

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Gabriel Pereira Vezono

**Análise visual do processo de espalhamento de  
doenças infecciosas por meio de redes  
temporais**

**Uberlândia, Brasil**

**2021**

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Gabriel Pereira Vezono

**Análise visual do processo de espalhamento de doenças  
infecciosas por meio de redes temporais**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Ciência da Computação.

Orientador: Jean Roberto Ponciano

Universidade Federal de Uberlândia – UFU

Faculdade de Ciência da Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2021

*Dedico esse trabalho à toda comunidade científica, em especial ao meu orientador Jean R. Ponciano e ao pesquisador Cláudio D. G. Linhares, que ,também, muito contribuiu para o sucesso desse trabalho.*

# Agradecimentos

Agradeço primeiramente a Deus, por tantas bênçãos, pela força, pela fé e pela perseverança. Agradeço aos meus pais, minha irmã, meus avós e toda minha família e amigos por todo apoio e incentivo. E, também, agradeço a cada um dos meus professores, que foram pessoas de grande valia para meu desenvolvimento, tanto no meio acadêmico, quanto pessoal, nessa fase tão única da vida.

*“Porquanto, ainda que a figueira não floresça, nem haja fruto na vide; o produto da oliveira minta, e os campos não produzam mantimento; as ovelhas da malhada sejam arrebatadas, e nos currais não haja vacas: Todavia eu me alegrarei no Senhor: exultarei no Deus da minha salvação.”, Habacuque 3:17-18*

# Resumo

As redes temporais são uma maneira útil de representar instâncias de dados e suas interações ao longo do tempo. As características dessas redes, tais como: alta concentração de interações e ociosidade, as tornam adequadas para a simulação e análise de processos dinâmicos, como por exemplo a transmissão de doenças infecciosas. Este trabalho visa simular e analisar, por meio de estratégias de visualização de redes temporais, diferentes cenários de espalhamento de doenças infecciosas, com intuito de ressaltar sua utilização neste contexto. Além disso, possibilita órgãos governamentais, bem como instituições de pesquisa uma melhor análise diante de um cenário epidêmico e, conseqüentemente o auxílio na tomada de decisões. Por meio de dois estudos de caso com redes temporais reais coletadas a partir de ambientes escolares, concluímos que, além de demonstrarem o potencial de estratégias de visualização para esse tipo de estudo, corroboram com estudos que defendem a alta eficiência de grandes níveis de isolamento social e de medidas preventivas pessoais, como o uso de equipamentos de proteção individual.

**Palavras-chave:** Redes temporais, Detecção de comunidades em redes complexas, Visualização da informação, Espalhamento de doenças infecciosas.

# Lista de ilustrações

Figura 1 – Relação dos conceitos abordados nesta pesquisa. . . . .	13
Figura 2 – Exemplo visual de uma rede. Pontos representam nós e linhas representam arestas (interações entre os nós). . . . .	14
Figura 3 – Rede com 13 nós e 3 comunidades representadas. Figura adaptada de Ponciano et al. (2020c). . . . .	15
Figura 4 – Exemplo de uma rede temporal contendo 3 instantes de tempo e as interações que ocorrem em cada um deles. Figura adaptada de Linhares et al. (2016) . . . . .	16
Figura 5 – Exemplos dos tipos de layouts de rede. Sendo (a) a tabular, (b) a estrutural e (c) o MSV. Figura adaptada de Linhares et al. (2017) . . . .	18
Figura 6 – Exemplos de layouts temporais, onde em (a) mostra o MSV, tendo nós circulares e arestas à mostra. Em (b) também, porém com esquema de cores. E, em (c), aplicado o MAT. Os números representados por cores dizem respeito à quantidade de conexões que o nó realizou naquele instante. Em (a) e (b), as arestas são exibidas curvadas para fins didáticos. Figura adaptada de (LINHARES et al., 2016). . . . .	20
Figura 7 – Exemplo de poluição visual no MSV, com legenda representando níveis de atividade no decorrer do tempo. Figura adaptada de (LINHARES et al., 2019a) . . . . .	21
Figura 8 – Exemplo de poluição visual no layout estrutural não animado. Figura adaptada de (LINHARES et al., 2019a) . . . . .	21
Figura 9 – Layout MSV de uma rede de hospital (VANHEMS et al., 2013), ordenado de 3 maneiras diferentes. (a) por nascimento, (b) lexicograficamente e (c) por vizinhos recorrentes. Figura adaptada de (LINHARES et al., 2017). . . . .	22
Figura 10 – Exemplo didático do funcionamento do algoritmo <i>Community-based Node Ordering</i> . Figura adaptada de (LINHARES et al., 2019b) . . . . .	23
Figura 11 – <i>Community-based Node Ordering</i> aplicado em uma rede de hospital, usando as técnicas Louvain e vizinhos recorrentes. (a) exibindo apenas arestas intercomunitárias. (b) exibindo apenas arestas intracomunitárias. Figura adaptada de (LINHARES et al., 2019b) . . . . .	23
Figura 12 – Ilustração dos modelos do tipo espalhamento de doença considerados neste trabalho. S = “Suscetível”, I = “Infectado” e R = “Recuperado”	24

Figura 13 – Figura representando os modelos apresentados pelo Sistema de Informação Geográfica. O mapa de pontos (à esquerda) os utiliza para representar uma certa medida ou característica exibida sobre um mapa geográfico. No meio, é representado um mapa coroplético, que usa cores ou sombras graduadas para indicar os valores de alguma medida agregada em áreas especificadas. À direita, tem o mesmo tipo de mapa, todavia com o objetivo de derivar estatisticamente um gradiente de risco espacial. Figura adaptada de (CARROLL et al., 2014) . . . . .	25
Figura 14 – Demonstração da árvore de transmissão. Figura adaptada de (LINHARES et al., 2019a) . . . . .	26
Figura 15 – Mapa de Atividade Temporal ilustrando os diferentes modelos de espalhamento de doenças. (a) SI, (b) SIS, (c) SIR, (d) SIRS. A cor azul indica o estado de Suscetível, vermelho indica Infectado e, amarelo indica Recuperado. Figura adaptada de (LINHARES et al., 2019a) . . . . .	26
Figura 16 – Simulação de epidemia de pneumonia utilizando autômatos celulares. Figura adaptada de (SAITO; SILVA; ALVES, 2008) . . . . .	27
Figura 17 – Gráfico representando a evolução do número de infectados ao variar o valor de $\beta$ de 0 a 1 . . . . .	34
Figura 18 – Recorte do MAT para a situação de <i>full lockdown</i> ( $\beta = 0$ ) . . . . .	34
Figura 19 – MAT da combinação de 90% de distanciamento social ( $\beta = 0,1$ ) . . . . .	35
Figura 20 – MAT da combinação onde não há distanciamento social ( $\beta = 1$ ) . . . . .	35
Figura 21 – Recorte do MAT para a combinação ( $\beta = 0,2, \lambda = 0,7$ ) . . . . .	37
Figura 22 – Recorte do MAT para a combinação ( $\beta = 0,8, \lambda = 0,7$ ) . . . . .	38
Figura 23 – Recorte do MSV para a combinação ( $\beta = 0,2, \lambda = 0,05$ ), destacando apenas os infectados e o caminho da transmissão da doença. A cor preta destaca a turma 2BIO2, a cor verde destaca a turma 2BIO3 e a cor azul destaca a turma 2BIO1. O recorte mostra apenas os três primeiros dias da rede. . . . .	38
Figura 24 – Recorte do MSV para a combinação ( $\beta = 0,8, \lambda = 0,05$ ), destacando apenas os infectados e o caminho da transmissão da doença. A cor preta destaca a turma 2BIO2, a cor verde destaca a turma 2BIO3 e a cor azul destaca a turma 2BIO1. O recorte mostra apenas os dois primeiros dias da rede. . . . .	39
Figura 25 – Representação do último <i>timestamp</i> das redes utilizadas por meio do layout estrutural, mostrando os nós Suscetíveis e Infectados. Ambas as representações estão para a combinação ( $\beta = 0,2, \lambda = 0,05$ ). (a) <i>PrimarySchool</i> . (b) <i>HighSchool</i> . . . . .	40



# Lista de tabelas

Tabela 1 – Propriedades das redes. ‘APT’ se refere à quantidade de arestas por <i>timestamp</i> . ‘RT’ se refere à resolução temporal (em minutos). . . . .	31
Tabela 2 – Dados quantitativos da rede <i>Primary School</i> considerando a média de dez execuções para cada nível de comunicação intercomunitária. ‘ATI’ se refere à média do tempo de infecção. ‘MTI’ se refere à mediana do tempo de infecção. . . . .	33
Tabela 3 – Dados quantitativos da rede <i>High School</i> considerando a média de 10 execuções para cada combinação $(\beta, \lambda)$ . ‘ATI’ se refere à média de tempo de infecção. ‘MTI’ se refere à mediana do tempo de infecção. . .	36

# Lista de abreviaturas e siglas

MAT	Mapa de atividade temporal
MSV	<i>Massive Sequence View</i>
CNO	<i>Community Node-ordering</i>
SI	Suscetível-Infetado
SIS	Suscetível-Infetado-Suscetível
SIR	Suscetível-Infetado-Recuperado
SIRS	Suscetível-Infetado-Recuperado-Suscetível
RFID	<i>Radio Frequency Identification</i>
MP	<i>mathematics and physics</i> - matemática e física
PC	<i>physics and chemistry</i> - física e química
PSI	<i>engineering</i> - engenharia
BIO	<i>biology</i> - biologia
APT	Arestas por <i>timestamp</i>
RT	Resolução temporal
ATI	<i>Average time infection</i> - Média do tempo de infeccção
MTI	<i>Median time infection</i> - Mediana do tempo de infeccção
Inf.	Infetado
Não-Inf.	Não infetado

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>1.1</b>	<b>Objetivos</b>	<b>12</b>
<b>1.2</b>	<b>Estrutura dos próximos capítulos</b>	<b>12</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>13</b>
<b>2.1</b>	<b>Redes Complexas</b>	<b>14</b>
2.1.1	Comunidades	14
2.1.2	Redes temporais	15
<b>2.2</b>	<b>Visualização da informação</b>	<b>16</b>
2.2.1	Visualização de Redes Temporais	17
2.2.2	<i>Visual Clutter</i> (poluição visual)	19
<b>2.3</b>	<b>Modelos de espalhamento de doenças</b>	<b>24</b>
<b>2.4</b>	<b>Trabalhos correlatos</b>	<b>25</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>29</b>
<b>3.1</b>	<b>Metodologia</b>	<b>29</b>
<b>3.2</b>	<b>Estudos de Caso</b>	<b>30</b>
3.2.1	Redes temporais utilizadas	30
3.2.2	Objetivos e definição de parâmetros	31
3.2.3	<i>Primary School</i>	32
3.2.4	<i>High School</i>	36
3.2.5	Discussão e Limitações	39
<b>4</b>	<b>CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS</b>	<b>41</b>
	<b>REFERÊNCIAS</b>	<b>43</b>

# 1 Introdução

Entender a dinâmica pela qual ocorre o espalhamento de uma doença ajuda a saber como preveni-la ou regulá-la, por exemplo, por meio de isolamento social (DARSIE; WEBER, 2020), esquemas de vacinação, identificação da sua origem (país ou paciente zero) e rastreamento de contatos (visando maior testagem na população). Isso se torna ainda mais importante em períodos de pandemia, como a vivenciada atualmente com o coronavírus (COVID-19) e que é objeto de estudo de vários pesquisadores, tanto no Brasil (FRANCÊS et al., 2020), quanto em outros países (FANELLI; PIAZZA, 2020).

Do ponto de vista computacional, diferentes abordagens podem ser utilizadas para o estudo do espalhamento de doenças e comportamentos de epidemias, como, por exemplo, análises envolvendo autômatos celulares (MELOTTI, 2009), utilização de modelo estocástico híbrido (MARQUES, 2015) e até por meio de equações diferenciais ordinárias (BARROS, 2007). Outra abordagem, útil nesse contexto, é a utilização de redes complexas (ROCHA; BLONDEL, 2013; GEMMETTO; BARRAT; CATTUTO, 2014).

As redes complexas (FIGUEIREDO, 2011) visam a representação das instâncias do domínio em questão e o relacionamento entre elas na forma de nós e arestas, respectivamente. Assim, no nosso contexto de espalhamento de doenças, pode-se estudar as comunicações (também chamadas de interações ou conexões) entre os indivíduos da rede e o impacto causado por elas (NEWMAN, 2018). Alguns exemplos de redes complexas são: a interação entre alunos de uma escola (como o caso da nossa pesquisa), entre médicos num hospital, relação entre usuários de uma rede social, etc. Assim, redes que descrevem interações cara-a-cara (*face-to-face*) entre os indivíduos permitem observar como o espalhamento da doença ocorre.

Porém, para estudar o espalhamento de doenças é necessário mais do que saber quem se conectou com quem, mas, também, quando essas conexões ocorreram. Redes que possuem essa informação temporal são chamadas Redes Dinâmicas ou Redes Temporais (HOLME; SARAMÄKI, 2012). A rede temporal possui por característica sua estrutura dinâmica, isto é, ela possui entrada e saída de nós e arestas ocorrendo ao longo do tempo e de forma irregular, com períodos de ociosidade e períodos com alta concentração de conexões. Por ser capaz de representar padrões reais de interações entre indivíduos, ela representa uma importante ferramenta para o estudo de espalhamento de doenças (ROCHA; BLONDEL, 2013). Além de podermos saber quais foram os participantes do espalhamento da doença, também será de conhecimento onde a doença surgiu e qual caminho (sequência de conexões) foi tomado por ela.

Assim, esse contexto pode ser estudado de diferentes maneiras, sendo a análise

visual uma das principais (LINHARES et al., 2019a). A análise visual segue os princípios da Visualização da Informação (WARE, 2019), que usufrui de ferramentas como imagens, gráficos, estruturas etc, para auxiliar na interpretação dos dados apresentados, já que os dados por si só são abstratos, além de oferecer ferramentas interativas para que o usuário seja ativo no processo de exploração e validação dos dados. Logo, aplicando essas técnicas em redes temporais, temos uma otimização para identificar padrões e, assim, melhorar tomadas de decisão (LINHARES et al., 2019b; ZHAO et al., 2018; ELZEN et al., 2013b). No nosso caso, todos esses benefícios citados, afetam positivamente na análise do espalhamento de doenças (LINHARES et al., 2019a).

## 1.1 Objetivos

O objetivo geral deste trabalho é utilizar redes temporais para simular diferentes graus de isolamento entre indivíduos e verificar, utilizando técnicas de visualização da informação, como ocorre o espalhamento de doenças em cada cenário. Para isso, tem-se os seguintes objetivos específicos:

1. Simular diferentes cenários epidemiológicos (diferentes probabilidades de transmissão e graus de isolamento) e analisar visualmente como ocorre o espalhamento de doenças.
2. Demonstrar, utilizando estudos de caso com dados reais, como os diferentes graus de isolamento e probabilidades de transmissão afetam o espalhamento de doenças.

## 1.2 Estrutura dos próximos capítulos

- O Capítulo 2 apresenta uma fundamentação teórica para se contextualizar na pesquisa. Serão apresentadas definições envolvendo as Redes Complexas, estratégias que as abordam por meio da Visualização da Informação e, por fim, como o espalhamento de doenças pode ser simulado e analisado neste contexto.
- O Capítulo 3 apresenta o desenvolvimento da pesquisa. Sendo assim, apresentando a metodologia utilizada e os estudos de caso. Nos estudos de caso, são definidos os parâmetros para os experimentos, as redes com dados reais que foram utilizadas e, por fim, a apresentação dos resultados obtidos.
- O Capítulo 4 conclui este trabalho. Nele, são retomados os problemas abordados, depois são apresentadas as soluções propostas para tais, destacando os objetivos específicos, e, por fim, apresenta-se os trabalhos futuros.

## 2 Fundamentação Teórica

Este capítulo aborda os temas: Redes Complexas, Visualização da Informação e Modelos de Espalhamento de Doenças, dando suas definições e teorias desses assuntos importantes para essa pesquisa, como ilustrado na Figura 1. A Seção 2.1 define e exemplifica redes complexas, já a Seção 2.2 define a Visualização da Informação e mostra recursos para visualização do contexto desta pesquisa. Na Seção 2.3, é abordada a definição de alguns dos modelos de espalhamento de doenças existentes na literatura e como eles procedem. Por fim, na Seção 2.4, são apresentados trabalhos correlatos a esta pesquisa.

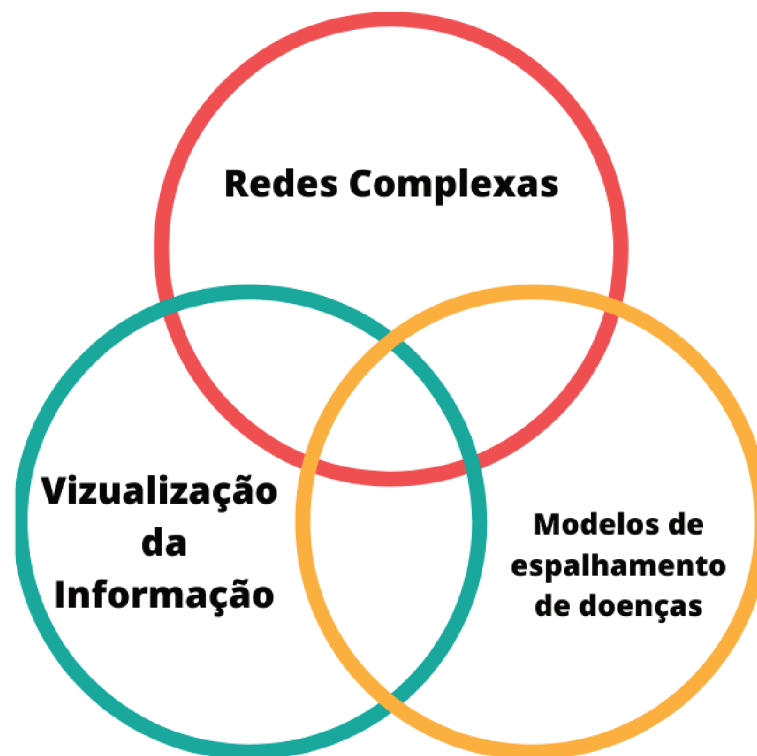


Figura 1 – Relação dos conceitos abordados nesta pesquisa.

Como ilustrado na Figura 1, os temas abordados neste trabalho têm intersecções entre eles, ou seja, ao aplicar seus recursos em conjunto, tem-se a delimitação do escopo dessa pesquisa. Nesse caso, necessita-se das **redes complexas**, para representar os dados a serem analisados, da **visualização da informação** para que a exploração desses dados seja feita visualmente e de forma interativa e, também, dos **modelos de espalhamento de doenças**, para que seja possível simular e analisar como a dinâmica de infecções ocorre sobre os dados considerados.

## 2.1 Redes Complexas

Redes complexas são definidas por um conjunto  $R(N,A)$ , onde  $N$  representa o conjunto de nós (instâncias) e  $A$  representa o conjunto de arestas (interações) que ligam esses nós entre si. Como exemplo, na Figura 2 pode-se observar que os pontos A, B, C, D e E são os nós da rede, enquanto as linhas que os ligam são as arestas. Assim, obtemos uma estrutura que, no caso deste trabalho, representa a interação entre instâncias de um sistema ou aplicação, por exemplo, rede de roteadores transitando pacotes, alunos conversando no intervalo das aulas etc. Neste contexto, será adotado tal definição para representar as Redes Complexas.

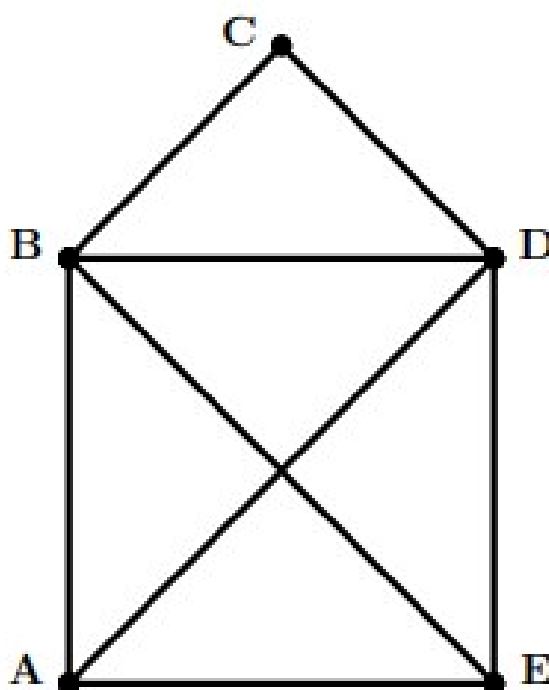


Figura 2 – Exemplo visual de uma rede. Pontos representam nós e linhas representam arestas (interações entre os nós).

Assim, para compreender a função do sistema representado, a interação entre os nós da rede servem de base (NEWMAN, 2018). Também, é válido ressaltar que, por definição, o grau de um nó se dá pelo número de interações que ele realiza. Partindo desse princípio de que as redes estabelecem interação entre instâncias de um sistema por meio dos nós e as arestas que os unem, o conceito de comunidade pode, agora, ser abordado.

### 2.1.1 Comunidades

Alguns nós dentro da rede interagem mais entre si do que com os demais, sendo assim, esse grupo é denominado **comunidade** (FORTUNATO; HRIC, 2016). Como pode-se observar na Figura 3, há 13 nós e 3 comunidades sendo representadas: a dos nós azuis,

a dos nós verdes e a dos nós vermelhos. As ligações entre nós de mesma comunidade são representadas por arestas de cor preta e as ligações entre nós de comunidades diferentes são representadas por arestas vermelhas. Assim, as ligações representadas, nesse exemplo, pelas arestas pretas serão chamadas de comunicação **intra**comunitárias, já as vermelhas, de **inter**comunitárias.

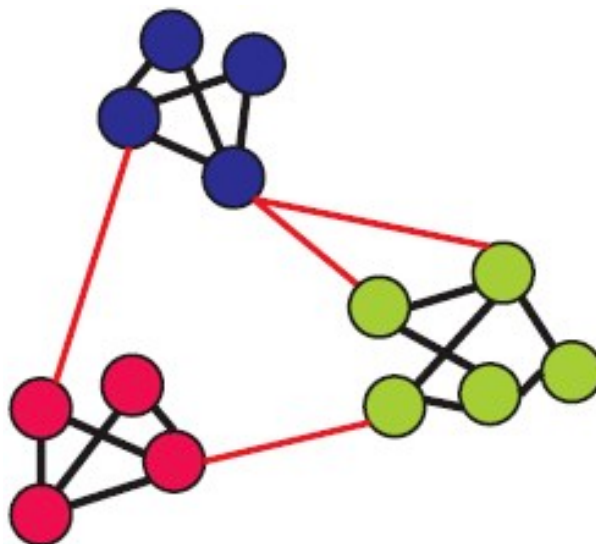


Figura 3 – Rede com 13 nós e 3 comunidades representadas. Figura adaptada de [Ponciano et al. \(2020c\)](#).

A etapa de detecção de comunidades é importante, porém difícil ([FORTUNATO; HRIC, 2016](#)). Na pesquisa de ([FORTUNATO; HRIC, 2016](#)), algoritmos são avaliados e comparados com diferentes tipos de detecção, sendo o **Infomap** ([ROSVALL; BERGSTROM, 2008](#)) e o **Louvain** ([BLONDEL et al., 2008](#)) dois dos métodos mais recomendados por apresentarem melhor qualidade nos resultados e tempo de resposta (complexidade computacional  $O(n)$ , em que  $n$  é o número de arestas da rede complexa) ([LINHARES et al., 2020](#)). As comunidades e a evolução das mesmas, no sistema em análise, são importantes, pois permitem estudar e analisar como essas mudanças afetam o sistema. E assim, para representar esses fenômenos evolutivos são utilizadas as redes temporais, que serão abordadas a seguir.

### 2.1.2 Redes temporais

Como dito anteriormente, em alguns casos é necessário estudar a evolução do sistema em questão. Para abordar essa questão, existem as redes temporais, que são estruturas dinâmicas em que suas mudanças (inserção e/ou remoção de nós e/ou arestas) ocorrem ao longo do tempo ([HOLME; SARAMÄKI, 2012](#)). Dessa forma, torna-se possível identificar alguns comportamentos, por exemplo, ritmo circadiano (que são um conjunto de fenômenos fisiológicos ou comportamentais que ocorrem sucessivamente a cada certo



período de tempo) e períodos de maior atividade (*burst interactions* - número expressivo de interações num determinado instante de tempo comparado a todos os outros do período analisado) (HOLME; SARAMÄKI, 2012).

Na Figura 4 podemos observar um exemplo de rede temporal representando 3 instantes de tempo. As comunicações são representadas por uma tupla  $(X,Y,t)$  de nós e um instante de tempo, ou seja, há uma aresta não-direcionada que liga X a Y no instante de tempo t, representando que aquela aresta está ativa naquele momento t.

Sendo assim, na Figura 4 (Tempo 1) pode-se observar que existem  $(A,B,1)$ ,  $(A,C,1)$  e  $(B,C,1)$ . Já na Figura 4(Tempo 2), ocorreram algumas mudanças, sendo elas: a adição de  $(A,D,2)$ ,  $(A,E,2)$  e  $(C,E,2)$ , ocasionando na aparição dos nós D e E no sistema, e a remoção do nó B, por não possuir conexões com os nós do sistema nesse instante. E, ao fim da rede - Figura 4 (Tempo 3), tem-se a estrutura formada pelos nós A, F e G com as arestas  $(F,G,3)$  e  $(A,G,3)$ .

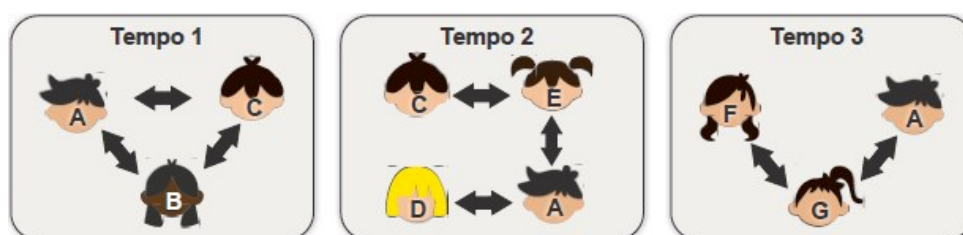


Figura 4 – Exemplo de uma rede temporal contendo 3 instantes de tempo e as interações que ocorrem em cada um deles. Figura adaptada de Linhares et al. (2016)

As redes temporais analisadas nas pesquisas atualmente, por exemplo (LINHARES et al., 2017; ROCHA et al., 2017; MELOTTI, 2009; FRANCÊS et al., 2020; FANELLI; PIAZZA, 2020), têm quantidades elevadas de nós e arestas, o que faz com que a análise, tabular ou por meio de análises visuais simples como a apresentada no exemplo anterior, seja inviável. Nesse sentido, técnicas de visualização mais robustas podem ser empregadas para construir *layouts* que irão auxiliar no estudo do comportamento desses dados. A seção a seguir apresenta, de maneira sucinta, a área de Visualização da Informação e algumas de suas técnicas.

## 2.2 Visualização da informação

Pesquisas com dados em grandes proporções, como uma rede de alunos de uma escola, ou funcionários de uma empresa multinacional e até um mapeamento de neurônios do cérebro humano podem ser complicadas para conseguir analisar sem um certo tipo de auxílio. Para isso, a **Visualização da Informação** (WARE, 2019) ajuda na interpretação desses dados abstratos, por meio de ferramentas de representação como: imagens, gráficos, estruturas etc. Ramos como Computação, Bioquímica, Medicina, Biologia, Matemática

etc, utilizam técnicas de Visualização da Informação aplicadas em redes de dados. No nosso contexto, esse conjunto de técnicas auxiliares proporciona meios para identificar padrões que possam existir em redes temporais relacionadas ao espalhamento de doenças infecciosas e, com isso, agentes do governo e profissionais da saúde, por exemplo, poderiam, a partir dos padrões e comportamentos encontrados, instruir melhor a população referente aos hábitos de prevenção recomendados.

### 2.2.1 Visualização de Redes Temporais

Nas Seção 2.1.2, foi apresentado que as redes temporais abordam problemas de evolução do sistema estudado. Porém, na Seção 2.2 nota-se que as redes de dados de grande escala necessitam de técnicas que construam layouts para melhor visualização dos dados. Assim, nesta Seção será abordado algumas técnicas de visualização da informação usadas em redes temporais, afim de ajudar a reconhecer padrões, comportamentos, anomalias etc.

As técnicas de visualização de redes temporais comumente se beneficiam das **Leis de Gestalt**, que mostram como os seres humanos percebem os elementos visuais (WARE, 2019). Três principais dessas leis são consideradas neste contexto. A primeira é a de **fechamento**, revelando que o cérebro identifica e forma imagens completas a partir de imagens inacabadas ou apenas silhuetas das mesmas, assim, em casos de presença mais densa de nós no layout, pode-se concluir uma taxa de conexão expressiva, mesmo tendo as arestas descartadas (lembrando que: os nós apenas aparecem na representação se manterem ao menos uma aresta). A segunda é a **proximidade**, dizendo que mesmo elementos sendo distintos, se forem posicionados próximos, tendem a ser enxergados juntos, assim, formando apenas uma unidade, logo, deixando um espaço entre cada comunidade, pode-se perceber que nós mais agrupados fazem parte da mesma comunidade. Por fim, a terceira é a **semelhança**, mostrando que objetos com aparência semelhante tendem a ser percebidos como um só, por exemplo, nós com cores diferentes representam atributos diferentes no sistema. Desse modo, será apresentado a seguir algumas técnicas para visualização de redes temporais.

Uma importante característica das técnicas de visualização de redes temporais é que elas devem preservar o mapa mental do usuário o máximo possível, isto é, elas devem facilitar ao máximo a percepção das mudanças que ocorrem na rede a cada tempo/período, sejam elas inserções ou retiradas de nós e/ou arestas. Para tal, as diferenças que ocorrerem entre cada tempo consecutivo devem ser o mais sutis possíveis (LIN; LEE; YEN, 2011). Assim, a posição dos nós, logo também as arestas, precisam ser preservada o máximo possível (SAFFREY; PURCHASE, 2008). Dito isso, será apresentado agora, algumas técnicas para visualização das redes temporais:

- A *tabular* consiste em representar os dados em uma tabela exibindo os nós que se conectam e o período de tempo em que houve essa interação. A desvantagem é a dificuldade para realizar análise, principalmente quando há muitos dados representados, já que é uma representação textual (Figura 5(a)).
- A *estrutural* (LINHARES et al., 2016) é uma representação por meio de diagramas de nós e arestas (Ex.: Figura 5(b)). Ela facilita a análise de maneira global (mostrando o resultado final do sistema) e, também, mostra ações ocorridas a cada instante de tempo quando utilizada de maneira animada (Figura 4). O layout estrutural não possui boa preservação do mapa mental, especialmente quando aplicado em redes grandes ou animações muito rápidas (LINHARES et al., 2019a).
- O *layout temporal* ou MSV (*Massive Sequence View*) (CORNELISSEN et al., 2007), mostra a rede de maneira cronológica, focando nos instantes de tempos por meio de um intervalo definido. Assim, em cada instante de tempo representado, pode-se analisar as instâncias ativas e as interações ocorridas naquele momento. Nessa representação os instâncias dos nós são colocados em paralelo ao eixo das abscissas (lembrando que, a posição dos nós não podem ser alteradas ao longo do tempo) e os instantes de tempo em paralelo ao eixo das ordenadas. Quando surge uma aresta na rede, uma linha vertical é desenhada, ligando os dois nós envolvidos (Ex: Figura 5(c)). Ressaltando que, em casos de múltiplas conexões no mesmo instante de tempo, pode haver sobreposição de arestas. Aliás, de acordo com (LINHARES et al., 2019a), o MSV apresenta melhor preservação do mapa mental do que o layout estrutural animado.

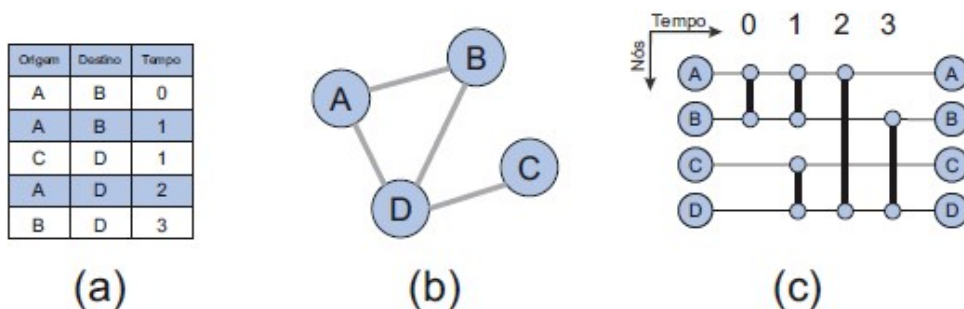


Figura 5 – Exemplos dos tipos de layouts de rede. Sendo (a) a tabular, (b) a estrutural e (c) o MSV. Figura adaptada de Linhares et al. (2017)

Algumas características podem ser implementadas tanto na técnica de visualização **estrutural** quanto no **MSV** na tentativa de otimizar a análise visual, que são as definições de **cor**, **tamanho** e **forma**, onde:

- as **cores** podem representar o grau de interação de cada nó, a comunidade que cada nó participa, etc;

- o **tamanho** refere-se à aumentar ou diminuir o símbolo que representa o nó;
- e a **forma** para escolher como representar o nó geometricamente na rede (geralmente são usados quadrados ou círculos).

Tendo isso, é válido apresentar agora o **Mapa de Atividade Temporal (MAT)**. Nele, cada nó é um **quadrilátero**, logo, a noção de **continuidade** (segunda das três Leis de Gestalt apresentadas no início da Seção 2.2.1) é destacada (WARE, 2019). A **cor** dos nós varia a cada instante de acordo com sua **participação** no sistema analisado em termos de número de arestas envolvendo ele naquele instante, assim, é possível haver uma percepção dos que realizam maior taxa de atividade. E as **arestas deixam de ser exibidas**, ou seja, o MAT foca na intensidade da **atividade dos nós**, conseqüentemente eliminando a poluição visual por arestas em excesso e suas sobreposições, como em períodos de alto índice de atividade.

A Figura 6 ilustra a correspondência entre os layouts MSV e MAT. Na Figura 6(a), está representado o layout **MSV** mostrando os nós e as arestas (lembrando que as arestas neste exemplo estão curvadas por motivos didáticos, assim, evitando sobreposições e deixando visível quais nós estão se comunicando). Já a Figura 6(b) mostra o mesmo layout, porém, com o esquema de cores nos nós, representando a atividade de cada um, isto é, o número de conexões que cada nó realiza naquele instante de tempo. E, por fim, a Figura 6(c) apresenta o **MAT**, onde as **arestas** foram ocultadas, o **formato** dos nós são quadriláteros e utiliza-se o esquema de **cores** para representar o grau de atividade dos nós.

Por fim, lembramos que, os layouts estrutural e MSV podem apresentar grande número de sobreposições de informação, no caso do MSV sendo sobreposição de arestas e no estrutural podendo ser tanto nós quanto arestas. A próxima Seção apresenta mais formalmente o que é a poluição visual (do inglês *Visual Clutter*) e algumas possíveis soluções.

### 2.2.2 *Visual Clutter* (poluição visual)

Dependendo do tamanho da rede sendo analisada, em termos de quantidade de nós e arestas, os layouts estrutural e MSV podem apresentar alto grau de poluição visual (*visual clutter*). No contexto de visualização de redes complexas, poluição visual refere-se ao grande número de informações sobrepostas em uma imagem (ELLIS; DIX, 2007), podendo ser grandes quantidades de arestas sobrepostas por exemplo e, assim, dificultando a análise que se baseia em qual nó realizou conexão e com quem.

Tanto na Figura 7, quanto na Figura 8, podemos notar que há uma poluição visual causada pela grande densidade de arestas em determinadas situações. Na primeira, devido a alta concentração de comunicação em um certo período da rede, enquanto na segunda é

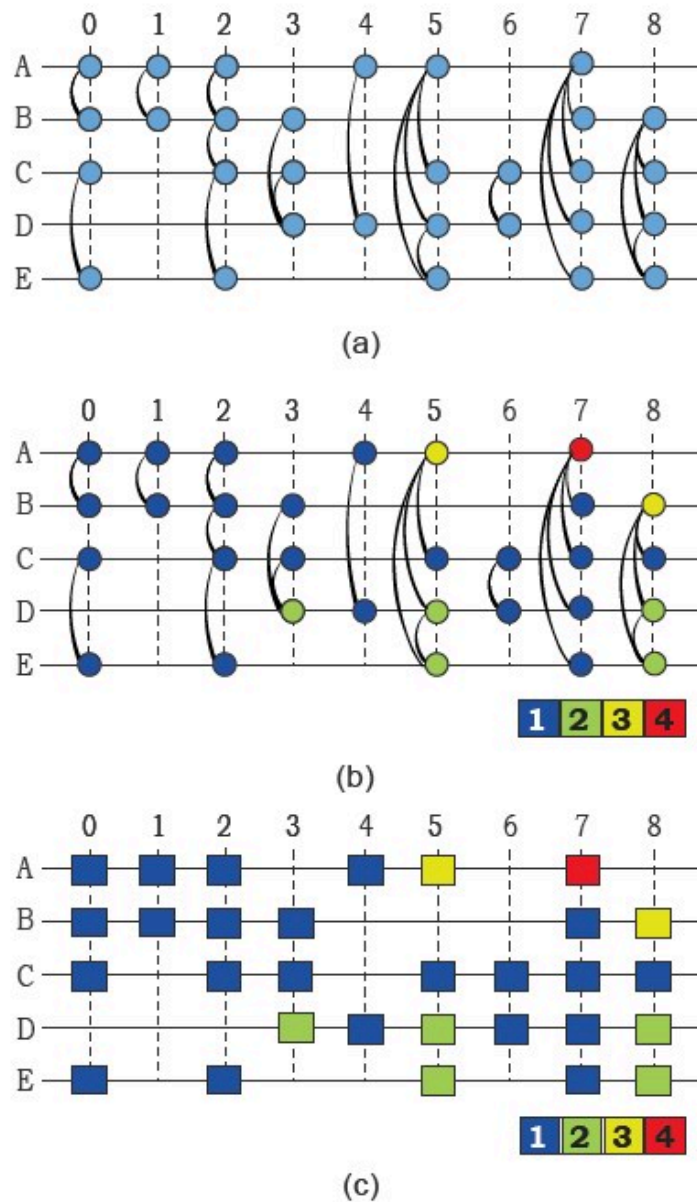


Figura 6 – Exemplos de layouts temporais, onde em (a) mostra o MSV, tendo nós circulares e arestas à mostra. Em (b) também, porém com esquema de cores. E, em (c), aplicado o MAT. Os números representados por cores dizem respeito à quantidade de conexões que o nó realizou naquele instante. Em (a) e (b), as arestas são exibidas curvadas para fins didáticos. Figura adaptada de (LINHARES et al., 2016).

devido ao posicionamento dos nós no layout, que, mesmo sendo de maneira esparsa, ainda gera uma grande sobreposição.

Para reduzir a poluição visual, diferentes estratégias podem ser adotadas. Para o layout estrutural, estratégias envolvem reposicionar os nós seguindo uma abordagem **circular**, que organiza os nós formando um polígono regular (BAUR; BRANDES, 2004), **hierárquica**, onde os nós são colocados verticalmente e suas arestas conectando-os de



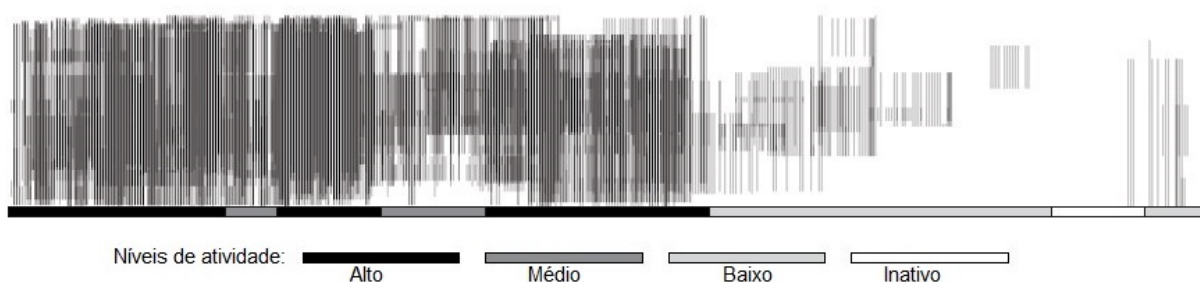


Figura 7 – Exemplo de poluição visual no MSV, com legenda representando níveis de atividade no decorrer do tempo. Figura adaptada de (LINHARES et al., 2019a)

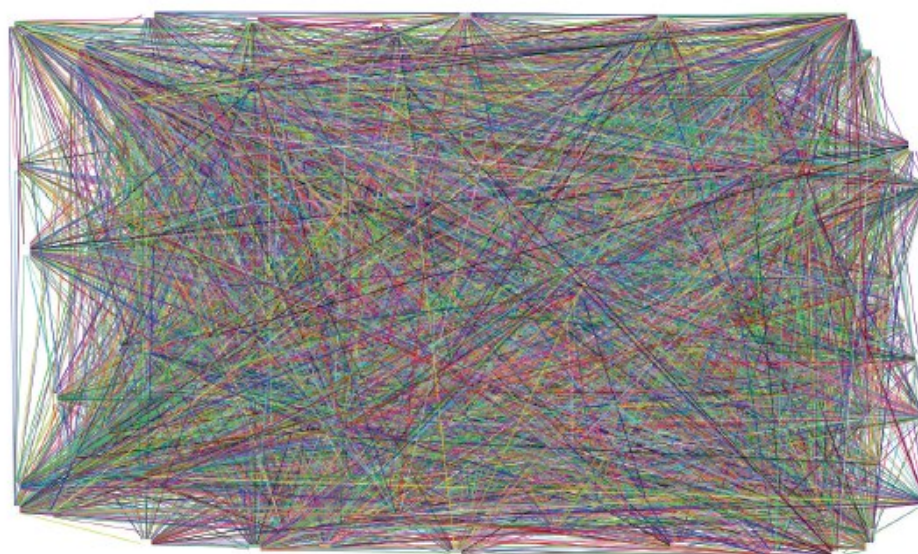


Figura 8 – Exemplo de poluição visual no layout estrutural não animado. Figura adaptada de (LINHARES et al., 2019a)

cima para baixo (BASTERT; MATUSZEWSKI, 2001) ou baseada em **força**, em que os nós são como anéis e as arestas são como molas, onde as molas atraem os anéis se estiverem longe, ou repelem se estiverem perto e, a posição final dos nós representa o sistema de molas em equilíbrio (ZHANG et al., 2005).

Para MSV incluem amostragem de arestas (de forma a exibir apenas arestas consideradas relevantes) (ZHAO et al., 2018; LINHARES et al., 2019b). Em outras, o objetivo é reposicionar os nós de modo a reduzir o tamanho das arestas e suas sobreposições (ELZEN et al., 2013b; LINHARES et al., 2019b), assim, reduzindo a poluição. Alguns dos métodos de reordenação de nós no MSV disponíveis na literatura são:

- **nascimento** (ELZEN et al., 2013b): os nós são posicionados de acordo com a ordem que nascem (o primeiro a aparecer na rede fica em cima e os últimos em baixo)(Figura 9(a));

- **lexicográfico** (ELZEN et al., 2013b): os nós são posicionados em ordem crescente de acordo com o tipo de identificação (por exemplo, numericamente ou alfabeticamente etc.)(Figura 9(b));
- **por grau** (ELZEN et al., 2013b): nós com menor grau na rede agregada são colocados na parte superior da rede, enquanto os de maior grau na parte inferior;
- **por vizinhos recorrentes** (LINHARES et al., 2016): o nó com maior grau fica no centro, então dois nós que mais se conectam a ele são posicionados acima e abaixo dele, por ordem de quem conectou primeiro a ele. Depois, o algoritmo é repetido nas partes superiores e inferiores do grafo, sempre buscando nós com maior conectividade com o nó em questão. Caso não haja conexões com o mesmo, o próximo nó que tiver maior grau de conectividade é selecionado (Figura 9(c)).

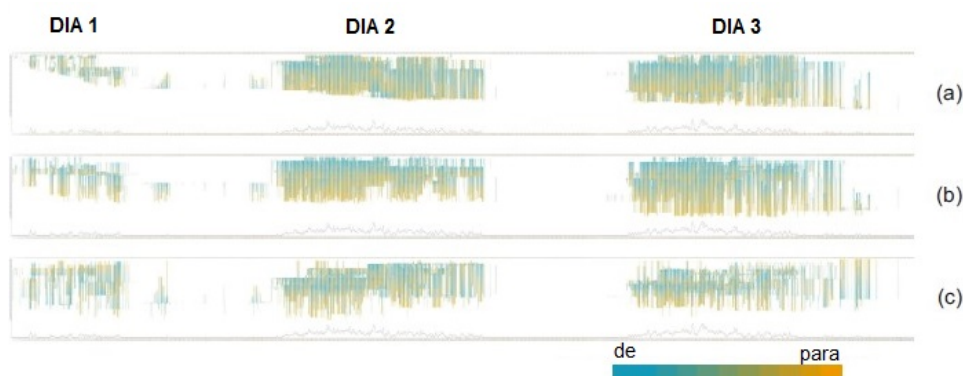


Figura 9 – Layout MSV de uma rede de hospital (VANHEMS et al., 2013), ordenado de 3 maneiras diferentes. (a) por nascimento, (b) lexicograficamente e (c) por vizinhos recorrentes. Figura adaptada de (LINHARES et al., 2017).

Nenhuma das estratégias mencionadas, no entanto, são visualmente escaláveis. Em outras palavras, todas possuem limitações quando aplicadas em redes de larga escala (em termos do número de nós e arestas). O método *Community-based node ordering* (CNO) (LINHARES et al., 2019b) apresenta-se como um método visualmente escalável e pode, portanto, ser aplicado em redes de variados tamanhos. O CNO realiza uma combinação de um algoritmo de detecção de comunidades com o de reordenação dos nós para assim organizar os nós no layout de forma a facilitar a identificação de MSV ou MAT. Este método é composto por três principais etapas:

1. **Detectar as comunidades** na rede agregada (considerando todos os instantes de tempo) utilizando algum método de detecção existente, como o Louvain ou o Infomap.
2. **Organizar a ordem das comunidades**, utilizando um dos algoritmos de reordenação de nó, porém cada comunidade é considerada como se fosse um nó. Assim,

reduzindo o tamanho das arestas entre as comunidades, logo, destacando-as de maneira mais limpa.

3. **Organizar a ordem dos nós dentro das comunidades**, também utilizando um dos algoritmos de reordenação de nós, com o propósito de melhorar o layout apresentado, aproximando os nós que tem mais conexões entre si dentro da comunidade.

A Figura 10 mostra um exemplo didático do funcionamento do algoritmo **CNO**. Essa representação mostra apenas as arestas, com fim didático de ser possível perceber a reorganização dos nós e a formação das comunidades ao longo das etapas. A Figura 10(a) apresenta o *layout MSV* original. Já as Figuras 10(b-d) ilustram a aplicação recursiva do CNO, decompondo as comunidades em  $N$  níveis (o valor de  $N$  é o número máximo possível de decomposições ou é escolhido pelo usuário respeitando esse limite). Por fim, a Figura 10(e) mostra apenas as arestas intercomunitárias (interação entre nós de comunidades diferentes) e a Figura 10(f) apenas as arestas intracomunitárias (interação entre nós de mesma comunidade).

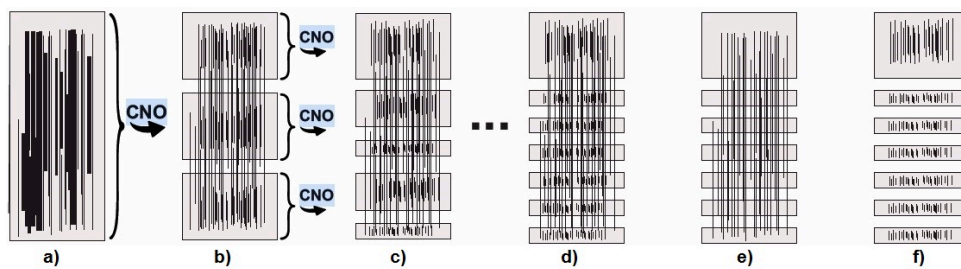


Figura 10 – Exemplo didático do funcionamento do algoritmo *Community-based Node Ordering*. Figura adaptada de (LINHARES et al., 2019b)

A Figura 11 mostra como é o resultado final aplicado em uma rede real de hospital. A Figura 11(a) mostra apenas as arestas intercomunitárias enquanto a Figura 11(b) exibe as arestas intracomunitárias.

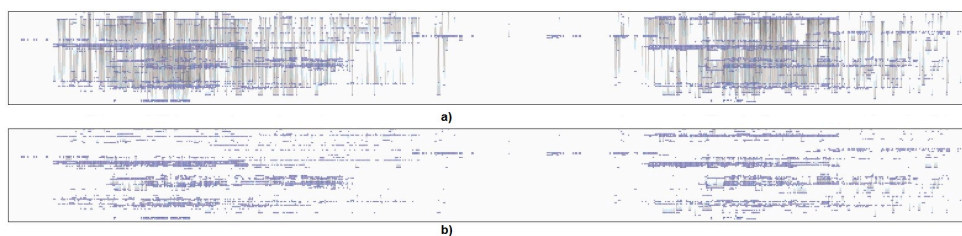


Figura 11 – *Community-based Node Ordering* aplicado em uma rede de hospital, usando as técnicas Louvain e vizinhos recorrentes. (a) exibindo apenas arestas intercomunitárias. (b) exibindo apenas arestas intracomunitárias. Figura adaptada de (LINHARES et al., 2019b)

Neste trabalho, utilizaremos os layouts MSV e MAT, juntamente com métodos de ordenação de nós como o CNO, para simular e analisar o processo de espalhamento



de doenças. A próxima seção apresenta diferentes modelos de espalhamento de doenças e introduz como a visualização de redes temporais é útil nesse contexto.

## 2.3 Modelos de espalhamento de doenças

As redes temporais são capazes de representar padrões reais de interações entre indivíduos, sendo assim uma importante ferramenta para o estudo do espalhamento de doenças infecciosas. Em complemento a isso, na literatura são abordados modelos para representar esse espalhamento, sendo alguns deles: SI, SIS, SIR e SIRS, em que S representa **Suscetível**, I de **Infectado** e R de **Recuperado**. Nessa pesquisa será utilizado o modelo SI.

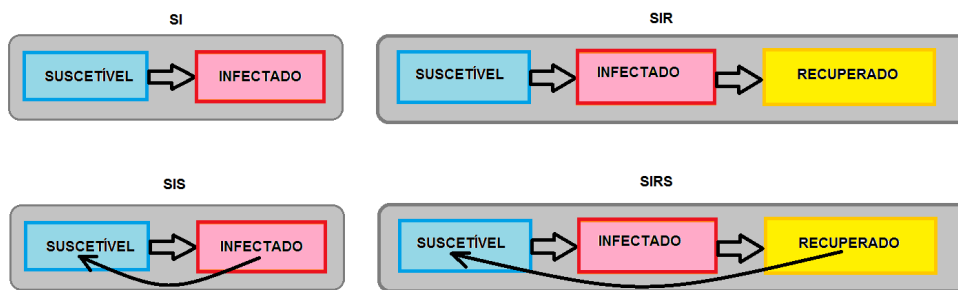


Figura 12 – Ilustração dos modelos do tipo espalhamento de doença considerados neste trabalho. S = “Suscetível”, I = “Infectado” e R = “Recuperado”

Explicando os modelos de espalhamento, o SI e SIS são dois modelos em que há dois estados de nó, podendo estar **suscetível** ou **infectado**. No **SI** (Figura 12 - SI), os nós iniciam em estado suscetível e têm a possibilidade de serem alterados para o estado infectado caso realizem conexão com nós que estejam em tal estado, ou seja, não basta apenas realizar conexão com infectados, pois há uma probabilidade de serem infectados ou não quando realizam essa interação. Além disso, nesse modelo, uma vez que o estado do nó foi alterado para infectado, ele permanece assim permanentemente (simulando uma doença sem cura). Já no **SIS** (Figura 12 - SIS), quando o nó for para o estado infectado, depois de um período determinado ele volta para o estado de suscetível, tornando-se um modelo em ciclo, ou seja, podendo repetir os estados de maneira ordenada.

Os modelos SIR e SIRS possuem um detalhe a mais, que é a aparição do estado **recuperado**. No **SIR** (Figura 12 - SIR) os nós iniciam no estado suscetível, possuem a chance de passarem para o estado infectado e, por fim, caso sejam infectados, depois de um período determinado passam para o estado recuperado e permanecem assim permanentemente (imunidade permanente). Enquanto no **SIRS** (Figura 12 - SIRS) depois de passar pelos mesmos processos que no SIR, no fim, em vez de obter imunidade permanente, o nó pode voltar ser suscetível novamente após um período de tempo (imunidade temporária), assim, tornando-se mais um modelo em ciclo.

## 2.4 Trabalhos correlatos

Ultimamente, novas ferramentas no intuito de prever futuras epidemias e criar visualização com a informação das taxas de espalhamento de infecções têm surgido, por exemplo, o painel interativo proposto por [Dong, Du e Gardner \(2020\)](#), que rastreia casos de COVID-19 relatados. Além disso, [Carroll et al. \(2014\)](#) apresenta o Sistema de Informação Geográfica (*Geographic Information Systems*), que identifica casos particulares, tendências espaciais e *clusters* de doença (Figura 13).

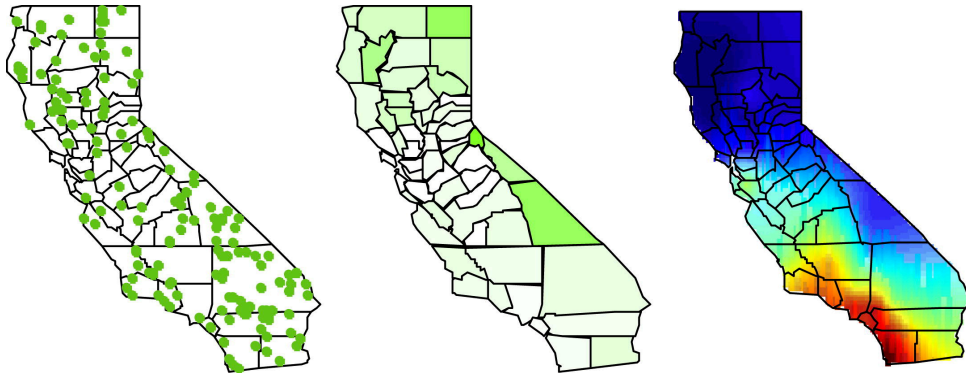


Figura 13 – Figura representando os modelos apresentados pelo Sistema de Informação Geográfica. O mapa de pontos (à esquerda) os utiliza para representar uma certa medida ou característica exibida sobre um mapa geográfico. No meio, é representado um mapa coroplético, que usa cores ou sombras graduadas para indicar os valores de alguma medida agregada em áreas especificadas. À direita, tem o mesmo tipo de mapa, todavia com o objetivo de derivar estatisticamente um gradiente de risco espacial. Figura adaptada de ([CARROLL et al., 2014](#))

[Linhares et al. \(2019a\)](#) mostram como a visualização dos processos dinâmicos na rede auxiliam a visualizar os caminhos do espalhamento das doenças infecciosas (Figura 14), incluindo o destacamento das principais transmissões e, também, na evolução dos estados dos nós (Figura 15). Nesse trabalho, é abordado problemas como a poluição visual causada por alta densidade de nós e sobreposições de arestas nos layouts e, por fim, chegando a algumas técnicas de reordenação de nós, exclusão de arestas e coloração do nós, atingindo o layout MAT, que foca na atividade dos nós e destaca seus respectivos estados ao longo do tempo (como explicado na Seção 2.2.1).

Em 2015, [Dunne et al. \(2015\)](#) propuseram o VoroGraph, uma ferramenta de análise visual que considera diferentes estratégias de visualização, se baseando em redes e dados geográficos visando auxiliar pesquisadores da área de epidemiologia. Além disso, [Gan e Thiebaut \(GAN; THIEBAUT, 2017\)](#) propuseram um layout radial, em que o usuário identifica indivíduos envolvidos no espalhamento da infecção, rastreia contatos e analisa como uma intervenção influencia os cenários avaliados. Mais recentemente, a visualização de rede temporal também foi usada no estudo do risco de transmissão do COVID-19

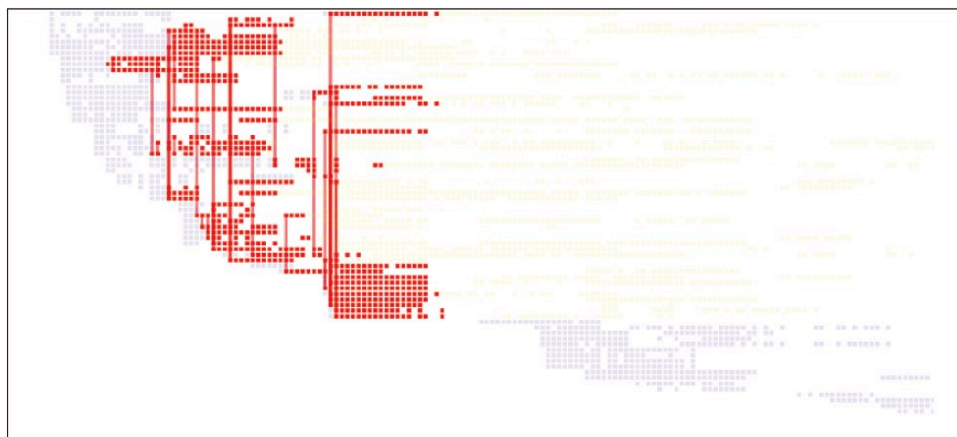


Figura 14 – Demonstração da árvore de transmissão. Figura adaptada de (LINHARES et al., 2019a)

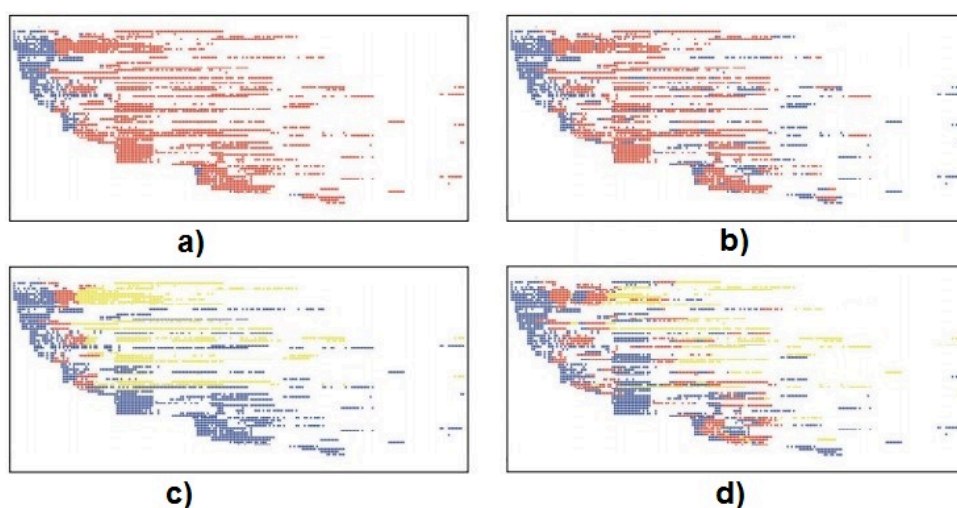


Figura 15 – Mapa de Atividade Temporal ilustrando os diferentes modelos de espalhamento de doenças. (a) SI, (b) SIS, (c) SIR, (d) SIRS. A cor azul indica o estado de Suscetível, vermelho indica Infectado e, amarelo indica Recuperado. Figura adaptada de (LINHARES et al., 2019a)

em diferentes países (por exemplo, Coreia do Sul e China), muitas vezes por meio de visualizações espaciais e/ou pelo layout estrutural (PARK, 2020; SO et al., 2020).

Como dito na Seção 2.2.1, MSV e MAT são utilizados para auxiliar a identificar padrões tanto estruturais quanto temporais em redes de diversos domínios, por exemplo, hospitais (LINHARES et al., 2019b; LINHARES et al., 2017), escolas (PONCIANO et al., 2020a; PONCIANO et al., 2020b), e outros (PEREIRA et al., 2019; ELZEN et al., 2013a). Mas, recentemente esses layouts foram considerados visando a análise de espalhamento de doenças infecciosas (LINHARES et al., 2019a), porém, não foram aplicados em cenários que consideram diferentes níveis de distanciamento social.

Para Chang et al. (2020), o distanciamento social acontece quando contatos envol-

vendo indivíduos fora do grupo familiar são restringidos enquanto interações domésticas permanecem inalteradas. Além dos membros de uma família, outros grupos baseados em fatores sociológicos podem ser considerados, por exemplo, indivíduos em um local de trabalho ou em uma classe escolar, e assim, serem observados como possibilidade em análise de espalhamento de doenças infecciosas. Grupos como esses representam informações úteis e podem ser modelados como comunidades de rede (KITCHOVITCH; LIO, 2010), como explicado anteriormente na Seção 2.1.1. E, mesmo que analisar a estrutura da comunidade da rede (por exemplo, a respeito do número e a da densidade de contatos inter e intra-comunitários) seja crucial para definir estratégias eficazes de imunização, poucos estudos consideram essa informação (GHALMANE; HASSOUNI; CHERIFI, 2019).

Na Seção 2.3 foram apresentados alguns modelos de espalhamentos e, utilizando um deles (SIS), Saito, Silva e Alves (2008) desenvolveram uma pesquisa baseada em autômatos celulares para identificar o espalhamento e evolução da pneumonia em uma população (Figura 16), onde dado os parâmetros de probabilidades, os nós mudam de estado de acordo com seu estado atual. No trabalho foram utilizados quatro redes contendo 2500, 1600, 900 e 400 indivíduos a fim de relatar um estudo envolvendo pequenas comunidades. Além disso, teve um período de 400 dias a fim de buscar um equilíbrio do sistema observado por meio de visualização e análise.

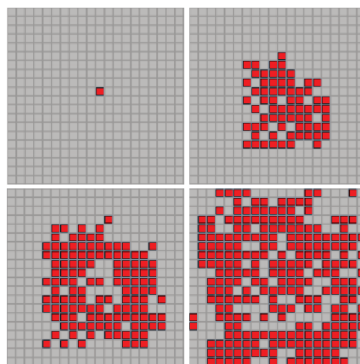


Figura 16 – Simulação de epidemia de pneumonia utilizando autômatos celulares. Figura adaptada de (SAITO; SILVA; ALVES, 2008)

Assim, analisando esses trabalhos correlatos, podemos identificar um foco em comum deles com o presente trabalho de conclusão de curso nos aspectos de identificação de comunidades e abordagem no que elas podem oferecer à análise de rede (CHANG et al., 2020), representações visuais para facilitar a análise e reconhecimento de padrões (DONG; DU; GARDNER, 2020; CARROLL et al., 2014; LINHARES et al., 2019a; DUNNE et al., 2015), uso de modelo de espalhamento aplicado a doenças infecciosas (SAITO; SILVA; ALVES, 2008), dentre outros.

Todavia, nosso estudo se faz importante no sentido de aplicar abordagens utilizando diferentes técnicas visuais (MSV e MAT) e cenários variados de infecção. São con-

siderados dois principais aspectos para nossa abordagem, sendo eles: os diferentes níveis de distanciamento social (por exemplo, se os indivíduos, de certa forma, evitam interações com os demais), em que no nosso experimento esse nível é definido por meio do número de interações intercomunitárias, e as diferentes probabilidades de se transmitir uma infecção (por exemplo, se medidas de proteção como o uso de equipamentos de proteção individual são adotadas pelos indivíduos). Assim, demonstraremos a importância dessa abordagem por meio de estudos de caso sobre duas redes, qual representam interações do mundo real.

## 3 Materiais e Métodos

Nesse capítulo será apresentado a metodologia que foi utilizada para a produção do experimento e em seguida os estudos de caso com seus devidos resultados, discussões e limitações para o estudo.

### 3.1 Metodologia

Utilizamos algumas técnicas de visualização citadas anteriormente. Dentre elas, usufruímos do MSV e também do MAT, juntos do algoritmo de ordenação de comunidades CNO. E, para aplicar essas técnicas e desenvolver os experimentos, foi utilizado uma ferramenta gratuita e *open-source* chamada *DyNetVis* (Linhares et al., 2020), que foi estendida para simular diferentes níveis de isolamento por meio da remoção de arestas inter-comunitárias obtidas com o CNO. Por fim, aplicamos essas técnicas de visualização nos layouts gerados pela ferramenta aderindo ao modelo de espalhamento de doença SI (Suscetível-Infetado).

Como dito anteriormente, foram utilizados dois principais aspectos para analisar e entender como uma doença infecciosa se espalha por meio dos nós de uma rede temporal. Sendo o primeiro a respeito das interações entre os nós existentes, tomamos por base a abordagem de Chang et al. (2020) e realizamos uma associação entre o número de interações com o distanciamento social.

Em nosso estudo, isso é demonstrado mantendo as condições estabelecidas pelo CNO para as interações intracomunitárias e apenas alterando o nível de comunicação intercomunitária. Para isso, é definido uma variável  $\beta$  (variando de 0 a 1) que controla quantas arestas intercomunitárias são mantidas, sendo elas escolhidas de maneira aleatória, enquanto as demais são removidas. Por exemplo, caso  $\beta$  seja 0, todas as arestas intercomunitárias são removidas e, assim, seria um caso de isolamento total (também conhecido pelo termo *full lockdown*), o que é algo que foge da prática no mundo real devido aos serviços que são essenciais para a sociedade (CANABARRO et al., 2020). Por outro lado, caso todas as arestas intercomunitárias sejam mantidas, ou seja, não há qualquer nível de isolamento social, então  $\beta = 1$ .

Já o segundo aspecto se refere à probabilidade que um nó no estado Suscetível tem de ser infectado ao interagir com outro no estado Infectado, definida por meio de uma variável  $\lambda$  (com valores de 0 a 1, onde 1 representa certeza de transmissão ao interagir).

Assim, depois de aplicar o algoritmo CNO na rede e ter definido a quantidade de arestas intercomunitárias que permanecerão existentes, deve ser escolhido um paciente

zero e um instante de tempo, representando, respectivamente, o indivíduo que iniciará o processo de espalhamento e o momento em que isso ocorrerá. A escolha do paciente zero pode ser arbitrária ou por meio de métricas de centralidade, como grau, proximidade (*closeness*), ou intermediação (*betweenness*) (PEREIRA; AMO; GAMA, 2016).

Logo após, é aplicado o modelo de espalhamento SI (Suscetível-Infectado) iniciando a partir do nó escolhido e com a probabilidade de transmissão ( $\lambda$ ) também definida. Por fim, os layouts produzidos (tanto MSV como MAT) permitem identificar, por exemplo, quais foram os nós que transmitiram a infecção dentro da comunidade e quais transmitiram para outras comunidades, verificar se há algum comportamento relevante durante o espalhamento (como uma elevação expressiva na taxa de infecções em um pequeno espaço de tempo, comparado aos demais), entre outros.

Além disso, a fim de obter uma avaliação quantitativa mais precisa, os nossos resultados se baseiam na média de dez execuções do procedimento acima descrito, utilizando as mesmas medidas adotadas para os dois aspectos, quantidade de arestas intercomunitárias habilitadas ( $\beta$ ) e probabilidade de transmissão ( $\lambda$ ) e, também, sempre partindo do mesmo paciente zero (nó inicial).

## 3.2 Estudos de Caso

Nessa Seção, serão apresentados dois casos de redes reais, sendo elas a *Primary School* (GEMMETTO; BARRAT; CATTUTO, 2014) e a *High School* (MASTRANDREA; FOURNET; BARRAT, 2015). Sendo assim, na Seção 3.2.1, são apresentadas as redes que foram utilizadas. Enquanto na Seção 3.2.2, são definidos e apresentados os parâmetros que serão utilizados para os experimentos e os resultados obtidos a partir dos mesmos. Nesses experimentos, buscamos resultados como: identificar padrões de comportamento, comparativos de quantidade de infectados (seja em um determinado período da rede ou no total), possíveis padrões de espalhamento, dentre outros. Por fim, na Seção 3.2.5 discutimos os resultados obtidos e, também, são apresentadas as limitações para esse trabalho.

### 3.2.1 Redes temporais utilizadas

Para esse experimento foram utilizadas duas redes temporais, a *Primary School* e a *High School*. Ambas serão apresentadas a seguir. Além disso, as principais características dessas redes podem ser observadas na Tabela 1.

A rede *Primary School* foi coletada na cidade de Lyon, França (GEMMETTO; BARRAT; CATTUTO, 2014) e consiste em interações de proximidade (coletadas por meio de sensores RFID que estavam em crachás, além disso, a proximidade era considerada de 1 a 1,5 metros entre os participantes) que ocorreram entre 242 participantes (nós), sendo



232 estudantes e 10 professores. Além disso, nessa rede há cinco séries escolares e cada uma delas divididas em duas salas (A e B), sendo um professor designado para cada uma dessas salas. Também, nessa rede há um total de 55.046 interações (arestas) distribuídas ao longo de dois dias.

A rede *High School* foi coletada na cidade de Marseille, França (MASTRANDREA; FOURNET; BARRAT, 2015) e também consiste em interações de proximidade, todavia com um total de 327 estudantes. Além disso, nessa rede há nove classes, divididas em quatro categorias: MP (*mathematics and physics* - matemática e física), PC (*physics and chemistry* - física e química), PSI (*engineering* - engenharia) e BIO (*biology* - biologia). Também, nessa rede há um total de 47.586 arestas distribuídas ao longo de 5 dias.

Para essas duas redes foi adotada uma resolução temporal de cinco minutos. A resolução temporal considera todas as arestas que ocorreram nesse intervalo, produzindo assim um *timestamp* para cada cinco minutos de duração da rede. Com isso, na *Primary School* há um total de 389 *timestamps* produzidos, enquanto na *High School* há um total de 1.211 *timestamps*.

Tabela 1 – Propriedades das redes. ‘APT’ se refere à quantidade de arestas por *timestamp*. ‘RT’ se refere à resolução temporal (em minutos).

Rede	#Nós	#Arestas	#Tempo	APT	#Salas	RT
<b>Primary Sch.</b>	242	55.046	389	141,5	10	5
<b>High Sch.</b>	327	47.586	1.211	39,29	9	5

### 3.2.2 Objetivos e definição de parâmetros

Nossos dois estudos de caso possuem objetivos diferentes. No primeiro, exploramos a rede *Primary School* para analisar como a velocidade de transmissão se comporta quando variamos o grau de distanciamento social ( $\beta$ ) para uma mesma probabilidade de infecção ( $\lambda$ ). Para essa análise, fixamos  $\lambda = 0,05$  (5% de chance de transmissão ao interagir um nó Infectado com um Suscetível - valor escolhido empiricamente) e variamos  $\beta$  entre 0 e 1 (passo 0,1), indo de *full lockdown* ( $\beta = 0$ ) para a situação onde não há isolamento ( $\beta = 1$ ).

Já em relação ao segundo estudo de caso, nosso objetivo é analisar combinações envolvendo diferentes graus de isolamento social e probabilidades de infecção e entender como elas influenciam o processo de espalhamento. Trabalhos como (CANABARRO et al., 2020), (FERGUSON et al., 2020), (CHANG et al., 2020) e (DAGHRIRI; OZMEN, 2020) defendem que um isolamento social adequado requer uma redução de, no mínimo, 75% ou 80% na quantidade de interações envolvendo pessoas de diferentes núcleos de convivência (os quais consideramos como comunidades em nosso contexto). Baseando-nos nesses trabalhos, para este estudo de caso, adotaremos uma taxa de isolamento social



de 80% (isto é,  $\beta = 0,2$  - lembre-se que  $\beta$  refere-se à quantidade de interações intercomunitárias que são mantidas), e, como contra-ponto,  $\beta = 0,8$  (redução de 20% no nível normal de interação). Por fim, consideraremos  $\lambda = 0,05$  e  $\lambda = 0,7$ , valores escolhidos empiricamente que visam simular diferentes níveis de proteção durante as interações na rede, como por exemplo, o uso (ou não) de máscaras.

Além disso, para cada caso analisado, o paciente zero foi escolhido como o nó mais ativo (com maior número de interações) da comunidade mais ativa, utilizando o algoritmo CNO. E, para a *Primary School*, o primeiro passo (detecção de comunidades) do CNO adotou o algoritmo *Infomap*, enquanto para a *High School* foi adotado o de *Louvain*, pois de acordo com Linhares et al. (2020) esses algoritmos foram os melhores para cada rede visando relacionar as comunidades com as turmas em cada sala de aula. Também, para definir tanto o posicionamento das comunidades, quanto dos nós na rede, foi adotada a técnica de Vizinhos Recorrentes (LINHARES et al., 2016), apresentada na Seção 2.2.2.

### 3.2.3 *Primary School*

Para obter resultados quantitativos referentes à rede *Primary School*, foram realizados experimentos com dez execuções para cada combinação  $(\beta, \lambda)$ , onde  $\beta$  variou de 0 a 1 com passo de 0,1 e  $\lambda$  foi mantida em 0,05 para todas as ocasiões. Nosso objetivo com esse experimento é analisar, dada uma probabilidade  $\lambda$  de infecção, como diferentes graus de distanciamento social afetam a evolução do espalhamento.

A Tabela 2 mostra os valores de uma média calculada para dez execuções de cada nível de isolamento ( $\beta$ ) na rede *Primary School*. A saber, para cada uma dessas dez execuções para cada  $\beta$ , foi utilizada uma *seed* diferente, mas essas mesmas dez *seeds* foram usadas em todos os níveis de  $\beta$ , para assim termos uma comparação mais fidedigna do experimento. E, para cada valor  $\beta$  considerado, apresenta-se: a média (ATI) e mediana (MTI) do instante de tempo de infecção e número total de infectados ( $\#Inf.$ ) e não infectados ( $\#N\grave{a}o-Inf.$ ), em que esse número total de infectados e não infectados foram medidos no último *timestamp*.

Como foi dito na Seção 3.1,  $\beta$  se refere à quantidade de arestas intercomunitárias que são mantidas, ou seja, quanto maior o  $\beta$ , menor o isolamento social. Os valores ATI e MTI são medidos em *timestamps* e cada *timestamp* representa 5 minutos da rede para a resolução temporal adotada. Analisar a variação dessas variáveis é útil no contexto estudado, pois nos dá uma noção da velocidade em que a infecção se espalha. O número de infectados ( $\#Inf.$ ) atingiu um máximo de 241 nós (tanto para  $\beta = 0,9$  quanto para  $\beta = 1$  - cenários onde quase não há e onde não há isolamento social, respectivamente) e o número de não infectados ( $\#N\grave{a}o-Inf$ ) atingiu um máximo de 219 nós (para  $\beta = 0$  - cenário de *full lockdown*)

Aliás, para observar unicamente a diferença no número de infectados ao variar o  $\beta$ , é apresentada a Figura 17, onde seu eixo vertical representa números de infectados totais na rede ( $\#Inf$ ) e o eixo horizontal representa os valores de  $\beta$  (nível de arestas intercomunitárias que foram mantidas - isolamento social indo de *full lockdown* para nenhum isolamento).

Tabela 2 – Dados quantitativos da rede *Primary School* considerando a média de dez execuções para cada nível de comunicação intercomunitária. ‘ATI’ se refere à média do tempo de infecção. ‘MTI’ se refere à mediana do tempo de infecção.

$\beta$	ATI	MTI	#Inf.	#Não-Inf.
0	70,51 $\pm$ 14,35	45	23	219
0,1	240,32 $\pm$ 31,33	272	202	40
0,2	162,77 $\pm$ 27,41	109	232	10
0,3	126,09 $\pm$ 43,1	101	237	5
0,4	115,40 $\pm$ 34,22	98	239	3
0,5	97,84 $\pm$ 14,09	70	238	4
0,6	89,73 $\pm$ 9,79	66	240	2
0,7	88,48 $\pm$ 13,6	65	240	2
0,8	81,09 $\pm$ 7,37	59	240	2
0,9	77,88 $\pm$ 9,77	55	241	1
1	78,77 $\pm$ 5,68	55	241	1

Por meio dos dados na Tabela 2 e do gráfico apresentado na Figura 17, ao comparar a situação de *full lockdown* ( $\beta = 0$ ) com quando tem pelo menos 10% das comunicações intercomunitárias habilitadas ( $\beta \geq 0,1$ ) notamos uma diferença mais expressiva no número de infectados quando comparamos cenários entre os intervalos de  $\beta = 0,1$  a  $\beta = 1$ . Isso mostra que o cenário de *full lockdown*, por mais que seja inviável na prática, seria um método importante para evitar a abrangência do espalhamento da doença.

Analisando-se a Figura 18 em paralelo com a Tabela 2, pode-se notar que houve um total de 23 infectados quando  $\beta = 0$ , que corresponde à totalidade da comunidade do paciente zero e somente ela. Tal comunidade representa a turma 3B da *Primary School*. Houve a contaminação de todos os alunos logo no primeiro dia, exceto do aluno identificado pelo número 1744 (indicado pelas setas na Figura 18), que faltou ao primeiro dia, mas, logo no início do segundo dia também foi infectado. Aliás, esse número total de infectados representa cerca de 9,5% do número total de participantes na rede (242).

Já observando a coluna de  $\beta$  da Tabela 2, variando os valores de 0,1 (Figura 19) a 1 (Figura 20), o número de infectados ( $\#Inf.$ ) varia de 202 a 241, ou seja, cerca de 83% a praticamente 100% dos nós da rede. Assim, percebemos que para situações em que o distanciamento social é quase completo ( $\beta = 0,1$ ) ou simplesmente não havê-lo ( $\beta = 1$ ), pelo menos 80% da rede será infectada, em que a diferença no número total de infectados

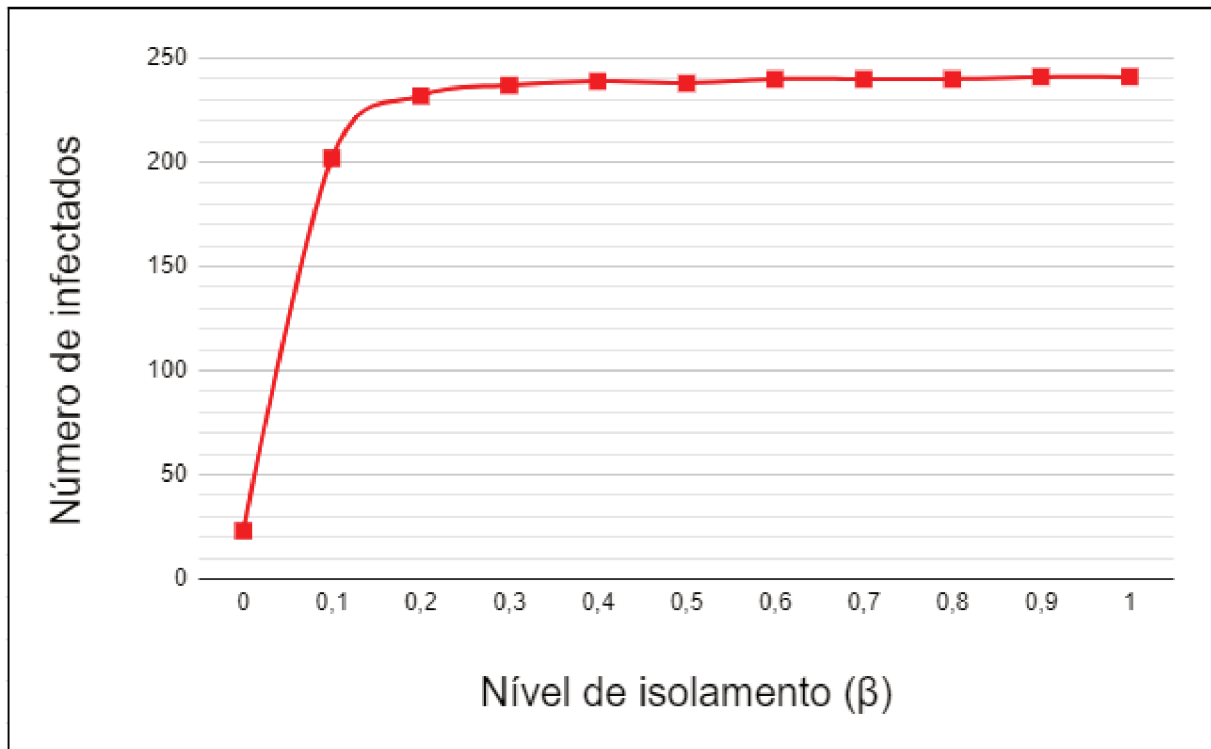


Figura 17 – Gráfico representando a evolução do número de infectados ao variar o valor de  $\beta$  de 0 a 1

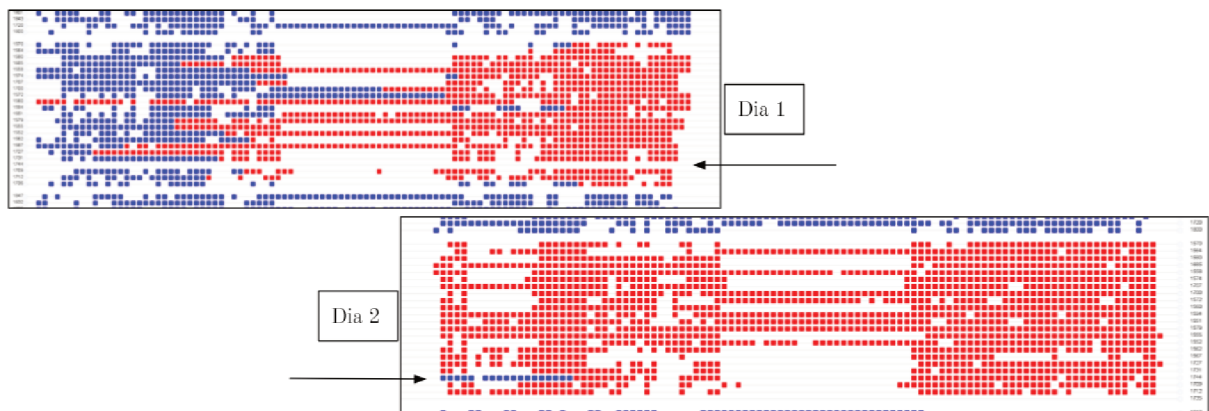


Figura 18 – Recorte do MAT para a situação de *full lockdown* ( $\beta = 0$ )

de ( $\beta = 0,1$ ) para ( $\beta = 1$ ) é cerca de 16% dos nós da rede.

No entanto, ao observarmos a coluna MTI (*Median time infection*) da Tabela 2, percebemos um retardo importante no tempo de infecção. Quando o valor de  $\beta = 0,1$ , mesmo que o número de infectados ( $\#Inf.$ ) seja quase o mesmo, o MTI é cerca de 5 vezes menor de quando  $\beta = 1$ , ou seja, quanto menos interações inter-comunitárias ( $\beta = 0,1$ ), mais a doença infecciosa demorará para se espalhar. Também podemos notar isso melhor dentro do intervalo indo de  $\beta = 0,6$  para  $\beta = 0,8$ , onde o número de infectados é o mesmo (240), porém o MTI reduz de 66 para 59 (a velocidade de transmissão aumenta). Essa

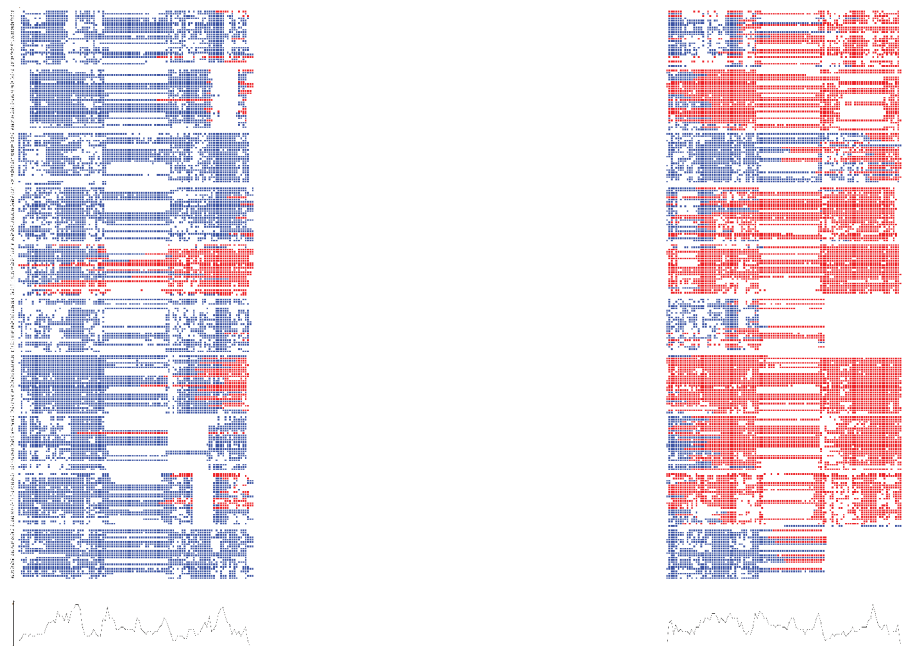


Figura 19 – MAT da combinação de 90% de distanciamento social ( $\beta = 0,1$ )

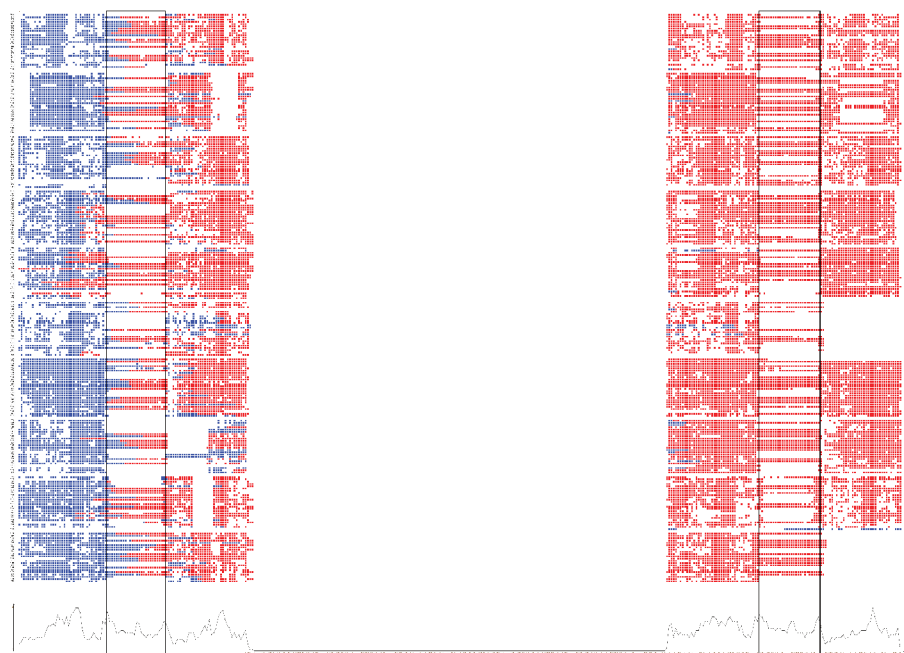


Figura 20 – MAT da combinação onde não há distanciamento social ( $\beta = 1$ )

percepção é importante em cenários reais, pois quanto mais doença demorar para se espalhar, maiores são as chances das entidades governamentais e da saúde tomarem medidas para reduzir ao máximo a continuidade do contágio, por exemplo, elaborando campanhas de vacinação e reforçando as campanhas de distanciamento. Com isso, auxiliando contra a lotação dos sistemas de saúde, mostrando assim, a importância do distanciamento social mesmo quando não há o *full lockdown*.

Outro fenômeno que podemos observar nessa rede é o aumento expressivo de in-

fecções que ocorreram no intervalo do 39º ao 64º *timestamp*, comparado com os demais *timestamps* da rede, a partir dos experimentos em que  $\beta \geq 0,3$ , já que, logo após esse intervalo todas as comunidades são afetadas. Esse intervalo corresponde ao horário de almoço e, assim, teve um número maior de comunicações entre indivíduos de classes distintas do que acontece eventualmente na rede. Esse mesmo nível de comunicação ocorre no dia 2 entre o 327º e o 353º *timestamp*, porém, para esses casos, a maioria dos nós já se encontram infectados. Ambos os intervalos foram destacados e podem ser observados na Figura 20.

### 3.2.4 High School

Para a *High School* também foram realizadas dez execuções para cada combinação  $(\beta, \lambda)$ , onde  $\beta$  recebeu os valores 0,2 e 0,8 e  $\lambda$  os valores 0,05 e 0,7. Os dados quantitativos a partir da média das dez execuções podem ser observados na Tabela 3, que segue as mesmas classificações de coluna da Tabela 2, porém, com a adição da coluna  $\lambda$  já que agora a mesma possui valor variável.

Tabela 3 – Dados quantitativos da rede *High School* considerando a média de 10 execuções para cada combinação  $(\beta, \lambda)$ . ‘ATI’ se refere à média de tempo de infecção. ‘MTI’ se refere à mediana do tempo de infecção.

$\beta$	$\lambda$	ATI	MTI	#Inf.	#Não-Inf.
<b>0,2</b>	<b>0,05</b>	680 ± 88,71	689	141	186
<b>0,2</b>	<b>0,7</b>	69,41 ± 27,95	29	327	0
<b>0,8</b>	<b>0,05</b>	722,06 ± 80	773	260	67
<b>0,8</b>	<b>0,7</b>	36,13 ± 3,01	21	327	0

Inicialmente, observando a coluna ‘#Inf.’ na Tabela 3, para as duas combinações em que  $\lambda = 0,7$ , todos os participantes (os 327 nós) foram infectados. Esse resultado era esperado, visto que a nova probabilidade de transmissão adicionada para o experimento nessa rede ( $\lambda = 0,7$ ) é 14 vezes maior que a outra ( $\lambda = 0,05$ ), então requer menos interações para ocorrer a transmissão. Aliás, para essas combinações, quase todos os nós passaram para o estado Infectado logo no primeiro dia (Figura 21 e Figura 22).

Por outro lado, quando adota-se  $\lambda = 0,05$  e variamos o  $\beta$ , notamos que reduzir o valor de  $\beta$  ocasiona na diminuição do número total de infectados (Tabela 3:#Inf.), indo de 260 para 141, sendo quase 2 vezes menor no número de infectados (essa diferença representa cerca de 37% dos nós da rede, ou seja, possivelmente significativa).

Também, na Tabela 3, mantendo o valor de  $\lambda$  e variando  $\beta$ , o MTI sofre uma diferença relativamente baixa comparado ao número total de *timestamps* que a rede possui (1211). Visto que, para  $\lambda = 0,05$ , o valor de MTI sofre uma diferença de 84 *timestamps* (de 773 para 689 - cerca de 7% do total) e para  $\lambda = 0,7$ , uma diferença de 8 *timestamps* (de 29 para 21 - cerca de 0,7% do total). Assim, levantamos uma hipótese de que comparar



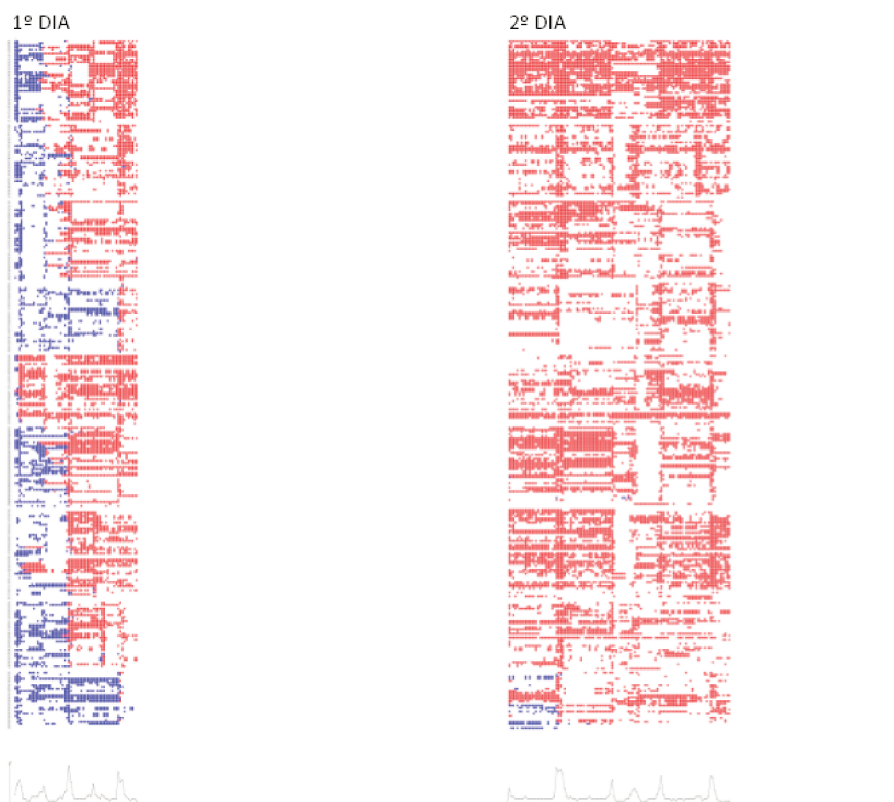


Figura 21 – Recorte do MAT para a combinação  $(\beta = 0,2, \lambda = 0,7)$

cenários (observando a tabela de dados) onde o valor  $\lambda$  é mantido, a observação da mediana do tempo de infecção (MTI), ou seja, visar a velocidade de transmissão, pode ser pouco relevante.

Se fixarmos o valor de  $\beta$  e variarmos apenas o  $\lambda$ , para ambos os casos, quanto menor o valor de  $\lambda$ , menor o número de infectados e menor a velocidade de transmissão também, lembrando que quanto maior o MTI, mais a doença demorou para se espalhar. Na prática, menores valores de  $\lambda$  podem ser associados ao uso de medidas preventivas de transmissão, como o uso de máscaras faciais.

Observando a coluna MTI na Tabela 3, podemos notar que para  $\beta = 0,2$ , ao variar  $\lambda$  de 0,7 para 0,05, a velocidade de transmissão diminuiu cerca de 24 vezes (indo de MTI = 29 para MTI = 689) e, para  $\beta = 0,8$ , a velocidade diminuiu cerca de 37 vezes (indo de MTI = 21 para MTI = 773).

Por fim, outro fenômeno que observamos é que estudantes de uma determinada matéria tendem a se comunicar mais com outros da mesma, sendo de turmas diferentes ou não. Analisando as Figuras 23 e 24, que correspondem às combinações  $(\beta = 0,2, \lambda = 0,05)$  e  $(\beta = 0,8, \lambda = 0,05)$ , utilizando um layout MSV que destaca os infectados e o caminho da transmissão da doença infecciosa, observamos que o paciente zero se encontra na turma

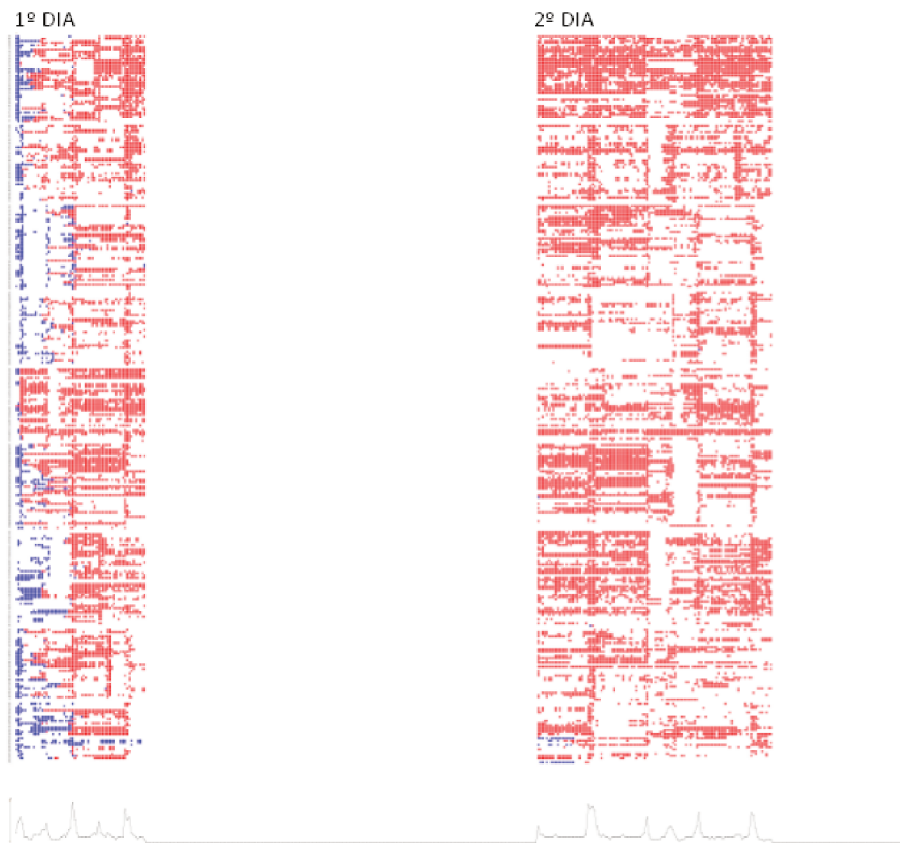


Figura 22 – Recorte do MAT para a combinação  $(\beta = 0,8, \lambda = 0,7)$

2BIO2 e que as turmas que se infectam na sequência são 2BIO1 e 2BIO3, variando apenas a ordem com que elas são infectadas dependendo da combinação  $(\beta, \lambda)$  analisada.



Figura 23 – Recorte do MSV para a combinação  $(\beta = 0,2, \lambda = 0,05)$ , destacando apenas os infectados e o caminho da transmissão da doença. A cor preta destaca a turma 2BIO2, a cor verde destaca a turma 2BIO3 e a cor azul destaca a turma 2BIO1. O recorte mostra apenas os três primeiros dias da rede.

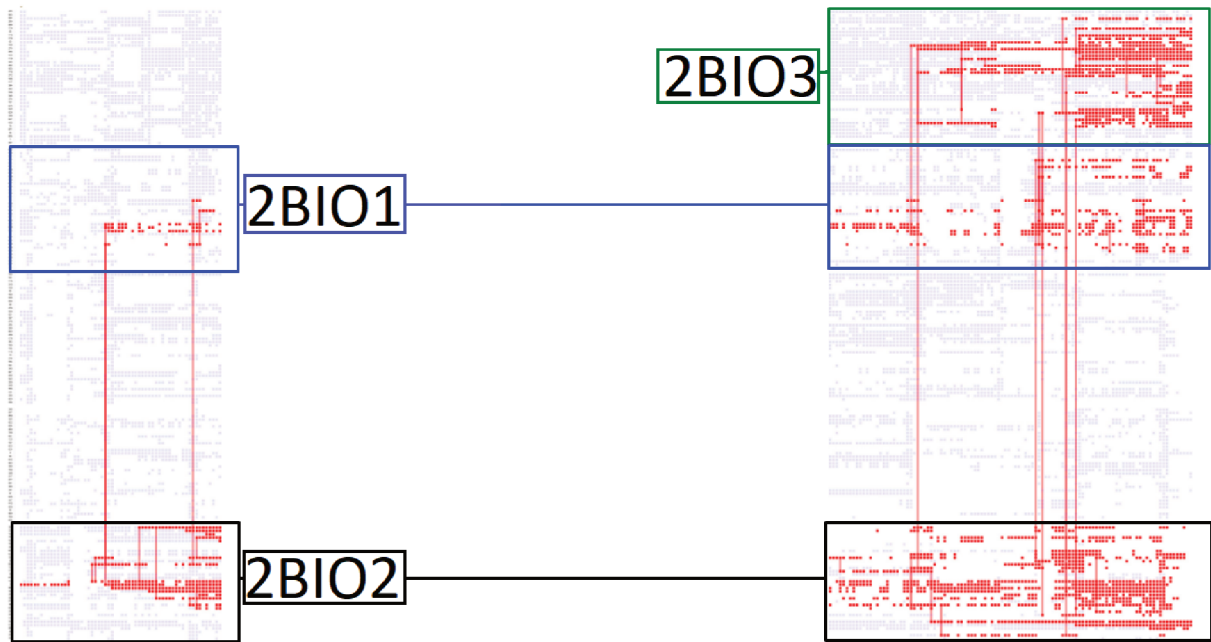


Figura 24 – Recorte do MSV para a combinação  $(\beta = 0,8, \lambda = 0,05)$ , destacando apenas os infectados e o caminho da transmissão da doença. A cor preta destaca a turma 2BIO2, a cor verde destaca a turma 2BIO3 e a cor azul destaca a turma 2BIO1. O recorte mostra apenas os dois primeiros dias da rede.

### 3.2.5 Discussão e Limitações

Realizamos experimentos baseados em transmissão de doenças infecciosas em duas redes reais. Do ponto de vista da visualização, as representações utilizadas são importantes para ter maior compreensão e gerar mais *insights* em relação aos dados apresentados, uma vez que permitem reconhecer padrões inesperados, fenômenos, anomalias que seriam difíceis de serem identificados de outra forma.

Analisando essas visualizações, percebemos que a propagação da doença primeiro se expande entre os alunos de turmas relacionadas entre si, seja turmas de uma mesma série (*Primary School*) ou de disciplinas semelhantes (*High School*). Além disso, a visualização permitiu identificar a ordem de infecção das turmas e até turmas que não foram infectadas.

Ao analisar esses caminhos que a doença infecciosa tomou para se espalhar, especialistas podem utilizar esses dados para retardarem e até evitar a continuidade da transmissão. Já que, a partir do layout que mostra os nós infectados e o caminho que a infecção tomou para se espalhar, é possível identificar quem infectou quem e quando ocorreu a infecção. Em ambientes escolares e até de trabalho, por exemplo, os coordenadores do local podem definir horários alternados para intervalos entre turmas diferentes, assim, evitando que a doença se espalhe entre nós de turmas distintas.

Mesmo nossas análises utilizando principalmente o MAT, outra possibilidade seria por meio da visualização estrutural mostrando o estado final dos nós (Figura 25). Os nós



que tiveram menos interações na rede são posicionados na região periférica do layout, caso seja adotado um algoritmo baseado em força (ZHANG et al., 2005). Esses nós têm menor chance de serem infectados até o fim da rede quando comparados aos que apresentam maior número de interações. Além disso, é mais fácil identificar a quantidade de nós infectados com o layout estrutural do que no temporal, assim, adotar diferentes layouts para visualização agrega na melhor análise do experimento.

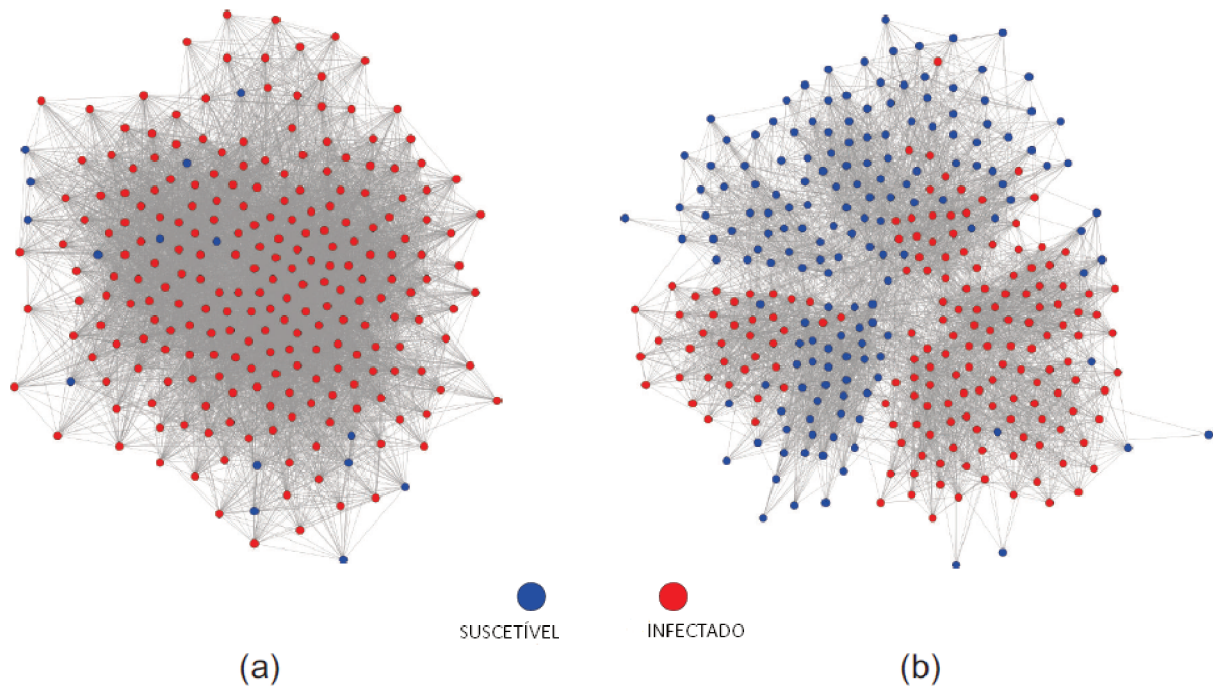


Figura 25 – Representação do último *timestamp* das redes utilizadas por meio do layout estrutural, mostrando os nós Suscetíveis e Infectados. Ambas as representações estão para a combinação ( $\beta = 0,2, \lambda = 0,05$ ). (a) *PrimarySchool*. (b) *HighSchool*.

Por fim, como vimos na Seção 3.2.4, comparando as combinações ( $\beta = 0,8, \lambda = 0,05$ ) e ( $\beta = 0,2, \lambda = 0,05$ ) para a *High School*, notamos que para a primeira combinação a velocidade de transmissão foi menor (maior ‘MTI’) que na segunda, mesmo ela apresentando menor grau de isolamento ( $\beta = 0,8$ ). Acreditamos que isso possa ter ocorrido devido ao grau médio dos nós após a filtragem das arestas intercomunitárias, podendo acarretar em mudanças significativas na estrutura da rede. Porém, essa questão não foi validada para esse trabalho. Outra hipótese é de que essa mudança não seja significativa, já que a diferença na velocidade de transmissão representa um intervalo de apenas 7% do total de *timestamps* da rede. Essas análises para validar tais suposições serão adotadas como propostas para um trabalho futuro.

## 4 Considerações Finais e Trabalhos Futuros

A análise de uma grande quantidade e variedade de dados pode ser complexa dependendo da maneira em que for feita. Uma das abordagens para esse problema é a utilização de redes complexas, em que os dados são representados pelas suas instâncias (nós) e interações (arestas). E, para auxiliar nesse contexto, este trabalho também apresentou estratégias de visualização da informação por meio de layouts temporais e com ferramentas interativas. Dessa forma, o usuário pode explorar e validar os dados a serem analisados. Além disso, usufruímos de diferentes métodos bem estabelecidos na literatura, como é o caso do *Community Node Ordering* (CNO) para ordenação de nós e *Louvain* e *Infomap* para detecção de comunidades.

Este trabalho teve como principal objetivo a simulação e análise de diferentes cenários epidemiológicos. Utilizamos redes temporais para representar as interações que poderiam (ou não) resultar em espalhamento de infecções entre os indivíduos envolvidos. Para a análise visual, tivemos um foco maior no Mapa de Atividade Temporal (MAT), mas também em conjunto do *Massive Sequence View* (MSV) para ter conhecimento de certos dados não mostrados pelo MAT, por exemplo, os caminhos de transmissões da doença infecciosa. Esses layouts possuem características que permitem abordar o contexto de infecção de uma forma útil, auxiliando assim na visualização e análise da informação. No caso desse experimento, no MAT, os estados dos nós são destacados em colorações definidas (vermelho para Infectado e azul para Suscetível, por exemplo), juntamente com os algoritmos CNO para melhor posicionar os nós no layout baseado em suas interações. Assim, principalmente com esses dois layouts, foi possível reconhecer melhor os grupos da rede, quem transmitiu a infecção para quem e quando essas transmissões ocorreram, de tal maneira que pudemos, também, reconhecer o caminho tomado pela doença infecciosa.

Logo, escolhida a maneira para representar os dados, o próximo objetivo foi simular e aplicar as estratégias de análise. Para a experimentação, primeiramente definimos certos graus de isolamento e probabilidades de transmissão da doença infecciosa. Após isso, identificamos as principais características para se analisar o espalhamento de uma doença transmissível, assim, os aspectos focados foram os estados dos nós (Suscetível ou Infectado) e as proporções de espalhamento ao longo do tempo, tudo isso combinado com os fatores de isolamento social. Como dito anteriormente, utilizamos uma ferramenta gráfica, *DyNetVis*, para simular os experimentos propostos. Nela, definimos os parâmetros de simulação, como o grau de isolamento social e a probabilidade de transmissão da doença infecciosa ao ocorrer uma interação entre dois nós. Com isso, a ferramenta fornecia o MAT para cada uma das combinações de parâmetros escolhidas, para que assim pudéssemos analisá-los.

Como os principais aspectos destacados foram os estados dos nós e suas interações, levantamos sugestões e hipóteses diante dos resultados, visto que ao longo do desenvolvimento da pesquisa, nos estudos de caso, foi mostrado a importância e notoriedade do uso de equipamentos de proteção individual e do alto grau de isolamento social diante de uma situação de epidemia.

Por fim, como trabalhos futuros, o objetivo será ampliar a nossa abordagem para novos níveis de probabilidade de transmissão, adotar diferentes modelos de espalhamento (SIS, SIR, SIRS) e validar situações de hipótese que foram levantadas, por exemplo, como o grau médio dos nós podem afetar o espalhamento por meio de mudanças na estrutura da rede e, qual o nível de proporção de *timestamps* pode ser considerado para que a mudança na velocidade de propagação seja considerada válida.

## Referências

- BARROS, A. M. R. de. Modelos matemáticos de equações diferenciais ordinárias aplicados à epidemiologia. *Revista de Ciências Exatas e Tecnologia*, v. 2, n. 2, p. 62–67, 2007. Citado na página 11.
- BASTERT, O.; MATUSZEWSKI, C. Layered drawings of digraphs. In: *Drawing graphs*. [S.l.]: Springer, 2001. p. 87–120. Citado na página 21.
- BAUR, M.; BRANDES, U. Crossing reduction in circular layouts. In: SPRINGER. *International Workshop on Graph-Theoretic Concepts in Computer Science*. [S.l.], 2004. p. 332–343. Citado na página 20.
- BLONDEL, V. D. et al. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, IOP Publishing, v. 2008, n. 10, p. P10008, 2008. Citado na página 15.
- CANABARRO, A. et al. Data-driven study of the covid-19 pandemic via age-structured modelling and prediction of the health system failure in brazil amid diverse intervention strategies. *medRxiv*, Cold Spring Harbor Laboratory Press, 2020. Citado 2 vezes nas páginas 29 e 31.
- CARROLL, L. N. et al. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *Journal of biomedical informatics*, Elsevier, v. 51, p. 287–298, 2014. Citado 3 vezes nas páginas 7, 25 e 27.
- CHANG, S. L. et al. Modelling transmission and control of the covid-19 pandemic in australia. *arXiv preprint arXiv:2003.10218*, 2020. Citado 4 vezes nas páginas 26, 27, 29 e 31.
- CORNELISSEN, B. et al. Understanding execution traces using massive sequence and circular bundle views. In: IEEE. *15th IEEE International Conference on Program Comprehension (ICPC'07)*. [S.l.], 2007. p. 49–58. Citado na página 18.
- DAGHRIRI, T.; OZMEN, O. Quantifying the effects of social distancing on the spread of covid-19. *Available at SSRN 3696983*, 2020. Citado na página 31.
- DARSIE, C.; WEBER, D. L. Doença e controle espacial: questões sobre dispersão e isolamento em tempos de pandemia. *Journal of Infection Control*, v. 9, n. 2, p. 47–48, 2020. Citado na página 11.
- DONG, E.; DU, H.; GARDNER, L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, Elsevier, v. 20, n. 5, p. 533–534, 2020. Citado 2 vezes nas páginas 25 e 27.
- DUNNE, C. et al. Vorograph: visualization tools for epidemic analysis. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. [S.l.: s.n.], 2015. p. 255–258. Citado 2 vezes nas páginas 25 e 27.

- ELLIS, G.; DIX, A. A taxonomy of clutter reduction for information visualisation. *IEEE transactions on visualization and computer graphics*, IEEE, v. 13, n. 6, p. 1216–1223, 2007. Citado na página 19.
- ELZEN, S. van den et al. Dynamic network visualization with extended massive sequence views. *IEEE transactions on visualization and computer graphics*, IEEE, v. 20, n. 8, p. 1087–1099, 2013. Citado na página 26.
- ELZEN, S. van den et al. Reordering massive sequence views: Enabling temporal and structural analysis of dynamic networks. In: IEEE. *2013 IEEE Pacific Visualization Symposium (PacificVis)*. [S.l.], 2013. p. 33–40. Citado 3 vezes nas páginas 12, 21 e 22.
- FANELLI, D.; PIAZZA, F. Analysis and forecast of covid-19 spreading in china, italy and france. *Chaos, Solitons & Fractals*, Elsevier, v. 134, p. 109761, 2020. Citado 2 vezes nas páginas 11 e 16.
- FERGUSON, N. M. et al. Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. imperial college covid-19 response team. *Imperial College COVID-19 Response Team*, p. 20, 2020. Citado na página 31.
- FIGUEIREDO, D. R. Introdução a redes complexas. *Atualizações em Informática*, p. 303–358, 2011. Citado na página 11.
- FORTUNATO, S.; HRIC, D. Community detection in networks: A user guide. *Physics reports*, Elsevier, v. 659, p. 1–44, 2016. Citado 2 vezes nas páginas 14 e 15.
- FRANCÊS, C. R. L. et al. Análise da evolução da pandemia de covid-19 no brasil—o estado do Pará. *Laboratórios de Tecnologias Sociais. Universidade Federal do Pará. Acesso em: <https://portalamazonia.com/images/p/27016/NOTA-TECNICA—Analise-da-evolucao-da-pandemia-de-COVID-19-no-Brasil-O-Estado-do-Para.pdf>*, 2020. Citado 2 vezes nas páginas 11 e 16.
- GAN, J.; THIEBAUT, D. Data visualization of agent-based simulation of an infectious spread. 2017. Citado na página 25.
- GEMMETTO, V.; BARRAT, A.; CATTUTO, C. Mitigation of infectious disease at school: targeted class closure vs school closure. *BMC infectious diseases*, BioMed Central Ltd, v. 14, n. 1, p. 695, dez. 2014. ISSN 1471-2334. Citado 2 vezes nas páginas 11 e 30.
- GHALMANE, Z.; HASSOUNI, M. E.; CHERIFI, H. Immunization of networks with non-overlapping community structure. *Social Network Analysis and Mining*, Springer, v. 9, n. 1, p. 45, 2019. Citado na página 27.
- HOLME, P.; SARAMÄKI, J. Temporal networks. *Physics reports*, Elsevier, v. 519, n. 3, p. 97–125, 2012. Citado 3 vezes nas páginas 11, 15 e 16.
- KITCHOVITCH, S.; LIO, P. Risk perception and disease spread on social networks. *Procedia Computer Science*, Elsevier, v. 1, n. 1, p. 2345–2354, 2010. Citado na página 27.
- LIN, C.-C.; LEE, Y.-Y.; YEN, H.-C. Mental map preserving graph drawing using simulated annealing. *Information Sciences*, Elsevier, v. 181, n. 19, p. 4253–4272, 2011. Citado na página 17.

- LINHARES, C. D. et al. Visualisation of structure and processes on temporal networks. In: *Temporal Network Theory*. [S.l.]: Springer, 2019. p. 83–105. Citado 8 vezes nas páginas 6, 7, 12, 18, 21, 25, 26 e 27.
- LINHARES, C. D. et al. A scalable node ordering strategy based on community structure for enhanced temporal network visualization. *Computers & Graphics*, Elsevier, v. 84, p. 185–198, 2019. Citado 6 vezes nas páginas 6, 12, 21, 22, 23 e 26.
- LINHARES, C. D. et al. Visual analysis for evaluation of community detection algorithms. *Multimedia Tools and Applications*, Springer, p. 1–23, 2020. Citado 2 vezes nas páginas 15 e 32.
- LINHARES, C. D. et al. Análise temporal de uma rede de contato hospitalar utilizando técnicas de visualização de informação. In: SBC. *Anais Principais do XVII Workshop de Informática Médica*. [S.l.], 2017. Citado 5 vezes nas páginas 6, 16, 18, 22 e 26.
- LINHARES, C. D. G. et al. Técnicas de análise visual de redes temporais. Universidade Federal de Uberlândia, 2016. Citado 6 vezes nas páginas 6, 16, 18, 20, 22 e 32.
- Linhares, C. D. G. et al. Dynetvis - an interactive software to visualize structure and epidemics on temporal networks. In: *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. [S.l.: s.n.], 2020. p. 933–936. Citado na página 29.
- MARQUES, F. S. *Modelo híbrido estocástico aplicado no estudo de espalhamento de doenças infecciosas em redes dinâmicas de movimentação de animais*. Tese (Doutorado) — Universidade de São Paulo, 2015. Citado na página 11.
- MASTRANDREA, R.; FOURNET, J.; BARRAT, A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PloS one*, Public Library of Science, v. 10, n. 9, p. e0136497, 2015. Citado 2 vezes nas páginas 30 e 31.
- MELOTTI, G. Aplicação de autômatos celulares em sistemas complexos: Um estudo de caso em espalhamento de epidemias. Universidade Federal de Minas Gerais, 2009. Citado 2 vezes nas páginas 11 e 16.
- NEWMAN, M. *Networks*. [S.l.]: Oxford university press, 2018. Citado 2 vezes nas páginas 11 e 14.
- PARK, J. Y. Spatial visualization of cluster-specific covid-19 transmission network in south korea during the early epidemic phase. *medRxiv*, Cold Spring Harbor Laboratory Press, 2020. Citado na página 26.
- PEREIRA, F. S.; AMO, S. de; GAMA, J. Detecting events in evolving social networks through node centrality analysis. In: *STREAMEVOLV@ ECML-PKDD*. [S.l.: s.n.], 2016. Citado na página 30.
- PEREIRA, F. S. F. et al. Uma análise sobre a evolução das preferências musicais dos usuários utilizando redes de similaridade temporal. *iSys-Revista Brasileira de Sistemas de Informação*, v. 12, n. 3, p. 94–115, 2019. Citado na página 26.
- PONCIANO, J. R. et al. An online and nonuniform timeslicing method for network visualisation. *arXiv preprint arXiv:2009.11422*, 2020. Citado na página 26.



- PONCIANO, J. R. et al. Visual analysis of contact patterns in school environments. *Informatics in Education*, Vilnius University Institute of Data Science and Digital Technologies, v. 19, n. 3, p. 455–472, 2020. Citado na página 26.
- PONCIANO, J. R. et al. Improving visual analysis of streaming networks. Universidade Federal de Uberlândia, 2020. Citado 2 vezes nas páginas 6 e 15.
- ROCHA, L. E. C.; BLONDEL, V. D. Bursts of vertex activation and epidemics in evolving networks. *PLOS Comput. Biol.*, v. 9, p. e1002974, 2013. Citado na página 11.
- ROCHA, W. A. A. d. et al. Aspectos de redes complexas com aplicações em neurociência. [sn], 2017. Citado na página 16.
- ROSVALL, M.; BERGSTROM, C. T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 105, n. 4, p. 1118–1123, 2008. Citado na página 15.
- SAFFREY, P.; PURCHASE, H. The "mental map" versus "static aesthetic" compromise in dynamic graphs: a user study. In: CITESEER. *Proceedings of the ninth conference on Australasian user interface-Volume 76*. [S.l.], 2008. p. 85–93. Citado na página 17.
- SAITO, Y.; SILVA, M. A. A. da; ALVES, D. Autômato celular para estudar espalhamento e evolução da pneumonia por pneumococos em uma população. 2008. Citado 2 vezes nas páginas 7 e 27.
- SO, M. K. et al. Visualising covid-19 pandemic risk through network connectedness. *International Journal of Infectious Diseases*, Elsevier, 2020. Citado na página 26.
- VANHEMS, P. et al. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PloS one*, Public Library of Science, v. 8, n. 9, p. e73970, 2013. Citado 2 vezes nas páginas 6 e 22.
- WARE, C. *Information visualization: perception for design*. [S.l.]: Morgan Kaufmann, 2019. Citado 4 vezes nas páginas 12, 16, 17 e 19.
- ZHANG, Q.-G. et al. Drawing undirected graphs with genetic algorithms. In: SPRINGER. *International Conference on Natural Computation*. [S.l.], 2005. p. 28–36. Citado 2 vezes nas páginas 21 e 40.
- ZHAO, Y. et al. Eod edge sampling for visualizing dynamic network via massive sequence view. *IEEE Access*, IEEE, v. 6, p. 53006–53018, 2018. Citado 2 vezes nas páginas 12 e 21.