
Técnicas de Análise Visual de Redes Temporais

Cláudio Douglas Gouveia Linhares



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2016

Cláudio Douglas Gouveia Linhares

Técnicas de Análise Visual de Redes Temporais

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Bruno Augusto Nassif Travençolo

Coorientador: José Gustavo de Souza Paiva

Uberlândia

2016

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

L755t Linhares, Cláudio Douglas Gouveia, 1990-
2016 Técnicas de análise visual de redes temporais / Cláudio Douglas
Gouveia Linhares. - 2016.
74 f. : il.

Orientador: Bruno Augusto Nassif Travençolo.
Coorientador: José Gustavo de Souza Paiva.
Dissertação (mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Ciência da Computação.
Inclui bibliografia.

1. Computação - Teses. 2. Redes de computadores - Teses. 3.
Processamento de dados - Teses. I. Travençolo, Bruno Augusto Nassif.
II. Paiva, José Gustavo de Souza, 1979-. III. Universidade Federal de
Uberlândia. Programa de Pós-Graduação em Ciência da Computação.
IV. Título.

CDU: 681.3

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada "**Técnicas de Análise Visual de Redes Temporais**" por **Cláudio Douglas Gouveia Linhares** como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**

Uberlândia, 18 de fevereiro de 2016

Orientador: _____
Prof. Dr. Bruno Augusto Nassif Travençolo
Universidade Federal de Uberlândia

Coorientador: _____
Prof. Dr. José Gustavo de Souza Paiva
Universidade Federal de Uberlândia

Banca Examinadora:

Prof. Dr. Anderson Rodrigues dos Santos
Universidade Federal de Uberlândia

Prof. Dr. Luis Enrique Correa Rocha
Karolinska Institute - Suécia

*A todos os professores do curso, que foram tão importantes
na minha vida acadêmica e no desenvolvimento desta dissertação.*

Agradecimentos

Aos meus pais, Enivaldo e Elcione, e meu irmão, Hélder, pelo amor, incentivo e apoio incondicional, pois sem eles, nada disto seria possível.

A minha linda companheira, Sara Fernandes, que sempre me fortaleceu e me apoiou em todas as minhas decisões. Sua dedicação me inspira e foi de suma importância para a conclusão deste trabalho.

Aos meus amigos Jean e Everton que fizeram parte de toda a minha caminhada acadêmica, sempre sendo solícitos e parceiros.

Aos professores Bruno Travençolo, José Gustavo e Luis Rocha, pela orientação, apoio e confiança.

A todos que direta e indiretamente fizeram parte da minha formação, e meu muito obrigado.

*“A educação do homem deve começar pela poesia, ser fortificada pela conduta justa e
consumar-se na música.”
(Confúcio)*

Resumo

Existem diversas propostas na literatura para utilizar Visualização da Informação em redes, mas ainda existem vários desafios na área, como a falta de técnicas que melhorem a interpretação de grandes quantidades de dados e relacionamentos. Este trabalho apresenta o estudo e aplicação de técnicas de visualização para análise e exploração de redes temporais. A ideia é identificar a presença de padrões e tendências que ajudem a explicar fenômenos e comportamentos em diversos cenários. Este trabalho apresenta também o sistema computacional *DyNetVis*, que pode ser utilizado para comprovar a eficácia das estratégias propostas, bem como tornar essas estratégias disponíveis à comunidade científica. Finalmente, são apresentados resultados da aplicação das técnicas propostas em diversas redes temporais, utilizando o sistema desenvolvido. Com a utilização das técnicas propostas, é possível que se reconheça padrões que antes eram difíceis de serem identificados.

Palavras-chave: Redes Temporais. Redes Complexas. Visualização. Grafos. Redes de contato.

Abstract

There are many proposals in the literature that use Information Visualization to visualize networks, but there still exist several challenges in the area, as the lack of techniques that improve the interpretation of big data and their relationships. This thesis presents the study and application of data visualization techniques in Temporal Networks. The goal is to identify patterns that will help to explain phenomena and behavior from various scenarios. As a result of this study, we present a system, called DyNetVis, which can be used to test the effectiveness of the proposed techniques, well be make these techniques available to the scientific community. Finally, we present results of the application of the proposed techniques in temporal networks using the developed system. With the utilization of the propose techniques, it is possible to recognize patterns that were difficult to identify without the techniques.

Keywords: Temporal Networks. Complex Networks. Visualization. Graphs. Contact Networks.

Lista de ilustrações

Figura 1 – Ilustração do conceito de grafos.	18
Figura 2 – Ilustração do conceito de diagrama de grafos.	19
Figura 3 – Exemplificação de Redes Temporais.	20
Figura 4 – Exemplificação de Visualização de Redes.	21
Figura 5 – Ilustração dos princípios de Gestalt aplicados a grafos.	22
Figura 6 – Ilustração do modelo da visão temporal proposto por (HOLME; SAMÄKI, 2011).	24
Figura 7 – Ilustração do modelo da visão temporal e da ordenação das linhas proposto por (ELZEN et al., 2013).	25
Figura 8 – Dez formas de ordenar as linhas na visão temporal aplicadas a dados reais.	26
Figura 9 – Visualização da base de dados de e-mails da empresa Enron.	27
Figura 10 – Representação circular de uma rede temporal.	29
Figura 11 – Ilustração da propagação de doença em uma rede temporal.	30
Figura 12 – Diferentes visões dos dados.	32
Figura 13 – Passo a passo do esquema do trabalho.	33
Figura 14 – Construção da rede estrutural considerando o tempo acumulado.	34
Figura 15 – Construção da rede estrutural considerando o instante de tempo.	34
Figura 16 – Diferenciação dos vértices pelo tamanho e cor.	35
Figura 17 – Diferenciação das arestas pelo tamanho, cor e forma.	36
Figura 18 – Seleção por profundidade na visão estrutural.	37
Figura 19 – Diferentes estratégias de ordenação para o <i>layout</i> temporal.	40
Figura 20 – Rede temporal para exemplificar o funcionamento do algoritmo de cores na visão temporal.	41
Figura 21 – Exemplificação do Mapa de Atividade Temporal (MAT).	43
Figura 22 – Representação das interações na visão temporal.	44
Figura 23 – Ilustração do conceito de resolução aplicado a visão temporal.	44
Figura 24 – Comunicação entre as visões estrutural e temporal.	45

Figura 25 – Telas de filtros e customização do sistema <i>DyNetVis</i>	48
Figura 26 – Visão estrutural detalhada da tela inicial do <i>software</i>	49
Figura 27 – Visão temporal detalhada da tela inicial do <i>software</i>	50
Figura 28 – Visão geral de dados reais de uma Conferência.	54
Figura 29 – Visão geral de dados reais de uma Escola.	54
Figura 30 – Visão geral de dados reais de um Hospital.	54
Figura 31 – Visão geral de dados reais de um Museu.	54
Figura 32 – Visão geral de dados reais de uma Rede Sexual.	56
Figura 33 – Visão geral de dados reais de um Museu.	57
Figura 34 – Visão geral dos dados da Escola (FOURNET; BARRAT, 2014).	58
Figura 35 – Visão específica do nó 1207 selecionado na visão temporal do Hospital.	59
Figura 36 – Visões específicas da visão temporal da Conferência.	60
Figura 37 – Visão específica de dois nós na visão estrutural e temporal dos dados do Museu (ISELLA et al., 2011).	62
Figura 38 – Visão específica de dados reais da Conferência.	63
Figura 39 – Representação da propagação da epidemia em uma rede temporal.	68

Lista de tabelas

- Tabela 1 – Tabela com as redes de contatos, com suas respectivas quantidades de vértices, arestas e resolução original da rede. 53
- Tabela 2 – Tabela com o desempenho relacionando os vértices e arestas com o tempo médio gasto para iniciar o sistema. 65

Lista de siglas

CEO *Chief Executive Officer*

DyNetVis *Dynamic Network Visualization*

MAT Mapa de Atividade Temporal

RFID *Radio Frequency Identification Devices*

Sumário

1	INTRODUÇÃO	15
1.1	Objetivos	16
1.2	Organização da Dissertação	17
2	FUNDAMENTAÇÃO TEÓRICA	18
2.1	Grafos	18
2.2	Redes Complexas	20
2.3	Visualização de Redes	21
2.3.1	Ferramentas para visualização de redes	22
2.4	Trabalhos Correlatos	23
3	ANÁLISE VISUAL DE REDES TEMPORAIS	31
3.1	Visão estrutural	33
3.1.1	Organização dos vértices	33
3.1.2	Atribuição de cor, tamanho e forma	34
3.1.3	Interação do usuário	36
3.2	Visão temporal	36
3.2.1	Organização dos vértices	37
3.2.2	Atribuição de cor, tamanho e forma	38
3.2.3	Mapa de Atividade Temporal	39
3.2.4	Interação do usuário	42
3.3	Resolução temporal da rede	42
3.4	Coordenação entre visões	45
3.5	Sistema <i>DyNetVis</i>	45
3.5.1	Funcionalidades do sistema	46
4	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	51
4.1	Redes de contato	51

4.2	Experimentos e Avaliação dos Resultados	53
4.2.1	Visão geral	53
4.2.2	Visão específica	59
4.2.3	Desempenho do sistema	63
5	CONCLUSÃO	66
5.1	Contribuições	67
5.2	Trabalhos Futuros	67
5.2.1	Análise e exploração de características associadas aos indivíduos	67
5.2.2	Organização dos vértices	68
5.2.3	Aplicações em Epidemiologia	68
5.2.4	Melhoria da escalabilidade do sistema	69
	REFERÊNCIAS	70

Introdução

A Visualização da Informação é uma área de pesquisa que estuda a construção de representações visuais de um espaço do domínio abstrato utilizando gráficos, imagens, animações, estruturas, entre outros (WILLIAMS; SOCHATS; MORSE, 1995). Essas representações, apoiadas por técnicas de interação computacionais, auxiliam na compreensão dos dados e na extração de conhecimento do resultado de simulações ou medições (McCORMICK, 1988).

A Visualização da Informação tem sido usada com bastante êxito em diversas áreas, dentre elas finanças, sociologia e transportes (ELZEN et al., 2013). Quando o espaço de domínio a ser visualizado é representado por uma estrutura de grafos, abordagens específicas são necessárias para tornar efetivo o processo de visualização (HERMAN; MELANCON; MARSHALL, 2000a). Encontrar uma representação adequada não é uma tarefa trivial, e continua sendo um desafio para os pesquisadores (SPENCE, 2007).

Têm-se observado um grande crescimento no uso da representação de grafos para modelagem de diversos fenômenos sociais e da natureza. Essa abordagem resultou em uma nova área científica chamada de redes complexas (NEWMAN, 2003). É possível encontrar, na literatura, exemplos de uso de modelagem por redes complexas para estudos nas áreas de epidemiologia, genética, computação, economia, neurociência, linguística, pesquisas com redes sociais, redes de terrorismo, entre outras (COSTA et al., 2011).

Para descrever uma rede complexa utiliza a representação clássica de grafos (BENDER-DEMOLL; MCFARLAND, 2006), a qual contempla a existência de um conjunto de itens, chamados de vértices ou nós, com conexões entre eles chamadas de arestas. Como a área de redes complexas é posterior a área de grafos, muitos dos conceitos e trabalhos direcionados a análise, quantificação e visualização de grafos são utilizados na análise de redes complexas. Dessa forma, para a visualização de redes complexas pode-se utilizar técnicas já estabelecidas para visualização de grafos.

A representação mais comum de rede complexa é a chamada na literatura de *Node-link diagram* (diagrama de nós e arestas) e será chamada neste trabalho de **estrutural**, que dispõe os nós em um plano, com arcos ou retas conectando-os (WILSON, 2012).

Entretanto, em uma rede com grande quantidade de vértices, a extração de informações pelo olho humano é dificultada, tornando inadequado o uso exclusivo dessa estratégia de visualização (NEWMAN, 2003). Apesar de existirem diversas propostas na literatura para visualização de redes, ainda existem vários desafios para a área de Visualização da Informação no tocante à proposição de técnicas que efetivamente melhorem, sob o aspecto cognitivo, a tarefa de interpretar grandes quantidades de dados e seus relacionamentos (HERMAN; MELANCON; MARSHALL, 2000a).

Dentre os tipos de redes complexas existentes, existe um tipo particular chamado de **rede temporal** (ou dinâmica) que, por definição, são redes nas quais as mudanças estruturais se dão de acordo com o tempo, isto é, vértices ou arestas são adicionadas ou removidas ao longo do tempo (ZAIDI; MUELDER; SALLABERRY, 2014). A sua visualização pode ser dada de forma que os nós são organizados em linhas horizontais e suas conexões são representadas, ao longo do tempo, por ligações entre essas linhas (HOLME; SARAMÄKI, 2011). Em razão de a área de visualização de redes temporais ser uma área recentemente explorada (ZAIDI; MUELDER; SALLABERRY, 2014), grande parte dos trabalhos que abordam o tema não disponibilizam ferramentas para realizar a visualização temporal (CROUSER; FREEMAN; CHANG, 2011). O uso das técnicas de Visualização de Informação ajuda a compreender melhor a estrutura e dinâmica de redes temporais, possibilitando a identificação de padrões que auxiliam na compreensão do fenômeno modelado.

Neste trabalho, as representações estruturais e temporais são analisadas aplicando técnicas de análise visual, que possibilitam alternativas visuais para interpretar os dados. Para comprovar estas técnicas, foi desenvolvido um sistema computacional que permite demonstrar as referidas formas de visualização de redes temporais com dados reais. Desta forma, espera-se que seja possível identificar quais padrões e tendências podem ser obtidos observando e interagindo com tal sistema de visualização.

1.1 Objetivos

O objetivo principal deste trabalho é desenvolver uma metodologia de visualização de redes temporais. Tem-se como hipótese que o uso das técnicas de Visualização de Informação ajuda a compreender melhor a estrutura e dinâmica de redes temporais, possibilitando a identificação de padrões que auxiliam na compreensão do fenômeno modelado. Como objetivos específicos tem-se:

1. Investigar estratégias de visualização e exploração da estrutura de conexões da rede, de forma a construir modelos que permitam identificar padrões e tendências dos dados em diferentes cenários;
2. Investigar estratégias de visualização interativa de conexões ao longo do tempo,

adicionais às conhecidas da literatura, e construir, utilizando essas e novas técnicas, outras formas de analisar as informações de redes temporais;

3. Propor novas estratégias de coordenação entre as visões construídas, possibilitando a análise da rede sob diversas perspectivas simultaneamente;
4. Desenvolver um sistema computacional dotado de um conjunto de ferramentas interativas, de forma a validar a metodologia proposta e disponibilizar para a comunidade as técnicas propostas.

1.2 Organização da Dissertação

O Capítulo 1 de introdução discute os desafios da área de Visualização de Informação relacionados a análise de redes temporais e detalha como os objetivos deste trabalho podem contribuir para a área.

O Capítulo 2 de fundamentação teórica trata dos conceitos básicos da área para que o trabalho possa ser melhor compreendido, bem como analisa os diversos trabalhos relacionados à área de redes temporais.

O Capítulo 3 dispõe sobre as técnicas de análise visual de redes temporais detalhadas, analisando as diferenças entre os trabalhos demonstrados no segundo capítulo, abordando detalhes da rede estrutural e temporal e das diversas técnicas para auxiliar na identificação de padrões, como atribuir cor, tamanho e forma aos vértices e arestas de maneira eficaz. Esse capítulo finaliza com a demonstração do *software* desenvolvido.

O Capítulo 4 apresenta a análise dos resultados aplicando as técnicas que foram propostas no terceiro capítulo, trabalhando com algumas redes temporais que modelam interações entre pessoas em diferentes contextos (em uma escola, um hospital, um museu, durante uma conferência e em um fórum para avaliar encontros sexuais).

O Capítulo 5 apresenta uma conclusão a respeito das técnicas propostas e dos resultados, bem como propõe trabalhos futuros.

Fundamentação Teórica

Neste capítulo são apresentados os conceitos básicos para melhor compreensão deste trabalho. Além disso, propostas com objetivos semelhantes são analisadas de forma crítica, servindo de inspiração para as técnicas desenvolvidas.

2.1 Grafos

Matematicamente, um **grafo** G é definido como uma tupla $G(V, E)$, em que o conjunto V são os vértices ou nós e o conjunto E são as interações representadas por arestas (ZAIDI; MUELDER; SALLABERRY, 2014). Segundo (WILSON, 1986), considerando a Figura 1(a) parte de um circuito elétrico e Figura 1(b) parte de um mapa de ruas, a Figura 1(c) pode ser a representação, por meio de um diagrama, que utiliza pontos e linhas, do circuito e da rua. Os pontos P , Q , T , S e R são chamados de **vértices** (ou nós), e as linhas que conectam esses vértices são chamadas de **arestas**.

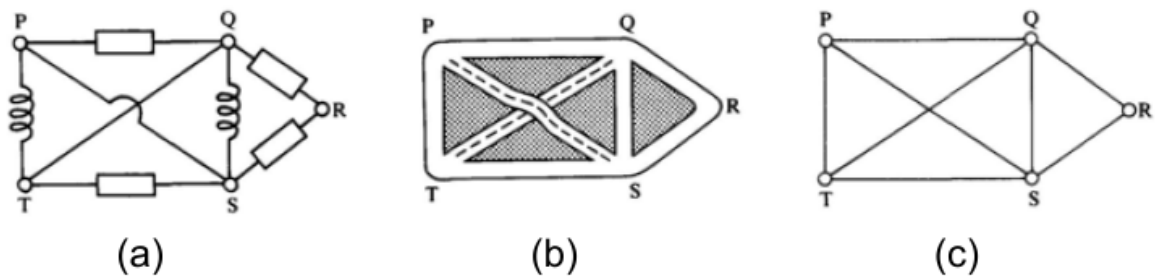


Figura 1 – Ilustração do conceito de grafos. Em (a) é representado um circuito elétrico, em (b) é representado um mapa de ruas e em (c) é representado um diagrama de vértices e arestas – um grafo – que pode ser usado tanto para modelar o circuito elétrico quanto as ruas do mapa. Figura adaptada de (WILSON, 1986) ©1986 Prentice-Hall.

Um importante conceito relacionado a grafos é o grau do vértice. O conceito de **grau**

do vértice é o número de arestas que tem aquele vértice como um ponto de chegada e saída; por exemplo, observando a Figura 1(b), ele consiste no número de ruas de uma intersecção, portanto, o grau do vértice Q é quatro.

Um grafo pode ser representado por uma **lista de arestas**. Para representar uma aresta, basta indicar o par de nós que forma a aresta, como por exemplo, para uma aresta que liga os nós P e Q é representado por $[P, Q]$. A lista de arestas da Figura 1 é exemplificada como $[[P, Q], [P, T], [P, S], [T, Q], [T, S], [S, Q], [S, R], [Q, R]]$.

Outra definição importante de grafos são os Vizinhos de um vértice. Vértices que se conectam diretamente uns com os outros são chamados de **Vizinhos**. Por exemplo, na Figura 1(c), os vizinhos do vértice P são representados por Q , T e S . Nota-se que o vértice R não se conecta diretamente ao vértice P , portanto, ele não faz parte dos seus vizinhos. Em contrapartida, todos os vértices são vizinhos de Q e S , pois todos os vértices se conectam diretamente a eles.

Um grafo pode ser representado por meio de um diagrama. Esses diagramas são formados por pequenos círculos que são conectados por linhas, sendo que em alguns casos, esses círculos não possuem um posicionamento definido no plano. Na Figura 2(a) tem-se um diagrama representado no formato de um pentágono e na Figura 2(b) tem-se um diagrama representado no formato de uma estrela. Os dois diagramas representam o mesmo grafo – observe, por exemplo, os nós que se conectam ao nó 1 são os nós 2 e 5 em ambas figuras; o mesmo padrão de conexão ocorre para todos os pontos desses dois diagramas. Portanto, diferentes diagramas podem representar um mesmo grafo. A forma de representação pode ter impacto na compreensão dos dados e do fenômeno modelado. A representação de grafos utilizando diagramas neste trabalho é denominada **visão estrutural**, a fim de diferenciar das outras formas de visualização que são propostas.

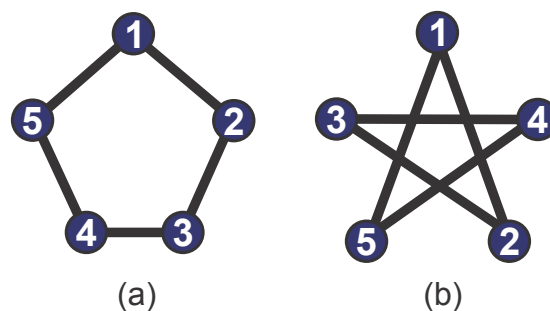


Figura 2 – Ilustração do conceito de diagrama de grafos. Em (a) é representado o diagrama no formato de um pentágono e em (b) é representado o diagrama no formato de uma estrela. Os dois diagramas referem-se ao mesmo grafo.

2.2 Redes Complexas

Tem-se notado um crescimento no uso da representação de grafos para modelagem de diversos fenômenos da natureza e da sociedade. Essa abordagem resultou em uma nova área científica chamada **Redes Complexas** (ALBERT; BARABÁSI, 2002). O termo Redes Complexas é referenciado neste texto somente como redes.

As **Redes Temporais** (ou Dinâmicas) são redes nas quais as mudanças estruturais se dão de acordo com o tempo, isto é, adicionam ou removem vértices ou arestas ao longo do tempo (HOLME; SARAMÄKI, 2011). Em Redes Temporais o tempo tem um papel importante na sua evolução e no seu entendimento. Matematicamente, Redes Temporais são definidas como um grafo $G = (V, E)$, com os vértices sendo um conjunto V e as arestas um conjunto E , onde o domínio de $E \subseteq V \times V \times T$ com as tuplas (v_a, v_b, e_t) representando o par de vértices (v_a, v_b) e um atributo de tempo $e_t \in [t_{min}...t_{max}]$ para cada aresta (ELZEN et al., 2013).

Aplicações de redes temporais podem ser encontradas, por exemplo, em redes de transporte aéreo (ROCHA, 2009), tráfego de pacotes de dados da internet, redes sociais (HEER; BOYD, 2005), entre outras. De acordo com (FREEMAN, 2004), as técnicas de visualização tem se tornado um componente essencial para o estudo de Redes Temporais.

Na Figura 3 é mostrado um exemplo que pode ser modelado como uma rede temporal. Nesse exemplo, é indicada a comunicação entre pessoas em três instantes de tempo. No instante de tempo 1 tem-se a pessoa A comunicando com a pessoa B e C e também a comunicação entre a pessoa B e C. No instante de tempo 2 acontece a adição de mais duas novas pessoas, D e E, com a pessoa B tendo saído de cena, e as pessoas A e C deixando de se comunicarem. No instante de tempo 3, todas as pessoas que estavam se comunicando saíram de cena com exceção da pessoa A, e mais duas novas pessoas aparecem, sendo a F e G. Observe que a pessoa A está ativa em todos os instantes de tempo. Esse exemplo demonstra as características de Redes Temporais, que são adição e remoção de vértices (consideradas as pessoas nesse exemplo) e arestas (consideradas as comunicações entre as pessoas) ao longo do tempo.

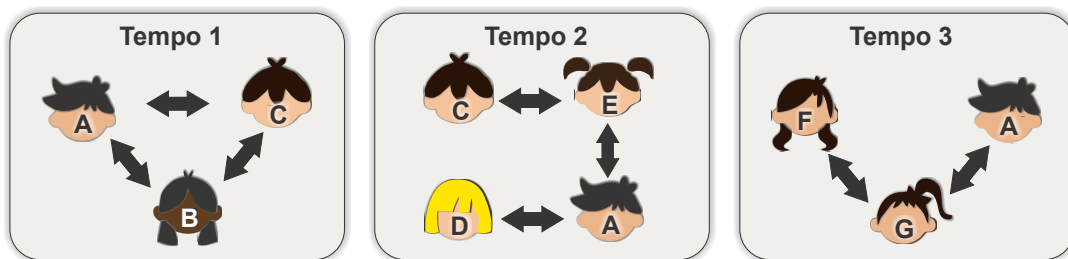


Figura 3 – Exemplificação de Redes Temporais. São mostrados três instantes de tempo, com a indicação dos contatos entre as pessoas em cada instante.

2.3 Visualização de Redes

Visualização de Redes tem muitas áreas de aplicação como na computação, em mapas de *site*, ou na biologia e química, em mapas moleculares e genéticos (HERMAN; MELANCON; MARSHALL, 2000b). O objetivo da Visualização é otimizar aplicações para ajudar a realizar um trabalho cognitivo mais eficientemente. A visualização é utilizada para resolver problemas de forma mais rápida e eficaz, ou para aprender algo novo (WARE, 2004). Um exemplo de aplicação da Visualização de Redes é apresentado na Figura 4, que demonstra a topologia de quatro redes obtidas a partir de termos descritos na *Wikipedia*, referentes a termos das disciplinas de Biologia, Matemática, Física e Medicina (SILVA et al., 2011).

Nas Figuras 4(a) e 4(d) são apresentadas redes de Biologia e Medicina, respectivamente, em que suas topologias contêm um padrão modular. Na Figura 4(b) é apresentada uma rede de Matemática, que é bem esparsa, e a Figura 4(c) apresenta uma rede de Física, que contém um núcleo denso. É notável que com a Visualização de Redes é possível identificar topologias e padrões que seriam dificultadas sem a visualização.

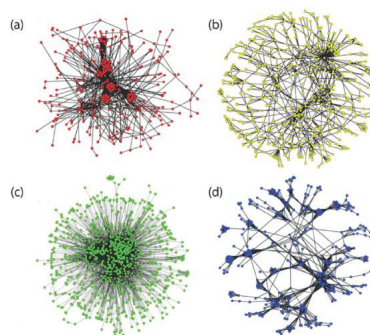


Figura 4 – Exemplificação de Visualização de Redes. A figura apresenta a topologia de redes das quatro principais categorias da *Wikipedia*: (a) Biologia, (b) Matemática, (c) Física e (d) Medicina. Figura adaptada de (SILVA et al., 2011).

Para auxiliar a identificação de padrões e tendências visuais em redes temporais utilizam-se os conceitos das **Leis de Gestalt**, que é um conjunto de regras para descrever a maneira que os seres humanos enxergam padrões em representações visuais (WARE, 2004). Os princípios mais relevantes das Leis de Gestalt que se aplicam neste trabalho são: fechamento, proximidade e similaridade.

O princípio do **fechamento** diz que “*tende-se a perceptivelmente fechar ou completar, objetos que não estão de fato, completos*” (WARE, 2004). Na Figura 5(a), os nós 1 e 3 estão parcialmente sobrepostos mas são distintos, pois devido ao fechamento do contorno do círculo o cérebro separa visualmente os elementos, da mesma maneira que na Figura 5(b) as retas cinzas são vistas como blocos. O princípio da **proximidade** diz que “*quando observa-se um conjunto de objetos, tende-se a ver objetos próximos uns dos outros como um grupo*” (WARE, 2004), tais como os nós 1 e 3 na Figura 5(a) e os blocos

cinzas identificados na Figura 5(b). O princípio de **similaridade** diz que “*tende-se a agrupar objetos baseado em suas similaridades*” (WARE, 2004), como por exemplo os nós 2 e 5 da Figura 5(a). Mesmo estando espacialmente mais distantes, por terem a mesma cor o cérebro tende a separá-los dos demais em um grupo, da mesma forma para os nós 1, 3 e 4. Na Figura 5(b) as retas também são visualmente segregadas pela diferenciação da cor azul e cinza.

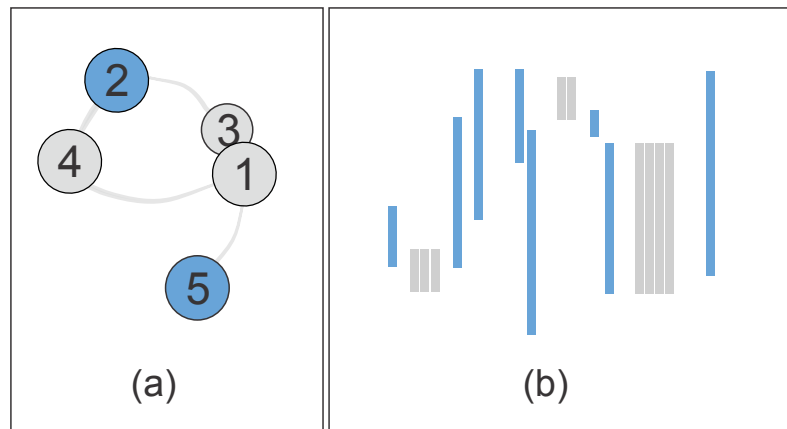


Figura 5 – Ilustração dos princípios de Gestalt aplicados a grafos. Em (a) representa-se um grafo, separado por cores e com sobreposição de nós e em (b) tem-se um conjunto de retas com comprimentos, cores e posicionamentos diferentes e com algumas retas mais próximas e outras mais separadas.

2.3.1 Ferramentas para visualização de redes

Existem alguns sistemas e bibliotecas para visualizar redes complexas, como *Gephi* (BASTIAN; HEYMANN; JACOMY, 2009), *Cytoscape* (SHANNON et al., 2003), *Graph-tool* (PEIXOTO, 2014), *Pajek* (BATAGELJ; MRVAR, 2004), *igraph* (CSARDI; NEPUSZ, 2006), *NetworkX* (HAGBERG; SCHULT; SWART, 2008), *Orion* (HEER; PERER, 2014), entre outros (BRANDES; WAGNER, 2003; PIGNÉ et al., 2008; LESKOVEC; SOSI, 2014; O'MADADHAIN et al., ; WONGPIROMSARN et al., 2011; BORGATTI; EVERETT; FREEMAN, 2002; CARLEY et al., 2012). Cumpre destacar a funcionalidade de suporte para diversos arquivos de entradas do *Cytoscape*, sendo configurável o formato do arquivo dentro do sistema. O *Graph-tool* apresenta um desempenho superior em relação ao tempo de execução, se comparado com *igraph* e *NetworkX* (PEIXOTO, 2014).

Todos os sistemas citados possuem uma vasta gama de funcionalidades para redes estruturais, como por exemplo: construir grafos; adicionar interativamente nós e arestas e gerar grafos aleatórios; utilizar algoritmos de força e hierarquia para separar os nós; customizar as cores, tamanhos e formas de arestas e vértices; dentre inúmeras outras funções. Entretanto, nenhum dos sistemas citados apresentam uma representação visual de redes temporais (HOLME; SARAMÄKI, 2011).

2.4 Trabalhos Correlatos

Nesta seção são discutidos os principais trabalhos existentes na área de visualização de redes complexas, bem como as inspirações para o desenvolvimento deste trabalho.

De acordo com (ZAIDI; MUELDER; SALLABERRY, 2014), apenas recentemente começou-se a perceber o real potencial de análises de redes temporais e como elas são úteis. Os autores afirmam, ainda, que a dimensão temporal tem poucos estudos na área, fato este justificado pela imaturidade dessa área de pesquisa e pela falta de bancos de dados que trabalhassem com a variável de tempo. Somando-se o crescimento das redes sociais, os autores citam que, hodiernamente, existem bancos de dados que incluem a dimensão temporal, facilitando pesquisas e gerando um crescimento na área.

Os autores citados classificam as pesquisas na área de redes dinâmicas em duas categorias: a parte analítica e a de visualização. Eles ressaltam que grafos dinâmicos analíticos são o método mais comum para lidar com aspectos temporais, aglomerando toda a rede em uma única rede estática, mantendo também a opção de exibir subgrafos do grafo original, exibindo nesse subgrafo o grafo correspondente a uma porção de tempo delimitada. O trabalho cita a parte de visualização em grafos dinâmicos como um desafio para os pesquisadores. Os autores afirmam que métodos tradicionais para visualização de grafos estáticos, que utilizam espaços bidimensionais para posicionar os vértices, não podem ser aplicados em redes dinâmicas, devido à dificuldade de representar uma nova dimensão nesse espaço. A aproximação simples para visualizar um grafo dinâmico seria uma imagem estática, porém esta falha em representar mudanças ocorridas em redes de acordo com o tempo (MOODY; MCFARL; BENDER-DEMOLL, 2005). Vale destacar também que em animações, mudanças podem ser visualizadas ao longo do tempo, mas sem interação por parte do usuário.

Em (HOLME; SARAMÄKI, 2011), os autores demonstram o termo redes temporais aplicado a diversas disciplinas, categorizando tipos de redes temporais, e também apresentando as limitações de um grafo estrutural, salientando-se a necessidade de se utilizar uma rede temporal. Um dos tipos de redes temporais que é citado no trabalho consiste em dados coletados por proximidade física, a partir da utilização de sensores aparelhos de identificação por radio frequência, ou *Radio Frequency Identification Devices* (RFID) (CATTUTO et al., 2010).

Em (HOLME; SARAMÄKI, 2011) é explicado a importância da rede temporal para uma análise real dos dados, como por exemplo, em redes de comunicação via e-mail, mensagens de texto ou telefonemas. Nesse contexto, os vértices e as arestas que representam as conexões não continuam ativos durante todo o tempo. Aplicando esses dados a uma representação de um grafo estrutural, perde-se a informação temporal, que pode ser crucial na análise dos dados, o que demonstra a importância de se utilizar uma representação temporal. O trabalho apresenta ainda um esboço do que seria a representação temporal, conforme apresentado na Figura 6, correlacionando a visão do grafo estrutural com a visão

temporal.

A Figura 6(a) mostra a visão estrutural de um conjunto de dados, sendo que cada vértice está representado por uma cor aleatória. Por cima das arestas que conectam os nós, na Figura 6(a), está representado o instante de tempo que aqueles nós fazem uma conexão. Por exemplo, entre o nó A e o nó B estão representados os números 6, 7 e 11. Isso indica que nos tempos 6, 7 e 11 os nós A e B fazem uma conexão. A Figura 6(b) é a representação temporal equivalente da Figura 6(a). Observa-se também que cada linha da Figura 6(b) está com a mesma cor correspondente, indicando assim a representação do mesmo nó. Além disso, as linhas da Figura 6(b) estão ordenadas verticalmente por ordem alfabética. A última linha horizontal, considerada, de cima para baixo, representa o eixo temporal, iniciando no tempo 0 e terminando no tempo 14.

Uma possível crítica ao trabalho consiste na forma pela qual os autores correlacionam as visões (exibição do grafo estrutural e temporal). Eles utilizam, como recurso, a cor, sendo que cada cor representa um vértice. Tal recurso, em um ambiente de vários nós, não seria viável, tendo em vista a limitação da capacidade humana da distinção de diferentes cores (WARE, 2004). É importante ressaltar que (HOLME; SARAMÄKI, 2011) apresenta apenas o que seria o esboço da visão temporal, no entanto a visualização não foi aplicada a dados reais a fim de se perceber o comportamento da rede e qual seria o resultado dessa visão.

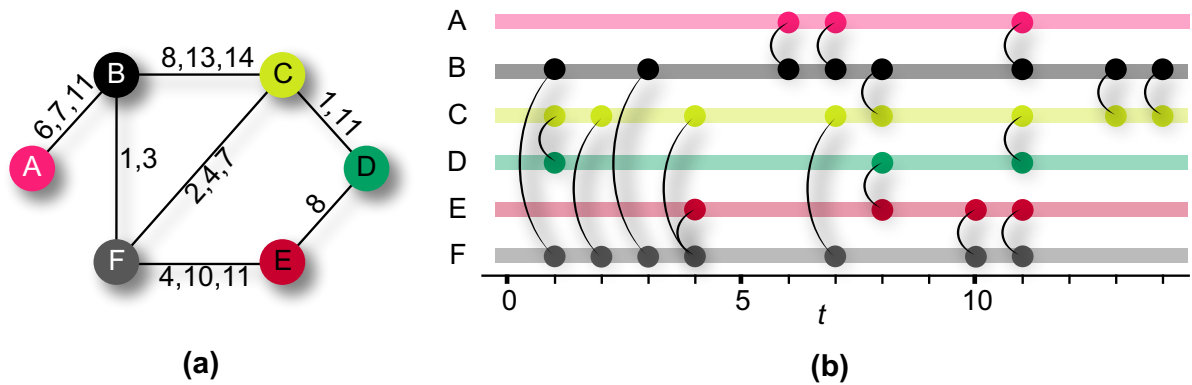


Figura 6 – Ilustração do modelo da visão temporal proposto por (HOLME; SARAMÄKI, 2011). Em (a) representa a visão estrutural e em (b) representa a visão temporal. Figura adaptada de (HOLME; SARAMÄKI, 2011) ©2011 arXiv.

Em (ELZEN et al., 2013) os autores trabalham com a análise das estruturas de redes temporais. A inovação do trabalho baseia-se na ideia de se atribuir funções dinâmicas a essa rede temporal, adicionando-se a interação de forma a auxiliar a visão e exploração dos dados naquele contexto. Os autores exploram a ideia de se realizar a alteração da ordem das linhas horizontais (que representam os vértices) reduzindo a sobreposição das linhas verticais (que representam as arestas), como pode ser visto na Figura 7, sempre se preocupando em se adequar as Leis de Gestalt.

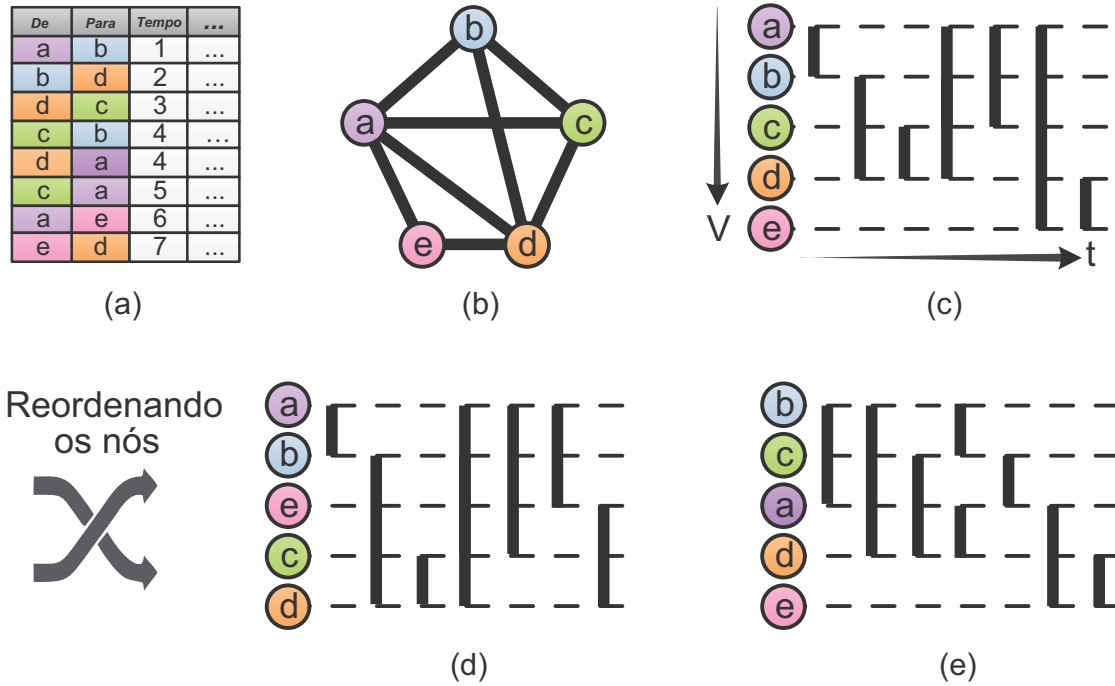


Figura 7 – Ilustração do modelo da visão temporal e da ordenação das linhas proposto por (ELZEN et al., 2013). Em (a) é apresentada uma tabela com os dados originais, a primeira coluna correspondendo a origem da conexão, a segunda coluna correspondendo ao destino da conexão e a terceira coluna referindo-se ao o tempo dessa conexão. Em (b) é demonstrada uma representação estrutural dessa rede correspondente em (a). Em (c) é ilustrada uma visão temporal com os nós ordenados pela ordem alfabética. Em (d) e em (e) são demonstrados os resultados da ordenação das linhas: a primeira (d) ordenando por arestas de maior tamanho e a segunda (e) as linhas estão ordenadas pelas arestas de menor tamanho. Figura adaptada de (ELZEN et al., 2013) ©2013 IEEE.

Seguindo essa linha, na Figura 8, é apresentado um conjunto de estratégias de reordenação das linhas que representam a atividade temporal dos nós, baseando-se em medidas como o grau e a atividade das arestas, e também baseando-se em problemas mais complexos de minimização de sobreposição de arestas. Os autores descrevem tais métodos, apresentando resumidamente os algoritmos utilizados para realizá-los.

Na Figura 8(a) tem-se a ordenação Lexicografia, que ordena os nós por ordem alfabética. Na Figura 8(b) tem-se os nós ordenados baseado em seus respectivos graus, que é normalmente um importante atributo, pois mostram os nós que mais se conectam na rede, separadamente dos que menos se conectam. Para revelar mais estruturas em nós de baixo e alto grau, divide-se em graus de entrada (Figura 8(c)) e graus de saída (Figura 8(d)). Na visualização da Figura 8(e) os nós são ordenados pela sua atividade temporal, com os nós mais ativos nos tempos iniciais da rede posicionados na parte superior do *layout*, nós que são mais ativos nos tempos finais posicionados na parte inferior e nós ativos ao longo de todo o tempo posicionados próximo ao centro.

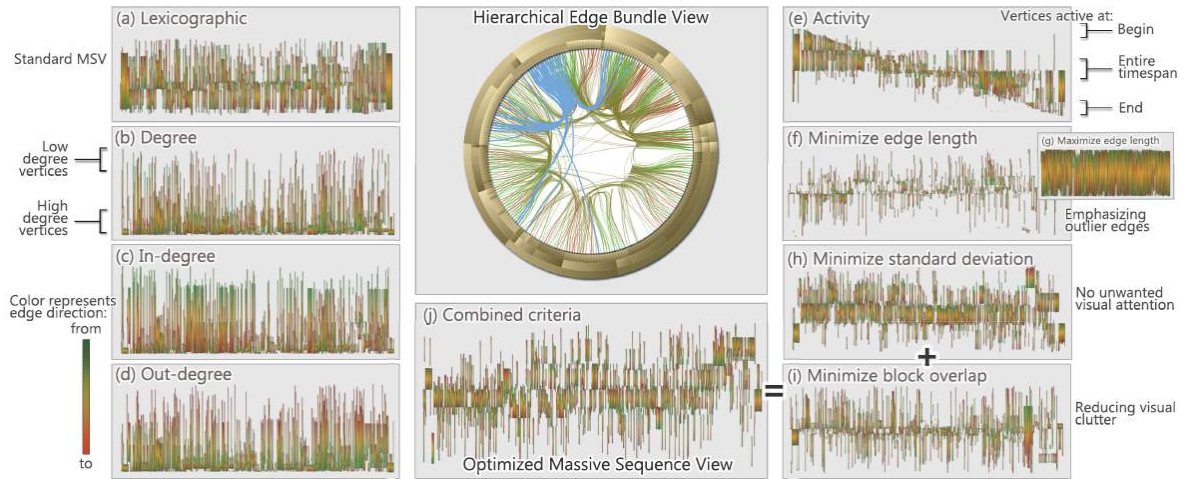


Figura 8 – Dez formas de ordenar as linhas na visão temporal aplicadas a dados reais. Cada ordenação provê uma visualização única em aspectos temporais e estruturais dos dados. Os dados reais são de uma transação financeira contendo 302 contas e 997 transações. Em (a) tem-se a ordenação Lexicografia, que somente ordena os nós por ordem alfabética. Em (b) tem-se os nós ordenados baseado em seus respectivos graus. Esses graus podem ser divididos em graus de entrada (c) e graus de saída (d). Em (e) os nós são ordenados pela sua atividade temporal, com os nós mais ativos nos instantes iniciais posicionados na parte superior e nós que são mais ativos nos instantes finais posicionados próximo a parte inferior do *layout*. Em (f) é utilizado o algoritmo de minimização adaptado para o comprimento das arestas chamado de *simulated annealing*, que busca organizar as arestas que mais aparecem de forma a terem seu tamanho reduzido. Em (g) é apresentado o conceito de maximização do tamanho das arestas, organizando as arestas que menos aparecem de forma a terem seu tamanho aumentado. Em (h) é utilizado o algoritmo de minimização do desvio padrão do comprimento das arestas, para prevenir de que arestas longas dominem a visualização. Em (i) é minimizado a sobreposição dos blocos de arestas, para reduzir uma confusão visual devido as sobreposições. Em (j) tem-se a combinação dos métodos de (h) e (i), gerando a visualização otimizada de uma sequencia densa, que é a proposta principal do trabalho. Figura retirada de (ELZEN et al., 2013) ©2013 IEEE.

Na Figura 8(f) é utilizado o algoritmo de minimização aproximada adaptado para o comprimento das arestas chamado de *simulated annealing*. O algoritmo de minimização busca organizar as arestas que mais aparecem de forma a terem seu tamanho reduzido, entretanto, arestas *outliers* são enfatizadas, devido ao crescimento de sua largura. No entanto, com essa minimização é difícil realizar a percepção de padrões. Na Figura 8(g) é apresentado o conceito de maximização do tamanho das arestas, organizando as arestas que menos aparecem de forma a terem seu tamanho aumentado. Essa maximização revela quais são os nós mais dominantes da rede, que são posicionados próximo da parte superior e da parte inferior do *layout*. Inversamente, as arestas de tamanho pequeno tendem a ser os *outliers* dessa visualização. Em geral, a maximização não gera bons resultados, pois a região fica muito densa e difícil de discernir. Na Figura 8(h) é utilizado o algoritmo de minimização do desvio padrão do comprimento das arestas, para prevenir que arestas longas dominem a visualização. A minimização por desvio padrão aumenta a legibilidade e simplifica o processo de análise, mas não leva em conta a sobreposição de blocos. Na Figura 8(i) é minimizada a sobreposição dos blocos de arestas, aumentando o seu comprimento, visando reduzir uma confusão visual devido as sobreposições. Na Figura 8(j) tem-se a combinação dos métodos de Figura 8(h) e Figura 8(j), gerando a visualização otimizada de uma sequencia densa, que é a proposta principal do trabalho. Para combinar as duas métricas elas devem ser normalizadas, senão uma métrica terá um peso maior do que a outra. A combinação de minimização da sobreposição de blocos e desvio padrão do comprimento das arestas auxilia a perceber toda a rede uniformemente, sem distrações erradas que retratam eventos temporais, reduzindo significativamente a poluição visual.



Figura 9 – Visualização da base de dados de e-mails da empresa Enron, correspondendo a 151 nós e 21.374 arestas (KEILA; SKILLICORN, 2005). Em (a) é mostrada a forma estrutural dessas conexões, em (b) demonstra-se a visão temporal com as linhas sem ordenação. Em (c) o resultado de combinar estratégias de reordenar as linhas é ilustrado e em (d) são destacados na visão temporal, somente os blocos que são considerados mais importantes. Figura retirada de (ELZEN et al., 2013) ©2013 IEEE.

A Figura 9 traz um exemplo de dados reais aplicados a visão estrutural e temporal,

utilizando a abordagem proposta de combinar as duas métricas. Essa imagem representa dados de um conhecido banco de dados de e-mails da empresa Enron (KEILA; SKILLICORN, 2005), que foram filtrados, removendo *spam* e duplicatas, contendo 151 nós e 21.374 arestas. Enron foi uma empresa de serviços de energia, conhecida pela maior falência de uma empresa americana devido à fraudes contábeis. Em Dezembro de 2000, Jeffrey Skilling se tornou o novo *Chief Executive Officer* (CEO) da Enron e juntamente com seus executivos, esconderam bilhões de débitos por fraude de contabilidade. Em Agosto de 2001, Jeffrey renunciou ao seu cargo de CEO. Em Outubro de 2001 as comissões de intercâmbio e segurança começaram a investigação da empresa Enron através de uma denúncia, iniciando a crise da empresa. Em Dezembro de 2001 a empresa foi decretada falida e diversos executivos foram sentenciados a prisão. Dois anos e meio de e-mails corporativos de funcionários da Enron foram disponibilizados durante as investigações legais.

Os autores apresentam a importância da reordenação dos nós para a percepção de padrões que antes não poderiam ser encontrados. Na Figura 9(c) é demonstrado a visão temporal com as técnicas combinadas, que são a proposta principal do artigo. Nota-se que com as técnicas aplicadas é possível identificar um acréscimo do número de conexões ao longo do tempo, iniciando com a designação do novo CEO e tendo o seu pico durante o processo de investigação da empresa, bem como um declínio das conexões após a quebra da empresa.

A primeira crítica em relação ao trabalho fica por conta da abordagem utilizada, pois os autores (ELZEN et al., 2013) somente observam os resultados a partir da exibição das arestas, desconsiderando uma possível relevância do vértice para demonstrar a atividade da rede. Outro problema identificado é que, mesmo após aplicar a reordenação dos nós, os padrões não estão evidentes, como quando é iniciado as investigações a região do *layout* continua densa, tendo em vista que uma grande quantidade de dados está sendo visualizada, ocorrendo diversas sobreposições de arestas, o que acarreta na perda de uma percepção real de entendimento.

Em (ELZEN et al., 2014) tem-se uma nova proposta de visualização para redes temporais, apresentando uma rede temporal de forma circular, como pode ser visto na Figura 10. Cada circunferência representa um instante de tempo, e cada porção colorida representa uma aresta entre dois nós (as arestas se sobrepõem dentro do circunferência). O tempo segue de dentro para fora, sendo que as menores circunferências representam os instantes iniciais, e as maiores circunferências representam os instantes finais. Uma crítica fica por conta da escalabilidade da rede em forma de circunferências, tendo em vista que ao aplicar os dados para uma rede com grande porção de tempo, as arestas das últimas circunferências apresentam visualmente um maior destaque do que as arestas dos tempos iniciais, podendo gerar assim uma maior relevância as últimas atividades que ocorreram naquela rede.

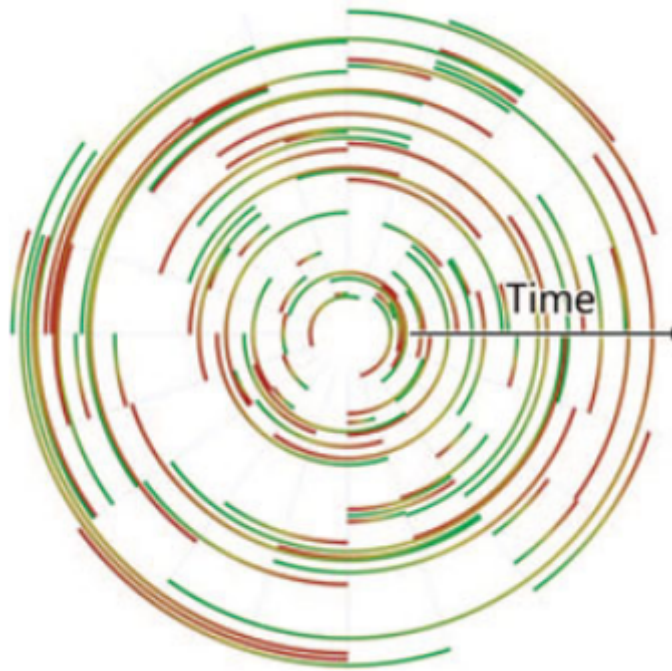


Figura 10 – Representação circular de uma rede temporal. As circunferências mais internas representam os tempos iniciais, e esse crescimento ocorre de dentro para fora. Figura retirada de (ELZEN et al., 2014) ©2014 IEEE.

Em (LEE et al., 2012) novos conceitos de interação com a visão temporal foram criados a partir da demanda de uma melhor interpretação do usuário. Tendo em vista que, nesse trabalho, o exemplo utilizado para a pesquisa, mostrado na Figura 11, consiste em um modelo de rede de propagação de doenças. Nesse contexto, existe uma relevância em saber, qual é a origem da doença, bem como qual é o ponto de início da propagação e, ainda, se existe algum indivíduo vacinado, pois ele não propagaria a doença mesmo que estivesse em contato com um indivíduo contaminado. O trabalho propõe, então, novas alterações das visões discutidas, dentre as quais a adição de uma barra vertical que irá representar o instante de início de propagação da doença, qual linha corresponde à origem da doença e quais são os indivíduos vacinados (vistos na Figura 11). Tais estratégias, no entanto, não foram aplicadas a dados reais, bem como a escolha de diferentes cores para representar os nós não ajudam na visualização.

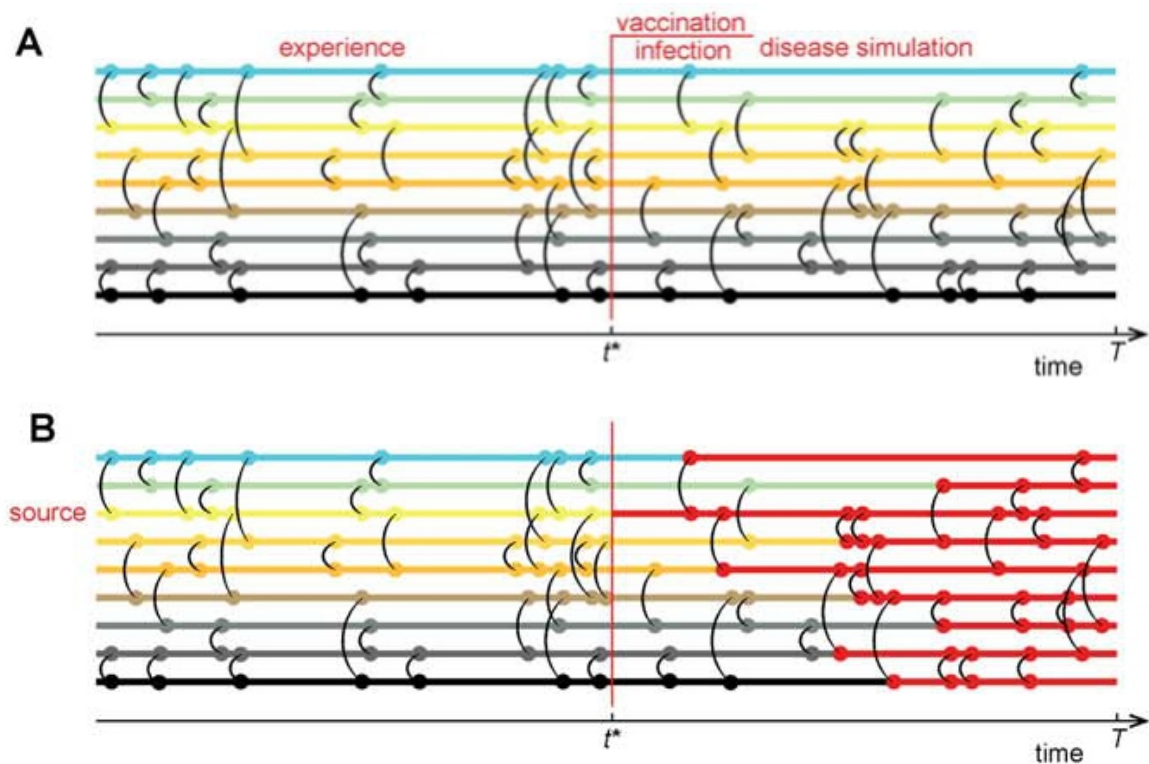


Figura 11 – Ilustração da propagação de doença em uma rede temporal. Em (a) é demonstrado a visão temporal original dos dados, com um divisor t^* indicando o ponto de início da infecção. Em (b) exibe o resultado da infecção da doença ao longo do tempo, representado nas linhas vermelhas, a partir do divisor t^* , considerando que a terceira linha é a origem desta infecção. Figura adaptada de (LEE et al., 2012) ©2012 Lee et al.

Análise Visual de Redes Temporais

O presente trabalho consiste no estudo de técnicas de Visualização de Informação para a análise e exploração de redes temporais, analisando as necessidades de interação em diversas situações. Um *software*, denominado **DyNetVis** foi desenvolvido de maneira a validar e tornar disponível à comunidade científica as técnicas discutidas (Seção 3.5). Neste capítulo serão detalhadas algumas técnicas que já existem na literatura, bem como técnicas que são propostas originais deste trabalho, como uma técnica de organização dos vértices, denominada de **Vizinhos Recorrentes** (Seção 3.2.1.3) e o **Mapa de Atividade Temporal** (Seção 3.2.3).

A Figura 12 mostra as três visões diferentes de uma rede que são utilizadas neste trabalho. A Figura 12(a) mostra a representação da rede em forma de tabela, sendo que cada linha corresponde a uma conexão da rede mostrada na Figura 12(b). Essa figura mostra a visão mais comumente utilizada, representando o grafo em sua forma estática, e que neste trabalho é referenciada por **visão estrutural**. Na Figura 12(c), o eixo das abscissas exibe o tempo e o eixo das ordenadas os vértices, possibilitando uma melhor percepção da relação dos vértices ao longo do tempo, chamada de **visão temporal**.

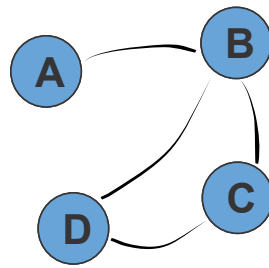
A Figura 12(b) fornece uma boa visualização do todo, uma visão geral do que representa os dados. Por exemplo, analisando a Figura 12(b), caso fosse necessário verificar a quem o vértice B se conecta, tem-se de maneira imediata que são os vértices A, C e D, cabendo ressaltar que essa estrutural não permite a visualização de quando as conexões ocorreram. Vale observar que essa informação não é vista de imediato na Figura 12(c), que concentra-se mais em uma percepção do instante do que do todo. Nesse sentido, a ideia do trabalho não é somente apresentar uma nova visualização, mas também coordenar entre as visões já conhecidas.

Diferentemente da representação temporal de (HOLME; SARAMÄKI, 2011), já ilustrada na Figura 6, este trabalho considera a forma da representação das arestas como uma reta (ELZEN et al., 2013), ilustrada na Figura 7. Foi escolhido manter as arestas como retas, devido a uma maior facilidade de interpretá-las como blocos e facilitar a identificação de padrões, como já foi demonstrado nos princípios de Gestalt (visto na seção 2.3).

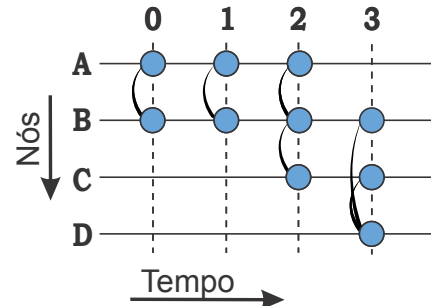
Estrutura de dados

Origem	Destino	Tempo
A	B	0
A	B	1
A	B	2
B	C	2
B	D	3
C	D	3

(a)

Estrutural

(b)

Temporal

(c)

Figura 12 – Diferentes visões dos dados. Em (a) representa o arquivo com os dados originais. Em (b) representa os dados na forma estrutural e em (c) demonstra a visão temporal. Todas as três visões são representações visuais dos mesmos dados.

Exclusivamente para facilitar a didática, algumas figuras neste trabalho apresentam as arestas da visão temporal como curvas.

Adicionalmente à outros trabalhos, foi incluída um nova tratativa que tem por objetivo adicionar a transparência nas arestas, na visão temporal, para que, havendo casos de sobreposição, seja possibilitado um ganho de informação, visto que regiões que forem densas ficam mais intensas devido à sobreposição de várias arestas transparentes. Com isso, há também uma diferenciação de regiões que têm poucas sobreposições, pois apresentará pouca concentração de cor, indicando uma tendência de pouca atividade naquela faixa de tempo.

Na Figura 13 pode-se verificar os passos da proposta. Para que os dados possam ser visualizados faz-se necessário convertê-los, simplificando os elementos e o instante de tempo que ocorre a conexão entre esses elementos. A etapa de pré-processamento é optativa, podendo filtrar os dados previamente antes de aplicá-los a visualização. Os dados podem ser filtrados no pré-processamento alterando-se o tempo inicial e final, bem como podendo realizar uma mudança da resolução temporal. Entretanto, outras operações podem ser realizadas nesse passo. A partir desses dados é então construído um grafo. Com o grafo, pode-se aplicar as técnicas propostas neste trabalho, bem como outras técnicas já conhecidas na literatura e que foram mencionadas na fundamentação. Como resultado da aplicação das técnicas, tem-se o *layout* com características de cor, tamanho e forma, em sua visão temporal e estrutural, prontos para serem analisados. As técnicas de visualização podem ser combinadas e alteradas para se realizar uma nova análise visual. As visões estrutural e temporal também possuem uma coordenação, tal que alterações na visão estrutural impactam diretamente na visão temporal, e vice versa.

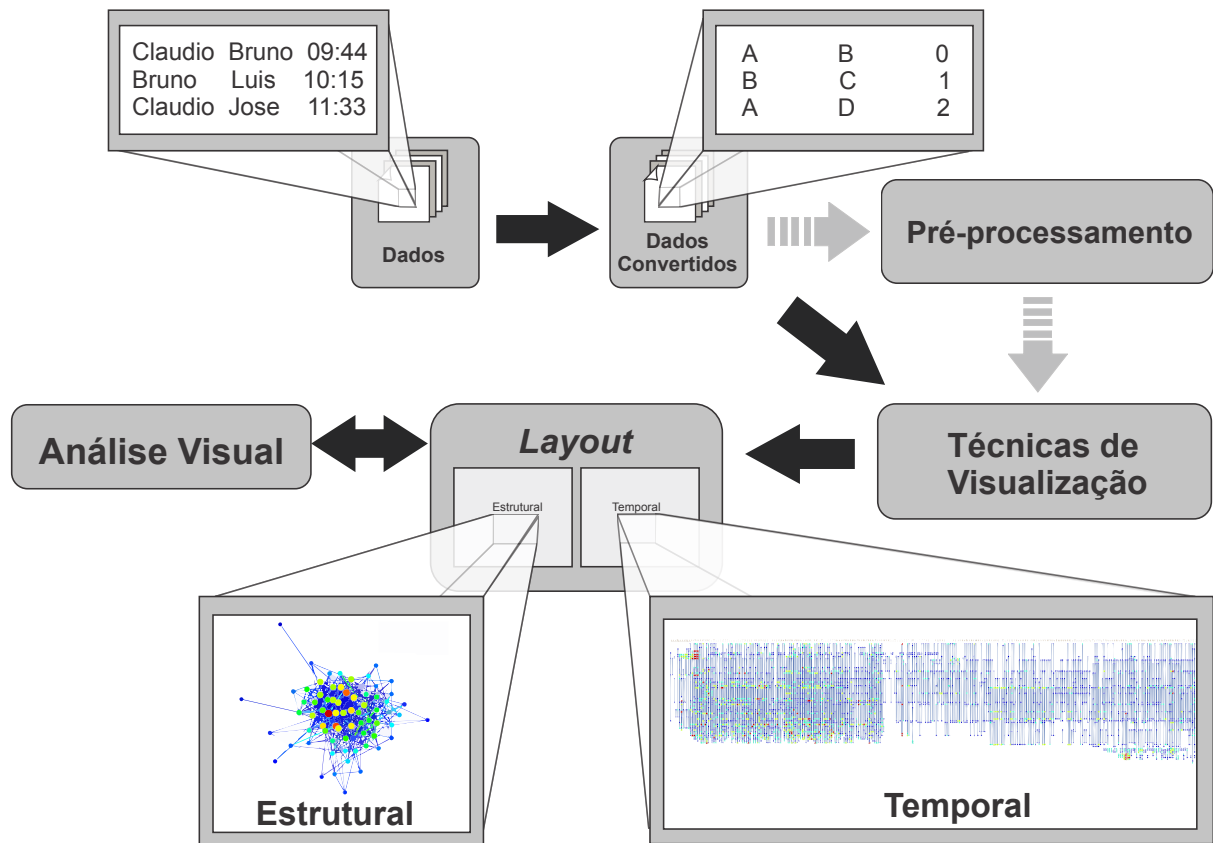


Figura 13 – Passo a passo do esquema do trabalho. Inicialmente são obtidos e convertidos os dados, realizando-se um pré-processamento nos mesmos e aplicando-se técnicas de visualização para realizar a análise visual nas visões estrutural e temporal.

3.1 Visão estrutural

Existem várias técnicas consolidadas de visão estrutural, apesar dessa área ainda ser um campo aberto de pesquisa. Nesta seção são detalhadas as técnicas utilizadas neste trabalho.

3.1.1 Organização dos vértices

Uma maneira de ajudar na percepção de padrões é alterar a disposição dos vértices em uma visão estrutural. Na visão estrutural existem algumas técnicas para organizar os vértices no espaço, dentre elas, uma consiste em dispor os vértices aleatoriamente, ou ainda utilizar algoritmos de força para separar os vértices (ZHANG et al., 2005). Algoritmos de força, como o Algoritmo de Mola, são baseados em um modelo físico, no qual os vértices são considerados anéis e as arestas são consideradas molas, nos quais as molas atraem os anéis caso eles estejam muito longe, ou repelem se eles estiverem muito perto e a posição final dos vértices representa o sistema de molas em equilíbrio. (BATTISTA et al., 1994).

Alguns outros algoritmos de ordenação dos vértices na visão estrutural são conhecidos na literatura, como o algoritmo Circular, que separa os vértices de maneira a formar um polígono regular (BAUR; BRANDES, 2005). Outra forma de ordenar os nós é utilizando o *layout* Hierárquico, em que os vértices são posicionados verticalmente, com as arestas conectando os respectivos vértices posicionadas de cima para baixo (BASTERT; MATUSZEWSKI, 2001).

Outra técnica usada por este trabalho é a visualização estrutural da rede ao longo do tempo, que são feitas de duas formas, uma considerando o tempo acumulado e outra considerando instantes de tempo isolados. A primeira proposta, ilustrada na Figura 14, exibe a rede construída considerando o tempo acumulado. A cada instante de tempo em que houver uma conexão, essa conexão será adicionada ao grafo já desenhado no instante anterior, obtendo-se no final a forma estrutural convencional. A segunda forma é demonstrada na Figura 15, na qual o grafo é exibido a cada instante de tempo, e somente são exibidas as arestas (com seus respectivos vértices) do respectivo tempo que está sendo exibido. Conforme o exemplo da Figura 15, no instante de tempo 5 somente são exibidos os vértices B e D com uma aresta conectando-os. Isso significa que no tempo 5 somente tal aresta realiza conexão, portanto as demais arestas e os vértices não são exibidos.

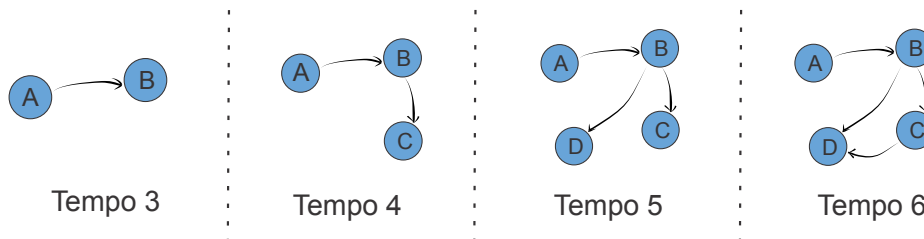


Figura 14 – Construção da rede estrutural considerando o tempo acumulado.

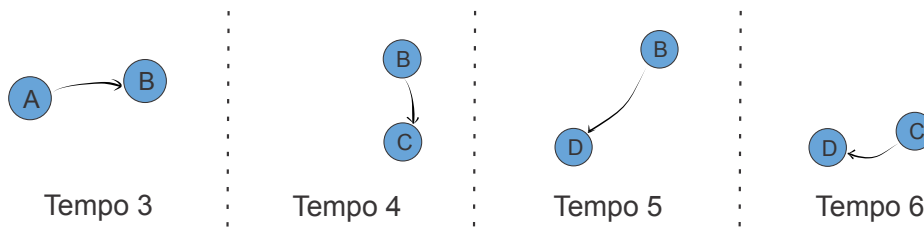


Figura 15 – Construção da rede estrutural considerando o instante de tempo.

3.1.2 Atribuição de cor, tamanho e forma

Para a diferenciação dos nós em uma visualização são atribuídos diferentes cores, tamanhos e formas a vértices e arestas. Segundo (WARE, 2004), atribuir várias características

(como por exemplo cor e tamanho) para representar um mesmo atributo é chamado de **codificação redundante**. A codificação redundante é um artifício que pode potencializar a percepção visual.

A Figura 16 apresenta alguns exemplos dessa diferenciação. Na figura, um atributo escolhido para definir o tamanho e a cor do vértice é o seu grau. As cores podem variar entre diferentes escalas de cores, utilizando a variação entre duas cores como azul e vermelho, visto na Figura 16(a), e as cores do arco-íris, visto na Figura 16(b), que são utilizadas em diversos exemplos deste trabalho. A Figura 16(c) trabalha apenas com tons de cinza. Nota-se que vértices de menor grau possuem círculos com tamanho menor e as cores à esquerda da escala de cores. Por outro lado, os vértices com maior grau possuem suas cores da direita da escala e cores, e um tamanho proporcionalmente maior.

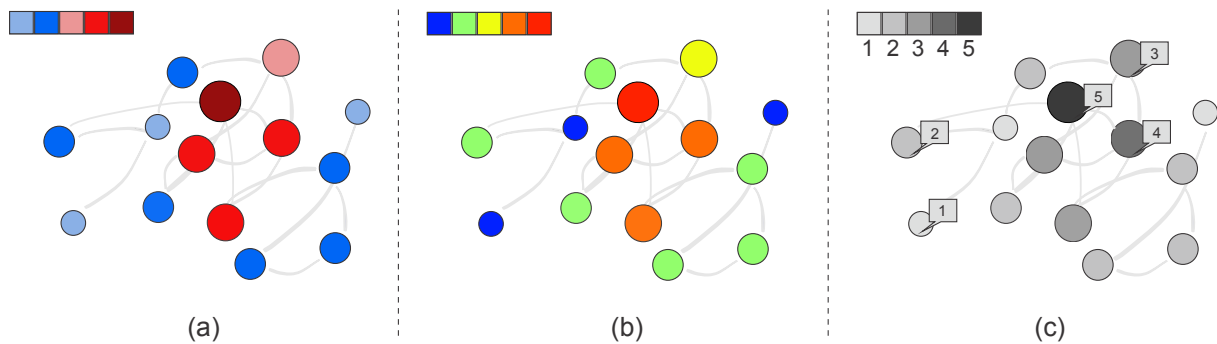


Figura 16 – Diferenciação dos vértices pelo tamanho e cor. Em (a) é aplicado uma variação de cores de azul até vermelho e em (b) é utilizado uma variação das escala de cores do arco-íris. Em (c) é utilizado tons de cinza e a caixa de texto com o número corresponde ao grau do respectivo vértice. Nota-se que em todos os casos o o tamanho e a cor dos vértices variam de acordo com o grau.

As arestas também podem ser modificadas para expressar propriedades das redes. A modificação mais usual em arestas é a adição de setas em grafos direcionados – essa representação é ilustrada na Figura 17(a). Outra modificação importante é a representação do peso da aresta, que pode ser usado para modificar a cor e o tamanho delas (BARRAT et al., 2004). Nas redes temporais usadas neste trabalho, as arestas não possuem pesos. Entretanto, pode-se considerar que na visão estrutural a quantidade de conexões entre um par de vértices pode ser interpretada como o peso em uma rede não temporal. Dessa forma, as características das arestas podem refletir essa medida.

Na Figura 17(b) tem-se a alteração da espessura e cor da aresta, representando a quantidade de conexões que existem entre aqueles dois vértices que são conectados por esta aresta. Nota-se que na Figura 17(b) existe a espessura da aresta e uma escala de cores para representar a quantidade de conexões, sendo que quanto mais grossa a espessura maior é a quantidade de conexões, e a escala de cores varia de azul até vermelho, sendo que o azul mais escuro representa uma conexão, e o vermelho mais escuro representa

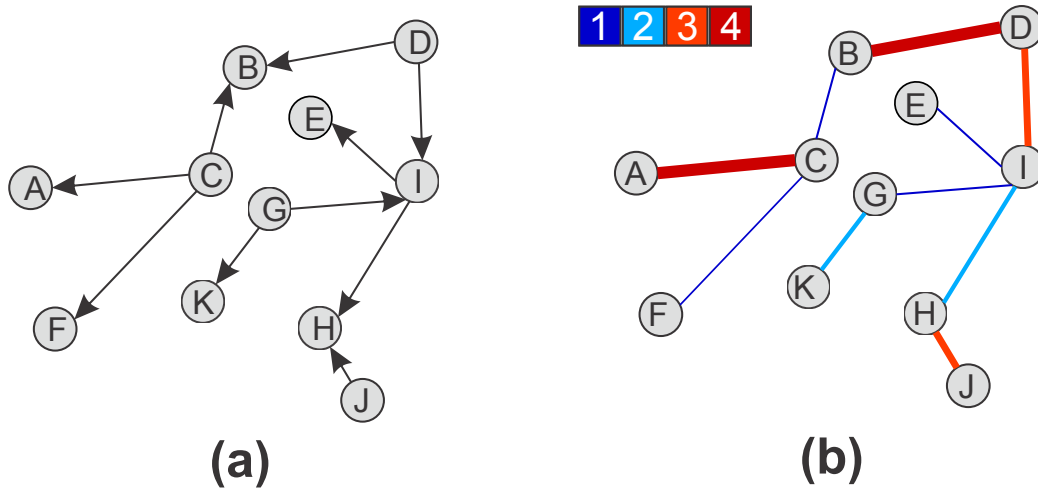


Figura 17 – Diferenciação das arestas pelo tamanho, cor e forma. Em (a) é representado um grafo direcionado adicionando-se a seta direcional nas arestas e em (b) é mostrado um grafo não direcionado em que é utilizada uma variação de cores de azul até vermelho e a espessura da aresta para representar a quantidade de conexões ao longo do tempo entre os dois vértices conectados.

quatro conexões. Portanto, estão novamente sendo utilizadas duas propriedades visuais para representar um mesmo atributo (codificação redundante), o que melhora a percepção visual e facilita a identificação das arestas.

3.1.3 Interação do usuário

Partindo do princípio de que dados reais irão gerar redes com alto número de conexões ao longo do tempo, aplicar-se-á diversas técnicas a fim de facilitar a visualização dos dados, como a diferenciação por cor, tamanho e transparência. Tendo em vista que essas redes contêm uma grande quantidade de dados, técnicas de interação foram desenvolvidas para visualizar contextos específicos, como explicitado na Figura 18. Nessa figura, está sendo representada uma seleção por profundidade, tendo selecionados o vértice F e todos os outros vértices que se conectam diretamente a ele. Tais vértices foram selecionados por serem vizinhos diretos do vértice F. Portanto, assim que o vértice F é selecionado, todos os vértices que estão diretamente conectados a ele também são destacados. Aos vértices que não foram selecionados aplica-se uma transparência.

3.2 Visão temporal

Diferentemente das técnicas da visão estrutural, não existem muitas técnicas de visualização para a visão temporal propostas na literatura. Nesta seção são descritas algumas dessas técnicas bem como são apresentadas novas propostas.

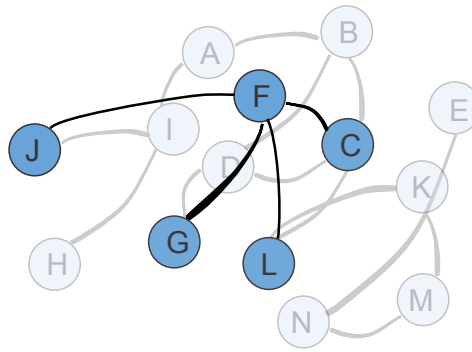


Figura 18 – Seleção por profundidade na visão estrutural. Considerando o vértice F e a distância 1, ficam destacados todos os vértices que se conectam diretamente com o vértice F.

3.2.1 Organização dos vértices

Na visão temporal, entende-se como organização dos vértices a ordem em que são dispostas as linhas horizontais. Uma maneira de ajudar na percepção de padrões é alterar a ordem dessas linhas. Em (ELZEN et al., 2013) são apresentadas diversas formas de alterar as linhas, já discutidas brevemente na Seção 2.4. Nesta seção serão descritas três técnicas de ordenação, a por grau, por nascimento e vizinhos recorrentes, detalhadas a seguir.

3.2.1.1 Nascimento

Nessa ordenação, os vértices são posicionados de acordo com a ordem em que nascem – por nascer, entende-se o instante de tempo em que o vértice se conecta com outro pela primeira vez na rede. Os vértices que aparecem nos primeiros instantes de tempo são os vértices que se localizam na parte superior do *layout* e os últimos vértices que nascem na rede são posicionados na parte inferior. Por exemplo, na Figura 12(c) a visão temporal já está ordenada pela ordem de nascimento, pois, de acordo com a tabela apresentada na Figura 12(a), tem-se que os dois primeiros vértices que nascem na rede são os vértices A e B, que são colocados nas duas primeiras linhas da visão temporal. Pela tabela, os próximos vértices que nascem são os vértices C e D, sendo colocados na terceira e quarta linha na visão temporal, respectivamente. Um outro exemplo de uma visão temporal ordenada pelo nascimento é apresentado na Figura 19(b).

3.2.1.2 Grau

Na ordenação por grau os vértices são ordenados de acordo com seus respectivos graus totais em relação a toda a rede, em que os menores graus se posicionam na parte superior e os maiores graus se posicionam na parte inferior. Na Figura 19(c) é exemplificado a ordenação pelo grau. Com essa ordenação, tem-se a sequência em que os vértices

são posicionados na visão temporal, com o vértice E sendo o vértice de menor grau e posicionado na primeira linha, e o vértice B sendo o vértice de maior grau que é então posicionado na última linha. Para vértices de mesmo grau, como no exemplo o vértice A e D, é escolhido aleatoriamente qual dos vértices vem primeiro na sequência de linhas.

3.2.1.3 Vizinhos Recorrentes

A técnica Vizinhos Recorrentes é uma proposta deste trabalho que tem como objetivo minimizar o tamanho das arestas na visão temporal, tendo em vista que, com a aproximação dos vértices que mais se conectam entre si, possivelmente haverá uma redução do tamanho das arestas. Em Algoritmo 1 são mostrados os passos dessa organização.

Inicialmente coloca-se o vértice de maior grau da rede no centro do *layout*, posteriormente seleciona-se os dois vértices que mais se conectam a este vértice de maior grau, colocando-se um vértice em cima e o outro embaixo, por ordem de chegada. Em seguida, repete-se o algoritmo para a parte de cima do grafo e para a parte de baixo do grafo, sempre selecionando o vértice que mais se conecta àquele que está sendo avaliado e ainda não tenha sido utilizado. Caso não exista mais nenhum vértice que se conecte àquele que está sendo avaliado, o algoritmo seleciona o vértice de maior grau que ainda não foi utilizado. O ponto de parada do algoritmo se dá quando todos os vértices do grafo já tenham sido posicionados.

A Figura 19(d) ilustra como fica uma ordenação utilizando a técnica de vizinhos recorrentes. O primeiro passo para organizar os vértices consiste em calcular os graus desses vértices. Do exemplo pode-se observar os graus conforme a Tabela 19(a). Portanto, o primeiro vértice escolhido é o B. Este vértice será posicionado no centro do *layout*, na posição em que se encontra originalmente o vértice C, quando usada a ordenação por nascimento (Figura 19(b)). O vértice B se conecta mais com os vértices A e D. O único vértice restante e que se conecta ao vértice A é o C, portanto ele é posto como vizinho, e, por fim, coloca-se o último vértice disponível (vértice E) na última linha.

3.2.2 Atribuição de cor, tamanho e forma

Na visão temporal, utiliza-se a cor para representar a medida de grau. A escala é feita de um até ao maior grau da rede. Para determinar o maior grau, verifica-se o grau de todos os vértices para cada instante de tempo separadamente, e seleciona-se o maior. Nota-se que um único valor é determinado para toda a rede e para todos os instantes de tempo.

Na Figura 20(a) criou-se uma rede para simular o funcionamento das cores na visão temporal. Nesse exemplo, o grau máximo é 4, que corresponde ao grau do vértice A no instante 7. Portanto, as cores irão variar de 1 a 4, como é visto na Figura 20(b). Nessa figura, foi utilizada a escala de cores do arco-íris, de acordo com a legenda da figura, em

Algoritmo 1: Técnica Vizinhos Recorrentes

```

1 início
2   Vetor de nós  $n$ ;
3   Seleciona o elemento de  $n$  com o maior grau para posicionar no centro do
     layout;
4   Seleciona os dois vértices que mais se conectam ao vértice de maior grau;
5   Posiciona o vértice que mais se conecta abaixo do centro do layout;
6   Posiciona o segundo vértice que mais se conecta acima do centro do layout;
7   Retira de  $n$  os vértices já posicionados no layout;
8   para toda a parte superior do layout faça
9     Considerando a posição central  $p$  do layout;
10    Seleciona o vértice que mais se conecta ao de maior grau em  $n$ ;
11    Posiciona o vértice que mais se conecta na posição  $p$  do layout;
12    Subtrai 1 em  $p$  para uma nova posição no layout;
13    Retira de  $n$  o vértice posicionado;
14  fim
15  para toda a parte inferior do layout faça
16    Considerando a posição central  $p$  do layout;
17    Seleciona o vértice que mais se conecta ao de maior grau em  $n$ ;
18    Posiciona o vértice que mais se conecta na posição  $p$  do layout;
19    Adiciona 1 em  $p$  para uma nova posição no layout;
20    Retira de  $n$  o vértice posicionado;
21  fim
22 fim

```

que a cor azul representa menos conexões e a cor vermelha representa os vértices com mais conexões no mesmo instante de tempo. Além da propriedade de cor, é possível alterar o tamanho e forma dos vértices na visão temporal, como é demonstrado na técnica de Mapa de Atividade Temporal, descrita na próxima seção.

3.2.3 Mapa de Atividade Temporal

Para demonstrar somente a atividade dos vértices na visão temporal é proposta uma nova visualização de redes temporais chamada de Mapa de Atividade Temporal (MAT), que, ao invés de vértices redondos, com ligações no mesmo tempo, é proposta uma outra visualização, em que cada vértice é um quadrilátero preenchido no espaço e sua cor varia de acordo com o grau (visto na seção 3.2.2). Nessa visualização são descartadas as arestas, focando-se somente na atividade do vértice.

Aplicando o MAT na mesma rede da Figura 20(b), tem-se a representação vista na Figura 20(c). As arestas são retiradas a fim de destacar a atividade dos vértices e quadrados são utilizados para demonstrar uma maior noção de continuidade (WARE, 2004). Nota-se que, na Figura 20(c), a identificação visual de quais são os vértices com maiores atividades fica facilitada, ao contrário do que é visto na Figura 20(b).

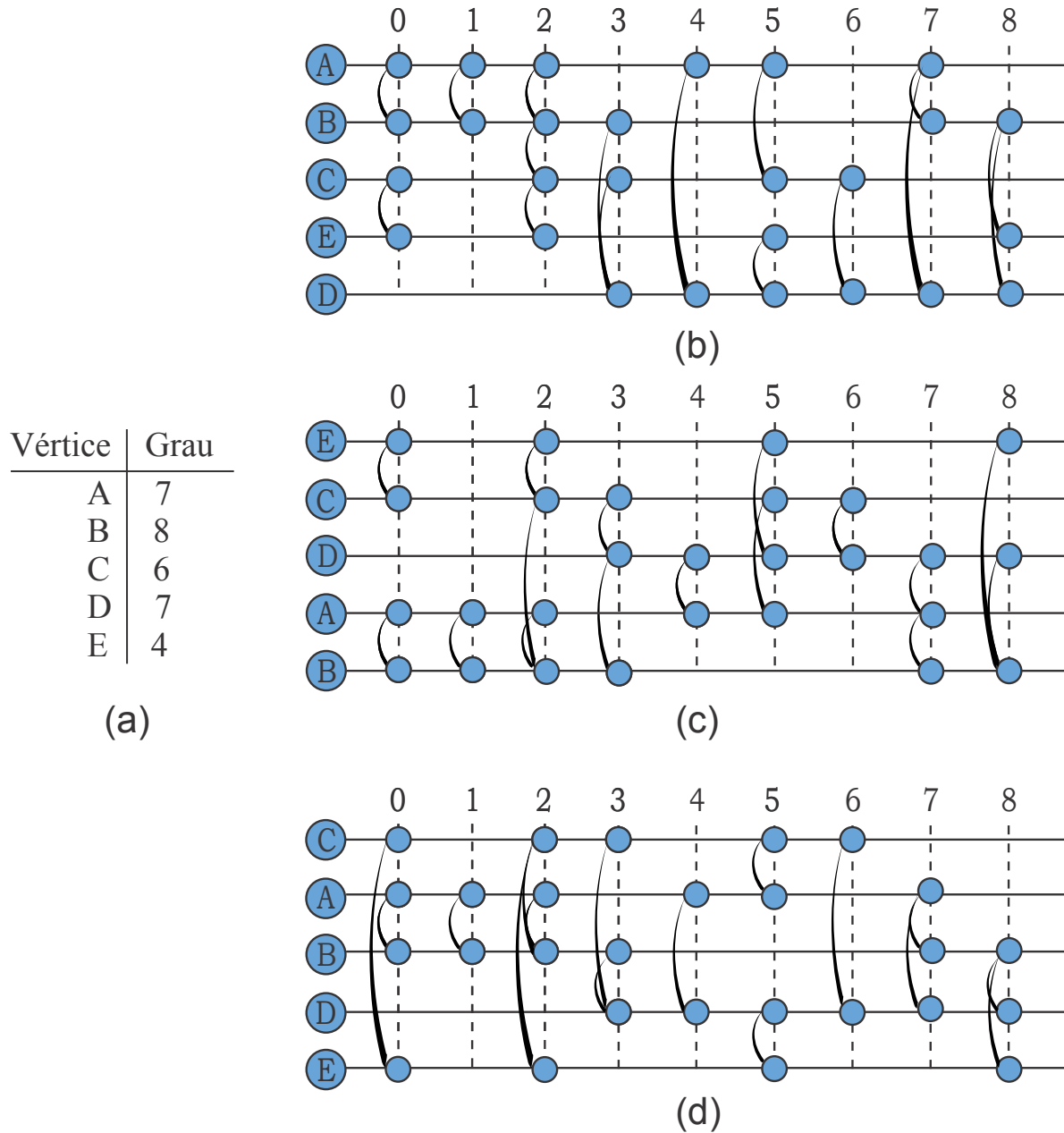
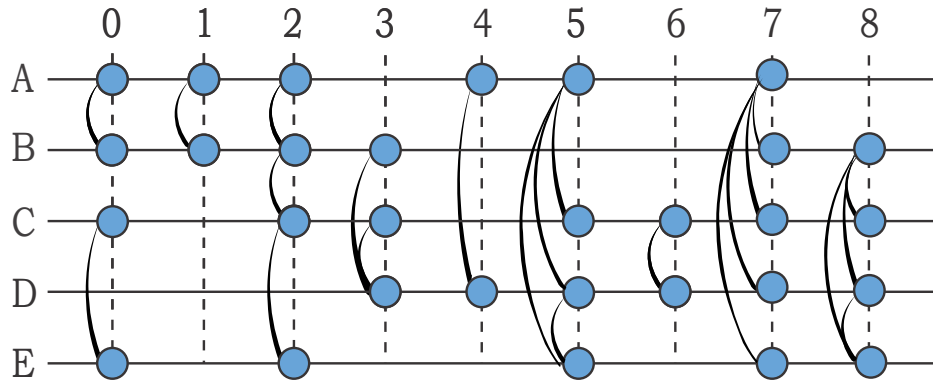
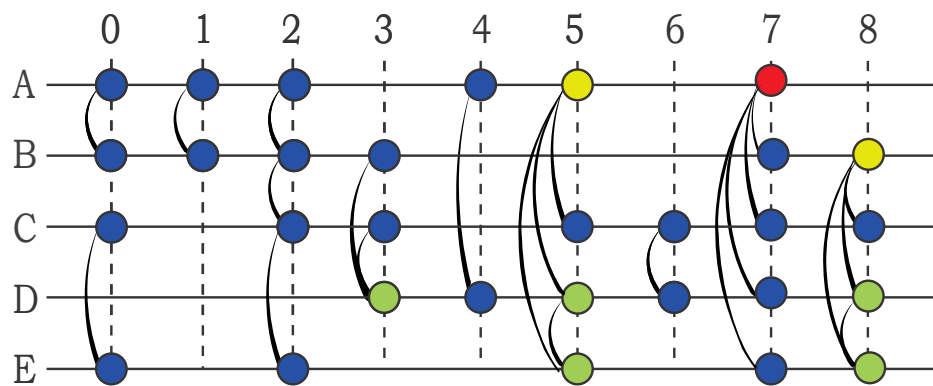


Figura 19 – Diferentes estratégias de ordenação para o *layout* temporal. Em (a) é representado a tabela indicando o grau dos nós da rede de (b). Em (b) é representado uma Rede temporal, em que as linhas são ordenadas de acordo com a ordem de nascimento dos vértices. Em (c) é representado as linhas ordenadas de acordo com o Grau do vértice. Em (d) é representado as linhas ordenadas de acordo com o algoritmo de Vizinhos Recorrentes.

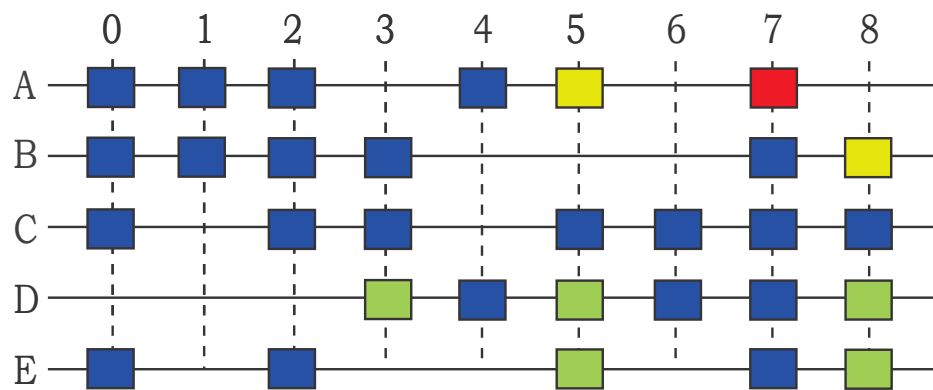
A Figura 21 mostra o MAT de uma rede que representa as atividades das pessoas que visitam um Museu (visto na seção 4.1). O MAT é representado de duas maneiras: primeiro, utilizando-se a cor de fundo como a menor cor da escala de cores, destacando-se assim, apenas os vértices que apresentam uma atividade atípica (Figura 21(b)); ou utilizando uma cor de fundo diferente da escala de cores representada, como por exemplo, a cor



(a)



(b)



(c)

Figura 20 – Rede temporal para exemplificar o funcionamento do algoritmo de cores na visão temporal. Em (a) representa o modelo sem o algoritmo de cores. Em (b) representa o modelo com o algoritmo de cores aplicado. As cores de (b) variam, de acordo com a legenda, de 1 à 4. Em (c) contém a representação do MAT correspondente aos dados demonstrados em (b). As arestas foram ocultadas e o formato dos vértices foi alterado de círculo para quadrado, a fim de demonstrar uma noção de continuidade.

branca, destacando todos os vértices na rede temporal (Figura 21(a)). Na Figura 21(b) existe perda de informação, pois vértices com baixa atividade na rede tenderão a desaparecer nesta visualização, destacando-se somente vértices com uma atividade maior que dos demais. Em contrapartida, na Figura 21(a), os vértices com menos atividades são mais visíveis, porém os vértices com maior atividade não são tão destacados quanto na Figura 21(b).

3.2.4 Interação do usuário

Um dos problemas da grande quantidade de dados é a dificuldade de visualizar todos os vértices e arestas ao mesmo tempo em tela, e isso não é diferente na visão temporal. Portanto, técnicas de seleção dos dados visando destacá-los dos demais são necessárias. Na Figura 22 são apresentadas três formas diferentes de interação nos dados na visão temporal.

Na Figura 22(a) é apresentada a seleção por linhas, sendo que, devido ao fato de cada linha representar um nó, ao selecioná-la gera destaque daquele respectivo nó em todos os instantes de tempo. Como no exemplo da Figura 22(a) é selecionada as linhas dos vértices A e B, então os vértices A e B são selecionados em todos os instantes de tempo, juntamente com as arestas que ligam estes dois vértices. Portanto, caso haja arestas entre dois ou mais vértices selecionados, estas arestas são destacadas. Na Figura 22(b) é demonstrado a seleção pela faixa temporal, com os instantes de tempo 1 e 2 selecionados, destacando conexões entre os nós A, B e C. A Figura 22(c) demonstra a seleção mais específica, selecionando um conjunto de vértices em uma porção do tempo. Nessa figura são destacado os vértices B, C e D nos instantes de tempo 1, 2 e 3. Essas formas de interagir com os dados na visão temporal visam selecionar um contexto específico para ser analisado.

3.3 Resolução temporal da rede

Um outro conceito abordado neste trabalho refere-se a possibilidade de se trabalhar com diferentes resoluções de tempo dos dados. Essa modificação da resolução consiste em amostrar novamente os dados temporalmente, agrupando os instantes de tempo para demonstrar tendências difíceis de serem visualizadas na resolução original. Para alterar a resolução utiliza-se a Equação 1, em que t'_i é o novo instante de tempo, t_i é um instante de tempo da rede original, t_s é o primeiro instante de tempo e f é o fator de resolução, isto é, a escala em que a resolução será alterada. Se, por exemplo, o valor de f for 2, isso implica que a resolução final será diminuída pela metade. Por exemplo, um intervalo temporal original de uma rede é de 20 em 20 segundos, ou seja, a cada 20 segundos são amostradas conexões, caso aplique a resolução na rede com um fator de resolução f igual

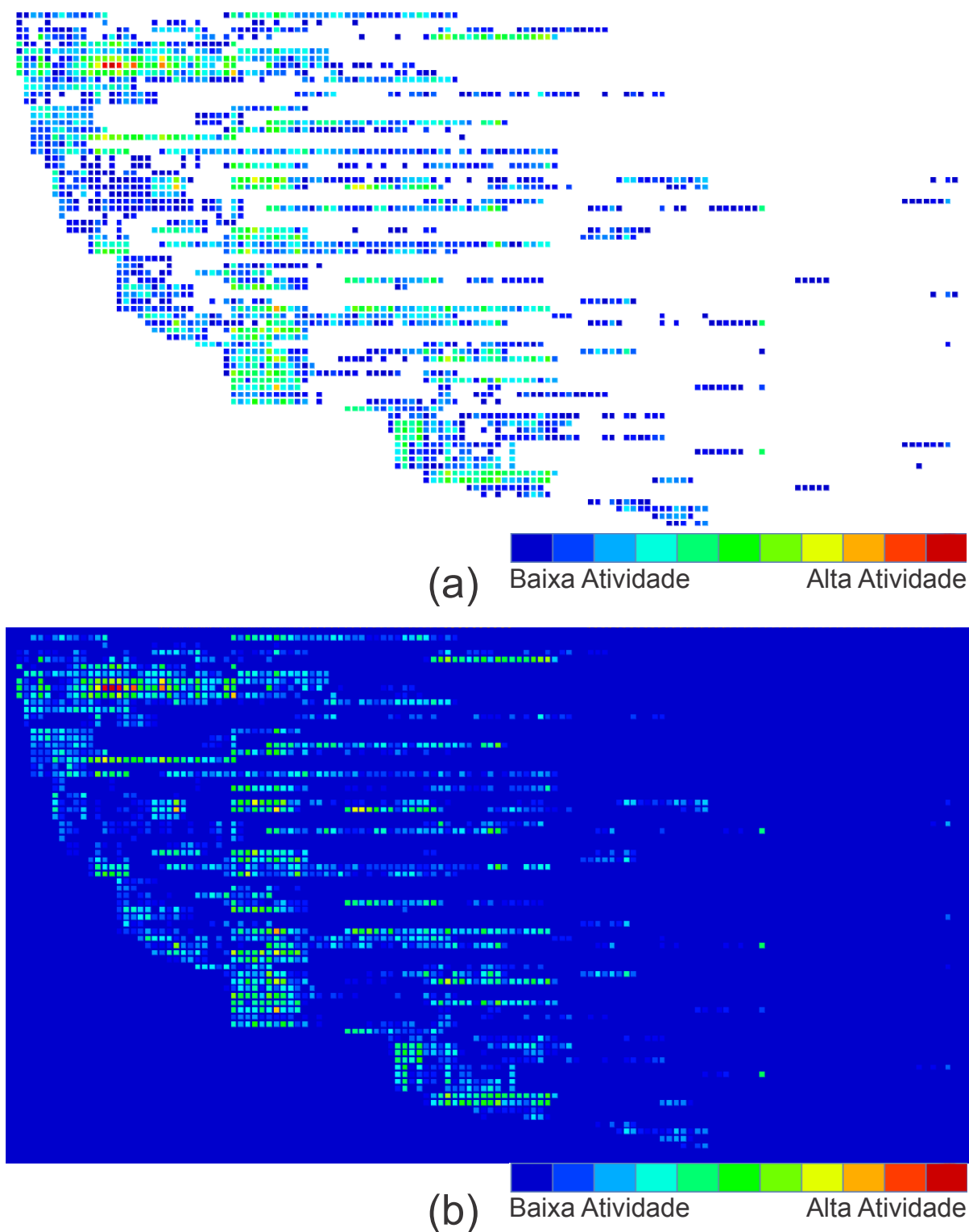


Figura 21 – Exemplificação do MAT. A escala de cores utilizada é as cores do arco-íris, como é visto no canto inferior direito. Em (a) é representado o MAT com o fundo branco. Em (b) é representado o MAT com o fundo correspondente à menor cor da escala.

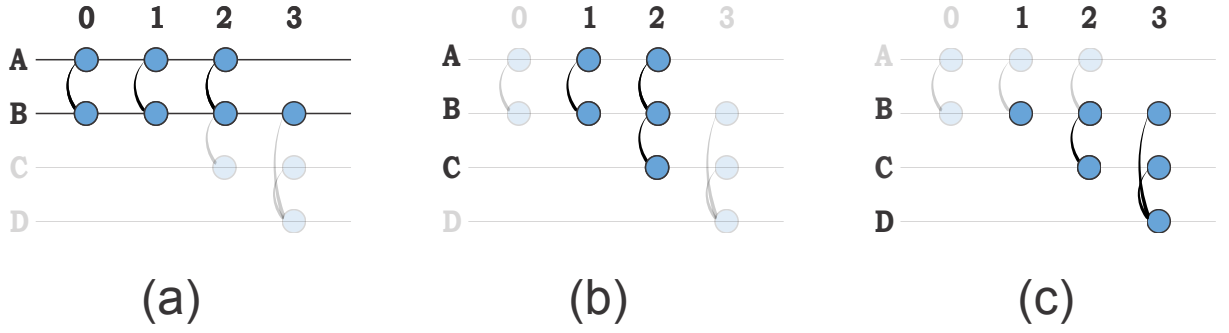


Figura 22 – Representação das interações por meio da seleção na visão temporal. Em (a) tem-se a seleção por linha, com os vértices A e B selecionados. Em (b) tem-se a seleção por faixa de tempo, com o tempo 1 e 2 selecionados. Em (c) tem a seleção específica dos nós B, C e D nos instantes de tempo 1, 2 e 3.

a 2, esse intervalo de resolução passa de 20 segundos para 40 segundos, ou seja, somente a cada 40 segundos são amostradas conexões.

$$t'_i = \left\lfloor \frac{t_i - t_s}{f} \right\rfloor f + t_s \quad (1)$$

A Figura 23 representa o impacto da alteração de resolução. Ela mostra a visão temporal de uma rede com seus dados originais, sendo que suas arestas estão com o atributo de tempo original. Já a Figura 23(b) representa as mesmas conexões da Figura 23(a), porém foi aplicado a resolução 2 ao tempo destas arestas. É importante observar na Figura 23(b) que o impacto da alteração da resolução gera a sobreposição de algumas arestas da rede. A alteração da resolução é bem aplicada em redes muito esparsas temporalmente, visando agrupar algumas conexões, auxiliando a identificação de padrões.

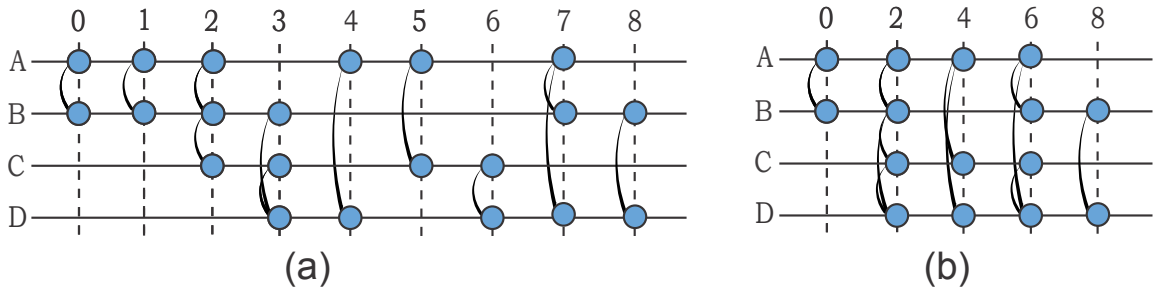


Figura 23 – Ilustração do conceito de resolução aplicado a visão temporal. Em (a) representa a visão temporal com a resolução inicial dos dados igual a 1 e em (b) representa a visão correspondente após a mudança de resolução utilizando um fator $f = 2$.

3.4 Coordenação entre visões

Um dos focos deste trabalho é possibilitar uma comunicação entre as representações de uma rede complexa, sendo possível, portanto, selecionar um vértice de uma visão e gerar destaque desse vértice em todas as visões. Assim que um conjunto de vértices em uma visão é selecionado, a seleção correspondente na outra visão é selecionada, como demonstrado na Figura 24. Nesse exemplo, os vértices B e D estão em destaque na visão estrutural e temporal. Devido à coordenação das visões, os vértices B e D ganharam destaque nas duas visões, assim como todas as conexões que existem entre eles.

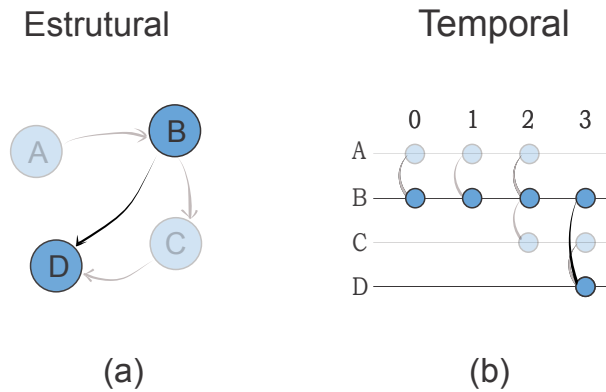


Figura 24 – Comunicação entre as visões estrutural e temporal. Coordenação entre a visão estrutural (a), e a visão temporal (b), considerando a seleção dos vértices B e D.

3.5 Sistema *DyNetVis*

Um sistema foi desenvolvido para comprovar a eficácia das estratégias apresentadas, tendo como finalidade auxiliar a análise e automatizar o processo de construção dos modelos de visualização. O sistema é denominado Sistema *Dynamic Network Visualization* (DyNetVis), desenvolvido na linguagem *Java*, visando ter suporte para diferentes plataformas.

O sistema utiliza uma biblioteca de grafos chamada *jgraphx* (ALDER, 2010), que fornece funcionalidades para visualização e interação em grafos, para a linguagem *Java*. Outros testes foram realizados com algumas bibliotecas, como *Java2D*, *GraphStream* e *JUNG*, bem como outras bibliotecas foram descartadas inicialmente por não apresentarem suporte as funcionalidades básicas do sistema. Utilizando a biblioteca *Java2D* o sistema apresentava lentidão para as funcionalidades básicas para todas as redes. Com as bibliotecas *JUNG* e *GraphStream* o sistema conseguia carregar as redes maiores, mas as funcionalidades básicas do sistema apresentavam lentidão para serem executadas.

A biblioteca *jgraphx* foi escolhida por fornecer suporte para as funcionalidades implementadas. Além disso, a biblioteca facilita a criação das visões, como algoritmos já implementados de organização dos nós para a visão estrutural, fornecendo suporte para *layouts* de algoritmos de força e hierarquias (ALDER, 2010). Testes foram realizados com a biblioteca *jgraphx* (como visto na seção 4.2.3) apresentando tempo de renderização satisfatório.

O sistema inicialmente recebe os dados de entrada em um arquivo no formato de origem, destino e tempo, que podem ser filtrados de acordo com a requisição do usuário; logo após os filtros serem selecionados, os dados são armazenados na estrutura de grafos original da biblioteca *jgraphx* e são exibidos na tela utilizando a biblioteca gráfica *jgraphx*, juntamente com os componentes de interface do *Java Swing*. As funções das janelas se dividem em estrutural e temporal.

3.5.1 Funcionalidades do sistema

O sistema DyNetVis possui diversas funcionalidades, entre elas as técnicas apresentadas neste trabalho aplicadas as visões estrutural e temporal. As funcionalidades explicadas nesta seção são divididas em funcionalidades comuns entre as visões temporal e estrutural, funcionalidades específicas da visão estrutural e funcionalidades específicas da visão temporal.

Nas Figuras 26 e 27 são detalhadas a tela inicial do *software* com suas respectivas funcionalidades. Na Figura 26 são listadas as funcionalidades em comum das visões estrutural e temporal (em verde) e as funcionalidades exclusivas da visão estrutural (em azul). Na Figura 27 são listadas as funcionalidades exclusivas da visão temporal (em amarelo).

3.5.1.1 Funcionalidades em comum das visões estrutural e temporal

Algumas funcionalidades são comuns das duas visões (estrutural e temporal). Na Figura 26 são listadas essas funcionalidades em caixas verdes. Inicialmente para carregar a rede no sistema é necessário abrir o arquivo com a rede correspondente. Após a rede estar carregada, é possível ver detalhes de suas propriedades, como quantidade de vértices e arestas, bem como alterar algumas configurações, como selecionar a qualidade de renderização e a cor de fundo de todo o sistema. Também é possível que seja alterada a resolução, bem como o tempo inicial e final da rede. O sistema é inicialmente definido sem nenhuma regra aplicada a variação de cores, sendo possível realizar alteração da cor, tamanho e forma. Alguns padrões de cores são pré-definidos, como escala de cinza, cores do arco-íris, cores variando de azul até vermelho, podendo também o usuário definir um novo intervalo.

Pela tela de alteração de cores, tamanho e forma (visto na Figura 25(b)) é possível selecionar *templates* pré-definidos, como os seguintes: deixar as cores, tamanhos e formas como definido inicialmente no sistema; selecionar as cores para vértices e arestas em relação ao grau na visão estrutural e sobre a atividade do vértice na visão temporal; selecionar o MAT com o fundo branco; selecionar o MAT com o fundo da menor cor da escala selecionada. Importante observar que por essa tela é possível selecionar separadamente se o tamanho e a cor dos vértices vão crescer conforme o grau do vértice, de forma análoga para as cores e a espessura da aresta, ou seja, é possível selecionar se os vértices e arestas contêm uma codificação redundante. Também é possível mostrar ou ocultar vértices e arestas em ambas visões, bem como aplicar o *zoom* para aproximar ou distanciar as visões da tela.

3.5.1.2 Funcionalidades específicas da visão estrutural

Algumas funcionalidades do sistema são específicas para visão estrutural, como é demonstrado na Figura 26. O sistema permite que os vértices sejam selecionados e movidos da posição em tela na visão estrutural, bem como exibe quantos e quais vértices foram selecionados. O sistema permite que seja selecionado se a visualização da rede é exibida direcionada, adicionando a direção da aresta, ou não direcionada, com somente uma linha para representar a aresta. Algumas formas de posicionar os vértices são disponibilizadas, como algoritmos de força, hierárquica e posicionamento aleatório. Algumas outras funcionalidades podem se aplicar a visão estrutural, como mostrar ou ocultar a identificação do vértice e mostrar ou ocultar o peso das arestas. O sistema permite que seja exibido uma animação da construção do grafo, variando de acordo com os parâmetros do tempo inicial e final, bem como se essa construção será feita de forma a considerar o tempo acumulado ou somente o instante de tempo. A seleção por profundidade não foi implementada neste trabalho.

3.5.1.3 Funcionalidades específicas da visão temporal

Algumas funcionalidades do sistema são específicas para visão temporal, como é demonstrado na Figura 27. Na visão temporal é possível somente selecionar os nós de acordo com as três formas que foram explicadas para interação do usuário na visão temporal (visto na seção 3.2.4). O sistema permite mostrar ou ocultar as linhas horizontais da visão temporal, que auxilia a manter uma visualização mais limpa quando é desejado uma visão mais geral da rede, bem como aumentar ou reduzir o espaço entre essas linhas, que auxilia a perceber os padrões quando o sistema é visualizado em uma visão geral da rede. Além das funcionalidades de *zoom* que são comuns entre ambas as visões, o sistema possui duas funcionalidades de *zoom* automático, uma que permite que toda a faixa temporal seja visualizada em tela ao mesmo tempo e outra para voltar ao *zoom* original. O

sistema implementa os algoritmos de ordenação dos vértices na visão temporal descrito na seção 3.2.1.

Na Figura 25(a) tem-se uma das telas iniciais do sistema, na qual o arquivo de entrada pode ser selecionado. O sistema já fornece uma pré-configuração de abertura do arquivo para o usuário, podendo ser alterada antes do mesmo ser carregado. O sistema permite um filtro inicial do tempo inicial e final do sistema, bem como a resolução na qual o arquivo será carregado.

Após o arquivo ser carregado, tem-se uma janela de customização de cores, tamanho e forma dos vértices e arestas, das visões estrutural e temporal, como pode ser visto na Figura 25(b). Existem alguns *templates* pré-definidos, porém a combinação das funcionalidades é permitida seguindo as opções que o sistema fornece. Em qualquer momento da navegação, no sistema, pode-se ter acesso a essas configurações, sendo possível que tais funcionalidades sejam alteradas diversas vezes em uma mesma análise.

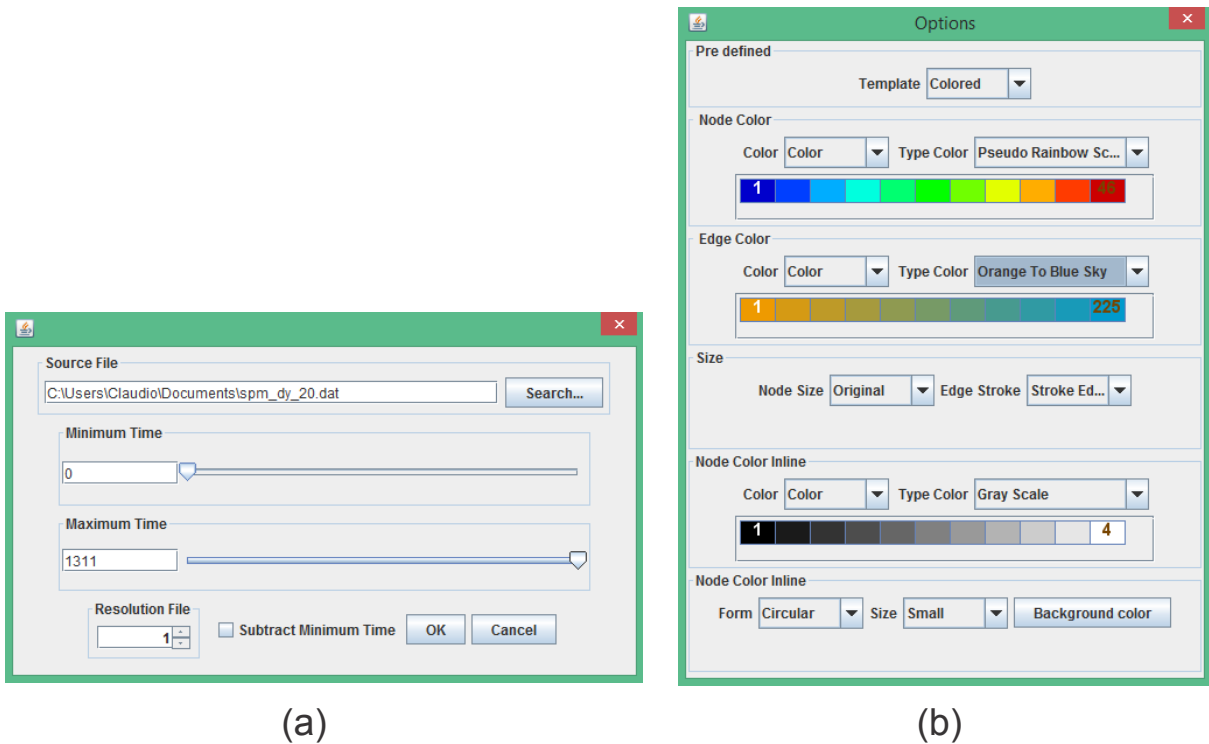


Figura 25 – Telas de filtros e customização do sistema *DyNetVis*. Em (a) tem-se a tela de selecionar um arquivo do sistema *DyNetVis* com seus respectivos filtros de alterar o tempo inicial e final e de alterar a resolução. Em (b) tem-se a tela de selecionar cores e formas do sistema *DyNetVis* nas visões estrutural e temporal.

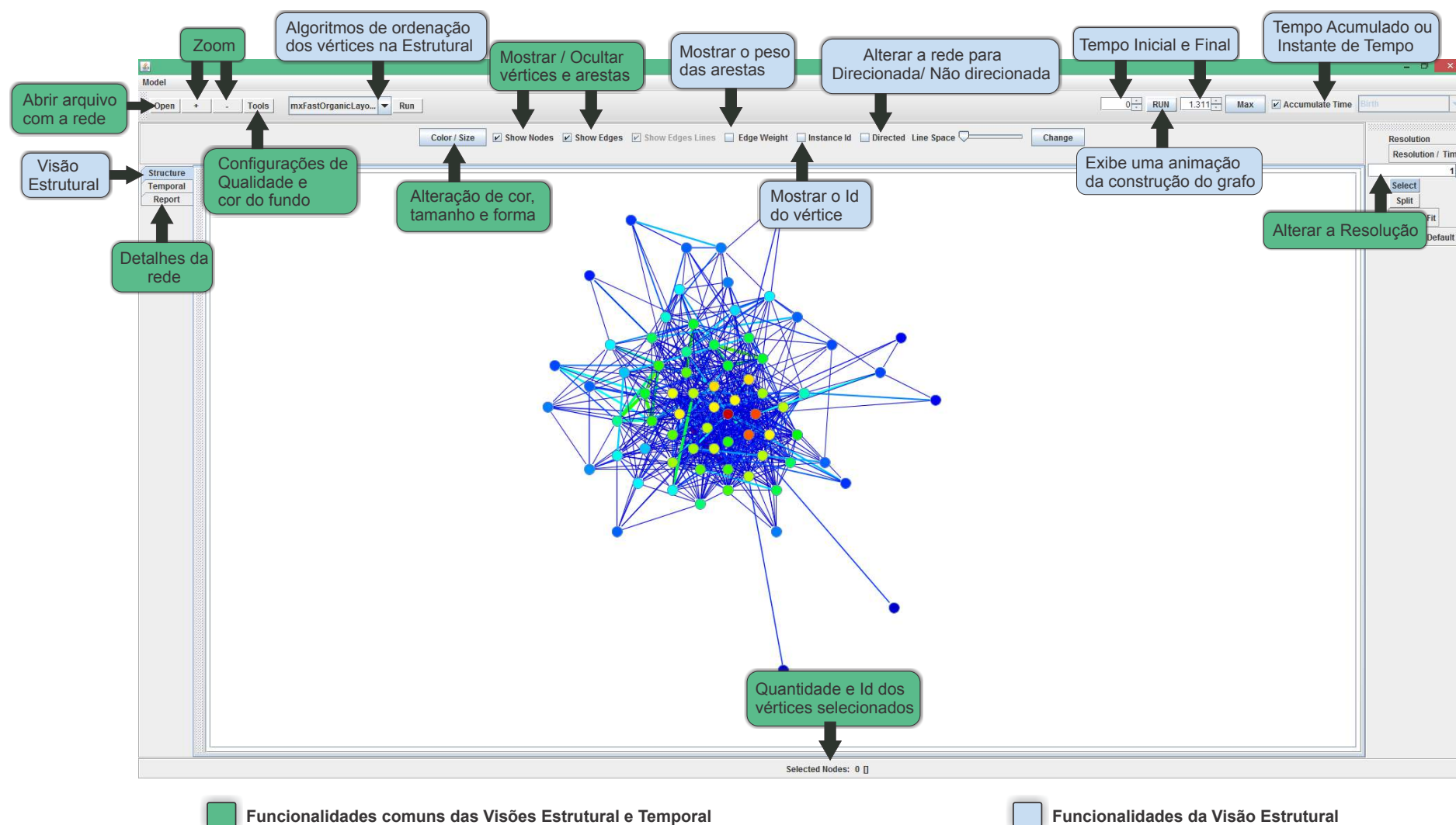


Figura 26 – Visão estrutural detalhada da tela inicial do *software*. As funcionalidades que são demonstradas em caixas verdes são funcionalidades comuns das visões estrutural e temporal, já as demonstradas em caixas azuis são funcionalidades específicas da visão estrutural.

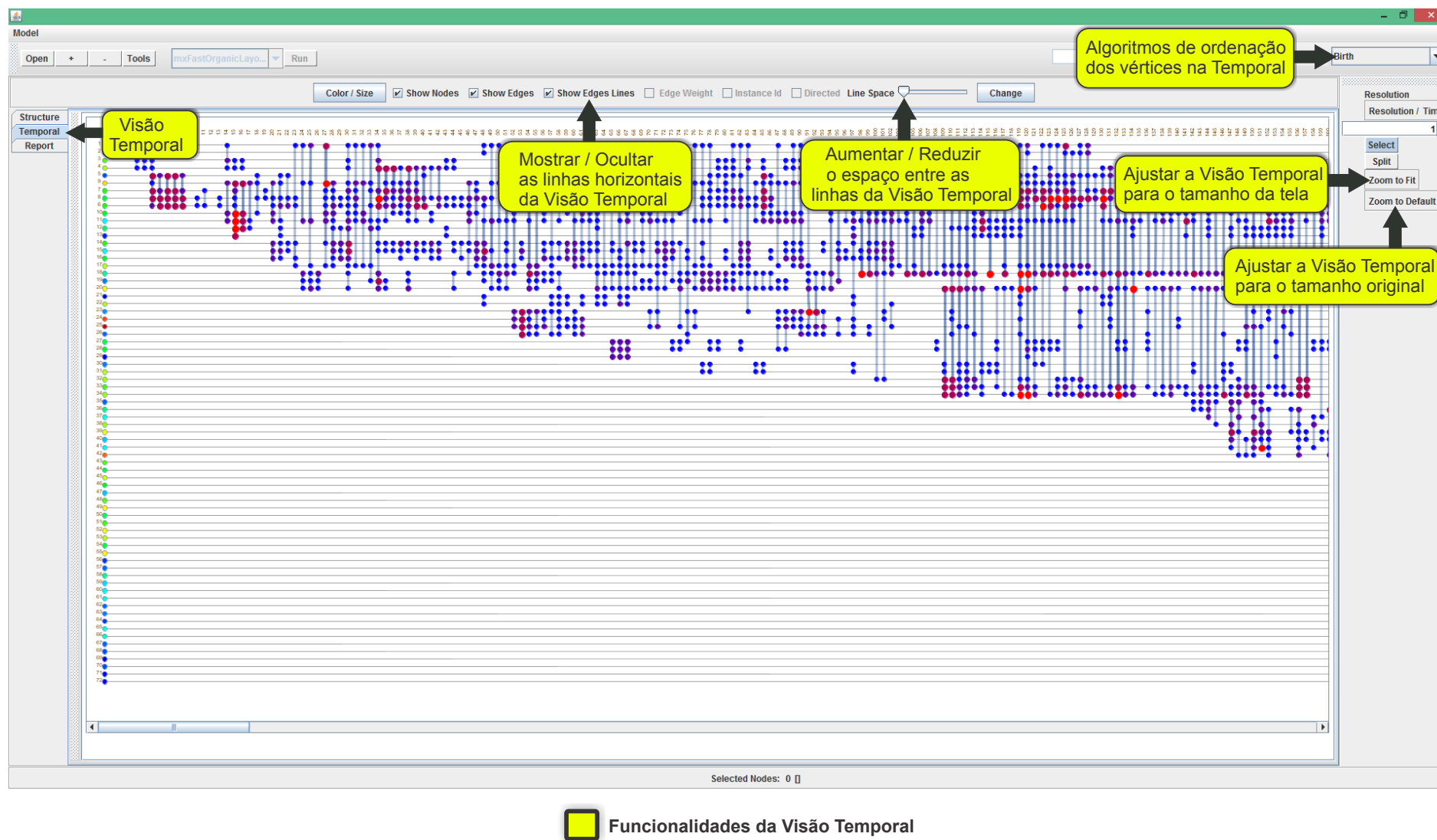


Figura 27 – Visão temporal detalhada da tela inicial do *software*. As funcionalidades mostradas são específicas da visão temporal.

Experimentos e Análise dos Resultados

Este capítulo apresenta experimentos de aplicação das técnicas propostas neste trabalho, implementadas no sistema *DyNetVis* para realizar a análise e exploração de redes temporais de diversas naturezas. Para realizar os experimentos, são utilizados **redes de contato**, ou também conhecida como redes de proximidade para redes coletadas por RFID, que são redes modeladas a partir da informação de quando, dentro de um grupo de pessoas, ocorre o contato entre duas pessoas.

Devido à forma de coleta dos dados analisados ser anônima, tem-se pouca informação (em alguns casos não existe informação precisa) do que acontece em cada instante de tempo, portanto faz-se necessárias suposições, ao analisar os dados, de fatos que podem acontecer naquele contexto. Para se ter uma maior precisão na identificação de padrões, é imprescindível um conhecimento sobre os dados para melhor entendimento dos comportamentos visualizados.

4.1 Redes de contato

Nesta seção são detalhadas as redes de contato que são utilizadas para os experimentos deste trabalho. Ao final desta seção, são sumarizados as informações estruturais das redes.

As redes de contato utilizadas neste trabalho foram obtidas de dados já publicados na literatura, e contêm dados de uma Conferência (ISELLA et al., 2011), uma Escola (FOURNET; BARRAT, 2014), um Hospital (VANHEMS et al., 2013), um Museu (ISELLA et al., 2011) e uma Rede Sexual (ROCHA; LILJEROS; HOLME, 2011).

Com exceção da rede Sexual, a coleta dos dados foi realizada em ambientes reais, com tecnologias de sensores cara-a-cara RFID. Dispositivos de RFID monitoram um contato de aproximação cara-a-cara de dois indivíduos distanciados de 1 até 1,5 metros um do outro, com o registro desse contato sendo gravado em uma faixa de 20 em 20 segundos. Dois indivíduos fazem contato caso seus dispositivos troquem informações com a distância permitida na janela temporal de 20 segundos, e esse contato é considerado interrompido

caso o dispositivo não troque informações em um intervalo de 20 segundos (FOURNET; BARRAT, 2014).

Os dados da rede do **Museu** (ISELLA et al., 2011) são da *Science Gallery*, em Dublin na Irlanda. O evento envolveu 72 indivíduos, com 6980 conexões entre eles.

Os dados da rede de **Conferência** (ISELLA et al., 2011), correspondem ao evento *Hypertext*, ocorrido em Turim na Itália, durante 3 dias entre 29 de junho e 1 de julho, em 2009. O evento envolveu 113 indivíduos, com 20818 conexões entre eles. Em um contexto da conferência, os participantes mantêm-se ativos – ou seja, há comunicação do participante com outros participantes – durante toda a duração do evento, e a movimentação se dá entre as áreas da conferência, como as salas, pausas, entre outras.

Os dados da rede do **Hospital** (VANHEMS et al., 2013) são da unidade geriátrica de um hospital universitário com quase 1.000 leitos, em Lyon na França, das 13h00m de uma segunda-feira, 6 de Dezembro até as 14h00m de uma sexta-feira, 10 de Dezembro, no ano de 2010. Os dados foram coletados de 46 funcionários (taxa de 92% de participação) e 29 pacientes (taxa de 94% de participação). Dos 46 funcionários, 27 são enfermeiras ou auxiliar de enfermagem, 11 médicos e 8 funcionários administrativos. Visitantes não foram incluídos no estudo. Os participantes foram monitorados durante todo o tempo em que permaneceram no hospital.

Os dados da rede da **Escola** (FOURNET; BARRAT, 2014) são da escola de Ensino Médio de Lycée Thiers, em Marseilles na França, no mês de Novembro de 2012. A coleta envolveu cinco turmas diferentes, monitorando 180 alunos. Essas turmas são chamadas de *classes préparatoires*, que são específicas do sistema escolar da França. Esse sistema dura dois anos e seleciona estudantes que já terminaram os estudos usuais do Ensino Médio. As turmas coletadas correspondem a estudantes do segundo ano do sistema correspondente. Os estudantes foram monitorados durante todo o tempo em que permaneceram na escola.

Os dados da rede **Sexual** (ROCHA; LILJEROS; HOLME, 2011) foram coletados a partir de resenhas um fórum (*web-forum*) público no Brasil, que é direcionado para homens homossexuais que avaliam e comentam sobre seus encontros sexuais com garotas de programa, no período entre Setembro de 2002 até Outubro de 2008, com um intervalo temporal de 1 dia, abrangendo 12 cidades brasileiras. A seção coletada é a de acompanhantes, por ser a seção mais organizada. Os dados são formados a partir de resenhas, em que os clientes e as garotas de programa são os vértices, as arestas são as conexões entre clientes e garotas de programa e a data de postagem das resenhas são os tempos dessas conexões. Foram coletados 50.185 contatos entre 6.642 acompanhantes e 10.106 clientes. Descartando os 1.000 primeiros dias disponíveis na base de dados originais correspondentes a um período transitório com poucos usuários e encontros escassos, restaram 12.157 acompanhantes e clientes com 34.060 conexões.

Na Tabela 1 tem-se um resumo de todas as redes de contato apresentadas nesta seção, bem como suas respectivas quantidades de vértices, arestas e resolução original da rede.

A rede do Museu apresenta a menor quantidade de vértices e arestas, a rede da escola apresenta o maior número de arestas, e a rede sexual o maior número de vértices.

Tabela 1 – Tabela com as redes de contatos, com suas respectivas quantidades de vértices, arestas e resolução original da rede.

Rede	Vértices	Arestas	Resolução Original
Museu	72	6980	20 s
Conferência	113	20818	20 s
Hospital	75	32424	20 s
Escola	180	45047	20 s
Sexual	12157	34060	1 dia

4.2 Experimentos e Avaliação dos Resultados

Os experimentos são apresentados e divididos, inicialmente, em uma visão geral das redes utilizadas, contendo, em seguida, alguns casos particulares com análise visual e crítica detalhada dos padrões identificados, seguido de um comparativo com os trabalhos apresentados atualmente na área.

4.2.1 Visão geral

Nesta seção são mostrados alguns experimentos analisando uma visão geral dos dados, aplicados as redes descritas nas seção 4.1.

De acordo com a Figura 28, que mostra o *layout* temporal da rede de Conferência utilizando a ordenação Vizinhos Recorrentes, é possível notar de imediato padrões que correspondem à atividade dos participantes durante o dia e durante a noite, sendo que o dia consiste em um período de atividade e a noite a quase total ausência dela. Pode-se perceber que esses dados estão sendo analisados em um total de 3 dias (contendo 2 noites). Nota-se, também, dois tipos de concentrações dos dados, uma menor seguida de uma mais densa ao longo dos 3 dias (conforme indicado na figura), sendo que a concentração mais densa de atividade em um contexto de conferência pode representar um intervalo entre palestras, já a concentração menor pode ser um período de perguntas durante as palestras.

Na Figura 29 é possível observar a visualização dos dados da Escola. A Figura 29(a) representa os dados originais e a Figura 29(b) representa os dados com a indicação de dias da semana e ciclos de dia e noite, para auxiliar na análise dos mesmos. Interessante destacar, neste exemplo, que fica evidente o espaço vazio representando o fim de semana.

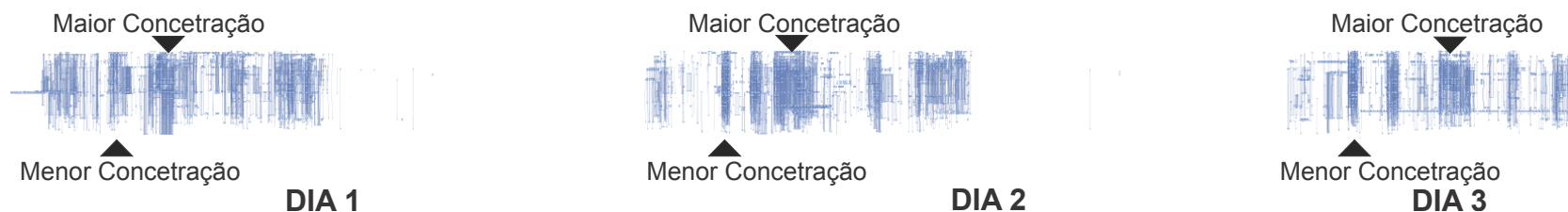


Figura 28 – Visão geral de dados reais de uma Conferência utilizando a ordenação Vizinhos Recorrentes. Percebe-se efeito de dia e noite, juntamente com um total de 3 dias, alternando entre menor e maior concentrações durante o dia.

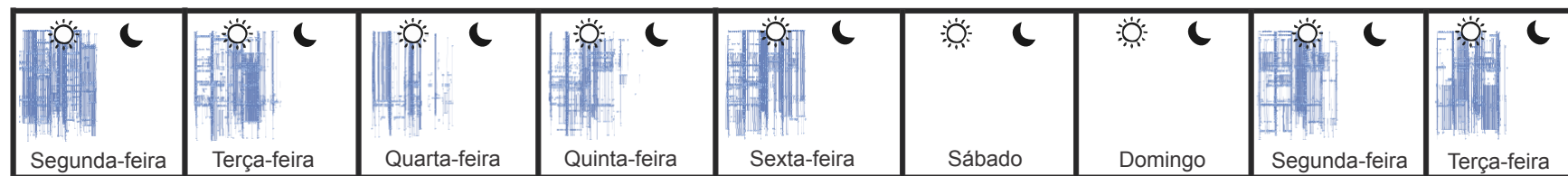


Figura 29 – Visão geral de dados reais de uma Escola utilizando a ordenação Vizinhos Recorrentes. Percebe-se um padrão de dia e noite, com destaque para ausência de atividade no fim de semana.



Figura 30 – Visão geral de dados reais de um Hospital utilizando a ordenação Vizinhos Recorrentes. Percebe-se que o ciclo de dia e noite não é nítido, pelo fato do hospital conter atividade noturna.

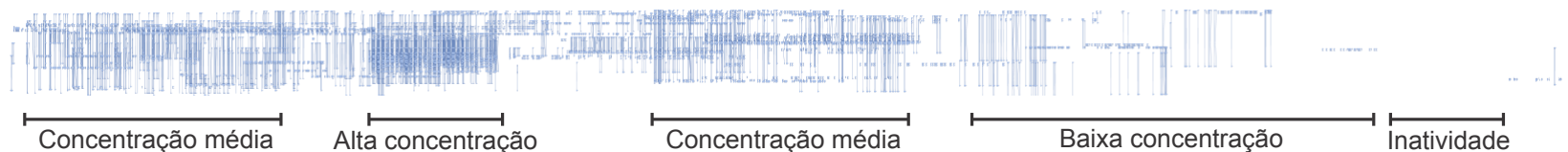


Figura 31 – Visão geral de dados reais de um Museu utilizando a ordenação Vizinhos Recorrentes. Devido aos dados corresponderem a um dia do museu, percebe-se que a atividade aumenta e diminui no decorrer do dia, com um período de atividade baixa na noite.

Na Figura 30 são representados os dados do Hospital. Pode-se observar que, diferentemente dos dois exemplos apresentados anteriormente (Conferência e Escola), neste contexto o ciclo dia e noite não é tão nítido, justamente pelo fato de em um hospital ocorrer atividade noturna; no entanto, fica evidente que essa atividade noturna não tem a mesma frequência da atividade diurna, fato que é visto pela densidade de arestas em um dia. É importante notar alguns padrões nessa Figura 30, como a quase ausência de atividade na segunda noite e um primeiro dia com baixa coleta de dados (indicados na figura).

Na Figura 31 são mostrados os dados do Museu utilizando a ordenação Vizinhos Recorrentes. Observa-se que logo no início da figura já existe uma concentração média de atividades, com destaque para um período de alta concentração logo em seguida (indicados na figura). Esse período de alta concentração pode indicar um horário de chegada de uma excursão no museu, informação que pode ser importante para o realocamento de funcionários para horários de maior atividade, caso esse seja um padrão identificado em vários dias da semana. Visualmente, a Figura 31 representa um dia de atividade no museu, com um período de baixa atividade ao fim do dia, chegando a ter um período de inatividade completa, em que provavelmente todos os visitantes já foram embora do museu.

Na Figura 32 são representados os dados da rede Sexual. No *layout* mostrado na figura ficam nítidas duas coisas: a primeira é que esta é uma rede com crescimento vertical, ou seja, muitas pessoas interagindo em um curto período de tempo, diferentemente de todas as outras imagens anteriores; em uma visão geral percebemos uma uniformidade das arestas, percebendo-se visualmente somente um bloco, fato que não gera ganho de informação.

Para mostrar a diferença entre os tipos de ordenação dos nós, na Figura 33(a) parte da rede do Museu é mostrada na visão temporal, com exibição apenas das arestas e ordenação por nascimento. É possível perceber uma grande concentração de atividade temporal ao longo do tempo, devido ao fato de existirem nós distantes que se conectam muito ao longo do tempo que acabam gerando arestas grandes visualmente. Assim, grandes linhas verticais acabam sendo desenhadas e atrapalham a identificação de padrões. Na Figura 33(b) tem-se a mesma parte dos dados na visão temporal, entretanto a ordenação dos nós está de acordo com a técnica de Vizinhos Recorrentes. Nota-se que na parte da visão mostrada da Figura 33(b) não há densidade no *layout*.

Uma das características da técnica utilizada consiste em centralizar as atividades no meio do *layout* e aproximar os nós que mais se conectam; portanto, é perceptível que no padrão com concentração de atividade naquela faixa temporal, que havia sido identificado na Figura 33(a), não correspondia ao correto, se comparando com a Figura 33(b) – ou seja, a primeira visualização indicava grande concentração de atividades quando na realidade a maioria das atividades estavam centradas em poucos pares de nós. Isso é importante



Figura 32 – Visão geral de dados reais de uma rede Sexual. Um exemplo de uma rede com crescimento vertical.

pois, no contexto em que os dados do Museu foram aplicados, um período de grande concentração de atividade pode determinar um acréscimo na quantidade de funcionários naquele horário, gerando custos extras desnecessários caso a informação esteja incorreta.

Na Figura 34 tem-se o MAT aplicado aos dados da Escola, com os dados filtrados para o primeiro dia de aula coletado. Nota-se que na Figura 34(a) com a resolução original (amostrada de 20 em 20 segundos), é quase imperceptível a variação de cores, devido ao fato do padrão da rede ser de baixas conexões em um mesmo instante de tempo, sabendo-se que o indivíduo que mais se relaciona em um mesmo instante para a resolução original se conecta 5 vezes. Na Figura 34(b) e Figura 34(c) tem-se a mesma rede com uma resolução reduzida em 5 e 10 vezes, respectivamente. Nota-se que, além da rede reduzir o tamanho horizontalmente e agrupar os instantes de tempo, visualmente tem-se um maior destaque aos nós que variam na rede. Portanto, a mudança de resolução leva a questionamentos de eventos, como por exemplo, em 34(b) e 34(c) é possível identificar padrões antes não vistos em 34(a). Portanto, a mudança de resolução permite encontrar

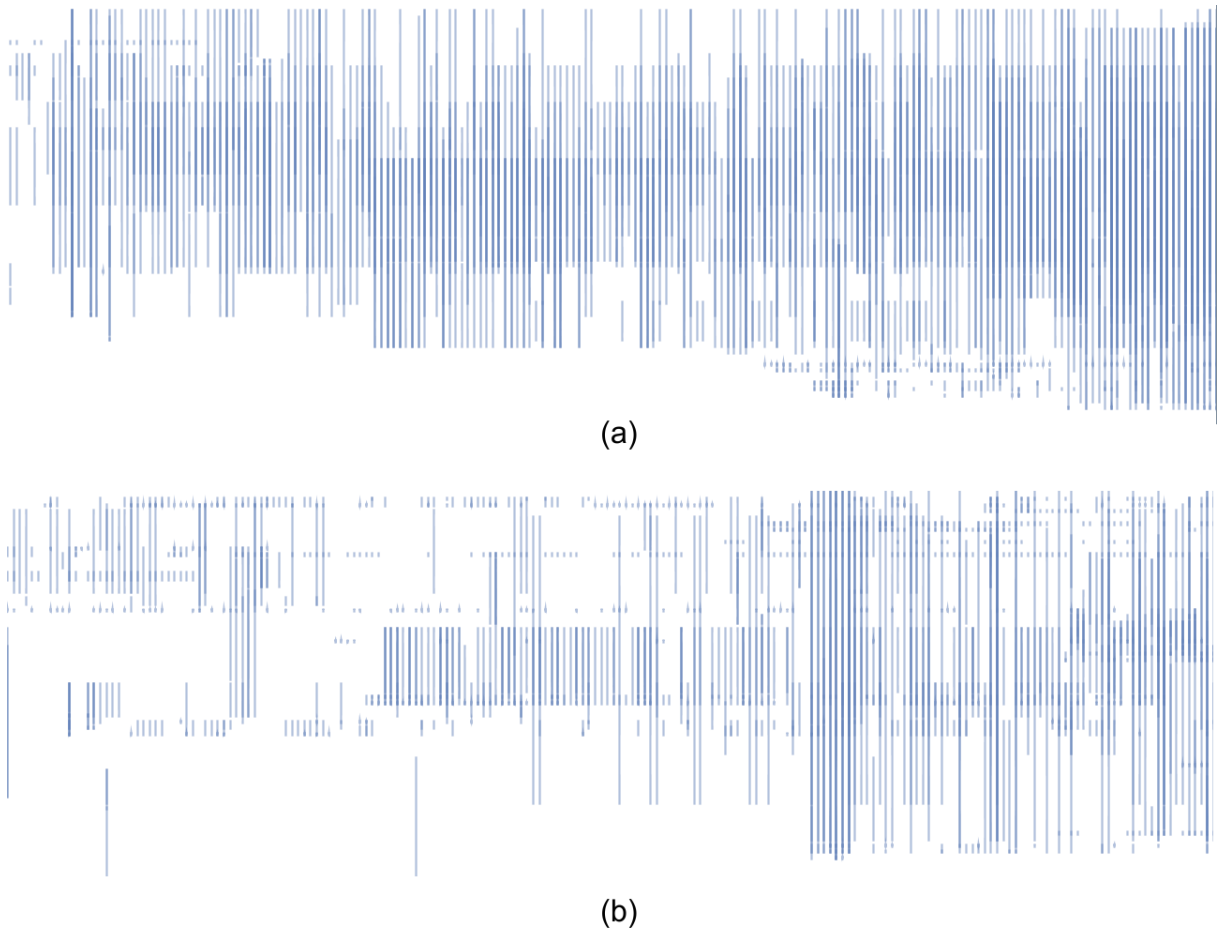


Figura 33 – Visão geral de dados reais de um Museu. Em (a) os dados são demonstrados ordenados pela ordem de nascimento, em (b) os dados são ordenados pela ordem de vizinhos Recorrentes. Importante citar que (a) e (b) representam exatamente a mesma faixa temporal dos dados.

padrões que são identificáveis somente em determinadas escalas.

Percebe-se que com as técnicas de alterar a resolução e com o MAT existem novas possibilidades de suposições, bem como é levado a novos questionamentos dos dados que, sem as técnicas, eram dificultados. Por exemplo, na Figura 34 pode-se supor que a alteração da resolução e o MAT podem auxiliar a identificar agrupamentos na rede, que no exemplo da escola podem ser atividades acadêmicas ou brincadeiras. A característica de alterar a resolução é agrupar os dados, e a característica do MAT com o fundo da menor cor da escala de cores é destacar os vértices que tem uma atividade atípica na rede.

Interessante ressaltar que, ao alterar a resolução, existe um ganho de informação visual no MAT, pois as conexões se agrupam em um mesmo instante de tempo e geram um maior destaque; no entanto, em contrapartida, existe uma perda de informação dos dados, pois não tem-se o real comportamento temporal da rede ao se agrupar algumas conexões.

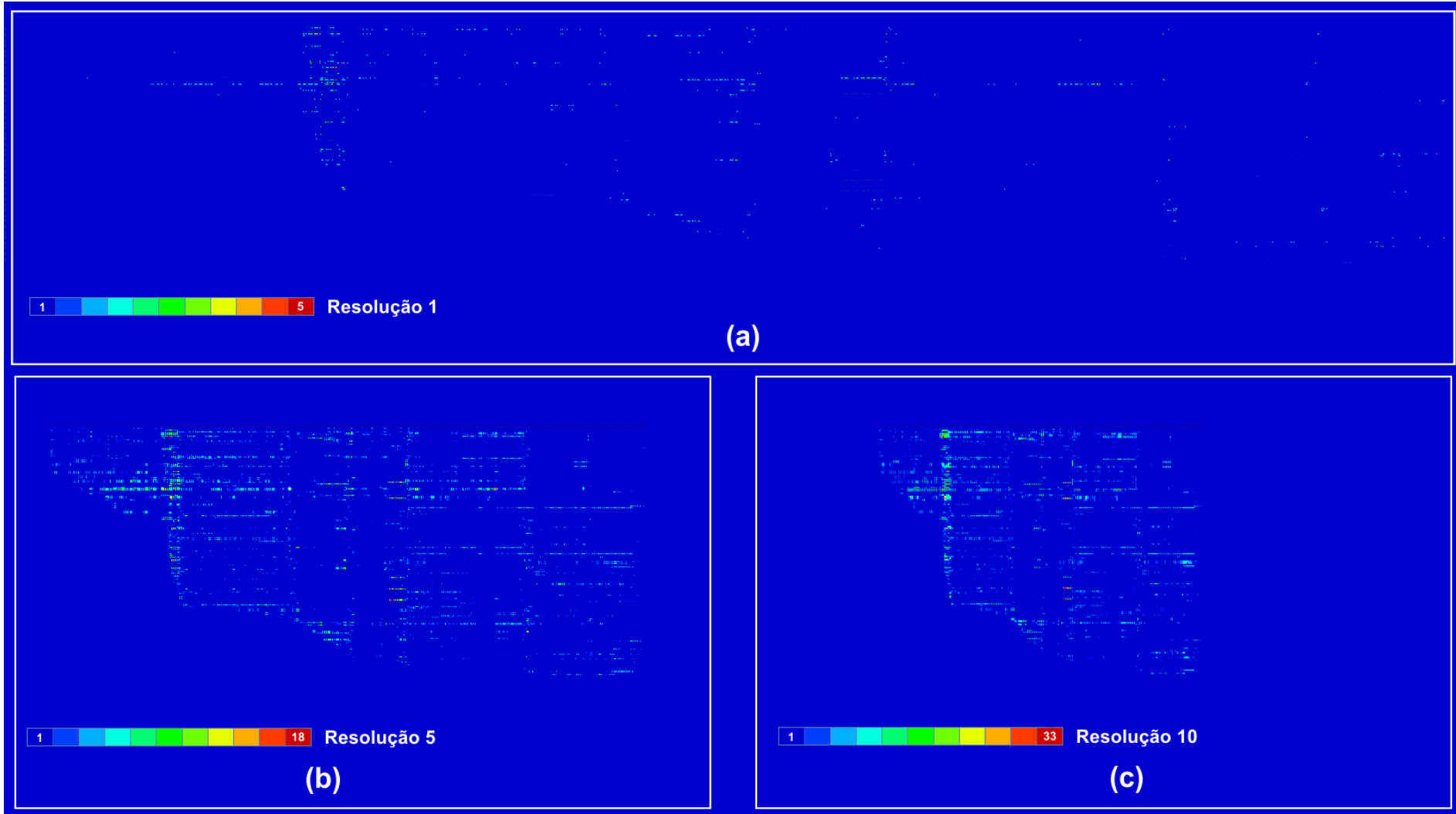


Figura 34 – Visão geral dos dados da Escola (FOURNET; BARRAT, 2014). Dados filtrados para a parte da manhã do primeiro dia da Escola. Em todos os três exemplos tem-se a escala de cores com seus respectivos valores variáveis, bem como o valor da resolução, tem-se também os dados aplicado ao MAT com o fundo com a menor cor da escala de cores. Em (a) a resolução é a original dos dados, com a escala de cores variando de 1 a 5. Em (b) tem-se uma resolução 5 vezes menor que a original, com a escala de cores variando de 1 a 18. Em (c) tem-se uma resolução 10 vezes menor que a original, com a escala de cores variando de 1 a 33.

4.2.2 Visão específica

Nesta seção alguns experimentos são analisados com uma visão específica dos dados, utilizando diversas redes e utilizando as técnicas propostas neste trabalho.

A Figura 35 mostra a visualização MAT aplicada à rede do Hospital, com destaque para o indivíduo 1207. Segundo (VANHEMS et al., 2013), esse indivíduo corresponde a uma enfermeira do hospital. Nota-se que a atividade desse nó é atípica dos outros, tendo grandes picos de atividade em um curto período de tempo, fato notado pela variação de cor dos nós selecionados. É de suma importância observar que, sem os artefatos de seleção com transparência nos demais nós, e o MAT, essa percepção se tornaria mais difícil.

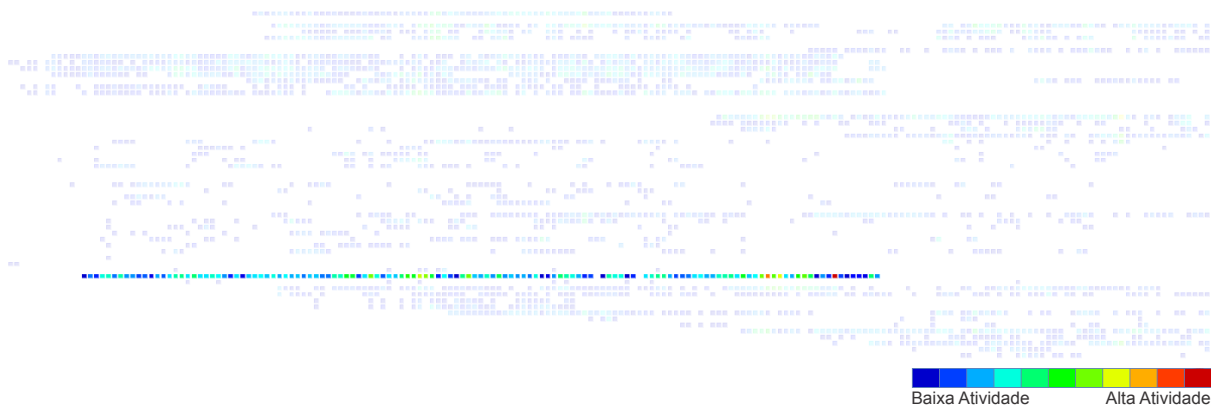


Figura 35 – Visão específica do nó 1207 selecionado na visão temporal do Hospital, demonstrando um padrão de atividades diferente dos demais nós da rede do Hospital. Este nó 1207 corresponde a uma enfermeira (VANHEMS et al., 2013).

Na Figura 36 são demonstradas cinco diferentes visualizações da mesma faixa temporal dos dados da Conferência, correspondendo ao último dia nos últimos instantes de tempo coletados. Na Figura 36(a) é representado a faixa temporal com os dados sem aplicar nenhuma técnica, demonstrando a dificuldade de encontrar padrões ao se analisar um contexto específico. Na Figura 36(b) são utilizadas as cores nos vértices da visão temporal, segundo o a atividade do nó (visto na seção 3.2.2). Nessa visão já são destacados alguns vértices dos demais, representados na cor vermelha. Na Figura 36(c) a faixa temporal em que esses nós que se destacam é selecionada, melhorando a percepção visual e destacando ainda mais quais são os vértices e conexões que acontecem naquele intervalo. Na Figura 36(d) é demonstrado o MAT com o fundo branco correspondente as demais visões. Nota-se que como no MAT são removidas as arestas, todos os nós daquela faixa temporal ganham mais destaque. Na Figura 36(e) é representado o MAT com o fundo da menor cor da escala de cores, destacando os nós que variam dos demais da rede. Das visões representadas, as visões das Figuras 36(c) e 36(e) são as visões que mais destacam os vértices com maior atividade.

Suposições sobre esses vértices destacados podem ser feitas. Sabendo que cada instante de tempo é de 20 em 20 segundos, e sabendo que este tempo representa o último dia nos últimos instantes da Conferência, pode-se supor que os nós que estão tendo muita atividade em um mesmo instante de tempo são pessoas que estão interagindo com muitas pessoas em um curto intervalo de tempo, provavelmente em um ato de despedida dos demais, com um ato de cumprimento apenas, devido ao fato de o nó que mais se conecta em um mesmo instante de tempo conecta 9 vezes em um mesmo instante, e de cada instante ter somente 20 segundos.

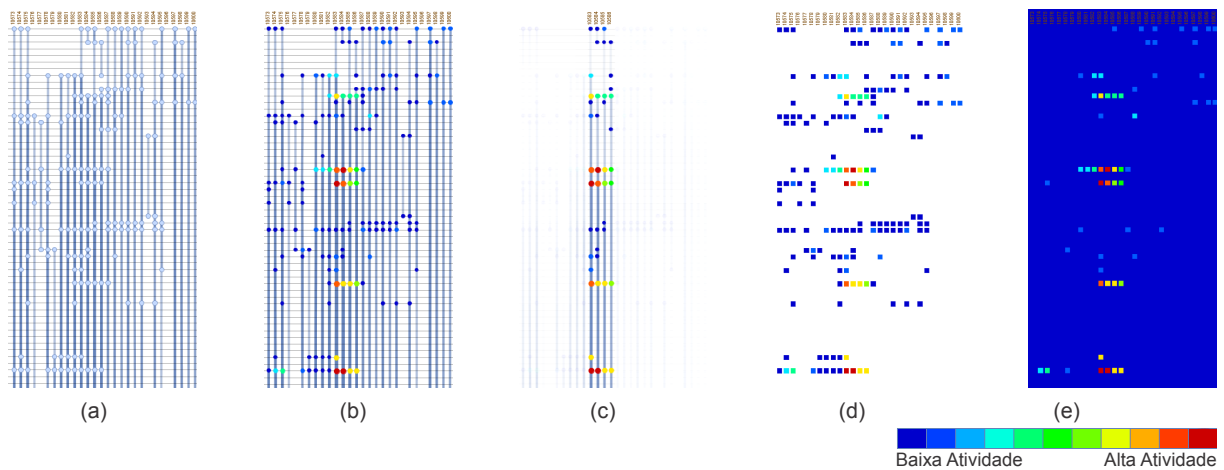


Figura 36 – Visões específicas da visão temporal da Conferência, correspondente ao último dia nos últimos instante de tempo coletados. Em (a) é demonstrado a visão temporal sem aplicar nenhuma técnica. Em (b) é utilizado as cores nos vértices da visão temporal. Em (c) a visão temporal é selecionada uma faixa temporal, para demonstrar a variação de atividade de alguns nós nos fim da conferência. Em (d) é demonstrado o MAT com o fundo branco. Em (e) é demonstrado o MAT com o fundo da menor cor da escala de cor escolhida. Todas as cinco visões representam a mesma faixa temporal.

A coordenação das visões tem grande importância para a análise visual dos dados, como é visto na Figura 37. Neste exemplo, diferentes técnicas foram utilizadas e os dados mostrados são o do Museu. Na visão estrutural da Figura 37(a), os nós estão organizados conforme o algoritmo de força, os vértices e arestas estão com as cores correspondentes ao grau e quantidade de conexões, respectivamente, e as arestas estão com a largura proporcional à quantidade de conexões entre os nós. Na Figura 37(b-c) tem-se a visão temporal correspondente aos nós selecionados da visão estrutural, com o intervalo temporal da Figura 37(b) ocorrendo antes que o intervalo da Figura 37(c). Em ambas as visões temporais estão aplicadas as cores da atividade do nó, bem como a ordem dos nós corresponde a ordenação por grau.

Os nós selecionados são dois dos nós com o maior grau na rede inteira e isso é perceptível na Figura 37(a) pela cor dos nós bem como na Figura 37(b), por eles estarem

localizados na parte inferior do *layout*. Visualmente, tem-se na Figura 37(a) uma percepção de dois nós com alta conectividade também se conectando algumas vezes entre si, mas não é possível perceber, como e quando essas conexões ocorrem. Na Figura 37(b) e Figura 37(c) já consegue-se perceber como são distribuídas essas conexões ao longo do tempo.

Importante verificar que na Figura 37(b), apesar de ambos os nós conectarem diversas vezes naquela faixa temporal com diversos outros nós na rede, eles se conectam entre si somente quatro vezes (representado pelas quatro arestas que estão destacadas entre os dois nós). Em compensação, na Figura 37(c) diversas conexões que esses nós realizam são entre si. Outro fator importante verificado é a necessidade da coordenação de ambas as visões para a análise desses dados, pois se por um lado, na visão temporal, é demonstrado como os dois nós se conectam ao longo do tempo, não tem-se a noção de como esta conexão é relevante, comparada à todas as outras conexões da rede, que pode ser muito bem visualizada pela cor e pela espessura da aresta na visão estrutural.

Outro exemplo de coordenação de visões é demonstrado na Figura 38, que representa os dados da Conferência com a resolução de 5 em 5 minutos. Na Figura 38(a) é representado a visão estrutural com as cores aplicadas aos vértices e arestas, e os vértices que mais se conectam entre si selecionados e deslocados dos demais da rede. Na Figura 38(b) é representado o primeiro dia de atividades da conferência. Nota-se que os dois vértices não se conectaram nenhuma vez entre si e nem com outras pessoas durante o primeiro dia, supondo-se que ambos não estavam presentes no primeiro dia de eventos, ou não se encontraram com ninguém. Na Figura 38(b) é apresentado o segundo dia de conferências, já demonstrando alguma atividade entre esses dois vértices e também com alguma atividade com os outros vértices da rede. Na Figura 38(c) é demonstrado o terceiro dia do evento, com muita atividade entre os dois vértices destacados.

Observa-se a importância da coordenação das visões, pois, somente na visão estrutural é possível identificar de imediato quais são os dois vértices que mais se conectam, identificados pela aresta com maior espessura e com a maior cor da escala selecionada, fato que não era percebido de imediato se fosse visto somente a visão temporal. Em contrapartida, observando somente a visão estrutural não é possível perceber como esses vértices se conectam ao longo do tempo, fato que pode ser identificado na visão temporal, bem como o fato de ambos não estarem ativos no primeiro dia de evento.

É possível supor que estes dois vértices selecionados da Figura 38 tenham se conhecido ou se encontrado apenas no segundo dia, mantendo alguma conexão ao longo do dia. Já no terceiro dia de evento os dois vértices se comunicaram ativamente durante todo o dia, estabelecendo uma relação de proximidade que não era vista no segundo dia, provavelmente tendo acompanhado um ao outro durante todo o terceiro dia de eventos na conferência. Interessante observar que como a resolução apresentada é de 5 em 5 minutos, quase em todos os instantes de tempos do terceiro dia esses dois vértices se conectam,

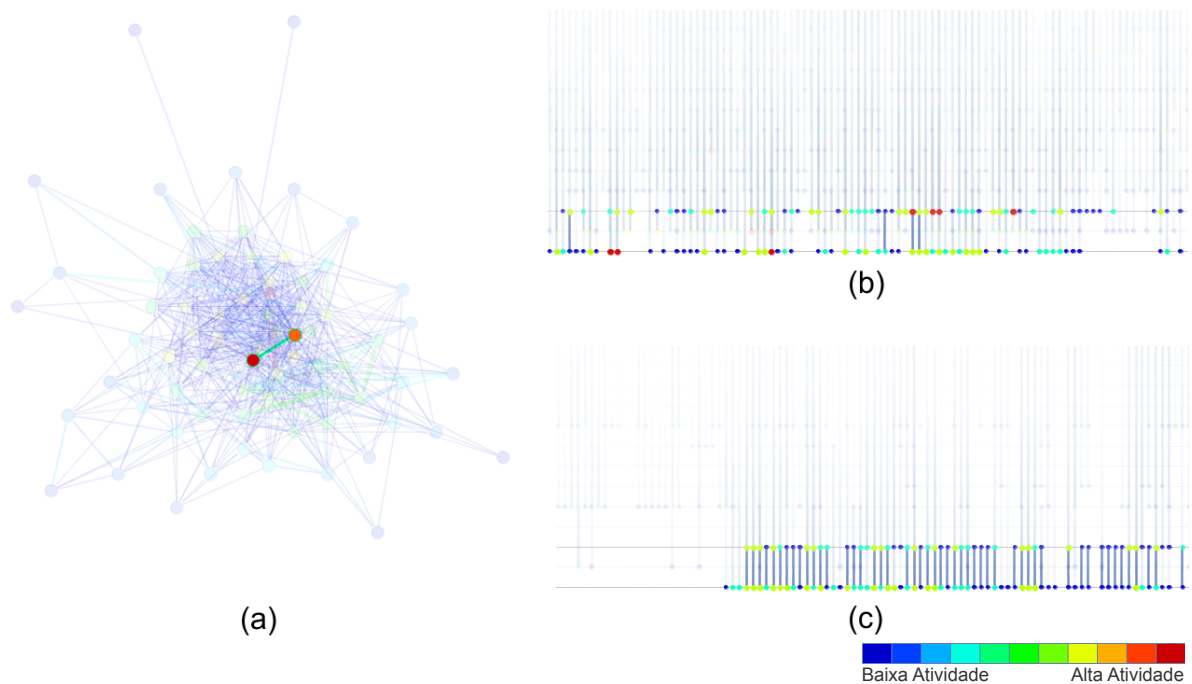


Figura 37 – Visão específica de dois nós na visão estrutural e temporal dos dados do Museu (ISELLA et al., 2011). As visões estão coordenadas, em (a) tem-se a visão estrutural com os dois nós selecionados, em (b) e (c) tem-se a visão temporal desses mesmos dois nós selecionados, ambas as visões temporais estão com a ordenação dos nós pelo grau. A diferença entre (b) e (c) é que a porção do tempo que é mostrada em (b) ocorre antes da porção do tempo de (c). A escala de cores utilizada em todas as visões é a *Pseudo Rainbow Scale*, como é visto no canto inferior direito. Importante ressaltar que, apesar de ser utilizada a mesma escala de cores para todas as medidas (grau do nó, conectividade da aresta, atividade), cada medida varia conforme seus próprios e independentes valores.

ou seja, pelo menos uma vez a cada 5 minutos esses vértices mantêm contato um com o outro.

Diversos outros experimentos foram realizados mas que não trouxeram visualizações significativas ou diferente das que já foram demonstradas. Importante citar que a eficácia das técnicas depende da estrutura, distribuição da atividade da rede, quantidade de vértices, e outros fatores. O objetivo das técnicas é fornecer mais opções para a análise de dados de uma maneira que ainda não havia sido percebida. Cada técnica deve ser escolhida dependendo do contexto que é requisitado para ser visualizado, e algumas podem ser combinadas para uma melhoria do resultado final. Por exemplo, caso o interesse seja visualizar a atividade dos dados sem se preocupar com detalhes, pode-se diminuir a resolução dos dados e aplicar o MAT (visto na Figura 34), porém, caso o interesse seja em verificar a atividade de um nó específico, esse nó deve ser selecionado e destacado, e a resolução pode ser mantida originalmente (visto na Figura 35), podendo-se realizar a

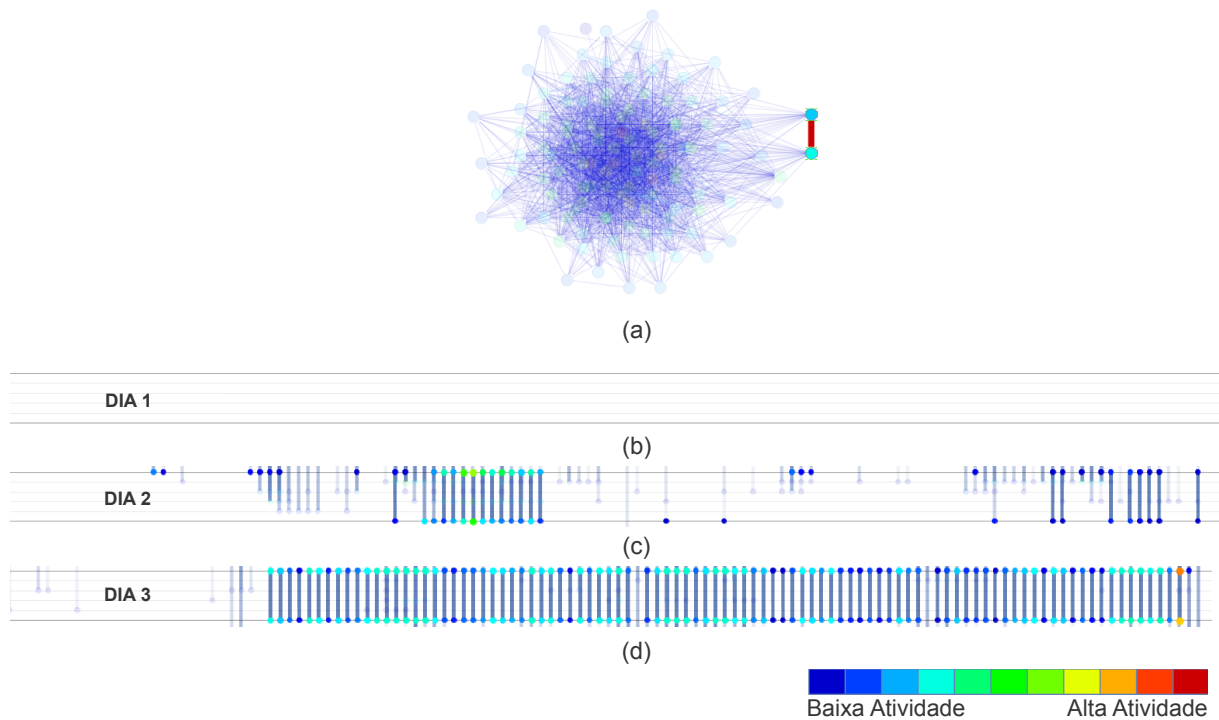


Figura 38 – Visão geral de dados reais da Conferência, com a resolução de 5 em 5 minutos. Em (a) é demonstrada a visão estrutural dos dados da Conferência, com os dois nós que mais se conectam entre si selecionados. Em (b) tem-se a visão temporal específica dos dois nós, no primeiro dia da Conferência, em que não foi possível identificar atividade entre esses dois nós. Em (c) tem-se a visão temporal específica dos dois nós, no segundo dia da Conferência, com alguma atividade entre os dois nós. Em (d) tem-se a visão temporal específica dos dois nós, no segundo dia de Conferência, com bastante atividade entre os dois nós. Todas as visões estão coordenadas e representam os mesmos dados selecionados.

construção da rede estrutural considerando o tempo acumulado para perceber como esse nó irá se comportar no decorrer do tempo.

Portanto, caso o interesse, na análise dos dados, seja identificar vértices com maior atividade, recomenda-se que seja utilizada a visualização com o fundo da menor cor da escala; entretanto, caso o interesse da análise consista em identificar todos os vértices da rede, incluindo os vértices de menor atividade, é viável que seja utilizado um fundo branco, ou um fundo de alguma cor mais clara que a cor associada ao menor valor da medida.

4.2.3 Desempenho do sistema

Alguns testes de desempenho foram executados em um computador Intel(R) Core(TM) i5-3330 CPU @ 3.00GHz e 8 GB de memória RAM. Na Tabela 2 tem-se o desempenho do

sistema para algumas redes. A tabela relaciona a quantidade de vértices, arestas, quais os tempos gastos para construir o *layout*, renderizar os vértices e as arestas e para carregar o arquivo, sendo que estes tempos computados são os tempos médios observados em cinco testes.

O tempo de construir o *layout* consiste em ler o arquivo de entrada, realizar o pré-processamento nos dados (caso seja necessário), bem como aplicá-los à estrutura de dados da biblioteca utilizada. O tempo para renderizar consiste no tempo gasto para gerar a visualização, utilizando os componentes gráficos para desenhar os vértices e arestas da biblioteca utilizada. O tempo total é a soma do tempo de construir o *layout* e o tempo de renderizar, resultando no tempo total para se executar um arquivo pela primeira vez. Importante observar que esse tempo de construção do *layout* gasto só é realizado na primeira vez em que cada arquivo é carregado.

Nota-se que, conforme cresce a quantidade de vértices e arestas, o tempo de execução também cresce. A rede Enron (vista na seção 2.4) utilizada é somente uma parte inicial da rede, filtrada para as cinquenta mil conexões iniciais da rede original (esse número foi escolhido para que a rede ficasse em tamanho semelhante ao das outras redes analisadas); esta rede é incluída nos testes somente para efeitos de verificação do desempenho do sistema.

Importante observar que o crescimento das arestas gera mais impacto no tempo de construção do *layout* do que o crescimento dos vértices, como pode ser observado na Tabela 2. Para os valores mostrados para as redes da Escola e Sexual, apesar de ter aumentado mais de 67 vezes o número de vértices, o tempo é o mesmo; em contrapartida, comparando os valores mostrados para as redes Sexual e Enron, houve um decréscimo do número de vértices e um acréscimo de aproximadamente 46% da quantidade de arestas, resultando em um aumento no tempo de construção do *layout*.

A rede Enron da Tabela 2 apresenta um tempo total de 65 segundos para carregar o arquivo, mas observa-se que no tempo gasto para construir e renderizar, tem-se 59 segundos gastos para construir o sistema e 6 segundos gastos para renderizar, o que é importante tendo em vista que esses 59 segundos só são computados na primeira vez em que o arquivo é aberto.

Algumas outras funcionalidades do sistema também apresentam um elevado tempo computacional para serem executadas, como a ordenação de arestas por Vizinhos Recorrentes, e o algoritmo de força para espalhar os nós na visão estrutural. Ambos os processos podem demorar alguns minutos para execução, dependendo do arquivo que está sendo analisado. Com o crescimento da quantidade de vértices e arestas, até funcionalidades mais básicas, como a seleção dos nós, podem ter um acréscimo em seu tempo de execução.

Portanto, uma das limitações do sistema é o fato da escalabilidade de vértices e arestas prejudicarem o desempenho de algumas funcionalidades citadas no sistema. O maior

Tabela 2 – Tabela com o desempenho relacionando os vértices e arestas com o tempo médio gasto para iniciar o sistema.

Rede	Vértices	Arestas	Tempo (s) construir <i>layout</i>	Tempo (s) renderizar	Tempo (s) total
Museu	72	6980	3	1	4
Conferência	113	20818	13	2	15
Hospital	75	32424	25	2	27
Escola	180	45047	51	4	55
Sexual	12157	34060	46	9	55
Enron (parcial)	7760	49980	59	6	65

arquivo testado contém aproximadamente cinquenta mil arestas, mas para dados com um milhão de arestas o sistema pode demorar um tempo considerável para executar determinados processos. Os filtros anteriores ao carregamento do arquivo, como o de limitar o tempo inicial e final, e o de diminuir a resolução auxiliam para amenizar, em parte, este problema.

Conclusão

Considerando o baixo número de técnicas para auxiliar a identificação de padrões na Visualização da Informação de redes temporais, este trabalho apresenta e discute as técnicas mais importantes consolidadas na literatura, bem como desenvolve, a partir das limitações dessas técnicas, novas abordagens para redes temporais. Conceitos de visualização para a visão estrutural já são amparados por várias propostas, tendo em vista que técnicas para melhorar essa visão já são desenvolvidas a algum tempo; entretanto, devido a recente aparição de redes temporais, poucas técnicas visuais foram desenvolvidas e demonstradas, assim como existem poucas aplicações de dados reais que utilizam essas técnicas.

Este trabalho aplica técnicas de visualização de redes temporais que foram desenvolvidas a diferentes cenários, como escola, museu, conferências e hospitais, com o intuito de mostrar a eficácia das técnicas e explorar cenários que outros trabalhos da área não identificaram. Conforme foi explorado nos experimentos, essas técnicas puderam ser aplicadas em diversos cenários diferentes, possibilitando identificar padrões visuais que sem as técnicas seriam difíceis de identificar. Foi desenvolvido um sistema, chamado DyNetVis, que serviu de base para a implantação das técnicas e para a realização de estudos de casos.

De acordo com o que foi apresentado nos experimentos, com a utilização da técnica do MAT a visão temporal perde informações das conexões dos vértices ao longo do tempo, mas tem-se um ganho visual na identificação dos vértices que são mais ou menos ativos na rede, bem como com a ordenação dos vértices pela técnica de Vizinhos Recorrentes, que centraliza no *layout* os nós mais ativos na rede e minimiza o tamanho das arestas, obtendo um *layout* mais limpo e possibilitando a identificação de padrões na rede.

Uma importância do trabalho é a forma de interagir com os dados e coordenar as visões apresentadas, de maneira dinâmica e que possa revelar e evidenciar padrões e tendências que não poderiam ser visualizados sem interação, coordenação e técnicas. Importante citar que a eficácia das técnicas depende da estrutura, distribuição da atividade da rede, quantidade de vértices, e outros fatores. O objetivo das técnicas é fornecer mais opções para a análise de dados de uma maneira que ainda não havia sido percebida. Cada

técnica deve ser escolhida dependendo do contexto que é requisitado para ser visualizado, e algumas podem ser combinadas para uma melhoria do resultado final.

5.1 Contribuições

Uma das contribuições deste trabalho é reunir o conhecimento disponível sobre redes dinâmicas temporais, bem como aplicar os conhecimentos da Visualização da Informação para auxiliar a análise de dados. As técnicas disponíveis na literatura são aplicadas a dados reais, juntamente com melhorias, de forma a potencializar a análise e a percepção dos fenômenos envolvidos. O trabalho também contribui com o desenvolvimento de um sistema computacional para validar todas as técnicas propostas.

5.2 Trabalhos Futuros

Diversas melhorias podem ser aplicadas no trabalho proposto, e novas estratégias podem ser propostas, permitindo a identificação de mais padrões nos dados, bem como a melhor compreensão dos padrões já identificados. A seguir são apresentadas algumas propostas.

5.2.1 Análise e exploração de características associadas aos indivíduos

Uma das formas de incrementar a metodologia proposta consiste em trabalhar com outras informações associadas aos nós. Por exemplo, nos dados analisados da escola (FOUR-NET; BARRAT, 2014), trabalhar com a informação de qual turma o aluno pertencia, bem como se o indivíduo era um aluno ou professor, dentre outras informações. Com essa informação, é possível que a cor ou forma possa diferenciar indivíduos de turmas diferentes, bem como alterar a ordem das linhas temporais de acordo com as turmas pré-definidas, colocando-se, sequencialmente, indivíduos de uma mesma turma. É importante explorar outras possibilidades de organização dos nós na tela, por exemplo por similaridade, aumentando o potencial de exploração da visão estrutural, por exemplo.

Outra forma de organizar os nós é adicionar propriedades extras dos nós, as quais podem variar de características geométricas de um nó (como altura, largura, comprimento) mapeando características de dados reais (como idade, sexo, nacionalidade). Estas características são importantes para que o agrupamento dos dados, em ambas as visões, seja baseado em informações do nó (meta-dados) e não somente em métricas, podendo gerar uma nova percepção e ajudando na identificação de padrões.

5.2.2 Organização dos vértices

Uma maneira de melhorar a percepção de padrões nas redes pode ser incorporar informações extras sobre essas redes como um todo, tais como medidas de comunidade, formando grupos que podem ser separados tanto na visão estrutural (de forma a espalhar os nós no plano conforme sua comunidade), conforme é descrito no trabalho (TRIER, 2008), quanto na visão temporal, na qual pode ser alterado a ordem das linhas, colocando sequencialmente nós que pertencem a uma mesma comunidade. Com novos atributos na rede, eles podem ser agregados para a identificação de comunidades, como por exemplo, o parâmetro de similaridade para detecção ser escolhido de acordo com as características do nó.

5.2.3 Aplicações em Epidemiologia

Uma das possíveis áreas de aplicações deste trabalho é na área de epidemiologia. Uma das formas de auxiliar a análise de dados de epidemias é acrescentar funcionalidades de interação para nós infectados e imunes, bem como um ponto de partida da infecção e, a partir de um determinado ponto, demonstrar como aquela infecção se propaga ao longo do tempo, na visão temporal. Em continuidade do trabalho de (LEE et al., 2012), espera-se que, com essas interações, seja possível visualizar o comportamento da propagação de uma infecção ou doença em uma rede, e também simular essa propagação, com o intuito de guiar a adoção de políticas de prevenção de epidemias. A Figura 39 demonstra a propagação da epidemia em uma rede temporal. Na Figura 39(a) nota-se que o nó 1 desta rede não é infectado ao longo do tempo, pois o nó 3 está imune e não transmite a doença. Já na Figura 39(b) todos os nós da rede que não são imunes são infectados ao longo do tempo.

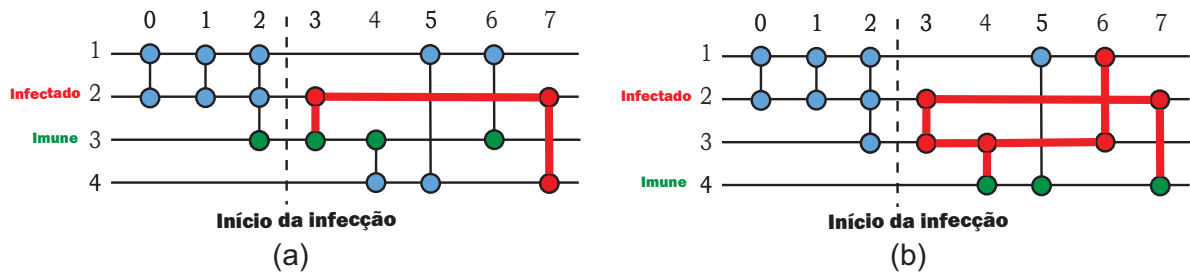


Figura 39 – Representação da propagação da epidemia em uma rede temporal. Em (a) o nó 2 está infectado e o nó 3 é imune à infecção. Em (b) o nó 2 está infectado e o nó 4 é imune à infecção. O caminho delimitado em vermelho representa como a doença se propaga na rede.

5.2.4 Melhoria da escalabilidade do sistema

Para algumas redes maiores, em termos de vértices e nós, o sistema DyNetVis demora mais de um minuto para criar a estrutura e construir o *layout*. Além das estratégias apresentadas neste trabalho, como filtros de tempo e alterar a resolução, outras propostas de melhoria para o desempenho do processo são necessárias, como a opção de carregar somente os nós mais importantes na rede, de forma automática, utilizando nesse processo métricas para identificar os nós importantes, ou de forma manual, dando ao usuário a opção de escolher quais nós não são relevantes para a rede.

Referências

- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, p. 47–97, jan. 2002.
- ALDER, G. **JGraphXn**. 2010. Disponível em: <<https://github.com/jgraph/jgraphx>>.
- BARRAT, A. et al. The architecture of complex weighted networks. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 11, p. 3747–3752, 2004. Disponível em: <<http://www.pnas.org/content/101/11/3747.abstract>>.
- BASTERT, O.; MATUSZEWSKI, C. Layered drawings of digraphs. In: KAUFMANN, M.; WAGNER, D. (Ed.). **Drawing Graphs**. Springer Berlin Heidelberg, 2001, (Lecture Notes in Computer Science, v. 2025). p. 87–120. ISBN 978-3-540-42062-0. Disponível em: <http://dx.doi.org/10.1007/3-540-44969-8_5>.
- BASTIAN, M.; HEYMANN, S.; JACOMY, M. **Gephi: An Open Source Software for Exploring and Manipulating Networks**. 2009.
- BATAGELJ, V.; MRVAR, A. Pajek — analysis and visualization of large networks. In: JÜNGER, M.; MUTZEL, P. (Ed.). **Graph Drawing Software**. Springer Berlin Heidelberg, 2004, (Mathematics and Visualization). p. 77–103. ISBN 978-3-642-62214-4. Disponível em: <http://dx.doi.org/10.1007/978-3-642-18638-7_4>.
- BATTISTA, G. D. et al. Algorithms for drawing graphs: an annotated bibliography. **Computational Geometry**, v. 4, n. 5, p. 235 – 282, 1994. ISSN 0925-7721. Disponível em: <<http://www.sciencedirect.com/science/article/pii/092577219400014X>>.
- BAUR, M.; BRANDES, U. Crossing reduction in circular layouts. In: HROMKOVIČ, J.; NAGL, M.; WESTFECHTEL, B. (Ed.). **Graph-Theoretic Concepts in Computer Science**. Springer Berlin Heidelberg, 2005, (Lecture Notes in Computer Science, v. 3353). p. 332–343. ISBN 978-3-540-24132-4. Disponível em: <http://dx.doi.org/10.1007/978-3-540-30559-0_28>.
- BENDER-DEMOLL, S.; MCFARLAND, D. A. The art and science of dynamic network visualization. **Journal of Social Structure**, 2006.
- BORGATTI, S.; EVERETT, M. G.; FREEMAN, L. C. **UCINET 6 for Windows: Software for Social Network Analysis**. Harvard, MA, Analytic Technologies. 2002.

- BRANDES, U.; WAGNER, D. Visone – analysis and visualization of social networks. In: **GRAPH DRAWING SOFTWARE**. [S.l.]: Springer-Verlag, 2003. p. 321–340.
- CARLEY, K. M. et al. Ora user’s guide 2012. **Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report**, Carnegie Mellon University, n. CMU-ISR-12-105, 2012.
- CATTUTO, C. et al. Dynamics of person-to-person interactions from distributed rfid sensor networks. **PLoS ONE**, Public Library of Science, v. 5, n. 7, p. e11596, 07 2010. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0011596>>.
- COSTA, L. d. F. et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329–412, 2011. Disponível em: <<http://dx.doi.org/10.1080/00018732.2011.572452>>.
- CROUSER, R.; FREEMAN, J.; CHANG, R. Exploring agent-based simulations using temporal graphs. In: **Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on**. [S.l.: s.n.], 2011. p. 271–272.
- CSARDI, G.; NEPUSZ, T. The igraph software package for complex network research. **InterJournal**, Complex Systems, p. 1695, 2006.
- ELZEN, S. van den et al. Reordering massive sequence views: Enabling temporal and structural analysis of dynamic networks. In: IEEE. **Visualization Symposium (PacificVis), 2013 IEEE Pacific**. [S.l.], 2013. p. 33–40.
- _____. Dynamic network visualization with extended massive sequence views. **Visualization and Computer Graphics, IEEE Transactions on**, v. 20, n. 8, p. 1087–1099, Aug 2014. ISSN 1077-2626.
- FOURNET, J.; BARRAT, A. Contact patterns among high school students. **PLoS ONE**, v. 9, p. e107878, 2014.
- FREEMAN, L. C. **The Development of Social Network Analysis: A Study in the Sociology of Science**. [S.l.]: BookSurge Publishing, 2004.
- HAGBERG, A. A.; SCHULT, D. A.; SWART, P. J. Exploring network structure, dynamics, and function using networkx. In: VAROQUAUX, G.; VAUGHT, T.; MILLMAN, J. (Ed.). **Proceedings of the 7th Python in Science Conference**. Pasadena, CA USA: [s.n.], 2008. p. 11 – 15.
- HEER, J.; BOYD, D. Vizster: visualizing online social networks. In: **Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on**. [S.l.: s.n.], 2005. p. 32–39.
- HEER, J.; PERER, A. Orion: A system for modeling, transformation and visualization of multidimensional heterogeneous networks. **Information Visualization**, SAGE Publications, v. 13, n. 2, p. 111–133, 2014.
- HERMAN, I.; MELANCON, G.; MARSHALL, M. Graph visualization and navigation in information visualization: A survey. **Visualization and Computer Graphics, IEEE Transactions on**, v. 6, n. 1, p. 24–43, Jan 2000. ISSN 1077-2626.

- _____. Graph visualization and navigation in information visualization: A survey. **Visualization and Computer Graphics, IEEE Transactions on**, v. 6, n. 1, p. 24–43, Jan 2000. ISSN 1077-2626.
- HOLME, P.; SARAMÄKI, J. Temporal networks. **Physics Reports**, Elsevier, v. 519, n. 3, p. 97–125, 2011.
- ISELLA, L. et al. What's in a crowd? analysis of face-to-face behavioral networks. **Journal of theoretical biology**, Elsevier, v. 271, n. 1, p. 166–180, 2011.
- KEILA, P. S.; SKILLICORN, D. B. Structure in the enron email dataset. **Comput. Math. Organ. Theory**, Kluwer Academic Publishers, Hingham, MA, USA, v. 11, n. 3, p. 183–199, out. 2005. ISSN 1381-298X. Disponível em: <<http://dx.doi.org/10.1007/s10588-005-5379-y>>.
- LEE, S. et al. Exploiting temporal network structures of human interaction to effectively immunize populations. **PLoS ONE**, Public Library of Science, v. 7, n. 5, p. e36439, 05 2012. Disponível em: <<http://dx.doi.org/10.1371/journal.pone.0036439>>.
- LESKOVEC, J.; SOSI, R. Snap.py: Snap for python, a general purpose network analysis and graph mining tool in python. In: . [S.l.: s.n.], 2014.
- MCCORMICK, B. H. Visualization in scientific computing. **SIGBIO Newsl.**, ACM, New York, NY, USA, v. 10, n. 1, p. 15–21, mar. 1988. ISSN 0163-5697. Disponível em: <<http://doi.acm.org/10.1145/43965.43966>>.
- MOODY, J.; MCFARL, D.; BENDER-DEMOLL, S. Dynamic network visualization. **American Journal of Sociology**, v. 110, p. 1206–1241, 2005.
- NEWMAN, M. E. J. The Structure and Function of Complex Networks. **SIAM Review**, v. 45, p. 167–256, jan. 2003.
- O'MADADHAIN, J. et al. **Analysis and Visualization of Network Data using JUNG**.
- PEIXOTO, T. P. The graph-tool python library. **figshare**, 2014.
- PIGNÉ, Y. et al. GraphStream: A Tool for bridging the gap between Complex Systems and Dynamic Graphs. **ArXiv e-prints**, mar. 2008.
- ROCHA, L. E. C.; LILJEROS, F.; HOLME, P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. **PLoS Comput Biol**, Public Library of Science, v. 7, n. 3, p. e1001109, 03 2011. Disponível em: <<http://dx.doi.org/10.1371/journal.pcbi.1001109>>.
- ROCHA, L. E. C. da. Structural evolution of the brazilian airport network. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2009, n. 04, p. P04020, 2009. Disponível em: <<http://stacks.iop.org/1742-5468/2009/i=04/a=P04020>>.
- SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research**, v. 13, n. 11, p. 2498–2504, nov. 2003.

SILVA, F. et al. Investigating relationships within and between category networks in wikipedia. **Journal of Informetrics**, v. 5, n. 3, p. 431 – 438, 2011. ISSN 1751-1577. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1751157711000344>>.

SPENCE, R. **Information Visualization: Design for Interaction (2nd Edition)**. [S.l.]: Prentice Hall, 2007.

TRIER, M. Research note—towards dynamic visualization for understanding evolution of digital communication networks. **Information Systems Research**, v. 19, n. 3, p. 335–350, 2008. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/isre.1080.0191>>.

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, 09 2013.

WARE, C. **Information Visualization: Perception for Design**. [S.l.]: Morgan Kaufmann Publishers Inc., 2004. 588 p. ISBN 1558608192.

WILLIAMS, J. G.; SOCHATS, K. M.; MORSE, E. Visualization. **Annual Review of Information Science and Technology**, p. 161–207, 1995.

WILSON, R. J. **Introduction to Graph Theory**. New York, NY, USA: John Wiley & Sons, Inc., 1986. ISBN 0-470-20616-0.

_____. **Introduction to Graph Theory (5th Edition)**. [S.l.]: Pearson, 2012.

WONGPIROMSARN, T. et al. Tulip: A software toolbox for receding horizon temporal logic planning. In: **Proceedings of the 14th International Conference on Hybrid Systems: Computation and Control**. New York, NY, USA: ACM, 2011. (HSCC '11), p. 313–314. ISBN 978-1-4503-0629-4. Disponível em: <<http://doi.acm.org/10.1145/1967701.1967747>>.

ZAIDI, F.; MUELDER, C.; SALLABERRY, A. Analysis and visualization of dynamic networks. **ArXiv**, abs/1409.5034, 2014. Disponível em: <<http://arxiv.org/abs/1409.5034>>.

ZHANG, Q.-G. et al. Drawing undirected graphs with genetic algorithms. In: WANG, L.; CHEN, K.; ONG, Y. (Ed.). **Advances in Natural Computation**. Springer Berlin Heidelberg, 2005, (Lecture Notes in Computer Science, v. 3612). p. 28–36. ISBN 978-3-540-28320-1. Disponível em: <http://dx.doi.org/10.1007/11539902_4>.