

Projeto de Pesquisa de Iniciação Científica

Modelagem de Câmaras de Eco e Polarização em Redes Sociais

PIBIC - Edital 04/2022

CMCC - UFABC, Campus Santo André

Início: Agosto de 2022 Duração: 12 meses

29 de Junho de 2022

Resumo

Pretendemos nesse projeto fazer a modelagem do surgimento de câmaras de eco e de polarização em redes sociais. A principal metodologia será a simulação em computador de modelos baseados em agentes para propagação de opiniões, em que os agentes estão distribuídos em uma rede livre de escala. Mais precisamente, pretendemos estender alguns resultados recentes da literatura que foram obtidos para o modelo do votante [1, 2, 3, 4], tentando entender como esses resultados se modificariam para outros tipos de modelo. Para tanto serão estudados principalmente o modelo do votante [5], o modelo Sznajd e as suas variantes [6, 7] e as redes que iremos usar para a interação entre os agentes vão ser principalmente as redes de Barabási-Albert [8] e de Watts-Strogatz [9]. O objetivo principal é que o aluno obtenha experiência com a modelagem e simulação de modelos baseados em agentes (principalmente na parte de programação e análise de dados), assim como com a metodologia científica em geral.

Palavras chave: Modelagem baseada em Agentes, Sistemas Complexos, Redes Livres de Escala

1 Introdução

1.1 Motivação para o estudo e resultados recentes

Nos últimos anos, redes sociais tem se tornado uma das principais formas como pessoas interagem umas com as outras. Isso levanta a questão de quais efeitos essa forma de comunicação pode ter para a diversidade de opiniões na sociedade como um todo e para o surgimento de polarizações em que opiniões extremas se tornam mais comuns. Um aspecto que tem gerado certa preocupação são as chamadas *Câmaras de Eco* [10, 11, 12, 13]. Essas são situações em que um usuário de uma rede social acaba sendo exposto desproporcionalmente mais a pontos de vista já alinhados com os seus.

A própria estrutura da maioria das redes sociais favorece o surgimento de câmaras de eco. Tipicamente os usuários vão consumir conteúdo (mensagens, imagens, áudios, vídeos, etc.) produzidos por determinados criadores (que em geral também são usuários da rede) e através de *seguimentos*, *subscrições* ou outros mecanismos cada usuário pode decidir de quais criadores ele vai consumir conteúdo. Desse ponto de vista, as câmaras de eco acabam sendo uma consequência dos usuários preferirem consumir conteúdos alinhados com seus pontos de vista.

Além dessa estrutura básica, a maioria das redes sociais tem mecanismos de recomendação de conteúdo através dos quais usuários podem consumir conteúdo de criadores que eles não conheçam, o que em princípio poderia expor o usuário a conteúdo fora da sua câmara de eco. Porém esses mecanismos de recomendação tem problemas com os chamados *Filtros de Bolha* [14, 15], uma situação em que o algoritmo de recomendação correlaciona as recomendações ao histórico do usuário e acaba expondo ele prioritariamente a conteúdo já alinhado com suas opiniões.

Recentemente alguns trabalhos tem usado modelos de propagação de opiniões, inicialmente propostos para modelar eleições e formação de consenso, para modelar o surgimento de câmaras de eco e os efeitos de algoritmos de recomendação [1, 2, 3, 4]. Apresentamos aqui um resumo dos resultados obtidos (a descrição dos modelos de propagação de opiniões e das redes está na

próxima seção):

- (*Helbing et al* [1]) usa uma variante do modelo do votante [5] (em que com uma probabilidade p pequena um agente adota uma opinião nova na rede) para estudar quais efeitos a rede tem na distribuição total de opiniões. Eles encontram que mantendo p fixo, a diversidade de opiniões na rede depende grosso modo da distância média entre os sítios da rede em que os agentes estão interagindo. Para distâncias médias pequenas temos uma situação em que uma opinião é compartilhada pela maioria esmagadora dos agentes (porém qual é essa opinião dominante muda com o tempo) enquanto que para distâncias médias grandes temos uma situação em que as opiniões presentes na rede detém mais ou menos a mesma quantidade de agentes. O tipo de distância média também é tal que podemos afirmar que redes de Watts-Strogatz [9] tem um tendência maior a apresentar diversidade do que redes Barabási-Albert [8].
- (*Castellano et al* [2]) usa o modelo do votante com a adição de uma influência externa (modelando um mecanismo de recomendação, que é seguido com uma probabilidade p ao invés das regras usuais do modelo) em que cada agente vai sofrer uma ação desse mecanismo externo que é personalizada e enviesada para ser similar ao histórico de opiniões do agente (modelando assim um filtro de bolha). Eles encontram que há uma transição brusca variando p , com o modelo atingindo uma situação de consenso para p baixo (correspondendo a uma influência pequena do sistema de recomendações) e a partir de um valor crítico atingindo uma situação de coexistência de diferentes opiniões, em que cada agente acaba sempre retornando para a opinião alinhada com o sistema de recomendação dele. Além disso, quando o tamanho da rede cresce, essa probabilidade crítica tende a 0.
- (*Bonchi et al* [3]) propõe um *framework* para tratar um sistema de recomendação de novas conexões em um modelo de propagação de opiniões. No contexto de uma rede social esses seriam mecanismos como

recomendações de usuários para serem seguidos ou *recomendações de amizade*. A relevância para o problema é que essas são dinâmicas que alteram a estrutura da rede social e logo quais conteúdos são consumidos por quais usuários. Alguns algoritmos de recomendação são apresentados, assim como métricas relevantes para identificar a formação de câmaras de eco. Os modelos de propagação de opiniões que são de fato usados são o modelo Deffuant [16] e um modelo epistemológico baseado em aprendizado Bayesiano [17]. Os resultados encontrados são que a introdução dos sistemas de recomendação aumentaram consideravelmente a formação de câmaras de eco comparado com os modelos simulados na ausência de recomendações.

1.2 Modelos de Propagação de Opiniões

Os seguintes são os principais modelos que pretendemos utilizar durante o projeto. Em todos esses modelos a rede social é representada por um grafo em que os agentes/pessoas são representados pelos vértices e as arestas definem quais interações entre os agentes são possíveis. Cada agente vai ter uma propriedade (tipicamente um número inteiro ou um número real) que é interpretada como uma opinião e que também vai influenciar em como acontecem as interações. Finalmente, as opiniões evoluem de acordo com um processo estocástico discreto no tempo.

1.2.1 Modelo do Votante

No modelo do votante [5] as opiniões são representadas por números inteiros e a cada passo de tempo realizamos o seguinte algoritmo:

- Sorteie de maneira uniforme um agente i , dentre todos os agentes na rede.
- Sorteie de maneira uniforme um agente j , dentre todos os agentes que estejam conectados com i diretamente (ou seja, o caminho indo de i a j cobre apenas uma aresta).

- O agente i adota a opinião do agente j .

A ideia central desse modelo é que as interações entre os agentes podem levar a um agente convencendo o outro. Desse ponto de vista, as arestas representariam relações entre colegas, amigos, familiares, etc.

1.2.2 Modelo Sznajd

No modelo Sznajd [6] as opiniões são representadas por números inteiros e a cada passo de tempo realizamos o seguinte algoritmo:

- Sorteie de maneira uniforme um agente i , dentre todos os agentes na rede.
- Sorteie de maneira uniforme um agente j , dentre todos os agentes que estejam conectados com i diretamente.
- Sorteie de maneira uniforme um agente k , dentre todos os agentes que estejam conectados com j diretamente.
- Se as opiniões dos agentes i e j forem iguais, o agente k adota a opinião deles. Caso contrário, nada acontece.

A principal diferença do modelo Sznajd para o modelo votante é que necessitamos de 2 agentes concordando para que um terceiro seja convencido. A ideia é que isso reflete o fato que grupos de pessoas teriam um poder de convencimento maior (o que é corroborado por alguns experimentos clássicos de psicologia [18, 19])

1.2.3 Modelo Deffuant

No modelo Deffuant [16] as opiniões são representadas por números reais no intervalo $[0, 1]$ e temos dois parâmetros $\varepsilon > 0$ e $\mu \in [0, 1/2]$. A cada passo de tempo realizamos o seguinte algoritmo:

- Sorteie de maneira uniforme um agente i , dentre todos os agentes na rede.

- Sorteie de maneira uniforme um agente j , dentre todos os agentes que estejam conectados com i diretamente.
- Sejam o_i e o_j as opiniões dos agentes i e j respectivamente. Se $|o_i - o_j| > \varepsilon$ então nada acontece. Caso contrário os agentes i e j passam a ter opiniões o'_i e o'_j respectivamente, onde

$$o'_i = o_i + \mu(o_j - o_i) \quad \text{e} \quad o'_j = o_j + \mu(o_i - o_j)$$

Esse modelo conceptualiza as opiniões como existindo em um contínuo entre 2 extremos e tem uma interação que segue a ideia geral do modelo do votante, porém com duas diferenças importantes. Primeiro, há uma simetria na interação dos agentes, ao invés de um agente convencer o outro, ambos se aproximam um da opinião do outro. Segundo, agentes com opiniões muito diferentes não conseguem interagir. Esse último ingrediente reflete (de forma simplificada) entre outras coisas a ideia de *Dissonância Cognitiva* [20].

1.3 Modelos de redes

Assim como temos modelos para como acontecem as interações dos agentes, temos modelos para a estrutura das redes sociais. Os modelos que iremos utilizar serão os seguintes

1.3.1 Modelo Barabási-Albert

Nesse modelo [8] a rede é criada a partir da adição gradual de vértices e arestas. Para a descrição do algoritmo k_i será o número de conexões que um vértice i possui e $m, N \geq 1$ são parâmetros inteiros.

- Começamos com uma rede completa contendo m vértices
- Enquanto a rede tiver menos do que N vértices:
 - Introduza um vértice novo, v .
 - Conecte v a m outros vértices **distintos** já da lista, de forma que a probabilidade que i e v se conectem seja proporcional a k_i .

O modelo Barabási-Albert consegue reproduzir algumas propriedades exibidas por redes complexas observadas experimentalmente, como uma distância média que escala logaritmicamente com o número de vértices N (a chamada *propriedade de mundo pequeno*) e uma distribuição $p(k)$ para o número de conexões k de um vértice escolhido aleatoriamente que segue uma lei de escala (é possível mostrar que no limite $N \rightarrow \infty$ a distribuição $p(k)$ se aproxima de k^{-3}). Porém o modelo falha em reproduzir certas propriedades (por exemplo, o número de ciclos e de cliques é muito menor do que nos casos experimentais)

1.3.2 Modelo Watts-Strogatz

Nesse modelo [9] começamos com um rede regular (ciclos e redes quadradas são escolhas bastante comuns) e para cada aresta fazemos um sorteio com probabilidade p (que é um parâmetro) para decidir se realocamos a aresta ou não. Caso decidamos realocar a aresta, os 2 sítios originalmente conectados são desconectados e sorteamos um par de vértices que não estejam já conectados para constituírem a nova aresta.

Dessa forma o modelo Watts-Strogatz pode ser pensado como uma interpolação entre uma rede regular ($p = 0$) e um grafo em que as arestas foram todas sorteadas ao acaso ($p = 1$, também conhecido como modelo de Erdős-Rényi [21]).

2 Objetivos e Metas

Do ponto de vista de aprendizado, o objetivo principal é que o aluno obtenha experiência com a modelagem e simulação de modelos baseados em agentes (principalmente na parte de programação e análise de dados), assim como com a metodologia científica em geral.

De um ponto de vista científico, nos trabalhos da literatura que citamos na seção 1.1 os modelos de opinião utilizados foram principalmente o modelo do votante e o modelo Deffuant. Isso faz sentido em uma exploração inicial uma vez que eles estão entre os modelos mais simples que são co-

mumente utilizados para propagação de opiniões. Por conta disso, um dos nossos primeiros objetivos vai ser avaliar a universalidade desses resultados (principalmente os de [1] e [2]), testando como os resultados se modificam se modelos de propagação de opiniões mais complexos são utilizados, focando no modelo Sznajd e em variações dele (como o modelo q -votante [7] e o modelo com a introdução de vieses na interação das opiniões [22]).

Em um segundo momento, pretendemos analisar o framework PROD para recomendações de usuários (apresentado em [3]) e como diferentes estratégias poderiam levar a cenários de consenso ou polarização e quando a formação de câmaras de eco é favorecida ou suprimida. Um aspecto importante que não foi abordado no trabalho original é como esses sistemas de recomendação poderiam se alinhar com os objetivos da empresa responsável pela rede social. Em geral esse objetivo é conseguir o maior lucro possível através de propagandas, o que correlaciona fortemente com maximizar o tempo gasto na rede social pelos usuários. Por causa disso, modelar como a interação do usuário com diferentes opiniões afeta esse tempo de uso poderia dar uma ideia melhor do quão viável um sistema de recomendação que suprima a formação de câmaras de eco seria.

3 Metodologia

A metodologia principal vai ser a simulação em computador dos diferentes modelos que pretendemos estudar e fazer a análise estatística dos resultados numéricos. Dada a experiência do aluno isso vai ser feito na linguagem Python.

Em alguns casos, as simulações em redes livres de escala (como a rede Barabási-Albert que pretendemos utilizar) podem ser aproximadas de forma razoável por aproximações de campo médio, descritas por sistemas de equações diferenciais (um exemplo recente da literatura é [4]). Como em geral essas equações são não-lineares, apenas resultados qualitativos podem ser obtidos de maneira analítica. Mesmo assim, a integração numérica das equações pode dar um ponto de vista alternativo e será considerada dependendo dos resultados que obtivermos nas simulações não-aproximadas dos modelos.

4 Viabilidade da Execução do Projeto

O projeto exige do aluno apenas conhecimento de programação para a criação dos códigos para simular os modelos e noções de estatística para fazer a análise de dados, sendo desse ponto de vista bastante acessível.

Sobre os problemas que estamos nos propondo a estudar, eles são variações novas de questões que foram abordadas recentemente na literatura. Isso nos dá um guia de como proceder em nossa pesquisa e acredito que isso vá nos ajudar a manter o cronograma que propomos na próxima seção.

5 Cronograma

Nessa seção apresentamos um cronograma tentativo para a execução do projeto. Dada a possibilidade de haverem alterações dependendo dos resultados que obtivermos, a estimativa do tempo que levaremos em cada parte foi feita de maneira conservadora

- (Agosto e Setembro de 2022)

Estudo dos resultados de literatura a respeito dos modelos que iremos utilizar, juntamente com a implementação do código a ser utilizado para as simulações.

- (Outubro e Novembro de 2022)

Experimentos numéricos, focando em modificações do cenário em Helbing et al [1]. Análise estatística dos resultados.

- (Dezembro de 2022 e Janeiro de 2023)

Experimentos numéricos, focando em modificações do cenário em Castellano et al [2]. Análise estatística dos resultados.

- (Fevereiro de 2023)

Elaboração do relatório parcial. Modificação do código para utilizar o framework PROD [3].

- (Março a Junho de 2023)
Experimentos numéricos, focando em diferentes estratégias de recomendação do framework PROD. Análise estatística dos resultados.
- (Julho de 2023) Elaboração do relatório final e preparação do pôster para o simpósio de iniciação científica.

Referências

- [1] Andrea Musso and Dirk Helbing. How networks shape diversity for better or worse, 2022. *arXiv:2201.09254*
- [2] Giulio Iannelli, Giordano De Marzo, and Claudio Castellano. Filter bubble effect in the multistate voter model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(4):043103, 2022.
- [3] Federico Cinus, Marco Minici, Corrado Monti, and Francesco Bonchi. The effect of people recommenders on echo chambers and polarization. *Proceedings of the International AAAI Conference on Web and Social Media*, 16(1):90–101, 2022.
- [4] Antonio F. Peralta, János Kertész, and Gerardo Iñiguez. Opinion formation on social networks with algorithmic bias: dynamics and bias imbalance. *Journal of Physics: Complexity*, 2:045009, 2021.
- [5] R. A. Holley and T. M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *Annals of Probability*, 3(4):643–663, 1975.
- [6] K. Sznajd-Weron and J. Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(6):1157–1165, 2000.
- [7] Cláudio Castellano, Romualdo Pastor-Satorras, and Miguel A. Muñoz. Nonlinear q-voter model. *Physical Review E*, 80(4):041129, 2009.
- [8] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

- [9] Duncan J. Watts and Steven Henry Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.
- [10] Pablo Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10):1531–1542, 2015.
- [11] Matthew Gentzkow and Jesse M Shapiro. Ideological segregation online and offline. Working Paper 15916, National Bureau of Economic Research, April 2010.
- [12] H. Akin Unver. Digital challenges to democracy: Politics of automation, attention, and engagement. *Journal of International Affairs*, 71(1):127–146, 2017.
- [13] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9):e2023301118, 2021.
- [14] Dominic DiFranzo and Kristine Gloria-Garcia. Filter bubbles and fake news. *XRDS*, 23(3):32–35, 2017.
- [15] H. Holone. The filter bubble and its effect on online personal health information. *Croat Med J.*, 57(3):298–301, 2016.
- [16] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3:87–98, 2000.
- [17] Venkatesh Bala and Sanjeev Goyal. Learning from Neighbours. *The Review of Economic Studies*, 65(3):595–621, 07 1998.
- [18] Solomon Eliot Asch. Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow, editor, *Groups, leadership and men*. Carnegie Press, Pittsburgh, PA, 1951.

- [19] Muzafer Sherif. A study of some social factors in perception. Master's thesis, Columbia University, 1935.
- [20] Leon Festinger. *A theory of Cognitive Dissonance*. Stanford University Press, 1957.
- [21] Paul Erdos and Alfréd Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.
- [22] André Martin Timpanaro and Carmen Pimentel Cintra do Prado. Generalized sznajd model for opinion propagation. *Physical Review E*, 80(2):021119, 2009.