

**CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA
FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA”
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**Implementação de Técnicas de Visualização Interativa para promover a
melhoria da qualidade da informação**

RODOLFO ROZA MIRANDA

Marília, 2013

**CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA
FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA”
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**Implementação de Técnicas de Visualização Interativa para promover a
melhoria da qualidade da informação**

Monografia apresentada ao
Centro Universitário Eurípides de
Marília como parte dos requisitos
necessários para a obtenção do
grau de Bacharel em Ciência da
Computação
Orientador: Prof. Leonardo
Castro Botega

Marília, 2013



CENTRO UNIVERSITÁRIO EURÍPIDES DE MARÍLIA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

TRABALHO DE CONCLUSÃO DE CURSO – AVALIAÇÃO FINAL

Rodolfo Roza Miranda

Desenvolvimento de Técnicas de Visualização Interativa para promover a melhoria da qualidade do conteúdo da informação.

Banca examinadora da monografia apresentada ao Curso de Bacharelado em Ciência da Computação do UNIVEM/F.E.E.S.R., para obtenção do Título de Bacharel em Ciência da Computação.

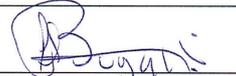
Nota: 8.0 (oito)

Orientador: Leonardo Castro Botega

1º. Examinador: Ildeberto de Gênova Bugatti

2º. Examinador: Adriano Bezerra







Marília, 02 de dezembro de 2013.

Sumário

Tabela de Siglas.....	7
Resumo.....	9
Abstract.....	10
Introdução.....	11
Motivação.....	12
Objetivos.....	13
Capítulo 1 - Sistemas de Fusão de Dados.....	14
1.1 O modelo de Fusão JDL.....	14
1.2 O papel do humano em sistemas de fusão.....	17
Capítulo 2 - Sistemas de Fusão de Dados e a Qualidade dos Dados e Informações.....	19
2.1 Dimensões e Taxonomias de Qualidade de Dados.....	19
2.2 Qualidade de Dados em Sistemas de Fusão.....	24
2.3 Trabalhos relacionados.....	25
2.4 Visualização Interativa e Transformações de Dados e Representações.....	26
Capítulo 3 - Metodologia.....	33
1. Levantamento Bibliográfico:.....	33
2. Prototipação.....	34
3. Análise de Resultados.....	40
Capítulo 4 - Resultados Obtidos.....	41
1. Componentes de Hardware.....	41
2. Componentes de Software.....	41
3. Estudo de caso em C2.....	41
4. Resultados das Implementações.....	42
Conclusões.....	52
Referências Bibliográficas.....	53

Lista de Figuras

Figura 1: Modelo de Fusão JDL.....	15
Figura 2: Taxonomia de ignorância de Smithson.....	21
Figura 3: Tipos de incerteza segundo Bouchon-Meunier e Nguyen.....	22
Figura 4: Taxonomia de incerteza proposta por Krause e Clark.....	22
Figura 5: Classificação da imperfeição segundo Smets.....	23
Figura 6: Ilustração do pipeline de visualização.....	28
Figura 7: Fluxo de atividades realizadas no trabalho.....	33
Figura 8: Janela de informações exibida pela interação de toque.....	36
Figura 9: Seleção de área de interesse através de toque.....	37
Figura 10: Filtro através da interação de drag no canto direito da tela.....	38
Figura 11: Realização de agregação entre dois objetos.....	39
Figura 12: Desagregação de um agrupamento.....	40
Figura 13: Interação de seleção de representações de posts.....	44
Figura 14: Interação de filtragem, valor 0,1.....	45
Figura 15: Interação de filtragem com valor de completude maior que 0,25.....	46
Figura 16: Passos realizados na interação de filtragem.....	47
Figura 17: Informações sobre o primeiro post.....	49
Figura 18: Informações sobre o segundo post.....	49
Figura 19: Nova representação gerada pela agregação dos posts.....	50

Lista de Tabelas

Tabela 1: Tabela de Siglas.....	7
Tabela 2: Transformação de dados por meio de interações.....	29
Tabela 3: Técnicas de Visualização Interativa implementadas.....	36

Tabela de Siglas

Tabela 1: Tabela de Siglas

VI	Visualização Interativa
C2	Comando e Controle
HCI (IHC)	Interação Humano-Computador
JDL	Joint Director of Laboratories
DoD	Department of Defense

Resumo

O processo de visualização de informações constitui-se em codificar dados em representações gráficas com o objetivo de ampliar a cognição humana em seu entendimento. A Visualização Interativa (VI) é uma técnica com origens na interação humano-computador (IHC), que permite que o usuário transforme diretamente os dados em busca de novas inferências sobre os mesmos, conduzindo conseqüentemente à mudanças na própria representação gráfica. O presente trabalho sustenta a hipótese de que a VI pode contribuir para a melhoria de qualidade dos dados, e conseqüentemente da informação percebida e compreendida pelo usuário, uma vez que o mesmo pode contribuir diretamente para a construção de seu conhecimento sobre uma situação. O objetivo do trabalho é provar tal hipótese através da implementação de técnicas de VI sobre dados de um sistema de comando e controle militar, onde a compreensão de um cenário é extremamente crítica.

Palavras-chave: Sistemas de Fusão de Dados, Comando e Controle, Visualização de Informações, Visualização Interativa, Interação Humano-Computador.

Abstract

The process of Information Visualization consists in encoding data in graphical representation aiming to extend the human cognition in their understanding. Interactive Visualization (IV) is a technique originated from Human-Computer Interaction (HCI), which allows the user to directly transform the data, searching for new inferences about them, consequently leading to changes in the graphical representation itself. This study supports the hypothesis that IV can contribute to the improvement of data quality, and consequently the perceived understood by the user, since it can contribute to building his knowledge about an situation. The objective is to prove this hypothesis by implementing VI techniques on military command and control systems, where the understanding of a scenario is extremely critical.

Keywords: Information Visualization, Interactive Visualization, Command and Control, Data Fusion Systems, Human-Computer Interaction.

Introdução

O processo de visualização de informações constitui-se em codificar dados em representações gráficas com o objetivo de ampliar a cognição humana em seu entendimento. A Visualização Interativa (VI) é uma técnica com origens na Interação Humano-Computador (IHC), que permite que o usuário transforme diretamente os dados em busca de novas inferências sobre os mesmos, conduzindo conseqüentemente à mudanças na própria representação gráfica. O presente trabalho sustenta a hipótese de que a VI pode contribuir para a melhoria de qualidade dos dados, e conseqüentemente da informação percebida e compreendida pelo usuário, uma vez que o mesmo pode contribuir diretamente para a construção de seu conhecimento sobre uma situação.

O Comando e Controle (C2) é uma aplicação de Sistemas de Fusão de Dados onde dados de diversas fontes são utilizados para gerar um conhecimento mais aprimorado sobre uma determinada situação, o qual não seria possível obter com uma única fonte de dados.

Esses dados são apresentados ao usuário do sistema por meio de um dispositivo de IHC, permitindo-o analisar e explorar os dados para se obter ciência da situação e identificar possíveis ameaças, tornando-se capaz de tomar decisões para solucionar essas ameaças.

Como se trata de uma aplicação de Sistemas de Fusão de Dados, sistemas de Comando e Controle também estão sujeitos à problemas de qualidade de dados. Esses problemas podem ser originados na fase de obtenção desses dados pelos sensores físicos, devido à geração ruídos, na fase de processamento dos dados pelo sistema, ou até mesmo na apresentação dos dados ao usuário, pelo fato do sistema não apresentar corretamente os problemas de qualidade ao usuário, ou não apresentar de forma alguma. Isso compromete o processo de tomada de decisão pois pode levar o usuário à tomar decisões incorretas sobre uma determinada ameaça.

O presente trabalho busca, por meio da utilização de técnicas de Visualização Interativa, apoiar o usuário durante a realização da atividade analítica, permitindo-o interagir com o sistema com o objetivo de obter conhecimento sobre a situação e dos objetos que a compõe. Além disso, o trabalho foca-se em demonstrar ao usuário o estado da qualidade dos dados, permitindo-o inferir sobre o sistema para realizar melhorias na qualidade das informações exibidas pelo sistema.

Motivação

Sistemas de Comando e Controle (C2) são baseados em Sistemas de Fusão de Dados. Tais sistemas utilizam dados vindos de diversas fontes com o objetivo de incrementar o conhecimento sobre uma situação. Por se tratar de sistemas onde o comandante deve tomar decisões críticas para a resolução de incidentes, o sistema deve ter um alto grau de confiança.

Por se tratar de sistemas que dispõem de dados gerados por várias fontes de dados, problemas de qualidade de dados podem ser agregados à esse processo. Tais problemas podem ser gerados por medições feitas por sensores físicos, devido à geração de ruídos, ou durante a fase de processamento dos dados pelo sistema, e também na apresentação dos dados ao usuário, já que os dados podem não ser apresentados apropriadamente, podendo acarretar em incertezas durante o processo de ciência da situação.

Por esse motivo são necessárias técnicas apropriadas para a análise desses dados, considerando também os problemas de qualidade de dados associados a esses dados, contribuindo no processo de aquisição de conhecimento do usuário para a ciência da situação, e consequentemente tornando-o capaz de tomar decisões mais eficientes para a solução de problemas.

Objetivos

O objetivo desse trabalho é desenvolver técnicas de Visualização Interativa que permitam ampliar o processo cognitivo do usuário durante a atividade analítica, podendo contribuir com uma melhoria durante a obtenção de conhecimento sobre a situação. As técnicas devem permitir também, que o usuário tenha ciência da qualidade dos dados exibidos, permitindo-o interagir com o sistema realizando transformações de dados e representações, contribuindo para a construção do próprio conhecimento sobre a situação, e conseqüentemente melhorando a qualidade das informações e do processo de tomada de decisão.

Capítulo 1 - Sistemas de Fusão de Dados

Sistema de fusão de dados é um processo que lida com a associação, correlação e combinação de dados e informações de um ou vários sensores para se obter posições refinadas e estimativas de identidade, e avaliações completas e atuais de situações e ameaças assim como sua significância [1].

O autor Khaleghi et al. [3] define a fusão de dados como o estudo de métodos eficientes de forma automática ou semi-automática para transformar a informação advinda de diferentes fontes e diferentes pontos no tempo em uma representação que ofereça suporte efetivo à tomada de decisão seja ela automatizada ou humana.

Os Sistemas de Fusão de Dados tiveram origem de aplicações militares. Essa tecnologia é utilizada em áreas do Departamento de Defesa (DoD) para reconhecimento automático de alvos, controle de veículos autônomos, etc.

Os autores Hall e Llinas comparam os conceitos de um sistema de fusão de fontes heterogêneas com as capacidades sensoriais dos seres humanos e animais, como: visão, audição, tato, paladar, etc. Esses sentidos nos dão um conhecimento mais preciso sobre o ambiente a nossa volta, enquanto a visão é limitada, a audição pode nos dar um alerta sobre possíveis ameaças e isso aumenta nossa chance de sobrevivência [19].

Com o surgimento de novos sensores, hardwares e técnicas com maior capacidade de processamento, várias aplicações de sistemas de fusão são utilizados atualmente, tanto militares quanto não militares.

As aplicações militares incluem: detecção automática de alvos, direção de veículos autônomos, sensoriamento remoto, vigilância de campo de batalha, sistemas de reconhecimento automático de ameaças, como sistemas de identificação aliado-inimigo (*Identification-Friend-Foe-Neutral* - IFFN). Aplicações não militares incluem: monitoramento de processos industriais, manutenção da condição de maquinários complexos, robótica e aplicações médicas [19].

1.1 O modelo de Fusão JDL

Para o desenvolvimento efetivo de tais sistemas de fusão, tornou-se necessário o desenvolvimento de uma terminologia unificada, para que desenvolvedores e pesquisadores

seguissem um padrão, ou um modelo, com o objetivo de melhorar a comunicação entre a comunidade de pesquisa e desenvolvedores de sistema. Com o objetivo de solucionar esse problema o *Joint Director of Laboratories (JDL) Data Fusion Workgroup* criou um modelo de processo para fusão de dados chamado JDL, demonstrado na Figura 1. Trata-se de um modelo conceitual de fusão de dados, de alto nível que foi adotado para o desenvolvimento de aplicações em diversas áreas [19].

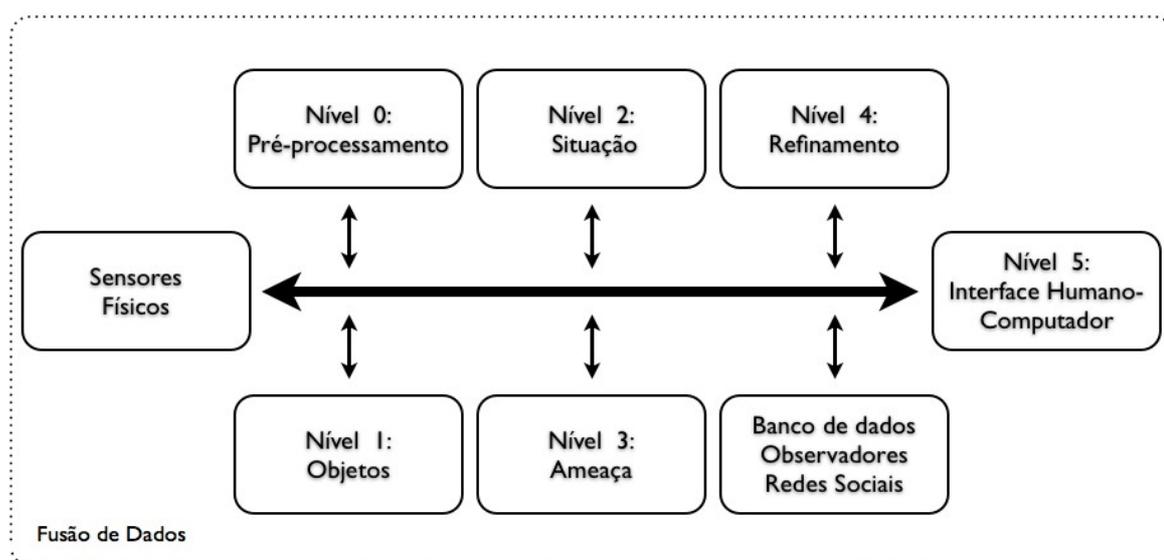


Figura 1: Modelo de Fusão JDL

Os níveis do modelo são:

- Sensores Físicos: são as fontes de dados físicas que alimentam o sistema, como: radares, câmeras de vigilância, sensores biométricos, etc [19], [18], [17].
- Nível 0 – Pré-processamento: o pré-processamento dos dados dos sensores tem como objetivo concentrar o processo de fusão no dado mais pertinente à situação atual. É responsável também por reduzir a carga de dados no processo de fusão distribuindo os dados para o nível de processamento apropriado. Os processos do nível 0 também envolvem filtragem e refinamento do dado bruto [19], [18], [17].
- Nível 1 – Identificação de Objetos: Esse processo combina informações locais, paramétricas e de identidade, para obter representações de objetos individuais, por exemplo, armamentos, unidades militares, plataformas. É baseado em 4 funções principais: 1) transforma os dados

dos sensores em conjuntos consistentes de unidades e coordenadas; 2) estima os atributos e posição do objeto; 3) atribui dados ao objeto para permitir estimações estatísticas; 4) refina a estimativa da classificação e identidade do objeto [19], [18], [17].

- Nível 2 – Identificação da Situação: o processamento do nível 2 utiliza os dados obtidos nos níveis anteriores para realizar tentativas, para que sejam traçadas as atuais relações entre os objetos e suas ações com o contexto do seu ambiente, esta relação busca identificar o quão relacionados estão determinados objetos [19], [18], [17].
- Nível 3 – Identificação de Ameaça: o processamento de nível 3 projeta a situação atual no futuro para obter deduções sobre ameaças inimigas, vulnerabilidades aliadas e inimigas, e oportunidades para operações [19], [18], [17].
- Nível 4 – Refinamento: o processamento de nível 4 desempenha 4 funções principais: 1) monitora a performance do processo de fusão para fornecer informações sobre controle em tempo real e performance em longo prazo; 2) identifica qual informação é necessária para melhorar o produto de fusão multinível, como posições e identidades; 3) determina qual os requisitos para coletar informações relevantes de uma fonte de dados específica; 4) aloca e direciona os sensores para atingir objetivos de missões [19], [18], [17].
- Nível 5 – Interface Humano-Computador (IHC): o nível 5 é o responsável por realizar a comunicação entre o sistema de fusão e o operador ou usuário do sistema. É através desse nível que o sistema se comunica com o operador por meio de alertas, visualizações, representações de informações de localização e identidade através de overlays dinâmicos em visualizações geográficas. Esse nível permite também que o usuário forneça dados de entrada para o sistema através de comandos, e apóia o operador no processo de raciocínio analítico [19], [18], [17].
- Banco de dados, Observadores, Redes sociais: é uma coleção de funções que provê acesso e administração de banco de dados de fusão, incluindo recuperação de dados, armazenamento, proteção, compressão. É composto por uma grande variedade de tipos de dados, como imagens, vetores,

dados textuais, por exemplo, relatórios vindos de operadores ou observadores [19].

O objetivo de um sistema de fusão é coletar dados de diversos sensores e refiná-los em informações significantes como posição e características de entidades, para completar atividades de planejamento estratégico e tático, e controle de ameaças. Essas informações são apresentadas para o usuário para que ele seja capaz de analisar e compreender a situação por completo, para que seja capaz de tomar decisões [18].

1.2 O papel do humano em sistemas de fusão

O papel humano em sistemas de fusão se baseia na interpretação de resultados gerados pelo sistema, tornando-o capaz de executar tarefas de raciocínio analítico, como a tomada de decisão. O nível 5 do sistema de fusão permite a interação entre o usuário e o sistema, onde através da interpretação e transformação dos resultados fornecidos pelo sistema, o usuário adquire conhecimento da situação sendo capaz de realizar a tarefa de tomada de decisão.

O sistema de fusão deve processar os dados obtidos dos diversos sensores e transformá-los em um formato que possa ser codificado, ou seja, representado graficamente, apresentando ao usuário os atributos mais relevantes do dado, bem como informações referentes à qualidade dos dados.

A representação gráfica de tais dados, quando feita de maneira eficiente, ajuda no processo cognitivo do usuário, permitindo uma análise mais eficiente sobre o estado do dado e da situação como um todo, permitindo o usuário a ter ciência da situação e ser capaz de tomar decisões mais eficientes.

Como se trata de um sistema que lida com uma quantidade muito grande de informações, uma representação estática dos dados do sistema não é suficiente para atender a execução da atividade analítica de forma eficiente. Um dos motivos é em relação à grande quantidade de informações para serem exibidas em uma área limitada do display. O usuário deve ser capaz de interagir com as representações em busca de informações adicionais ou completas sobre a situação sendo analisada [2].

Uma vez que o usuário tem conhecimento sobre os dados representados pelo sistema, ele é capaz de fazer julgamentos sobre a relevância dos dados para a atividade em

andamento, podendo filtrar os dados irrelevantes para a situação atual, permitindo-o diminuir a quantidade de informações exibidas para manter o foco apenas nas informações relevantes.

O usuário também pode atuar como entrada de dados para o sistema de fusão, podendo atribuir meta-dados, ou meta-informações às representações disponíveis baseadas no conhecimento que o usuário possui. Por exemplo, o usuário pode atribuir um valor de “validade” do dado de acordo com a fonte que o originou, como exemplo, atribuir a validade de dias para dados obtidos por satélite, horas para dados obtidos por aviões [12].

O recente trabalho de Wickramarathne [4] cita a atuação do humano como sensor em sistemas de fusão, fornecendo dados classificados como soft data, que se diferencia dos dados vindos de sensores físicos, que são classificados como hard data. Esses dados podem ser opiniões de especialistas, dicas, declarações de informantes. Essas informações tendem a ser mais qualitativas e subjetivas de interpretação. O autor cita também que essas informações (soft data) geralmente fornecem as evidências mais cruciais e talvez as mais complementares, comparadas com as vindas de sensores físicos (hard data).

A importância da utilização de humanos como sensores já se tornou reconhecida para a comunidade de pesquisa. No exemplo de sensoriamento participativo, observadores são encarregados de alimentar aplicações com informações, essas aplicações podem ser tanto de domínio militar quanto civil, como exemplos, planejamento urbano e saúde pública. Uma das formas em que o humano pode atuar como sensor é fornecendo dados através da Internet, por exemplo, em redes sociais como Flickr, Instagram, Facebook, Twitter, etc (Wickramarathne, 2012).

O capítulo seguinte descreve em mais detalhes de que forma o usuário é capaz de inferir sobre o sistema para colaborar com a qualidade do conteúdo de dados e informações, implicando no conhecimento adquirido durante a atividade de análise dos dados e conseqüentemente na eficiência do processo de tomada de decisão.

Capítulo 2 - Sistemas de Fusão de Dados e a Qualidade dos Dados e Informações

Sistemas de fusão de dados podem gerar uma grande quantidade de problemas na qualidade do dado. Por se tratar de um sistema complexo que lida com dados heterogêneos, no momento em que os dados de diferentes sensores são fundidos podem ocorrer problemas, assim como no momento da obtenção dos dados pelos sensores, pois qualquer sensor físico está sujeito à geração de ruídos, acarretando em imprecisão ou incerteza nas medições.

Antes de se discutir sobre a qualidade de dados é importante definir também o conceito de dado. Dados são representações digitais de objetos do mundo real, que podem ser armazenados, recuperados e manipulados através de um procedimento de software. Dados podem representar diferentes fenômenos como medições, eventos, características de pessoas, o ambiente, sons [36].

Informação, portanto, é o produto do processo de refinamento de dados realizados pelo sistema de fusão, ou melhor, a informação é o dado bruto refinado ou transformado em uma forma que transmita conhecimento para o usuário.

Em sistemas de fontes heterogêneas é gerada uma grande variedade de representação de dados, já que fontes diferentes produzem uma representação de dado de objetos diferentes. De acordo com a literatura os dados podem ser categorizados em três tipos:

- **Estruturado:** Cada dado possui uma estrutura fixa. Exemplos de dados estruturados são tabelas relacionais de banco de dados, onde cada dado da tabela pode possuir um determinado valor, por exemplo, inteiro, boolean, caractere ou uma cadeia de caracteres.
- **Semi-estruturado:** São dados estruturados, mas que possuem certa flexibilidade. Um exemplo de dado semi-estruturado é a linguagem XML, onde mesmo seguindo uma estrutura um dado pode ser representado de várias formas, por exemplo, uma data pode ser representada por um campo, por exemplo, data, ou vários campos, sendo dia, mês, ano.
- **Desestruturado:** São dados expressos em forma de linguagem natural.

2.1 Dimensões e Taxonomias de Qualidade de Dados

A medida de qualidade desses dados é feita através de dimensões ou atributos de qualidade. As dimensões de qualidade mais relevantes para esse trabalho são descritas abaixo de acordo com a definição de Batini [36].



Figura 2: Dimensões de qualidade de dados adaptada da taxonomia de Khaleghi [3]

Precisão (Accuracy): é definida como a diferença entre o valor do fenômeno do mundo real e o valor que se busca representar digitalmente. Como exemplo, podemos considerar um nome de uma pessoa $p = \text{João}$, e o valor de $p' = \text{Jão}$, que é incorreto. A precisão pode ser também dividida em dois tipos, sendo eles precisão sintática e precisão semântica.

Precisão sintática é definida como a proximidade entre um valor v e o valor dos elementos de um domínio D . Seguindo o exemplo anterior, o domínio D é caracterizado por nomes de pessoas, tomando como exemplo dois valores $p = \text{João}$ e $p' = \text{Maria}$, mesmo que o valor de p' seja diferente do valor de p , p' é considerado sintaticamente correto, pois também pertence ao domínio D .

A precisão semântica é definida com a proximidade entre um valor v e o valor verdadeiro de v' . Continuando com o exemplo anterior, se temos um valor $p = \text{Maria}$ e um valor $p' = \text{João}$, p é considerado um erro de precisão semântica.

Incompletude (Completeness): a completude se refere à quão completa está a informação para a execução da tarefa atual.

Atualidade (Timeliness): representa o quão atual é o dado em relação à tarefa sendo executada no momento.

Granularidade: define o quão detalhada está a informação.

No trabalho de Khaleghi [3], o autor divide a imperfeição do dado em três classificações: incerteza, imprecisão e granularidade.

Um dado é incerto quando o grau de confiança associado, sobre o que é indicado pelo dado, é menor do que 1. Já um dado impreciso é o dado que se refere a vários objetos, ao invés de apenas um objeto. A granularidade do dado se refere à habilidade de distinguir entre objetos que são descritos pelo dado, sendo dependente do conjunto de atributos fornecidos [3].

Há também, outras taxonomias que buscam classificar problemas de qualidade como a imperfeição e incerteza, como a de Smithson [32], Bouchon-Meunier e Nguyen [33], Krause e Clark [34] e Smets [35].

Smithson [32] classifica a incerteza como uma forma de ignorância. A ignorância é dividida em duas categorias: o estado de ignorância (erro) e o ato de ignorar (irrelevância). A irrelevância é o ato de ignorar algo que é irrelevante para a situação atual, já o erro é causado por vários fatores como a distorção ou conhecimento incompleto. A Figura 3 demonstra a taxonomia de ignorância de Smithson.



Figura 3: Taxonomia de ignorância de Smithson

Bouchon-Meunier e Nguyen [33] caracterizam a incerteza como conhecimento imperfeito, e a divide em três domínios: incerteza probabilística; conhecimento incompleto (crença, leis, imprecisão); descrição vaga ou imprecisa. Essa classificação é demonstrada pela figura 4.

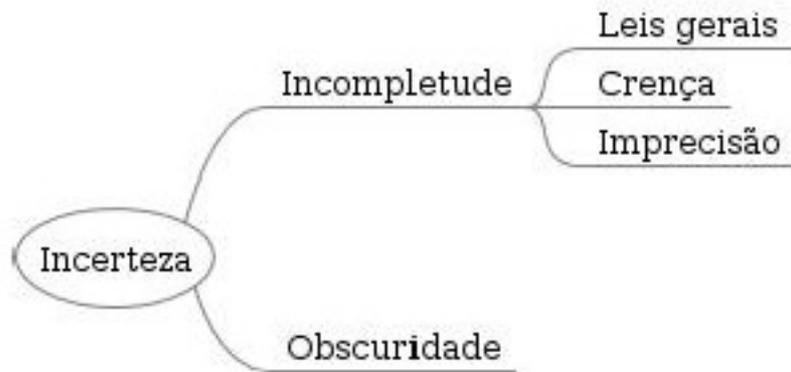


Figura 4: Tipos de incerteza segundo Bouchon-Meunier e Nguyen.

Krause e Clark [34] apresentaram uma taxonomia de incerteza unária e de conjunto. A incerteza unária é quando a incerteza é aplicada em proposições individuais, já a de conjunto é quando a incerteza é aplicada a um conjunto de proposições. Ambas as categorias levam ao conflito ou ignorância. A Figura 5 mostra tal taxonomia.



Figura 5: Taxonomia de incerteza proposta por Krause e Clark.

O autor Smets [35] propôs uma classificação da imperfeição da informação em três categorias, sendo elas:

- Imprecisão: relacionada ao conteúdo da informação;
- Inconsistência: afirmação sem manutenção;
- Incerteza: induzida pela falta de informação;

Smets considera a imperfeição como um termo geral, onde a incerteza é uma modalidade, dividida em objetiva ou subjetiva.

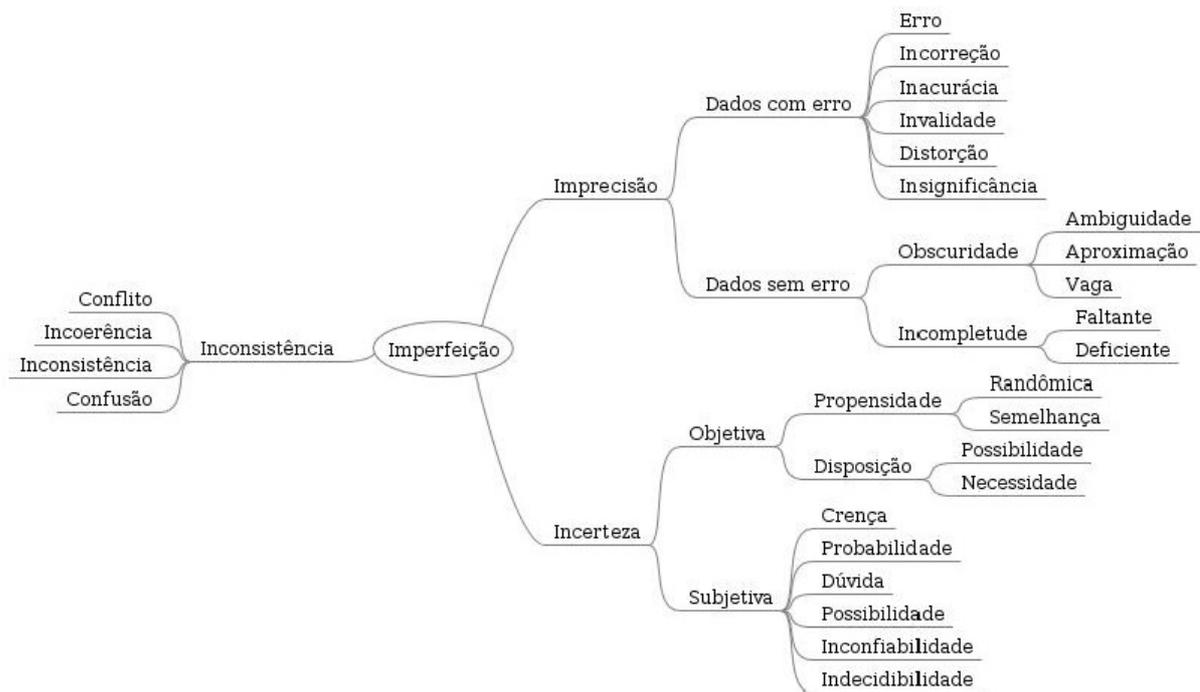


Figura 6: Classificação da imperfeição segundo Smets.

O presente trabalho busca resolver os problemas de qualidade de dados relacionados à imperfeição de dados, que é composta por: imprecisão; incompletude e incerteza, segundo Khaleghi et. al. [3]. Abaixo são mencionados alguns problemas de dados relacionados à imperfeição do dado, tomando como referência o trabalho de Wickramarathne[4].

O problema de incompletude do dado se refere à quão completa está a informação, ou seja, se certa parte, ou toda a informação está faltando. Como exemplo, considere uma base de dados contendo um registro com informações de uma unidade detectada em um campo de batalha, contendo as seguintes informações ou atributos: tipo da unidade, localização, velocidade. Se, por exemplo, o tipo da unidade estiver faltando no registro, então o registro é considerado incompleto.[4]

No caso da imprecisão, é analisado se a informação tem precisão suficiente para a execução da atividade sendo executada. Tomando como exemplo o mesmo cenário de campo de batalha do exemplo anterior, se a informação de velocidade da unidade detectada foi inserida com a possibilidade de estar entre 25 a 40 mph, essa informação pode não ser precisa o suficiente para tomar uma decisão de lançar um míssil não guiado, então nesse caso a informação é imprecisa [4].

Os casos de imprecisão e incompletude são dependentes de contexto, pelo fato de

que uma informação imprecisa ou incompleta pode ser insuficiente para um contexto, mas em outro contexto ela é suficiente, ou suficiente o bastante para se tomar uma decisão. Tomando o exemplo anterior, onde a identificação da velocidade de uma unidade inimiga é entre 25 e 40 mph, pode não ser precisa o suficiente para disparar um míssil não guiado, mas é suficientemente precisa para emitir um alerta de uma possível intrusão [4].

A incerteza é um caso de problema de dados onde a informação é precisa, é completa, mas é incerta, pela possibilidade de estar incorreta. Esse tipo de incerteza ocorre principalmente devido a imperfeições relacionadas à fonte da informação. Como exemplo o cenário de um campo de batalha, um soldado no comando, devido às más condições do clima, identificou o tipo da unidade como um tanque ou um jipe, com 75% de confiança. Nesse caso, mesmo a informação armazenada na base de dados sendo precisa e completa ela é considerada incerta [4].

Tais problemas de qualidade podem ser gerados durante o processo de obtenção dos dados, pelo fato de qualquer sensor físico estar sujeito à geração de ruídos, acarretando em dados incertos. Os problemas podem ser gerados também durante a fase de processamento dos dados, realizado pelo sistema de fusão, ou até mesmo no processo de apresentação dos dados ao operador, onde os problemas dos dados podem ser mal apresentados ou não serem apresentados, levando o operador a interpretações incorretas da real situação sendo analisada, comprometendo o processo de tomada de decisão.

2.2 Qualidade de Dados em Sistemas de Fusão

Sistemas de fusão falham ao não apresentar tais problemas de qualidade ao usuário durante a execução de atividades analíticas, o que pode levá-lo a interpretações incorretas sobre o dado. Para que a tomada de decisão seja feita de forma efetiva, o usuário deve ter total ciência da situação de qualidade do dado.

Para que a apresentação dos dados ao usuário seja feita de forma eficiente, é necessária a utilização de técnicas eficientes para a visualização e manipulação desses dados. Em sistemas onde há a necessidade de se analisar uma grande quantidade de informações, são adotadas técnicas de Visualização Interativa para permitir que o usuário seja capaz de analisar o dado com mais eficiência.

A Visualização Interativa é uma técnica originada da Interação

Humano-Computador (IHC), e permite ao operador manipular diretamente os dados em busca de novas informações, informações completas e até mesmo realizar transformações de dados, implicando em alterações nas representações gráficas dos mesmos. A Visualização Interativa é geralmente definida como o uso de computadores que suportam interação e representação visual de dados para ampliar o conhecimento [37].

Torna-se necessário então, representar não só o dado, mas também sua qualidade, além de permitir interações que dê ao operador a capacidade de realizar transformações de dados para ajustar as representações de acordo com a necessidade da atividade sendo executada no momento.

As técnicas de Visualização Interativa têm o objetivo de auxiliar o operador na execução de atividades de raciocínio analítico como tomada de decisões, ciência da situação (*sense-making*) e descoberta de conhecimento, por exemplo. A representação gráfica dos dados auxilia na execução de atividades cognitivas como categorização, abstração e dedução, que são processos dinâmicos e envolvem assimilação e reorganização constante de informações. Como as representações das informações são interativas, o operador é capaz de manipular e organizar as representações de acordo com as necessidades da atividade analítica sendo executada naquele momento [2].

2.3 Trabalhos relacionados

Alguns autores propõem soluções que incluem a participação do usuário na forma de representação das informações, dando-o o poder de manipular as informações através de interações.

Howard e MacEachren [38] descreveram um ambiente interativo chamado R-VIS, que foi desenvolvido para apoiar a exploração de incertezas. Eles propõem duas estratégias fundamentais para a visualização interativa da incerteza. Uma delas é exibir o dado e a incerteza juntos, tratando incerteza como uma segunda variável.

Hengl [39] apresentou um método utilizando *hue*, saturação e intensidade de cores para exibir incertezas. Seu método resultou no dado sendo exibido incrementalmente pálido, dependendo da magnitude da incerteza. Isso auxilia o operador a ter consciência da qualidade do dado apenas olhando para sua representação, através da transparência da representação do dado.

O trabalho de Summers et. al. [12] baseia-se na utilização de *overlays* em mapas terrestres. Através de um gerenciador de *overlay*, o operador é capaz de criar, alterar quais são visíveis e até mesmo reordená-los. Essa abordagem resolve o problema de complexidade de dados, isto é, o problema de muita informação sendo exibida ao mesmo tempo, o que pode confundir o operador. O operador é capaz de escolher qual informação deseja visualizar, deixando visíveis apenas informações que são relevantes para a atividade de tomada de decisão em andamento.

Cliburn et. al [40] propôs ajudar usuários a lidar com a complexidade de exibição de dados através de interatividade. Permitindo usuários clicar em uma única área ou em uma determinada região de dados e visualizar a parcela de incerteza daquela área ou região selecionada.

Os trabalhos citados acima possuem características importantes para a redução de problemas de qualidade de dados que influenciam no processo cognitivo do usuário durante a execução da atividade de tomada de decisão. Um fator é a exibição da qualidade do dado junto à representação gráfica do dado, utilizando variações na opacidade, cor e/ou saturação das representações para transmitir ao operador o fator de qualidade do dado. Outro fator é permitir que o usuário manipule as visualizações para atender as necessidades da tarefa de tomada de decisão através de Visualização Interativa.

2.4 Visualização Interativa e Transformações de Dados e Representações

O pipeline de visualização é dividido em três tarefas básicas: 1) Transformações de dados; 2) Transformações de representações; 3) Transformações de mapeamentos visuais.

A transformação de dados é a primeira fase do pipeline de visualização e tem o objetivo de transformar dados brutos em uma forma de dados abstratos que possam ser exibidos graficamente ao usuário. Os métodos de transformações de dados que são executados nessa etapa geralmente são métodos matemáticos como média, mediana, amostragem, entre outros métodos de limpeza e refinamento de dados brutos.

No trabalho de Thomas e Cook [11], os autores definem a transformação de dados como um procedimento computacional que converte entre representações de dados. As transformações de dados podem ampliar o dado derivando dados adicionais. Por exemplo, pode-se utilizar análise lingüística para atribuir significado às palavras de um documento de

texto. Outras transformações de dados podem determinar como apresentar o dado eficientemente, tal como a criação de uma representação 2D ou 3D do dado com centenas ou milhares de dimensões.

No trabalho de Côrtes et. al. [7], os autores citam a transformação de dados como um processo que possibilita a transformação ou consolidação dos dados no formato apropriado para o processo de mineração (mining), através de operações do tipo sumarização ou agregação, entre outras técnicas.

O autor Ed H. Chi [28] cita a transformação de dados como um processo que transforma os dados brutos em uma forma abstrata, que possa ser utilizada para a geração de uma representação visual.

A partir dos dados abstratos são geradas visualizações abstratas para se representar estes dados graficamente (transformações de representações). Essas transformações geralmente consistem em transformar os dados em uma estrutura como, por exemplo, Árvores de Discos (*Disk Tree*), Cones de Discos (*Cone Tree*) [28], ou componentes individuais como círculos, quadrados, retângulos, ícones, etc.

A tarefa de mapeamento de dados consiste em transformar os dados abstratos resultantes da tarefa de transformação de dados em uma forma abstrata de representação, ou seja, em um conteúdo visualizável. Já a transformação de visualizações permite que o operador modifique essa forma visualizável de informação, por exemplo, sua forma ou cor [28].

A transformação de mapeamentos visuais consiste em exibir graficamente ao usuário as informações que estão em um formato visualizável, ou seja, as informações resultantes do processo de transformações de representações. Uma vez que o usuário é capaz de visualizar graficamente as operações ele é capaz de realizar atividades de análise visual e obter conhecimento, ou ciência da situação que o sistema lhe esta fornecendo [28]. A Figura 7 demonstra como acontece a iteração entre essas três etapas.

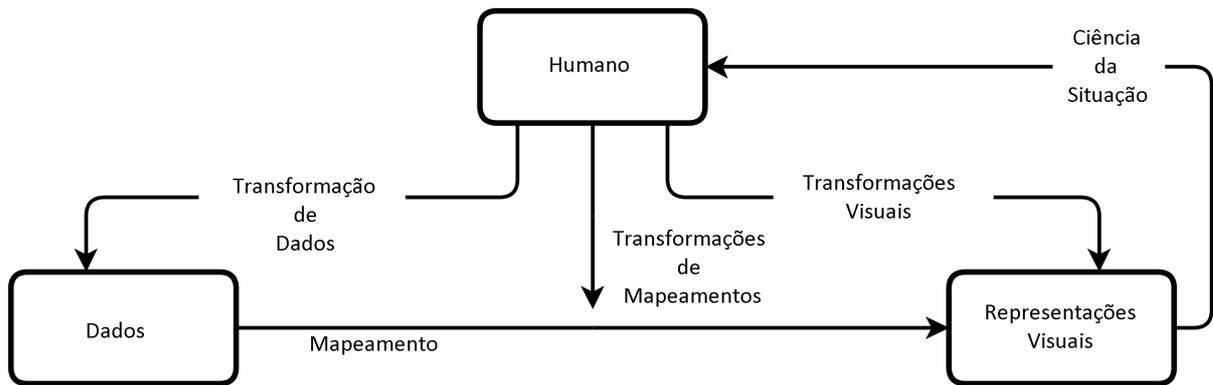


Figura 7: Ilustração do pipeline de visualização

Dessa forma o pipeline é um processo estático de apresentação de informações graficamente, o usuário deve também ser capaz de interagir com essas representações. Para que isso seja possível torna-se necessário o uso de interfaces interativas (mouse, teclado, interfaces tangíveis, etc.), bem como a utilização de técnicas eficientes de Visualização Interativa.

Através de interações o usuário é capaz de interagir com as representações gráficas, sendo capaz de manipulá-las, tornando o processo de análise visual muito mais interativo e intuitivo, já que o usuário pode escolher quais informações deseja visualizar, obter informações adicionais de um dado específico, bem como modificar as representações, podendo contribuir para a formação do próprio conhecimento.

Abaixo são listadas algumas transformações presentes no trabalho de Karen e Schreiber [20] que podem ser realizadas através de interações.

- Filtragem: essa técnica é de grande importância para a atividade de análise visual e permite ao operador reduzir a quantidade de dados exibidos. Alguns exemplos de filtragem são consultas dinâmicas utilizando range sliders ou escolhendo uma série de nós numa rede de visualização.
- Ordenação: ordena os dados de acordo com seus valores.
- Seleção: é geralmente utilizada como o avanço da filtragem. O objetivo é selecionar um objeto individual ou um conjunto de objetos para destacar, manipular ou filtrá-los.
- Navegação/Exploração: classe importante de técnicas de interação que tipicamente modificam o nível de detalhe seguindo o mantra “visualize primeiro, aproxime e filtre, detalhes sobre demanda (Shneiderman, 1996)”.

Abordagens bem conhecidas são foco&contexto, visão-global&detalhes, zoom&panning.

- Link: se baseia em ligar um conjunto de visualizações com o objetivo de permitir ao usuário descobrir relações entre as visualizações.
- Derivar: está fortemente relacionada ao processo de ciência da situação (sense-making), e baseia-se na realização de cálculos adicionais baseados no dado de entrada inicial. O usuário pode integrar resultados de cálculos estatísticos, como agregação e médias, ao dado que será visualizado.

Para que a atividade de análise visual seja executada de maneira eficiente o usuário deve ser capaz de interagir com o sistema através de técnicas de Visualização Interativa para que ele possa criar ciência da situação e colaborar com a melhoria da qualidade dos dados, e conseqüentemente a qualidade da tomada de decisão.

A Tabela 2 demonstra a relação entre transformações de dados e representações realizadas através de interações e qual dimensão de qualidade elas afetam.

Tabela 2: Transformação de dados por meio de interações

Interação	Dimensão de Qualidade	Transformação de Dados
Clique [31]	Incompletude	Seleção [31] [2] [20] [8]
Selecionar área de interesse [31]	Incompletude	Seleção / Amostragem [31] [20] [22]
Slider [31]	Imprecisão, Incompletude, Incerteza	Filtro[20] [31] [22]
Swipe sobre o mapa [22]	Incompletude	Navegação [20] [31] [22]
Agregar representações [10]	Incompletude, Granularidade, Imprecisão,	Agregação/desagregação [31] [22]

Uma das tarefas mais importantes no processo de análise visual é a seleção, que permite ao operador selecionar um ou mais objetos para serem manipulados, permitindo destacá-los, adicionar anotações, filtrá-los, ou obter detalhes sobre o objeto [31], [20], [2], [8].

A seleção de um objeto pode ser realizada através de interações como clique ou pela seleção de uma região retangular, elíptica ou forma livre. A seleção interativa está fortemente relacionada com a operação de filtro, já que é utilizada para a identificação de objetos, permitindo ao operador removê-lo, caso o objeto não seja relevante para a tomada de decisão em andamento [31].

Em um primeiro momento durante o processo de análise visual o usuário não tem total conhecimento sobre todos os objetos representados no display, é através de interações e transformações de dados/representações que ele é capaz de adquirir ciência da situação. Através de uma interação de seleção, o operador obtém informações adicionais sobre o objeto, fazendo com que seu conhecimento sobre esse determinado objeto seja mais completo.

À medida que o usuário obtém informações mais completas sobre situação sendo analisada, o problema de incompletude é amenizado, pois através da realização de seleções, obtendo detalhes sobre cada objeto, o operador tem informações mais completas para tomar uma decisão mais eficiente.

A utilização da seleção com o objetivo de selecionar um grupo de objetos através de uma região também é de extrema importância na análise visual. Um conjunto de dados inteiro pode conter partes desnecessárias para a tarefa analítica em andamento. Sendo o operador capaz de selecionar uma região de interesse, permite que ele se concentre apenas nos dados relevantes para a tarefa em andamento [8]. Isso faz com que a informação sendo analisada seja completa o suficiente para a realização da atividade analítica em execução, resolvendo o problema de incompletude e também a complexidade de dados exibidos no display. Isso é demonstrado na segunda linha da Tabela 2.

Além da realização de filtros através da seleção de uma área de interesse é possível filtrar dados utilizando *widgets* de consultas dinâmicas, que podem ser: caixas de seleção, botões de seleção (*radio button*), listas, caixas de pesquisa e *sliders*. Quando associados à atualizações de visualizações em tempo real, esses *widgets* permitem uma exploração rápida e reversível dos conjuntos de dados [31].

A terceira linha da tabela propõe a realização de filtros a partir de *sliders*, onde os objetos são filtrados de acordo com o seu valor de imprecisão, incompletude ou incerteza. Fazendo isso o operador pode escolher visualizar dados com mais, ou menos qualidade. Dessa forma as três dimensões de qualidade poderiam ser tratadas.

A quarta linha da tabela baseia-se em técnicas de navegação utilizadas na análise de dados. Geralmente, o início de uma atividade analítica é baseado na visualização global dos dados disponíveis, para a identificação de padrões, destaques ou possíveis problemas de qualidade. Essas tarefas de orientação podem levar a investigações mais específicas e detalhadas sobre um subconjunto de dados. Um exemplo comum seria um mapa geográfico, onde é exibido um território global, e a partir daí, são realizadas interações de navegação, como *zoom*, para focar em regiões de interesse. As operações de navegação como *panning* e

zoom, geralmente são realizadas a partir de eventos do mouse como o *scroll* e *drag*, o nível de *zoom* pode ser alterado por meio de um evento de *scroll* do mouse, ou pela manipulação de um *slider*. A navegação geralmente é feita através de eventos de *drag* ou por meio de barras de rolagem [31].

Através da navegação sobre o espaço de representações o operador pode concentrar-se apenas na região de dados relevantes para a atividade analítica em andamento, fazendo com que as informações exibidas sejam completas o suficiente para a realização dessa atividade, impedindo que o operador se confunda com informações desnecessárias.

A última linha da tabela demonstra a possibilidade de se utilizar o agrupamento ou agregação de objetos para se obter informações mais completas e/ou mais precisas.

O agrupamento é também uma forma eficiente de se explorar grandes conjuntos de dados a partir de pedaços menores. É difícil adquirir conhecimento significativo de uma base de dados muito ampla e desorganizada. Através do processo de agrupamento e agregação, as informações relevantes são obtidas, simplificadas, organizadas e rotuladas. Isso reduz a carga de trabalho de pesquisa e memória do operador, além de permitir que ele concentre sua atenção para encontrar fatos de alto nível e partir para novas descobertas de conhecimento [8].

Através da realização de interações de agregação ou desagregação o operador é capaz de alterar o nível de detalhe que deseja visualizar, alterando o nível de granularidade dos dados. Por meio de agregações os dados são agrupados de uma forma significativa para o operador. A desagregação de dados transforma um grande agrupamento de dados em grupos de dados menores [10].

A utilização dessa técnica é importante, pois o operador pode reduzir a quantidade de informações exibidas no display agrupando objetos relacionados. Se um objeto possui um problema de qualidade, por exemplo, a imprecisão, realizando o agrupamento de um ou mais objetos relacionados que têm uma precisão maior, resultará em uma visualização mais precisa, resultando em uma melhoria da qualidade da informação desses dados.

Essa técnica também é capaz de resolver o problema de muita informação para se exibir, devido à área limitada dos displays, ajustando a granularidade dos dados através dessa técnica o operador é capaz de agrupar objetos relacionados para gerar uma nova visualização, ocupando menos espaço do display. Se há a necessidade de analisar os dados mais detalhadamente, o operador também é capaz de desagrupar os objetos, podendo analisar as informações mais detalhadamente.

O capítulo seguinte descreve quais técnicas foram implementadas e como o operador é capaz de dispará-las através de interações. Também serão descritos os possíveis benefícios que tais técnicas trazem para a experiência do usuário e para a qualidade do conteúdo das informações.

Capítulo 3 - Metodologia

O presente trabalho foi baseado em três fases: 1) Levantamento bibliográfico; 2) Prototipação; 3) Análise de resultados.

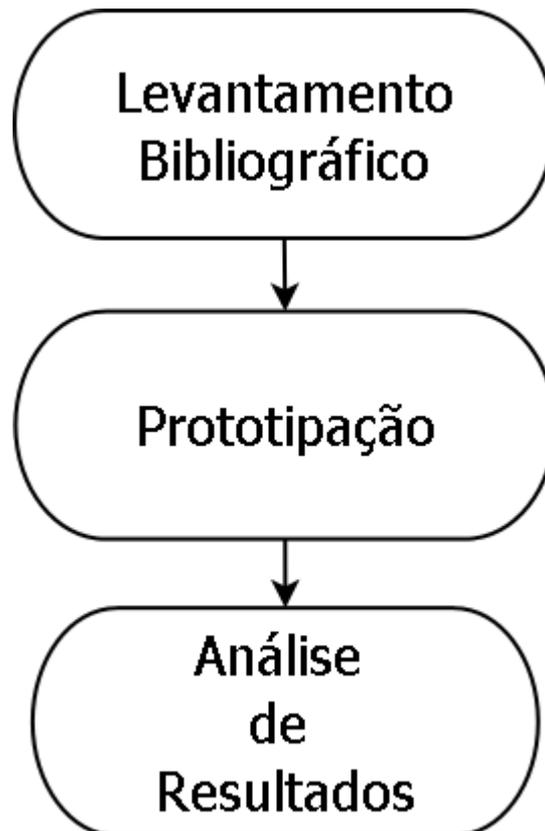


Figura 8: Fluxo de atividades realizadas no trabalho

1. Levantamento Bibliográfico:

Na primeira etapa do trabalho foi realizada a aquisição de documentos sobre as diferentes áreas que abrangem o trabalho como um todo, como Visualização Interativa, Interação Humano-Computador (HCI), Análise Visual, Sistemas de Fusão de Dados, Qualidade de Dados. Esses documentos são compostos basicamente de livros e artigos publicados em eventos da área de pesquisa.

Primeiramente foi realizado um estudo sobre a teoria de Sistemas de Fusão de Dados, quais os processos que o compõem, quais as suas finalidades e quais problemas existem em tais sistemas.

Entendendo o contexto de Sistemas de Fusão de Dados, foi realizado um estudo

mais aprofundado sobre os problemas que podem ser gerados em tais sistemas. Esse estudo consiste na pesquisa de dimensões de Qualidade de Dados e suas possíveis origens, e também formas de poder reduzir tais problemas, promovendo uma melhor qualidade da informação e consequentemente a eficiência da execução de atividades analíticas.

A terceira parte de estudo do trabalho foi aplicada na área de Visualização Interativa, Interação Humano-Computador (HCI) e Análise Visual. Essas áreas foram estudadas porque descrevem técnicas de visualização e análise de dados, bem como técnicas de interação com essas visualizações de dados.

Para que o usuário seja capaz de contribuir com a qualidade da informação ele deve ser capaz de inferir sobre o sistema através de técnicas de Visualização Interativa eficientes. O estudo realizado foi focado em encontrar técnicas que pudessem colaborar com a melhoria da qualidade da informação.

Para que essas técnicas de interação fossem implementadas foi necessário realizar uma pesquisa de APIs e frameworks que dessem suporte para que o operador fosse capaz de visualizar e interagir com as informações eficientemente. As ferramentas escolhidas para a implementação das técnicas foram:

- Google Maps Javascript API: a API Javascript do Google Maps foi utilizada para a criação de visualizações em um mapa geográfico, essas visualizações incluem: overlays, marcadores, ícones, etc.

- Hammer.js: essa biblioteca foi utilizada para implementar as técnicas de interação entre o operador e o sistema. A biblioteca é implementada em Javascript, e foi escolhida por permitir a integração entre os módulos de visualização e interação, já que os dois utilizam a mesma linguagem de programação. Além disso, a biblioteca Hammer.js permite interações multitouch, podendo ser utilizada em qualquer dispositivo que suporte interações multitouch.

- Google Fusion Tables, Facebook API e Twitter API: Essas bibliotecas foram utilizadas para a criação de massa de dados para que fosse possível a realização de testes.

2. Prototipação

Para aplicar os conhecimentos adquiridos nas etapas anteriores foi escolhido como estudo de caso do trabalho sistemas de fusão de dados multisensores, mais especificamente sistemas de Comando e Controle (C2).

Sistemas de Comando e Controle são utilizados por forças policiais para o gerenciamento de incidentes. O serviço policial atende muitos incidentes que tipicamente são resolvidos enviando recursos disponíveis naquele momento, um exemplo de recurso seria uma viatura, ou policiais. Estes incidentes podem estar relacionados à controle de grandes eventos, manifestações, saúde e bem estar público, eventos climáticos graves como inundações e epidemias [29].

O C2 é um sistema composto por sistemas que podem ser caracterizados com alto grau de complexidade, ambiente distribuído focado na distribuição de processos de tomadas de decisão sobre os agentes, alta demanda em confiança e, inevitavelmente, uma interface humano-computador implementada para apoiar a comunicação entre o operador e o sistema [30].

Os sistemas de C2 baseiam-se em um mapa geográfico onde são exibidas as representações de objetos como entidades inimigas, entidades aliadas e informações sobre terreno, por exemplo. Alguns exemplos desse tipo de sistema podem ser vistos nos trabalhos de: [12], [30], [15], [16].

Devido a esta demanda o presente trabalho busca, através das técnicas de Visualização Interativa, apoiar o usuário durante a atividade de análise visual, promovendo uma melhoria na qualidade do conteúdo da informação, o que implicará na geração de novas visualizações, demonstrando a melhoria de informação obtida, refletindo diretamente na qualidade da tomada de decisão.

Para a fase de desenvolvimento da prototipação do presente trabalho foram utilizadas técnicas de visualização já implementadas com o objetivo de demonstrar o estado de qualidade do dado, para que as interações implementadas fossem capazes de modificar tais visualizações em busca da melhoria da qualidade da informação, e em consequência gerando novas visualizações.

As técnicas de interação implementadas tem o objetivo de apoiar o usuário a manipular essas visualizações para que ele seja capaz de adquirir conhecimento sobre a situação e tomar decisões efetivas para o controle do incidente.

A Tabela 3 demonstra a relação entre as técnicas de Visualização Interativa implementadas com o objetivo de aprimorar a qualidade das informações, refletindo em novas visualizações e no processo de obtenção de conhecimento do usuário, que implicará na qualidade do processo de tomada de decisão:

Tabela 3: Técnicas de Visualização Interativa implementadas

Interação	Dimensão de Qualidade	Transformação de Dados
Toque sobre representação	Incompletude	Seleção
Selecionar área de interesse	Incompletude	Seleção / Amostragem
Slide sobre o mapa	Imprecisão, Incompletude, Incerteza	Filtro
Swipe sobre o mapa	Incompletude	Navegação
Agregar representações	Incompletude, Granularidade	Agregação/desagregação

A primeira técnica de interação implementada baseia-se na seleção de uma representação de um objeto no mapa através de um evento de toque, isso faz com que o sistema exiba uma janela de informações sobre o objeto selecionado, fornecendo ao operador informações mais completas sobre o objeto representado. A Figura 9 é um exemplo dessa interação.

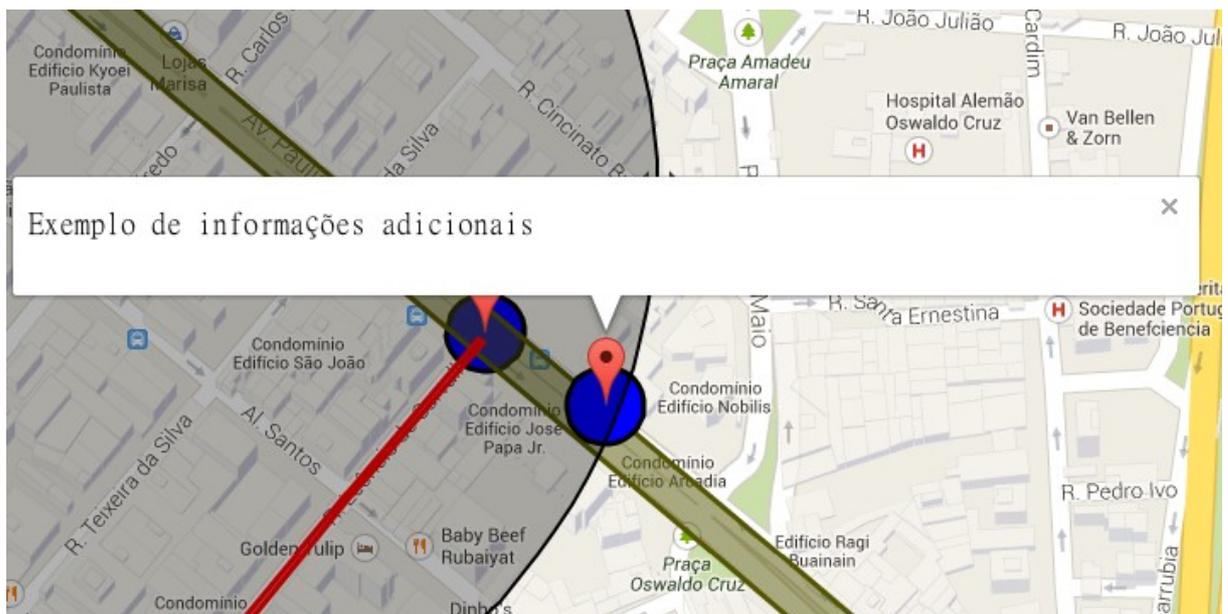


Figura 9: Janela de informações exibida pela interação de toque.

A segunda interação implementada permite que o usuário selecione uma área de interesse para obter somente informações presentes nessa região. Essa interação é feita através de um evento de toque sobre um ponto na tela, o sistema gera então uma representação de área (círculo cinza), que tem como centro, o ponto em que foi realizado o evento de toque. A transformação realizada pelo sistema nesse caso é a amostragem, onde será exibida apenas uma parcela das informações totais disponíveis no sistema. As informações da área selecionada podem ser suficientes para a execução da atividade analítica em andamento, resolvendo a questão de quão completa são as informações para a determinada tarefa. A complexidade dos dados também será reduzida, evitando que o operador se confunda com dados irrelevantes à situação em andamento. É possível também redimensionar a área de interesse para ajustar a quantidade de informação que é exibida, permitindo ao operador buscar por novas informações, caso as informações que são exibidas pelo tamanho padrão da área de interesse não sejam suficientes. Esse redimensionamento é realizado através do deslizamento do ícone ↔ na tela. A Figura 10 demonstra uma área de interesse gerada por um evento de toque no mapa.

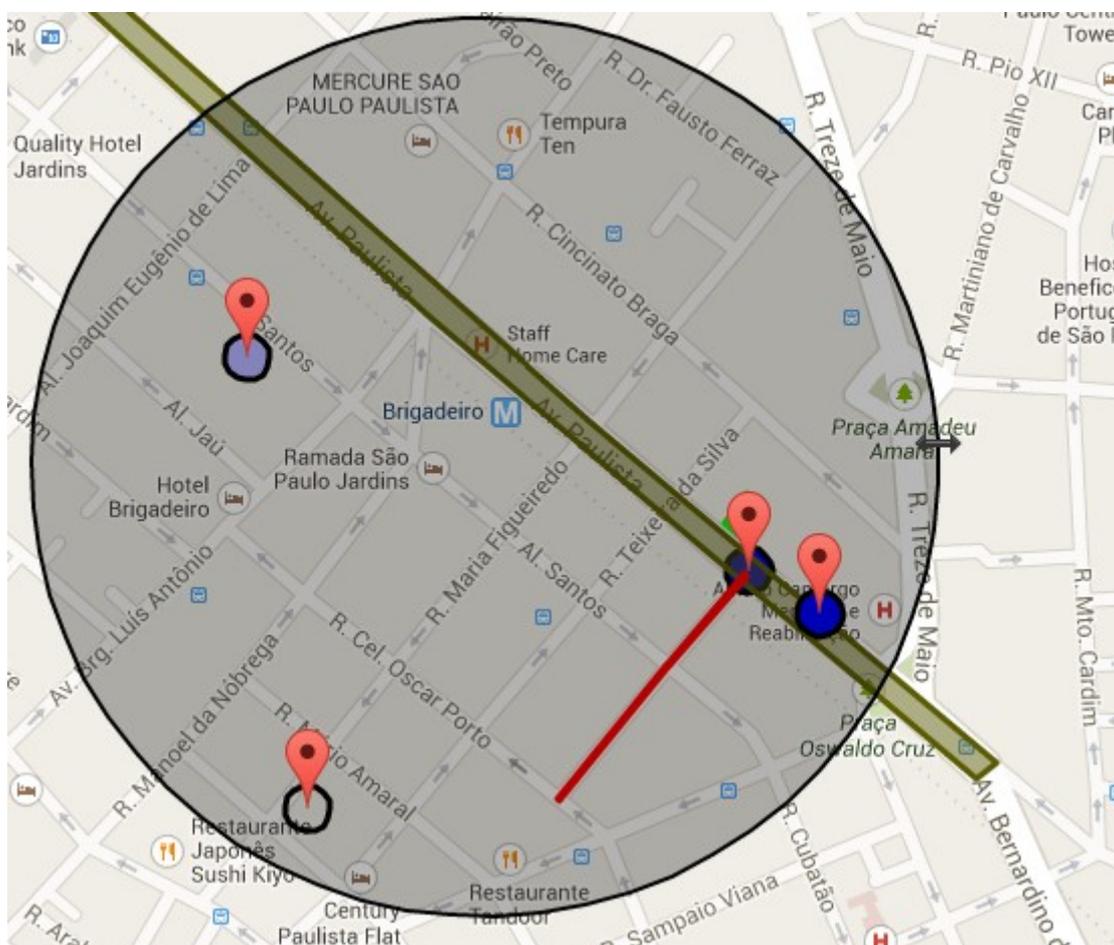


Figura 10: Seleção de área de interesse através de toque

A terceira interação permite que o operador filtre as representações de acordo com o valor de qualidade relacionado a elas, exibindo apenas dados com mais ou menos qualidade. Isso é feito através da interação de *drag* com dois dedos no canto direito da tela, simulando um *slider* na vertical. Realizando o *drag up*, o valor do filtro é incrementado o que implica na exibição das representações com mais qualidade. Realizando o *drag down*, o valor do filtro é decrementado, e as representações com menos qualidade são exibidas. A dimensão de qualidade que foi utilizada no exemplo foi a incompletude, o que implica em exibir dados mais ou menos completos de acordo com o valor do filtro.

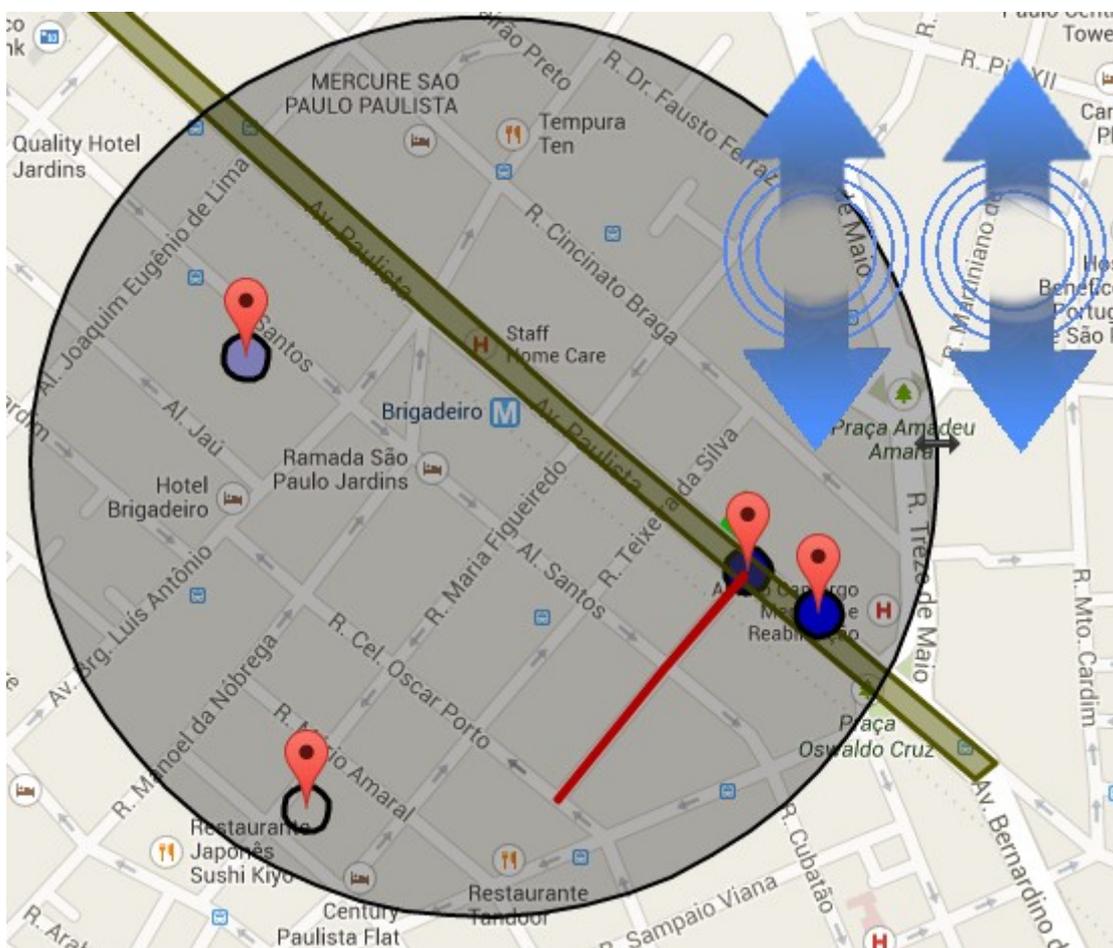


Figura 11: Filtro através da interação de drag no canto direito da tela

Através da quarta técnica implementada se baseia em técnicas de navegação, permitindo que o usuário explore o espaço de representações em busca de padrões, áreas de interesse e novas informações. Essa interação é realizada deslizando o mapa para direita,

esquerda, para cima ou para baixo. A partir dessa técnica o usuário é capaz de manipular as visualizações em busca de novas informações, tornando seu conhecimento sobre a situação em andamento mais completo.

A quinta interação baseia-se na agregação de duas ou mais representações com o objetivo de obter uma nova representação mais completa, ou mais precisa. Essa técnica também permite ao usuário alterar a granularidade dos dados, podendo criar grupos de dados através da técnica de agregação, ou repartindo grupos de dados por meio da desagregação. Essa técnica é realizada através da seleção de um objeto, ou seja, realizando um toque sobre a representação desse objeto, e arrastando-o até o outro objeto ou grupo de objetos ao qual se deseja agregá-lo, isso irá gerar uma nova visualização, ou seja, uma visualização de agrupamento, que pode ser visualizada na Figura 13. Já a desagregação é feita selecionando um grupo de objetos por meio de um toque e segurando por alguns segundos, isso irá dividir o grupo em um número menor de objetos. As Figuras 12 e 13 demonstram a realização das operações de agregação e desagregação, respectivamente.

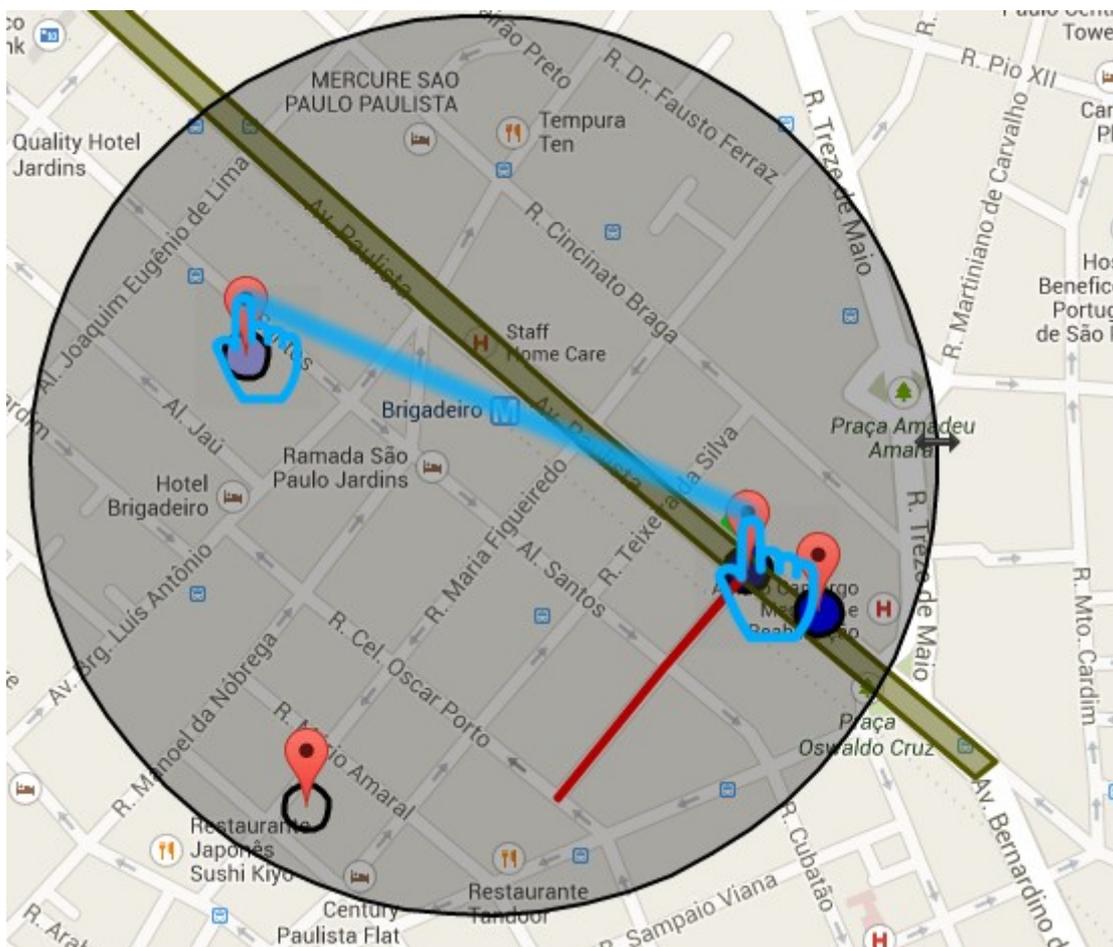


Figura 12: Realização de agregação entre dois objetos

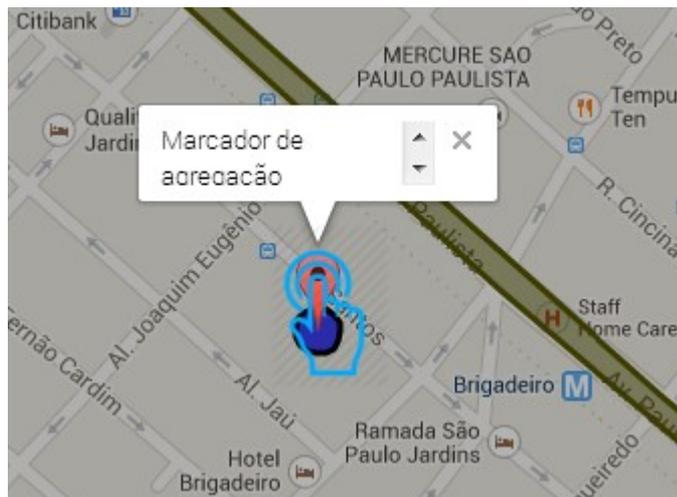


Figura 13: Desagregação de um agrupamento

3. Análise de Resultados

Para medir os resultados obtidos com as técnicas implementadas foram realizadas avaliações baseadas em conhecimento. Como citado no trabalho de Robertson et. al. [5]: “o propósito da visualização é conhecimento, não figuras”, e também: “o propósito da computação é conhecimento, não números”. Por esse motivo, para medir a eficiência das técnicas implementadas no presente trabalho, será estudado o quão eficiente são essas técnicas para a transmissão de conhecimento ao usuário, e o quanto elas colaboram para que ele tenha ciência da situação.

No capítulo seguinte serão demonstrados os resultados obtidos com a implementação das técnicas de Visualização Interativa, como elas ajudam o usuário durante as tarefas de análise visual, contribuindo para o processo de ciência da situação e tomada de decisão.

Capítulo 4 - Resultados Obtidos

Nesse capítulo será discutido o ambiente criado para a realização da avaliação dos resultados, tanto aspectos da arquitetura de hardware como de software utilizados. Também será descrito o cenário, ou seja, o contexto do estudo de caso em sistemas de Comando e Controle.

1. Componentes de Hardware

Para que o usuário fosse capaz de realizar interações com o sistema foi necessária a utilização de uma interface de interação. Para promover a experiência do usuário durante a realização da tarefa de análise visual, foi decidido utilizar um dispositivo de interface tangível, que permite uma forma de interação mais natural entre o usuário e o sistema.

A interface consiste em uma moldura de sensores infra-vermelho acoplada em uma TV. Através do protocolo TUIO, as informações de interações de toque são passadas para a aplicação.

2. Componentes de Software

Como citado na metodologia foram utilizadas bibliotecas em Javascript para o desenvolvimento da aplicação. Abaixo são descritas mais algumas informações:

- SO: Windows 7;
- Browser: Firefox;
- Bibliotecas Javascript: Google Maps, Hammer.js, Facebook API, Twitter API.

3. Estudo de caso em C2

A aplicação de C2 proposta por esse trabalho se baseia no contexto de gerenciamento de grandes eventos, como por exemplo, uma manifestação. A situação simulada nos exemplos foi o bloqueio de uma via principal utilizada para se chegar a um hospital, essa via foi bloqueada devido a uma manifestação no local. O sistema exibe a Avenida Paulista, que é o local onde a manifestação está acontecendo, na cor marrom. A rua contornada com a cor vermelha é a rua de acesso ao hospital que está bloqueada.

Através de interações do usuário o sistema exibe informações de redes sociais, como o Twitter, onde, por exemplo, uma pessoa está parada no trânsito porque estava a caminho do hospital e a rua foi bloqueada. Essa pessoa postou a informação de que a via teria sido bloqueada devido a manifestação, e através de uma cadeia de caracteres padrão, como por exemplo #HospitalSantaCatarina, #RuaLeoncioDeCarvalho, essas informações são obtidas e exibidas pelo sistema através de um ícone de marcador.

A medida da qualidade do dado, nesse caso, o quão completa está a informação sobre o bloqueio da rua é demonstrada por um círculo junto ao marcador, que varia sua opacidade de acordo com a variação da qualidade da informação, isto é, quanto maior a qualidade do dado, mais opaco é o círculo. Além disso, o valor de qualidade do dado é exibido em uma janela de informações adicionais sobre a representação do post, quando o usuário realiza uma interação de seleção de representação (item 4.1). A dimensão de qualidade do dado utilizada como teste foi a completude, que é medida pelo quão completas estão as informações dos posts referentes ao bloqueio da rua Leoncio de Carvalho, esse valor de qualidade pode variar entre 0 e 1, sendo 0 a informação menos completa, e 1 informações completa.

4. Resultados das Implementações

1. Selecionar área de interesse

Ao início da atividade de análise é fornecido ao usuário uma visão geral do mapa, e através da seleção de uma área de interesse, por meio de um evento de toque, o sistema exibe as informações na área desejada. Essas informações se remetem à localização pela qual foram enviados os tweets referentes ao bloqueio da rua de acesso ao Hospital. A partir daí o usuário é capaz de realizar novas descobertas sobre os dados exibidos.

Segundo Yi. et. al. [8] fornecer uma visualização global é o processo pelo qual o usuário entende a situação geral de um conjunto de dados. Além disso permite que ele

compreenda o que ele sabe, e o que não sabe sobre a situação, permitindo analisar quais áreas podem ser melhor investigadas e assim ganhar novos conhecimentos sobre o conjunto de dados.

A habilidade de selecionar de um conjunto de dados (área de dados, ou área de interesse) é uma técnica que apoia o usuário a explorar grandes conjuntos de dados. Um conjunto de dados completo pode conter partes não significativas para a situação sendo analisada atualmente, portanto, usuários geralmente ganham conhecimento apenas filtrando os dados desnecessários [8].

O principal benefício dessa técnica de interação é permitir que o usuário explore um conjunto de dados individualmente permitindo a descoberta novos conhecimentos sobre a situação sendo analisada. Além disso, reduz a quantidade de dados exibidos no display, permitindo que o usuário se concentre apenas nos dados relevantes a atividade analítica sendo executada no momento. A Figura 10 mostra um exemplo de seleção de uma área de interesse.

2. Seleção

A técnica de seleção implementada tem o objetivo de permitir ao operador obter informações adicionais sobre as representações. Essa técnica pode ser considerada como o próximo passo de técnicas de filtragem. Primeiramente o usuário escolhe um conjunto de dados para ser analisado (ou uma área de interesse), realizando a filtragem de dados irrelevantes e com a seleção da representação de um post (um marcador) ele obtém informações adicionais sobre o post, incrementando seu conhecimento sobre a situação. Além disso, o sistema exibe também informações sobre o valor de qualidade da informação do post, permitindo ao usuário tomar decisões mais eficientes. A demonstração dessa técnica em execução é mostrada pela Figura 14.

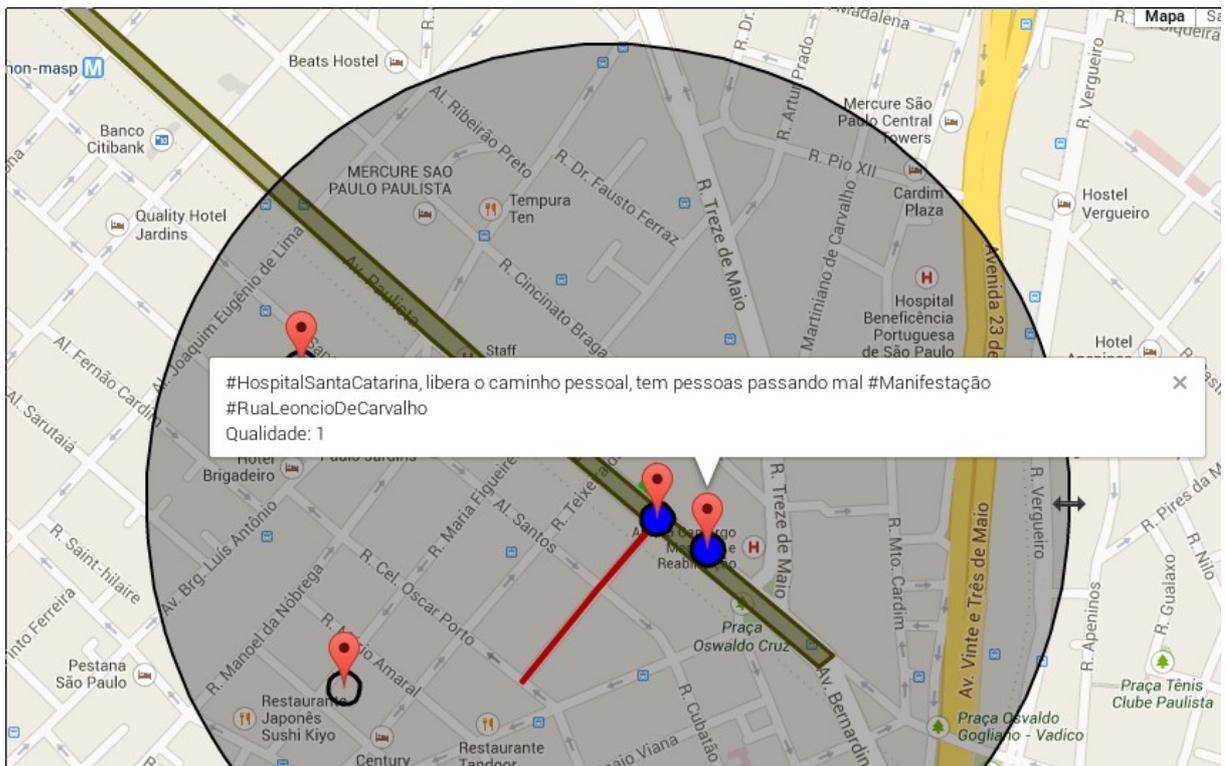


Figura 14: Interação de seleção de representações de posts

3. Filtro

Outra técnica que permite alterar o nível de detalhe de informações exibidas pelo sistema é através de filtros. A técnica de filtragem implementada permite ao usuário filtrar os posts exibidos pelo valor de qualidade associado ao dado, isto é, o usuário pode escolher visualizar posts com informações mais, ou menos completas. Muitas vezes, somente a exibição de dados completos podem não ser suficientes para o usuário obter total ciência da situação, por isso é importante que ele seja capaz de escolher quais informações deseja visualizar.

Através da visualização da qualidade do dado a partir da representação do círculo ou valor da janela de informações, o usuário obtém ciência da qualidade do dado, evitando que tome decisões incorretas sobre a situação, resultando em uma tomada de decisão ineficiente.

A filtragem dos dados é realizada através de interações de drag up/down com dois dedos, ou seja, deslizar dois dedos para cima ou para baixo, no canto direito da tela (Figura 11). Essa medida foi tomada para evitar que a aplicação confundisse a interação de filtragem com uma interação de navegação, que é feita com apenas um dedo.

Analisando a figura 10, é possível notar quatro representações de posts, um com o círculo representando a completude do post com qualidade 0 (círculo totalmente transparente), dois posts com qualidade 0 (círculos totalmente opacos na cor azul), e um post com o valor de completude 0.25 (círculo parcialmente transparente, cor roxa). Através de uma interação de drag up, o valor do filtro vai sendo incrementado até 0.1, o que provoca a filtragem do post com menor valor de completude, ou seja, o post com o círculo totalmente transparente, representando completude 0. A Figura 15 demonstra o estado da aplicação depois da aplicação do filtro.

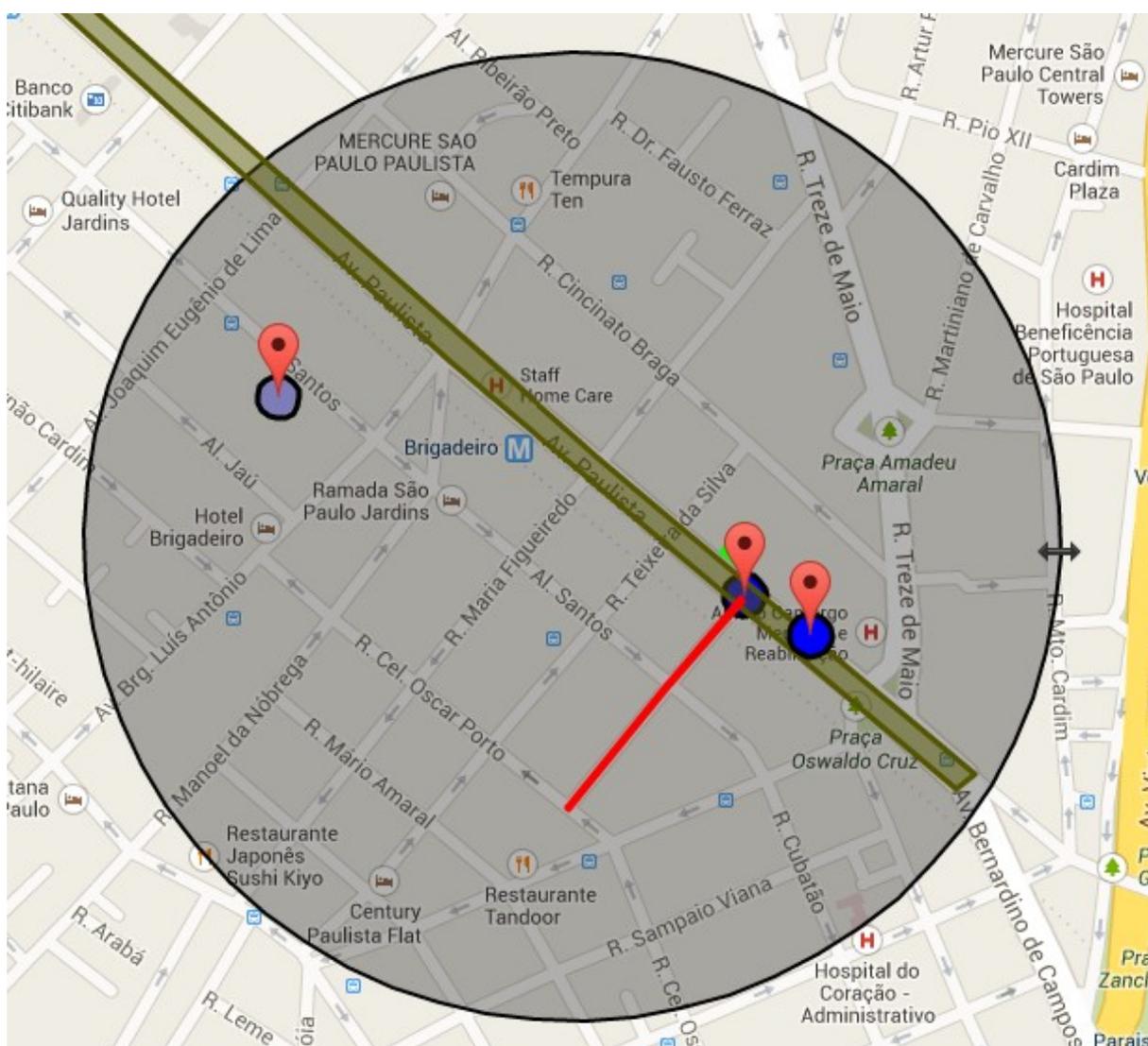


Figura 15: Interação de filtragem, valor 0,1.

Se realizarmos outra interação de drag up para incrementar o valor do filtro para um valor maior que 0.25, somente os posts mais completos, ou seja, os posts representados

pelos círculos azuis opacos, com valor de completude 1, serão exibidos ao usuário, como pode ser notado na Figura 16.

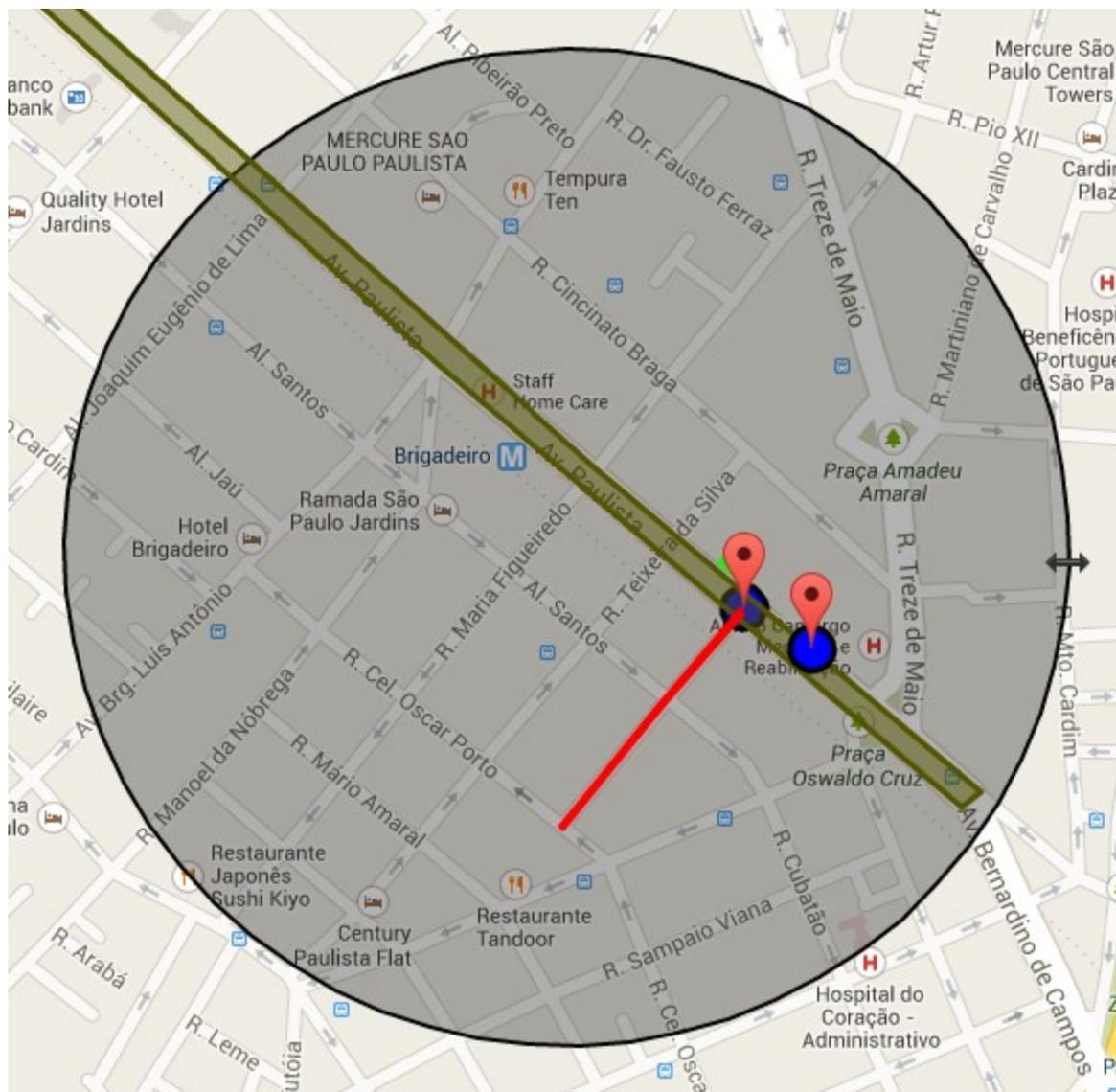


Figura 16: Interação de filtragem com valor de completude maior que 0,25.

Todos os passos realizados no processo de filtragem dos dados são demonstrados na Figura 17.

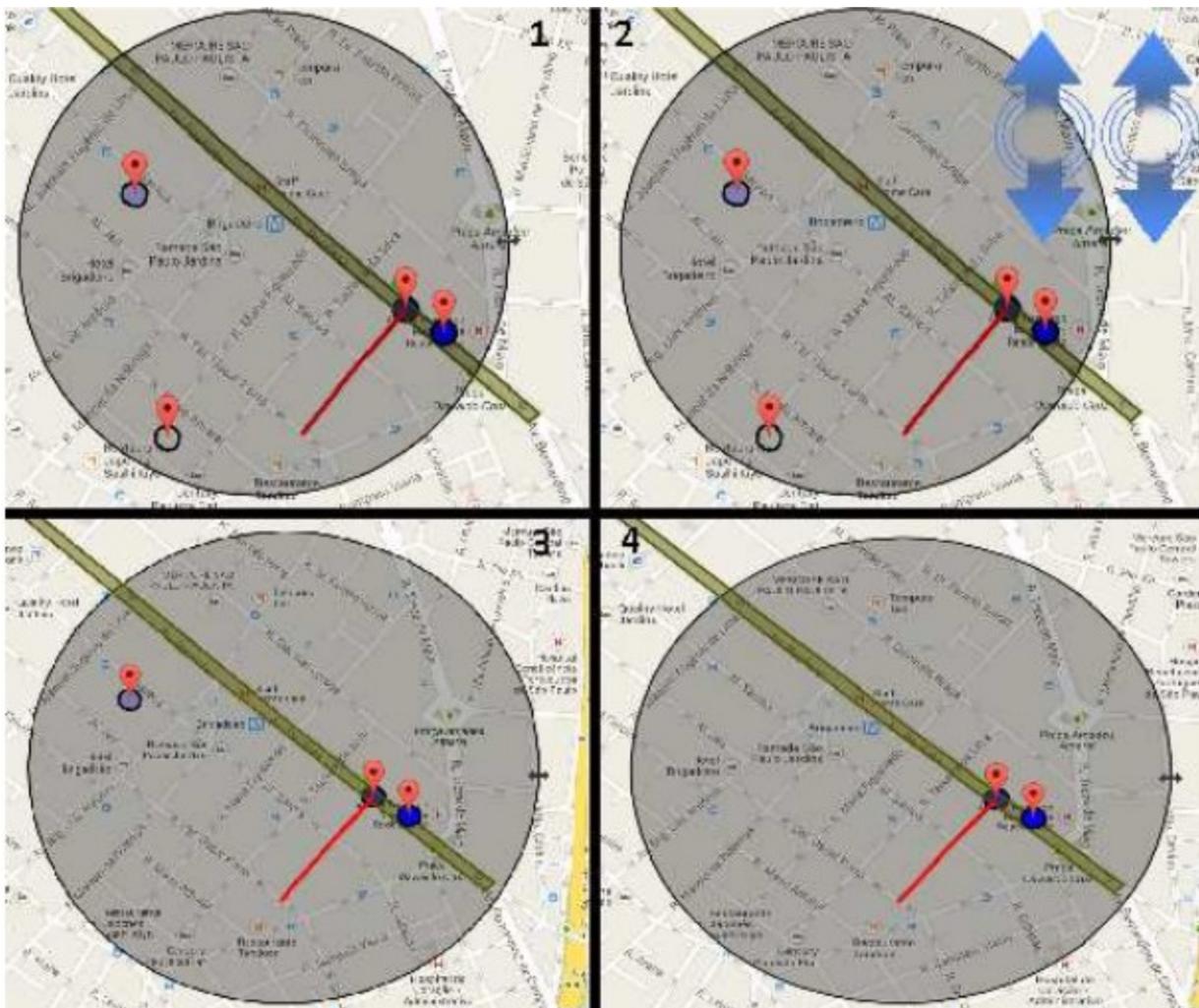


Figura 17: Passos realizados na interação de filtragem

4. Navegação

Através das técnicas de navegação implementadas o usuário é capaz de navegar sobre o mapa geográfico para descobrir novas áreas de conhecimento, podendo detectar padrões e destaques nas informações dos posts, que podem ser úteis para a resolução de uma situação.

A navegação sobre o mapa é realizada através de interações de swipe sobre o mapa, isto é, realizando uma interação de toque com um dedo no mapa e deslizando para esquerda, direita, para cima ou para baixo.

5. Agregação

Yi et. al. [8] cita a agregação como outra forma efetiva de se explorar um grande conjunto de dados, abstraíndo-o em pedaços menores. É difícil obter conhecimento

significativo através de uma grande quantidade de dados, através de processos de agrupamento e agregação, informações relevantes são agrupadas e organizadas, reduzindo o esforço de procura e de memória do usuário, além de permitir que ele descubra fatos de alto nível e procure por novos conhecimentos.

A técnica de agregação implementada consiste no agrupamento de um ou mais posts, gerando uma nova visualização, diminuindo conseqüentemente a quantidade de posts exibidos. Se o usuário agregar dois posts com valor de qualidade 0,25, e a informação de um complementa a do outro, a nova representação gerada irá demonstrar a melhoria de qualidade gerada, ou seja, a nova representação terá valor de qualidade 0,50.

A agregação pode ser realizada executando os passos descritos na Figura 12. As Figuras 18 e 19 mostram o estado da aplicação e dos posts antes da execução da interação de agregação.

Nota-se que existe um post com completude de valor 0.5, onde são mencionadas informações sobre a liberação da rua e que ambulâncias estão tentando passar, mas o post não se refere à manifestação ou o hospital. A qualidade desse post é demonstrada através do círculo abaixo do marcador, metade transparente na cor azul (Figura 18).

A Figura 19 mostra outro post, com completude de valor 0.25, onde as informações contidas no posts se referem à manifestação na avenida. A transparência do círculo desse está mais notável do que o post descrito anteriormente.

Através da agregação desses dois posts, o usuário obtém informações mais completas sobre a situação, além de reduzir a quantidade de dados apresentados pelos sistema, permitindo-o concentrar-se na análise de outros posts. Após a realização da agregação desses dois posts é gerada uma nova representação, como é mostrado na Figura 20, a qualidade da nova representação também é incrementada, como se pode notar pela cor e opacidade do círculo, e também pela informação de qualidade na janela de informações adicionais.



Figura 18: Informações sobre o primeiro post.

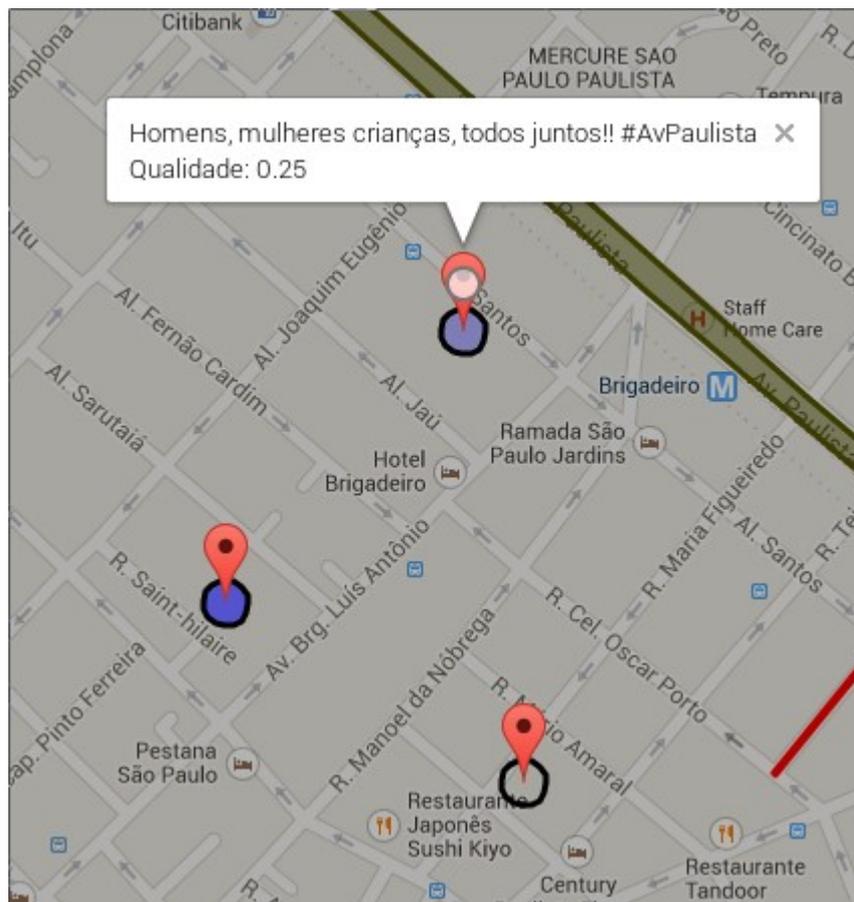


Figura 19: Informações sobre o segundo post.

	exibida.
Swipe sobre o mapa	Permite a navegação e descoberta de novas informações
Agregar representações	Diminui a quantidade de informações exibidas; melhora a qualidade das informações quando duas informações que remetem a um mesmo evento são agrupadas.

No capítulo seguinte serão discutidas as conclusões obtidas com a implementação das técnicas de Visualização Interativa, os resultados obtidos através delas, qual a contribuição do trabalho para a experiência do usuário durante o processo de tomada de decisão, e também possíveis trabalhos futuros.

Conclusões

Com a conclusão do desenvolvimento do trabalho conclui-se que as técnicas de Visualização Interativa implementadas atingiram o objetivo do trabalho, que é apoiar o usuário na obtenção de conhecimento sobre a situação apresentada pelo sistema por meio de interações, e também torná-lo capaz de contribuir com a qualidade das informações por meio dessas interações. As técnicas implementadas permitiram ao usuário interagir com as visualizações, obtendo informações adicionais sobre elas, também permitiram a manipulação dessas informações através de filtros e agregações.

As técnicas de Visualização Interativa implementadas permitiram ao usuário realizar transformações no espaço de representações permitindo a descoberta de conhecimentos e melhoria das informações apresentadas, por meio de filtragens, agregações e seleções. Porém, ainda é necessário explorar outras possibilidades de utilização de técnicas de Visualização Interativa para a resolução de outros problemas, tanto relacionados a maneira em que o usuário explora os dados, quanto para a resolução de problemas de qualidade de dados.

Referências Bibliográficas

- [1] White, Franklin E. **Data Fusion Lexicon**. The Data Fusion Subpanel of the Joint Directors of Laboratories, Technical Panel for C3. Naval Ocean System Center, San Diego. 1991.
- [2] Sedig, K. Parsons, P. Babanski, A. **Towards a Characterization of Interactivity in Visual Analytics**. Journal of Multimedia Processing and Technologies, Special Issue on Theory and Application of Visual Analytics, (2012), 3(1), 12-28.
- [3] B. Khaleghi et al., **Multisensor data fusion: A review of the state-of-the-art**, Elsevier. 2011.
- [4] Wickramaratne, T. L. **An Analytical Framework for Soft and Hard Data Fusion: A Dempster-Shafer Belief Theoretic Approach**. University of Miami. 2012.
- [5] Robertson, B. Czerwinski, M. Fisher, D. Lee, B. **Selected Human Factors Issues in Information Visualization**. Reviews of Human Factors and Ergonomics. 2009.
- [6] Steinberg, Alan N. Bowman, Christopher L. White, Franklin E. **Revisions to the JDL Data Fusion Model**. 1999.
- [7] Côrtes, Sérgio C. Porcaro, Rosa M. Lifschitz, S. **Mineração de Dados – Funcionalidades, Técnicas e Abordagens**. PUC-RioInf.MCC10/02. 2002.
- [8] Ji Soo Yi, Youn-ah Kang & John T. Stasko, Julie A. Jacko. **Understanding and Characterizing Insights: How Do People Gain Insights Using Information Visualization?** Florence, Italy. ACM. 2008.
- [9] Keim, Daniel A. **Information Visualization and Visual Data Mining**. IEEE Transaction On Visualization And Computer Graphics, Vol. 7, No. 1, 2002.
- [10] Goldstein, J. Roth, Steven F. **Using Aggregation and Dynamic Queries for Exploring Large Data Sets**. In Proceedings CHI'94 Human Factors in Computing Systems. ACM. 1994.
- [11] Thomas, James J. Cook, Kristin A. **Visualization Viewpoints – A Visual Analytics Agenda**. IEEE Computer Society. 2006.
- [12] Summers, Valerie A. Jones, Richard L. Flo, Robert. **Increasing Situational Awareness by Visualizing Uncertainty**. 10th International Command & Control Research & Technology Symposium. McLean VA. 2005.

- [13] Bakopoulos, Menelaos. Tsekeridou, Sofia. Giannaka, Eri. Tan, Zheng-Hua. Prasad, Ramjee. **Command and Control: Information Merging, Selective Visualization and Decision Support for Emergency Handling**. Proceedings of the 8th International ISCRAM Conference. Lisbon, Portugal. 2011.
- [14] Alan M. MacEachren , Anthony Robinson , Susan Hopper , Steven Gardner , Robert Murray , Mark Gahegan & Elisabeth Hetzler. **Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know**. Cartography and Geographic Information Science, 32:3, 139-160. 2013.
- [15] Hodick, Jan. Frantis, Petr. **Command and Control Systems and Data Processing**. Cybernetic Letters, Special Issue. 2009.
- [16] Gouin, Denis. Evdokiou, Peter. Vernik, Rudi. **A Showcase of Visualization Approaches for Military Decision Makers**. Massive Military Data Fusion and Visualisation: Users Talk with Developers, Halden, Norway. RTO-MP-105. 2002.
- [17] Hall, David L. Steinberg, Alan. **Dirty Secrets in Multisensor Data Fusion**. Pennsylvania State University Applied Research Laboratory. 2001.
- [18] Tang, Zhili. Zhang, An. **Human-Machine Cooperation for Control and Decision in Information Fusion System**. Chinese Control and Decision Conference. IEEE. 2008.
- [19] Hall, David L. Llinas, James. **An Introduction to Multisensor Data Fusion**. Proceedings of the IEEE, Vol. 85, No. 1, 1997.
- [20] Kerren, Andreas. Schreiber, Falk. **Towards The Role of Interaction in Visual Analytics**. Proceedings of the 2012 Winter Simulation Conference. IEEE. 2012.
- [21] Keim, Daniel A. Mansmann, Florian. Schneidewind, Jörn. Ziegler, Hartmut. **Challenges in Visual Data Analysis**. Proceedings of the Information Visualization (IV'06). IEEE. 2006.
- [22] Scholtz, Jean. **Beyond Usability: Evaluation Aspects of Visual Analytic Environments**. IEEE Symposium on Visual Analytics Science and Technology. Baltimore, MD, USA. 2006.
- [23] Correa, Carlos D. Silver, Deborah. Chen, Min. **Illustrative Deformation for Data Exploration**. IEEE Transaction on Visualization and Computer Graphics, Vol. 13, No. 6. 2007.
- [24] Wen, Zhen. Zhou, Michelle X. **Evaluation the Use of Data Transformation for Information Visualization**. IEEE Transactions on Visualization and Computer Graphics, Vol. 14, No. 6, 2008.

- [25] Jonker, David. Langevin, Scott. Bozowsky, Neil. Wright, William. **Aperture: An Open Web 2.0 Visualization Framework**. 46th Hawaii International Conference on System Sciences. IEEE. 2012.
- [26] Lalanne, Denis. Kohlas, Jürg. **Human Machine Interaction**. Springer. 2009.
- [27] Zhang, Qingyu. Segall, Richard. Cao, Mei. **Visual Analytics and Interactive Technologies**. Information Science Reference. 2011.
- [28] Chi, Ed H. **A Taxonomy of Visualization Techniques using the Data State Reference Model**. IEEE Symposium on Information Visualization (InfoVis'00). 2000.
- [29] Association of Chief Police Officers e National Policing Improvement Agency. **Guidance on Command and Control**. Reino Unido. 2009.
- [30] Hodicky, Jan. Frantis, Petr. **Low cost solution to control military command and control system in virtual environment**. International Conference on Computer Technology and Science. IACSIT Press, Singapore. 2012.
- [31] Heer, Jeffrey. Schneiderman, Ben. **Interactive Dynamics for Visual Analysis**. ACM. 2012.
- [32] Smithson, M. **Ignorance and uncertainty: Emerging paradigms**. New York: Springer Verlag, p.9. 1989.
- [33] Bouchon-Meunier, B. Nguyen, H. T. **Les incertitudes dans les systemes intelligents**. Que sais-je? Volume 3110. França. Paris. 1996.
- [34] Krause, P. J. Clark, D. A. **Representing Uncertainty Knowledge: An Artificial Intelligence Approach**. Intellect. Oxford. UK. 1996.
- [35] Smets, P. **Imperfect Information: Imprecision and Uncertainty**. Springer. US. 1997.
- [36] Batini, C. Scannapieca, M. **Data Quality – Concepts, Methods and Techniques**. Springer. 2006.
- [37] Card, S., J. Mackinlay, and B. Shneiderman. **Readings in Information Visualization: Using Vision to Think**. Morgan Kaufmann. 1999.
- [38] Howard, D., and A. M. MacEachren. **Interface design for geographic visualization: Tools for representing reliability**. Cartography and Geographic Information Systems 23(2): 59-77. 1996.
- [39] Hengl T. **Visualisation of uncertainty using the HSI colour model: computations with colours**. 7th International Conference on GeoComputation (CD-ROM), pp. 8. 2003.

[40] Cliburn, D. C., J. J. Feddema, J. R. Miller, and T. A. Slocum. **Design and evaluation of a decision support system in a water balance application.** Computers & Graphics (UK) 26(6): 931-49. 2002.