ .

É possível estabelecer uma conexão entre os dois cenários discutidos acima e a dinâmica observada sobre as curvas de entropia apresentadas na Figura 4.7. Tomando como exemplo a curva para  $\kappa = 15$  com  $\gamma = -3$  (é possível notar o mesmo para  $\kappa = 15$  e  $\gamma = 3$ , porém com maior dificuldade devido à escala do gráfico), vemos que a curva possui um ponto de inflexão, onde inicialmente há um crescimento mais acelerado da entropia, seguido de um crescimento desacelerado até o ponto de convergência. A primeira situação corresponde ao período de aparecimento de novas informações no sistema, enquanto que o segundo corresponde à etapa de homogeneização das distribuições. O aumento do parâmetro  $\kappa$  atua reduzindo a intensidade dessa oscilação inicial. Com um controle maior de distorção, há um período maior necessário para o aparecimento de novas variações, e ao longo deste in-

tervalo as informações já presentes podem ser distribuídas entre os indivíduos, reduzindo as diferenças entre eles. Como consequência, a oscilação da curva de proximidade é menos acentuada, podendo ser, para alguns valores de  $\kappa$ , uma simples queda até o ponto de equilíbrio, posto que os indivíduos não conseguem se afastar rápido o bastante para provocar uma queda acentuada da proximidade.

Como elaborado anteriormente, o parâmetro  $\gamma$  influencia o surgimento de subgrupos de indivíduos que passam a dominar a disseminação de informações no sistema. Esse comportamento se reflete nas curvas de proximidade, em que para  $\gamma = 3$  o ponto mínimo de proximidade é elevado em relação às curvas obtidas com  $\gamma = 0$ , e para  $\gamma = -3$  o ponto mínimo é reduzido. Para ambos os casos, a convergência se dá mais lentamente. Com  $\gamma$  positivo, a maioria dos indivíduos no sistema são ouvintes passivos de um subconjunto reduzido de indivíduos bem conectados, de modo que a composição de informações no sistema evolui de acordo com o que é produzido por estes disseminadores de informação. Nesse sentido, ocorre uma queda de proximidade na medida em que variações surgem no sistema, mas não se observa o comportamento característico de queda abrupta na proximidade média pelo fato de que os indivíduos, de forma geral, tendem a uma distribuição homogênea de informações sem nunca se afastarem demais da distribuição dos indivíduos controlando a propagação. Por outro lado, para  $\gamma$  negativo os indivíduos menos conectados dominam a disseminação, resultando em uma produção de ruído que demora a ser suavizado por conta da dificuldade de propagação de informações entre indivíduos com poucas conexões; em consequência, para  $\kappa$  pequeno observa-se uma queda mais acentuada e convergência mais lenta que a observada para  $\gamma = 0$ .

### 4.1.3 Comparação com a Literatura

O presente trabalho foi baseado principalmente no trabalho (Wang et al., 2017), introduzindo modificações que impossibilitam uma comparação direta com ele. Entretanto, é possível realizar observações acerca do comportamento do modelo comparando-o com os resultados obtidos até então na literatura.

A principal diferença entre o modelo proposto no presente trabalho e a proposta em (Wang et al., 2017) consiste no fato de que as memórias dos indivíduos na rede são inicializadas a partir de uma distribuição binomial, ao passo que é necessário um conteúdo de informação em cada indivíduo para que sejam calculadas os pesos nas arestas do grafo. Isso implica, em primeiro lugar, que observamos a variação da entropia do sistema a partir de um valor não-nulo, crescendo ou decaindo em função dos parâmetros do modelo. A partir disso, todos os indivíduos iniciam como propagadores das informações contidas em sua memória, diferindo do mecanismo usual de propagação em que um indivíduo é escolhido ao acaso para disseminar um dado rumor, e observa-se então como o rumor é difundido na rede, sofrendo eventuais modificações. Nesse sentido, o modelo proposto passa a enfatizar a dinâmica sobre as distribuições de informação no sistema, em contraponto à observação

de informações particulares.

A variação da entropia média da população em função dos parâmetros  $\kappa$  e  $\gamma$  guarda similaridades com o trabalho em (Wang et al., 2017). O parâmetro  $\kappa$  possui importância acentuada no controle de ruído do sistema, havendo um aumento da entropia para valores menores de  $\kappa$  até estabilizar em valores inferiores à entropia máxima, maior na medida em que  $\kappa$  decresce, e uma queda de entropia para valores grandes de  $\kappa$  (no modelo exposto em (Wang et al., 2017), a entropia permanece em zero). No modelo aqui proposto os valores de  $\kappa$  necessários para um controle mais assertivo do ruído são maiores que os observados em (Wang et al., 2017), isso porque o mecanismo de distorção aqui implementado introduz um ruído maior ao sistema ao aplicar uma probabilidade de mutação em cada bit transmitido, enquanto que anteriormente a chance de mutação correspondia a um único bit da informação transmitida, escolhido ao acaso com distribuição uniforme.

Ao contrário do que foi reportado em (Wang et al., 2017), conseguimos observar que a variação do parâmetro  $\gamma$  implica em mudanças na forma com que as informações são propagadas na rede, por afetar os padrões de preferências sociais entre os indivíduos. Isso se reflete na mudança da velocidade de variação da entropia e, negligenciando os efeitos de distorção e escolhendo valores suficientemente altos de  $\gamma$  em módulo, a entropia passa a cair com o controle de um subconjunto de indivíduos sobre a propagação de informações na rede.

## 4.2 Conclusão

Neste trabalho apresentamos um modelo de disseminação de informações em redes que, fazendo uso do conceito de Entropia e medidas de divergência baseadas em Teoria da Informação, passa a enfatizar a análise sobre a distribuição das informações propagadas no sistema e como se dá sua variação a partir de fatores que visam simular aspectos cognitivos atuantes nas interações sociais, como proximidade ideológica e distorções por imperfeições de memória, que buscamos introduzir ao modelo. O modelo introduzido traz resultados que, até onde se sabe, não foram observados em trabalhos anteriores, posto que as modificações introduzidas no presente modelo, em particular o uso de pesos nas arestas modelando a proximidade ideológica entre indivíduos a partir da medida de divergência de Jensen-Shanon, são inovadoras.

O modelo foi testado em diferentes topologias de redes, quais sejam: grafos completos, regulares e a rede de Barabasi-Albert (BA). Os resultados observados sobre as topologias mais simples, o grafo completo e regular, estão de acordo com o que intuitivamente se esperaria da dinâmica nestas condições, conferindo confiabilidade à definição e implementação do modelo. O estudo da dinâmica sobre topologias mais simples também nos permite observar mais facilmente a influência da variação do *fator de conservação*  $\kappa$  e o número de conexões dos vértices. O crescimento da entropia média do sistema, que equivale ao surgimento de novas variações de informações na rede, até um ponto de saturação em que a entropia é

máxima, é desacelerado com o aumento de  $\kappa$ , que dificulta a geração de novas informações por distorção, e pela queda no número de conexões, que dificulta a propagação dessas informações pela rede. Essa dificuldade de propagação influencia a evolução da proximidade ideológica média do sistema, tornando mais lenta a sua recuperação após uma queda inicial acentuada por conta do surgimento acelerado de novas informações no sistema, em particular para valores pequenos de  $\kappa$  - esse resultado é menos relevante no grafo completo, posto que qualquer nova informação é rapidamente propagada para toda a rede. Devido à homogeneidade da distribuição de graus dos vértices nestes grafos, o *fator de confiança*  $\gamma$  não altera a dinâmica do sistema.

O grafo de Barabasi-Albert oferece uma topologia mais adequada para simular interações sociais devido ao fato de que sua distribuição de graus dos vértices segue uma lei de potência, onde um pequeno conjunto de indivíduos possui um número elevado de conexões, e tal propriedade é comumente encontrada em redes empíricas. Por conta desta propriedade, o fator de confiança  $\gamma$  tem forte influência na dinâmica do modelo na rede BA. Ao analisarmos o número médio de informações transmitidas com sucesso em função do grau do vértice transmissor, observamos a possibilidade de descrever esta relação pela função  $y = a(\log \gamma)^b$ ,  $a, b > 0$  se  $\gamma > 0$ , e como  $y = a\gamma^b$ ,  $a > 0$ ,  $b < 0$ , se  $\gamma < 0$ . Para  $\gamma = 0$ , a relação é linear. Isso nos permite concluir que o parâmetro  $\gamma$  gera um padrão de dominância de um grupo de indivíduos sobre a propagação de informações na rede: com  $\gamma$  positivo, os indivíduos mais populares dominam a disseminação de informações, no sentido de que suas opiniões têm uma aceitabilidade maior que as opiniões de indivíduos menos populares, enquanto que, de forma análoga, com  $\gamma$  negativo, os indivíduos mais populares passam a ter suas opiniões rejeitadas, tornando-se ouvintes, e o grupo de indivíduos menos populares domina a disseminação de informações. Resultado similar é encontrado quando analisamos o número médio de informações recebidas em função do grau do vértice receptor, chegando a uma relação  $y = a(\log \delta)^b$ ,  $a, b > 0$  com  $\gamma$  negativo, entretanto para  $\gamma$  positivo não observamos nenhuma relação relevante.

Esse comportamento se reflete diretamente nas curvas de evolução da entropia média do sistema. A entropia cresce mais lentamente com  $\gamma \neq 0$ , posto que um grupo menor de indivíduos é responsável por propagar para a rede as novas variantes de informação. Com  $\gamma < 0$  o crescimento da entropia é ainda mais lento pois, ainda que o número de indivíduos disseminadores seja maior, estes são pouco conectados e, portanto, a informação se propaga mais lentamente no sistema. Com  $\kappa$  suficientemente grande, o efeito de distorção se torna pouco significativo e mudança na distribuição de informações do sistema passa a depender somente de fatores sociais, i.e. a percepção de popularidade, dependente de  $\gamma$ , e a proximidade ideológica. Nesse caso, observa-se que a entropia decresce, e analisando a evolução das probabilidades de cada informação vemos que há uma seleção de informações, em que algumas delas ganham mais saliência enquanto outras desaparecem do sistema. Essa seleção depende tão somente do grupo propagador, e pela maior facilidade de propagação por parte dos indivíduos mais conectados, vemos que a queda da entropia é mais rápida com

$\gamma > 0$ .

A análise da dinâmica do modelo com a rede BA exige um número alto de iterações. Por este motivo, não podemos concluir, por exemplo, que a entropia cairá para zero com  $\gamma \neq 0$  e  $\kappa \geq 30$ , pois o número de iterações utilizadas (5000) pode ser insuficiente para garantir essa afirmação. Com  $\gamma = 0$ , em estágios mais avançados da simulação observa-se um comportamento não-linear pouco previsível caso observássemos a dinâmica do modelo em um intervalo de tempo muito curto (e.g.  $T_{max} < 1000$ ). Neste contexto, o parâmetro  $\kappa = 30$  reduz o surgimento de variantes de informação, porém como  $\gamma = 0$ , nenhum grupo de indivíduos tem suas opiniões ativamente rejeitadas ou priorizadas e, portanto, todos os indivíduos têm uma alta probabilidade de terem suas opiniões aceitas pelos seus vizinhos. Tendo em vista que a probabilidade de distorção não é nula mas é baixa o suficiente para não eliminar imediatamente qualquer nova informação que surja no sistema, há um acúmulo gradual de novas informações. Essa heterogeneidade facilita o aumento da quantidade dessas novas variações (o número de mutações necessárias é menor), o que se reflete no crescimento mais acelerado da entropia até o seu ponto de convergência.

A necessidade de um número elevado de simulações traz mais desafios do ponto de vista computacional. Nas presentes condições de disponibilidade de recursos computacionais, o tempo de simulação e o espaço de armazenamento necessários para produzir um número de repetições adequado para a extração das médias das medidas de entropia e proximidade aumentam rapidamente ao elevarmos o número de iterações por repetição. Levando em conta as análises, testes e ajustes que eventualmente precisam ser feitas para validar os resultados e, enfim, apresentá-los, a produção destes resultados com um número elevado de iterações extrapola o escopo de tempo do presente trabalho.

Para o presente trabalho foi utilizada a linguagem Python para a implementação do modelo e para executar as simulações, e recorrendo principalmente aos pacotes Numpy para as operações numéricas e NetworkX para a manipulação das redes. O pacote Numpy é já bem conceituado pela comunidade por permitir um bom desempenho computacional para operações numéricas, por utilizar, subjacente à interface do Python, implementações já otimizadas de algoritmos numéricos em linguagens de maior desempenho para esta função (e.g. Fortran, C). O pacote NetworkX é também bastante utilizado, porém não faz uso extensivo de implementação otimizadas em baixo nível dos algoritmos sobre grafos e redes complexas, tal como as estruturas das redes são armazenadas em objetos próprios do Python, o que pode tornar o modelo menos eficiente computacionalmente caso aumentemos a escala da rede.

Uma forma de otimizar a velocidade das simulações consiste em reimplementar o modelo utilizando uma linguagem tal como a C++, comumente utilizada para a simulação de sistemas de grande escala. Neste caso, a dificuldade aumenta pela maior complexidade de se trabalhar com C++ em comparação à linguagem Python, mas pode ser um desafio necessário caso se queira ampliar ainda mais o escopo do modelo em termos de tamanho da rede e número de simulações. Para o presente caso, deve ser suficiente adaptar o código em Python

para tirar melhor proveito de paralelismo em uma máquina de maior poder computacional. Ademais, pacotes como o *graph-tool* oferecem implementações mais eficientes em C++ de algoritmos em redes, exigindo menos alterações no código já implementado do modelo, porém possivelmente incorrendo em dificuldades de compatibilidade do código ao passo que tais pacotes têm requisitos técnicos (e.g. sistema operacional) particulares.

#### 4.2.1 Trabalhos Futuros

Tendo em conta os resultados e análises apresentadas neste trabalho, a principal perspectiva de desenvolvimentos futuros envolve produzir resultados de simulação mais robustos para o modelo utilizando uma rede de Barabasi-Albert a fim de validar o comportamento não-linear de longo-prazo observado anteriormente. Com um número maior de simulações, a curva média da variação da entropia do sistema pode ser obtida com mais confiabilidade, reduzindo o efeito de anomalias que porventura apareçam em alguma das simulações do modelo.

A análise se deu a partir da variação dos parâmetros relativos às interações entre os indivíduos e distorções de informação, se mantendo fixados os parâmetros da construção da rede de Barabasi-Albert. Portanto, é cabível ampliar a análise variando também os parâmetros da rede a fim de verificar se há uma alteração significativa sobre a dinâmica do sistema.

Há o intuito de ampliar o modelo com a inclusão de uma dinâmica de polarização das informações. Partindo de uma função de atribuição a cada código binário uma polaridade (e.g. entre 0 e 1), os indivíduos do sistema poderiam ser viesados para um dos polos, tendo uma probabilidade maior de distorcer as suas informações para o seu polo de preferência antes de comunicá-la aos seus vizinhos. A proposta iria de encontro ao trabalho desenvolvido em (Wang et al., 2019), e dadas as diferenças na dinâmica do sistema, em relação ao modelo apresentado em (Wang et al., 2017), com as modificações introduzidas no presente trabalho, é razoável esperar que a inclusão do fator de polarização resulte em dinâmicas não observadas em (Wang et al., 2019) e, por consequências, contribuições importantes para a área.

Há ainda a possibilidade de estudar o comportamento do modelo utilizando diferentes topologias de redes complexas, inclusive topologias de redes empíricas obtidas a partir de, por exemplo, interações em redes sociais (e.g. Twitter). Pode-se adaptar as informações contidas nestas redes empíricas (e.g. as postagens dos usuários) para uma representação binária, de modo a obter uma condição inicial para iniciar a simulação do modelo. Seria interessante, além disso, estudar a aplicação das medidas de informação utilizadas para a caracterização dessas redes empíricas, medindo, por exemplo, a proximidade ideológica dos usuários da rede social, entre comunidades, etc.

## 5.1 Simulação de Variáveis Aleatórias

A fim de possibilitar a seleção aleatória de informações na memória dos indivíduos, precisamos primeiro definir sobre a memória uma distribuição de probabilidades e, em seguida, especificar um algoritmo para fazer amostragens a partir desta distribuição. Todos os métodos de geração de amostras aleatórias de uma distribuição qualquer partem da hipótese de que é possível simular variáveis aleatórias independentes e uniformemente distribuídas  $U_1, \dots, U_m$ . De fato, existem diversos algoritmos geradores de números *pseudo-aleatórios* que passam em testes de uniformidade. O problema de simular a distribuição de uma dada variável aleatória  $X$  é, pois, o de transformar a distribuição de uma variável aleatória uniforme na distribuição de  $X$ .

O caso em que  $X$  é uma variável aleatória discreta é bastante simples de ser trabalhado. Se  $X$  assume os valores  $x_1 < x_2 < \dots < x_k$ , então se  $U \sim \text{Uniforme}(0, 1)$ , podemos escrever que

$$\begin{aligned} P[F_X(x_i) < U \leq F_X(x_{i+1})] &= F_X(x_{i+1}) - F_X(x_i) \\ &= P(X = x_i). \end{aligned}$$

O algoritmo para gerar  $x_i \sim F_X(x)$  é direto,

1. Gere  $U \sim \text{Uniform}(0, 1)$ .
2. Se  $F_X(x_i) < U \leq F_X(x_{i+1})$ , faça  $X = x_{i+1}$ .

No contexto deste trabalho, as probabilidades  $f_x$  das informações  $x \in L_u$ , onde  $L_u$  é a memória de um dado indivíduo  $u$ , é simplesmente a frequência empírica dada por

$$f_x = \frac{\sum_{y \in L_u} \delta_{xy}}{|L_u|},$$

onde  $\delta_{xy}$  é o *Delta de Kroenecker*. As informações  $x$  são códigos binários de números inteiros entre 0 e  $2^m$ , onde  $m$  é um parâmetro do modelo, portanto é possível ordená-los a fim de calcular a distribuição acumulada  $F_X$ .

## 5.2 Algoritmos

---

**Algoritmo 2:** Inicialização do Modelo - É feita a criação do grafo e o cálculo dos parâmetros dos vértices e arestas.

---

**Entrada:**  $N, \mu, \kappa, \gamma$

Seja  $G = (V(G), E(G))$  uma rede BA de tamanho  $N$ ;

**para**  $u \in V(G)$  **faça**

$L_u \leftarrow$  Lista FIFO dada por  $\{x_i \mid i = 1, \dots, \mu; x_i = \varphi(X), X \sim B(2^m, 0.5)\}$ , onde  $\varphi(X)$  retorna a representação binária de  $X$ ;

$P_u \leftarrow \{f_x \mid f_x = \frac{\sum_{y \in L_u} \delta_{xy}}{|L_u|}, x \in L_u\}$ , onde  $\delta_{xy}$  é o Delta de Kroenecker;

$T_u \leftarrow [ ]$ ;

$u \leftarrow (L_u, P_u, T_u)$ ;

$H_u \leftarrow - \sum_{x \in L_u} P_u(x) \log P_u(x)$ ;

$\delta_u \leftarrow \frac{1}{\exp\left[\left(\frac{H_{\max} - H_u}{H_{\max}}\right) \cdot \kappa\right] + 1}$ ;

Seja  $\mathcal{D}_u(\delta_u)$  as probabilidades de distorção de bits  $0 \leftarrow 1$  e  $1 \leftarrow 0$ .

**fim**

**para**  $\langle u, v \rangle \in E(G)$  **faça**

$JSD(P_u; P_v) \leftarrow H\left(\frac{P_u + P_v}{2}\right) - \frac{H(P_u) + H(P_v)}{2}$ ;

Atribua a  $\langle u, v \rangle$  o peso  $JSD(P_u; P_v)$ ;

$\eta_{u \rightarrow v} \leftarrow \frac{1}{JSD(P_u, P_v)^{-1} + \left(\frac{\sigma(v)^\gamma}{\max_{w \in N(u) \cup \{u\}} \sigma(w)^\gamma}\right)^{-1}}$ , onde  $\sigma(u) = \sum_{w \in N(u)} JSD(P_u, P_w)$ ;

Análogo para  $\eta_{v \rightarrow u}$ ;

**fim**

---

---

**Algoritmo 3:** Simulação - O algoritmo percorre todas as arestas realizando troca de informações entre os vértices adjacentes. A atualização dos parâmetros utiliza as mesmas fórmulas aplicadas na inicialização

---

**Entrada:** O número de iterações  $n$

**para**  $1 \leq t \leq n$  **faça**

**para**  $\langle u, v \rangle \in E(G)$  **faça**

    Sejam  $x \in L_u$  e  $y \in L_v$  elementos tomados ao acaso segundo as distribuições  $P_u$  e  $P_v$ ;

$x, y \leftarrow \text{Corromper}(x, \mathcal{D}_u), \text{Corromper}(y, \mathcal{D}_v)$ ;

**se**  $v$  aceita  $x$  com probabilidade  $\eta_{u \rightarrow v}$  **então**

$T_v \leftarrow T_v \cup \{x\}$ ;

**fim**

**se**  $u$  aceita  $y$  com probabilidade  $\eta_{v \rightarrow u}$  **então**

$T_u \leftarrow T_u \cup \{y\}$ ;

**fim**

**fim**

**para**  $u \in V(G)$  **faça**

    Atualize  $L_u$  com  $T_u$  pelo mecanismo FIFO;

$T_u \leftarrow \{ \}$ ;

    Atualize os parâmetros  $\{H_u, \delta_u, \mathcal{D}_u(\delta_u)\}$

**fim**

**para**  $\langle u, v \rangle \in E(G)$  **faça**

    Atualize os parâmetros  $\{JSD(P_u; P_v), \eta_{u \rightarrow v}\}$ ;

**fim**

**fim**

---

**Algoritmo 4:** Corromper - O algoritmo percorre o código binário, fazendo mutações bit por bit a partir das probabilidades de transição de cada indivíduo, considerando efeitos de ruído e polarização.

---

**Entrada:** Um código binário  $x$ ; Uma distribuição  $\mathcal{D}$  das probabilidades de distorção

$0 \leftarrow 1$  e  $1 \leftarrow 0$

**Saída:** O código binário  $x$  potencialmente distorcido

**para** *Bit*  $b \in x$  **faça**

  Seja  $p \in [0, 1]$  tomado ao acaso com probabilidade uniforme ;

**se**  $b = 1$  e  $p \leq \mathcal{D}(0 \leftarrow 1)$  **então**

$b \leftarrow 0$

**senão se**  $b = 0$  e  $p \leq \mathcal{D}(1 \leftarrow 0)$  **então**

$b \leftarrow 1$

**fim**

---

## Referências

- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439), 509–512.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4-5), 175–308.
- Carrieri, V., Madio, L., & Principe, F. (2019). Vaccine hesitancy and (fake) news: Quasi-experimental evidence from italy. *Health economics*, 28(11), 1377–1382.
- Chen, P., & Redner, S. (2005). Majority rule dynamics in finite dimensions. *Physical review E*, 71(3), 036101.
- Csányi, G., & Szendrői, B. (2004). Structure of a large social network. *Physical Review E*, 69(3), 036131.
- Daley, D. J., & Kendall, D. G. (1965). Stochastic rumours. *IMA Journal of Applied Mathematics*, 1(1), 42–55.
- Di Mare, A., & Latora, V. (2007). Opinion formation models based on game theory. *International Journal of Modern Physics C*, 18(09), 1377–1395.
- Dorogovtsev, S. N., & Mendes, J. F. (2002). Evolution of networks. *Advances in physics*, 51(4), 1079–1187.
- Erdős, P. (1959). Alfréd, and rényi. “on random graphs”. *Publicationes Mathematicae, Debrecen*, 6, 290–297.
- Ferner, R. E., & Aronson, J. K. (2020). *Chloroquine and hydroxychloroquine in covid-19*. British Medical Journal Publishing Group.
- Galam, S. (2003). Modelling rumors: the no plane pentagon french hoax case. *Physica A: Statistical Mechanics and Its Applications*, 320, 571–580.
- Grein, T. W., Kamara, K., Rodier, G., Plant, A. J., Bovier, P., Ryan, M. J., ... Heymann, D. L. (2000). Rumors of disease in the global village: outbreak verification. *Emerging infectious diseases*, 6(2), 97.
- Hegselmann, R., Krause, U., et al. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation*, 5(3).
- Kimmel, A. J. (2004). *Rumors and rumor control: A manager’s guide to understanding and combatting rumors*. Routledge.
- Kosfeld, M. (2005). Rumours and markets. *Journal of Mathematical Economics*, 41(6), 646–664.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1), 79–86.
- Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... others (2018). The science of fake news. *Science*, 359(6380), 1094–1096.
- Lefevre, C., & Picard, P. (1994). Distribution of the final extent of a rumour process. *Journal of Applied Probability*, 244–249.

- Lin, J. (1991). Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1), 145–151.
- Maki-Thompson, D. (1973). *Mathematical models and applications, with emphasis on social, life, and management sciences*. NJ: Prentice-Hall, Englewood Cliffs.
- Moreno, Y., Nekovee, M., & Vespignani, A. (2004). Efficiency and reliability of epidemic data dissemination in complex networks. *Physical Review E*, 69(5), 055101.
- Nekovee, M., Moreno, Y., Bianconi, G., & Marsili, M. (2007). Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, 374(1), 457–470.
- Noymer, A. (2001). The transmission and persistence of ‘urban legends’: Sociological application of age-structured epidemic models. *Journal of Mathematical Sociology*, 25(3), 299–323.
- Richardson, M., & Domingos, P. (2002). Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 61–70).
- Rudin, W. (1987). *Real and complex analysis* (Third ed.). New York: McGraw-Hill Book Co.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379–423.
- Sudbury, A. (1985). The proportion of the population never hearing a rumour. *Journal of applied probability*, 443–446.
- Thomas, S. A. (2007). Lies, damn lies, and rumors: an analysis of collective efficacy, rumors, and fear in the wake of Katrina. *Sociological Spectrum*, 27(6), 679–703.
- Touret, F., & de Lamballerie, X. (2020). Of chloroquine and COVID-19. *Antiviral Research*, 177, 104762.
- Van Steen, M. (2010). Graph theory and complex networks. *An introduction*, 144.
- Wang, C., Koh, J. M., Cheong, K. H., & Xie, N.-G. (2019). Progressive information polarization in a complex-network entropic social dynamics model. *IEEE Access*, 7, 35394–35404.
- Wang, C., Tan, Z. X., Ye, Y., Wang, L., Cheong, K. H., & Xie, N.-G. (2017). A rumor spreading model based on information entropy. *Scientific Reports*, 7(1), 1–14.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442.
- Yeung, R. W. (2008). *Information theory and network coding*. Springer Science & Business Media.
- Zanette, D. H. (2002). Dynamics of rumor propagation on small-world networks. *Physical Review E*, 65(4), 041908.
- Zhang, Z.-L., & Zhang, Z.-Q. (2009). An interplay model for rumour spreading and emergency development. *Physica A: Statistical Mechanics and its Applications*, 388(19), 4159–4166.