

# **A FORMAÇÃO DE *CLUSTERS* NA LOGÍSTICA HUMANITÁRIA UTILIZANDO MINERAÇÃO DE DADOS.**

**Fabiana Santos Lima, Msc.**

**Daniel de Oliveira, Msc.**

**Mirian Buss Gonçalves, Dra.**

Programa de Pós Graduação em Engenharia de Produção.

Universidade Federal de Santa Catarina.

## **RESUMO**

Estudos feitos por organizações humanitárias internacionais tais como IASC (*Inter-Agency Standing Committee*), OCHA (*Office for the Coordination of Humanitarian Affairs*), WFP (*World food programme*), constata a necessidade de melhorar a eficácia da resposta humanitária, garantindo uma maior previsibilidade, responsabilização e parceria. Estes estudos referenciam a abordagem *clusters* com o objetivo de atender a mais beneficiários. Este trabalho tem como objetivo apresentar uma metodologia para formar *clusters* por combinação de frequência de desastre na região de Santa Catarina, utilizando mineração de dados. Pode-se assim auxiliar na coordenação estratégica, na definição de prioridades e no compartilhamento de experiências nas regiões. A proposta é verificar a frequência por tipo de desastre em determinada região e identificar regiões semelhantes para então, poder padronizar um método na prevenção e melhorar a tomada de decisão na logística humanitária. Para cada um dos *clusters*, as experiências e conhecimentos podem ser compartilhados, o que poderá servir de base para o planejamento integrado, visando à preparação e os processos logísticos na iminência dos desastres.

## **ABSTRACT**

Researches made by humanitarian organizations such as IASC (*Inter-Agency Standing Committee*), OCHA (*Office for the Coordination of Humanitarian Affairs*) and WFP (*World food programme*) show the necessity of a development in the efficacy of the humanitarian action that ensures a better forecast, accountability and partnership. These studies have used the *clusters* methodology as a reference to better assist in humanitarian action. The present study has as an objective to present a methodology to create *clusters* by using the disasters frequency combination in the Santa Catarina State, using data mining. In this way, the study can assist in the coordination strategy, in the definition of priorities and then share experience among the regions within the State. The purpose of this study is to verify the frequency of different types of disasters and then identify areas with similar cases to standardize the prevention method and improve the decision taking in the humanitarian logistics. This experience and knowledge can be shared among the *cluster*, being a base for the integrated planning of the humanitarian action in Santa Catarina, aiming for increased readiness of the logistics processes in the event of a disaster.

## **1. INTRODUÇÃO**

Organizações de ajuda humanitária coordenam bilhões de dólares anualmente no alívio a vítimas de catástrofes naturais, conflitos civis e guerras. Sua tarefa principal é a mobilização oportuna de financiamento e de bens de doadores internacionais e administração de alívio para os beneficiários vulneráveis em locais de desastres em todo o globo. Como tal, a logística é fundamental para suas atividades e estratégica para as suas missões.

Segundo pesquisa apresentada por Sapir (2011) no ano de 2010 o número total de desastres no mundo foi de 373 desastres naturais, com 296.800 mortes e um total de 207 milhões de pessoas afetadas, com um prejuízo estimado em 109 bilhões de dólares. Comparando o número de desastres naturais em 2010 tais como secas, terremotos, altas temperaturas, enchentes, movimentos de terra seca, movimentos de terra molhada, tempestades, vulcões e incêndios com a média da década de 2000 a 2009 constata-se que não houve crescimento quanto ao número de desastres naturais no mundo. Por outro lado, analisando um estudo mais amplo representativo de mortes, verifica-se um crescimento representativo nos últimos dez anos (CRED, 2011).

Segundo CREA (2011) o Brasil, em 2008, foi o 13º país mais afetado por desastres naturais, onde dois milhões de pessoas foram atingidas principalmente pela ação das chuvas e 1,5

milhão de pessoas sofreram com as catástrofes em Santa Catarina. Conforme estatísticas da Defesa Civil do Estado de Santa Catarina, somente nos meses de janeiro até março de 2011 foram registrados 152 eventos, como alagamentos, erosões, pragas, ressacas, vendavais e enxurradas sendo este último o evento predominante, com 143 ocorrências.

Em todos estes casos o atendimento aos desastres foi difícil e complicado. Muitos fatores, como o grande número de atores envolvidos na ajuda humanitária e a falta de recursos suficientes, contribuíram para as dificuldades de coordenação de socorro (Fenton, 2003 *apud* Balcik *et al*, 2010). Existem poucas histórias de sucesso de coordenação, de modo que “a coordenação continua a ser a fraqueza fundamental da ação humanitária” (Rey, 2001 *apud* Balcik, 2010).

O termo coordenação tem interpretações variadas dentro do ambiente de alívio ao desastre. Por exemplo, pode referir-se à coordenação de recursos e partilha de informação, centralização das decisões, a realização de projetos comuns, a divisão regional de tarefas, ou um sistema baseado em *clusters*, em que cada *clusters* representa uma área de setor diferente (por exemplo: alimentos, água, saneamento e tecnologia da informação). Trabalha-se também na coordenação de *clusters* por região, que permite analisar a frequência e o alcance geográfico dos eventos, resultando num perfil de classificação dos desastres (Almeida e Zarate, 2007; Souza, 2004).

Diante deste contexto, este artigo visa contribuir para a Logística Humanitária através de uma metodologia voltada a formar *clusters* por combinação de frequência de desastre com o intuito de auxiliar na coordenação estratégica, na definição de prioridades e no compartilhamento de experiências em determinadas regiões. Verificar a frequência por tipo de catástrofe, possibilita identificar regiões semelhantes para então poder padronizar um método na prevenção e resposta que pode assim, auxiliar os órgãos operacionais competentes na tomada de decisão referente ao tipo de capacitação necessária e investimentos em infra-estrutura. Esta metodologia será desenvolvida utilizando a técnica de Mineração de Dados (MD).

O artigo inicia-se apresentando alguns modelos associados à gestão em desastres. Na sequência, serão abordados os conceitos e características de um sistema *clusters* voltado à logística humanitária. Em seguida, contextualiza-se e referencia-se Mineração de Dados e destaca-se seu principal objetivo e características. Por fim apresenta-se a metodologia proposta com uma aplicação para o estado de Santa Catarina. Nos itens seguintes serão relatadas as considerações finais e bibliografia.

## **2. GESTÃO EM DESASTRES**

A gestão de desastres pode ser definida como o conjunto de atividades destinadas a manter o controle sobre as situações de desastre e emergência e fornecer uma estrutura para ajudar pessoas em situação de risco, para evitar ou recuperar os efeitos do desastre. A gestão de desastres trata de situações que ocorrem antes, durante e depois do desastre.

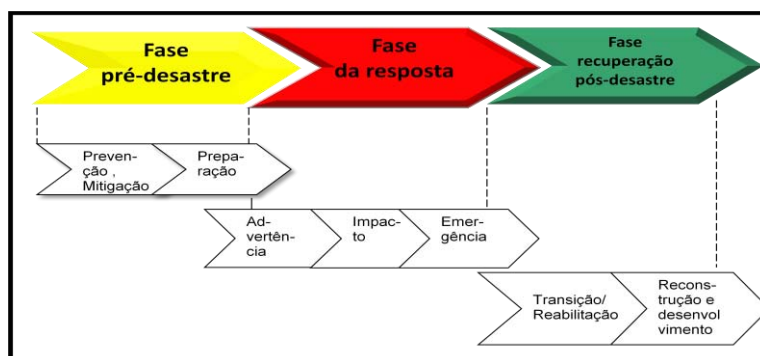
Os objetivos da gestão de desastres são: evitar ou reduzir as perdas humanas, físicas e econômicas sofridas pelas pessoas, pela sociedade e pelo país em geral; reduzir o sofrimento das pessoas; acelerar a recuperação da área afetada (Schultz, 2008).

### **2.1 Fases e Ciclos de alívio de desastres.**

Pesquisas referentes à gestão de desastres estão sendo desenvolvidas com o intuito de diminuir as dificuldades encontradas na coordenação dos diversos tipos de catástrofes.

Consideram-se três fases, ilustradas na Figura 1, que tradicionalmente envolvem a gestão em situações emergenciais:

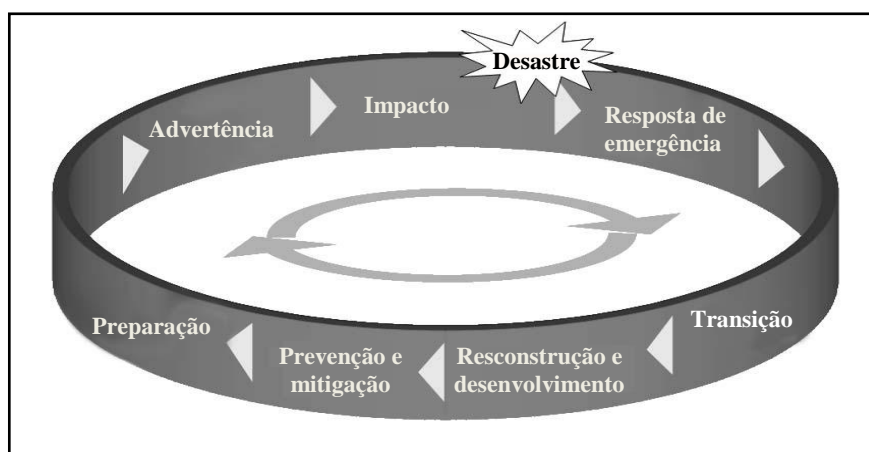
1. Pré-desastre (prevenção, mitigação e preparação);
2. Resposta (advertência, impacto e resposta de emergência);
3. Pós-desastre (transição, reabilitação e reconstrução).



**Figura 1:** Modelo das Três Fases (Adaptado de Tufinkgi,2006).

O desempenho de qualquer resposta ao desastre depende do nível de preparação, ou seja, da fase de pré-desastre. Segundo um modelo de simulação criado pelo *Massachusetts Institute of Technology (MIT)*, as pequenas despesas relativas à preparação para as catástrofes diminuí significativamente o tempo e o custo da logística de resposta (Gralla,2007 *apud* Schultz, 2008).

Uma vez que uma melhor prevenção e mitigação dos desastres podem e devem fazer parte da etapa de reconstrução e desenvolvimento, o Modelo das Três Fases pode ser convertido em um ciclo contínuo de gestão de desastres, como mostrado na Figura 2.

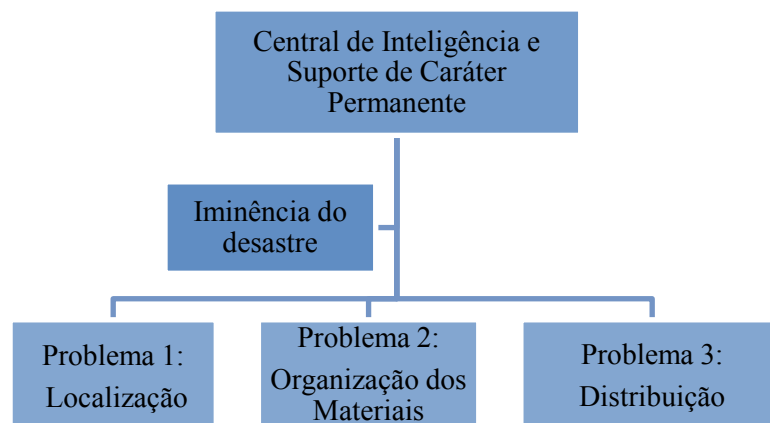


**Figura 2:** Ciclo Contínuo de Gestão de Desastres (Adaptado de Schultz, 2008).

Segundo Nogueira e Gonçalves (2010) aspectos ligados à infra-estrutura, localização de centrais de assistência, distribuição de recursos, coordenação de processos (pessoas, suprimentos, informações, materiais) merecem destaque nos processos logísticos sistematizados na logística humanitária.

Ainda para as autoras, no momento do desastre ou da iminência do mesmo, precisa-se considerar: onde localizar uma (ou várias) central (ais) de recebimento, controle e distribuição de materiais; como organizar e controlar o recebimento e a distribuição destes materiais;

como fazer a distribuição na rede. Assim é proposta a implantação de uma central de inteligência e suporte, de caráter permanente, que atue como indutora de políticas de preparação para situações de desastre. A Figura 3 mostra a metodologia proposta pelas autoras.



**Figura 3:** Principais problemas na iminência do desastre (Adaptado de Nogueira e Gonçalves,2010) .

Cabe ressaltar que o proposto neste trabalho está relacionado com o momento pré-desastre podendo ser uma das políticas de preparação para a situação de desastres da Central de Inteligência de Suporte de Caráter Permanente.

Trabalhos como este podem contribuir na iminência do desastre e também auxiliar na contribuição dos três problemas.

Considera-se que um mesmo desastre pode atingir diferentes municípios ao mesmo tempo, exigindo a coordenação entre eles para o desenvolvimento de ações de socorro ou auxílio. Além disso, o impacto de um desastre sobre um município específico pode exigir o auxílio de outro município, seja na forma de unidades de resgate, seja em empréstimo de equipamentos, assessoria de especialistas ou recolhimento de donativos. As operações nessas situações podem ter peculiaridades muito específicas. São características que podem tornar a coordenação dessas operações muito difícil se não houver um sistema padronizado, comum para o planejamento e implementação das ações.

Referencia-se neste artigo o trabalho desenvolvido por Tufinkgi (2006) que descreve processos de modelagem logística em desastres utilizando um Modelo de Processo Logístico de Socorro em Catástrofes Internacionais onde considera o gerenciamento logístico de desastres sob dois focos: de um lado, não específico (genérico), e de outro lado específico, considerando suas interdependências e referenciando a interação dos dois lados (genérico e específico). O Modelo proposto por Tufinkgi (2006) é definido em sete etapas, onde as quatro primeiras etapas estão relacionadas a casos genéricos e são pré requisitos para as três etapas restantes que estão relacionadas a casos específicos de desastres.

Este artigo volta-se para a região de *clusterização* e análise de risco proposto no primeiro item do Modelo de Processo Logístico de Socorro em Catástrofes Internacionais, o qual consiste de três funções:

- a) Estabelecer o perfil da extensão de cada desastre tomando como base a análise de dados obtidos a partir de catástrofes no passado. O perfil pode ser identificado através de uma análise histórica;
- b) Identificar regiões ameaçadas e principais ameaças, resultando em diferentes cenários com relação ao impacto dos diferentes tipos de desastres, os recursos e as necessidades previstas;
- c) A combinação das informações obtidas a partir do desempenho das duas primeiras funções permite criar perfis de necessidades de recursos, dependendo da região, bem como do tipo de catástrofe. Diferentes regiões com níveis similares de critérios serão reunidas para formar *clusters*.

Quando se trabalha com um sistema de *clusters* pretende-se auxiliar na coordenação estratégica, na definição de prioridades e no compartilhamento de experiências. Uma abordagem comum é a análise de risco, que leva a vantagem de que cada região ou município contribui com seu conhecimento e experiência para que, finalmente, um entendimento comum para o planejamento (conjunto) possa ser aplicado em várias regiões.

### **3. CLUSTERS NA LOGÍSTICA HUMANITÁRIA**

Crises humanitárias exigem uma rápida tomada de decisão. Perante a necessidade de agilidade decisória as organizações humanitárias não podem perder tempo com duplicidade de trabalhos e informações desencontradas.

Devido aos grandes problemas para fornecer cobertura suficiente nas operações de resposta no atendimento a grandes catástrofes como, por exemplo, o tsunami asiático em 2004 e a crise no Darfur em 2004/2005, os *clusters* começaram a ser utilizados como ferramenta estratégica na resposta ao desastre (Jahre e Jensen, 2010).

No fim de 2006, foi aprovado pelo Comitê Permanente Interagências (IASC – *Inter-Agency Standing Committee*), principal agência humanitária ligada a ONU, que todos os países com coordenadores humanitários implementarão a abordagem por *clusters* como ferramenta estratégica (ONERESPONSE, 2011).

Os *clusters* foram introduzidos na logística humanitária para melhorar a eficiência nas seguintes áreas: capacidade global suficiente para atender emergências atuais e futuras; liderança previsível a nível global e local; reforço de parcerias entre os organismos das Nações Unidas, ONGs e outras organizações internacionais; responsabilidade, tanto para a resposta como para a prestação de conta aos beneficiários; coordenação estratégica e a definição de prioridade em determinadas áreas de resposta, deixando a responsabilidade pela liderança e coordenação destas questões com os órgãos operacionais competentes (OCHA, 2007).

Considerando o conceito clássico definido por Porter (1998), *clusters* estão sendo utilizados para a concentração geográfica de indústrias do mesmo ramo e seus fornecedores. Incluem, por exemplo, fornecedores de materiais especializados, tais como componentes de máquinas e serviços e fornecedores de infra-estrutura especializada. Alguns *clusters* incluem instituições governamentais e outras instituições tais como universidades, instituições de formação profissional e associações comerciais que fornecem treinamento especializado, educação, informação, pesquisa e apoio técnico. Estas concentrações beneficiam cada membro, os quais passam a ter uma maior dimensão, permitindo assim trabalhar em conjunto sem que a sua flexibilidade seja afetada.

O conceito de *clusters* na logística humanitária vem sendo aplicado na organização da ajuda humanitária de acordo com um número de setores, com uma direção predefinida. Isto não significa que todos os *clusters* são ativos em todas as operações, pois podem ser mobilizados separadamente. No entanto, esses grupos que estão mobilizados devem cooperar a fim de servir melhor os beneficiários (Jahre e Jensen, 2010).

Segundo o Projeto de Avaliação e Classificação de Emergência (*Assessment and Classification of Emergencies -ACE project*) foram utilizadas ferramentas de avaliação humanitária para definir onze *Clusters* Globais que operam em 25 países: agricultura, coordenação e gestão, recuperação, educação, abrigo de emergência, telecomunicações de emergências, saúde, logística, nutrição, proteção, água e saneamento (ONERESPONSE, 2011a).

Este artigo visa a contribuir para a Logística Humanitária através de uma metodologia voltada a formar *clusters* por combinação de frequência de desastre na região de Santa Catarina, utilizando para tal, a técnica de mineração de dados. Com isso, pode-se verificar a frequência por tipo de desastre em determinada região e identificar regiões semelhantes, para então, poder padronizar um método na prevenção.

#### **4. MINERAÇÃO DE DADOS**

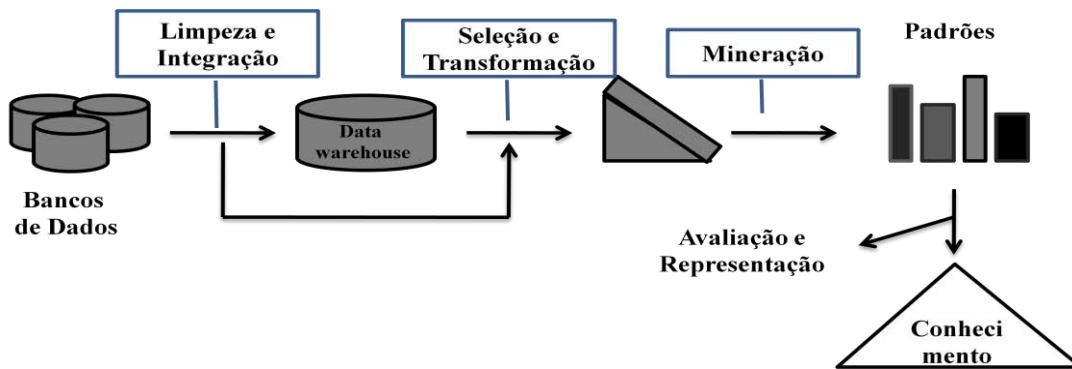
Segundo Fayyad (1996) Mineração de Dados (MD) é um processo de extração do conhecimento de Bases de Dados que tem o objetivo de encontrar o conhecimento a partir de um conjunto de dados para ser utilizado no processo decisório. A MD é uma atividade multidisciplinar e engloba o desenvolvimento de modelos aplicáveis a fenômenos em engenharia nas diversas escalas: espacial, temporal e multivariada.

A MD é considerada como sinônimo de *Knowledge Discovery in Databases* (KDD) ou Descoberta de Conhecimento em Banco de Dados. No entanto, o KDD é um processo mais amplo, onde a MD está inserida, e consiste das seguintes etapas as quais estão ilustradas na figura 4 (Souza,2004) :

- a) Limpeza dos dados: etapa onde são eliminados ruídos e dados inconsistentes;
- b) Integração dos dados: etapa onde diferentes fontes de dados podem ser combinadas produzindo um único repositório de dados;
- c) Seleção: etapa onde são selecionados os atributos que interessam ao usuário;
- d) Transformação dos dados: etapa onde os dados são transformados num formato apropriado para a aplicação de algoritmos de mineração (p.e. agrupamento de dados- *clustering*);
- e) Mineração: etapa essencial do processo, que consiste na aplicação de técnicas inteligentes a fim de se extrair os padrões de interesse;
- f) Avaliação ou Pós-processamento: nesta etapa são identificados os padrões interessantes de acordo com algum critério do usuário;
- g) Visualização dos Resultados: nesta etapa são utilizadas técnicas de representação de conhecimento a fim de apresentar ao usuário o conhecimento minerado.

Este artigo está focado nas técnicas frequentemente utilizadas na etapa de mineração do processo KDD. Considera-se que os dados foram selecionados e transformados, integrados num armazém de dados (*Data Warehouse*) e deles foram eliminados ruídos que possam afetar o processo de descoberta de conhecimento.

A MD reúne várias técnicas, entre as quais pode-se destacar (Fayyad et all,1996):



**Figura 4:** Etapas do Processo KDD.

- **Classificação e Predição:** Classificação é o processo de encontrar um conjunto de funções que descrevem e distinguem classes. Utiliza o modelo para prever a classe de objetos que ainda não foram classificados. Predição indica os valores ausentes nos dados.
- **Agrupamento de Dados (*clustering*):** Os agrupamentos de dados ou *clustering* são atividades em mineração de dados, que consistem em agrupar os dados dentro de classes ou *clusters* tais que os objetos dentro de uma classe tenham alta similaridade em comparação com um outro objeto dessa classe (Han e Kamber, 2000), mas têm baixa similaridade a objetos de outras classes. A árvore de agrupamento (ou árvore hierárquica) usa as dissimilaridades ou distâncias entre os objetos para formar as classes. O cálculo das distâncias Euclidianas é o método mais direto de calcular as distâncias entre os objetos num espaço multidimensional.
- **Análise de *Outliers* (exceções):** Um banco de dados pode conter dados que não apresentam o comportamento geral da maioria. Estes dados são denominados *outliers*.

A seguir descreve-se o método de agrupamento *k-means*, utilizado no presente trabalho.

#### 4.1. Método *k-means* para Análise de Clusters

O método *k-means*, para análise de clusters, recebe como *input* um banco de dados de objetos e um número *k* que representa o número de clusters que se deseja formar entre os objetos do banco de dados. (Souza, 2004)

O banco de dados é apresentado em forma de uma matriz de dissimilaridade entre os objetos. Nesta matriz o elemento da coluna *j* e linha *i* da matriz é o número  $d(i, j)$  que representa a distância entre os objetos *i* e *j*. Várias funções distância podem ser utilizadas, destacando-se: distância Euclidiana, distância de Manhattan e a distância de Minkowski (Souza, 2004). A técnica para construir a matriz de dissimilaridade depende do tipo de dados presentes no banco de dados.

Seja  $C = \{C_1, \dots, C_k\}$  uma partição do banco de dados em *k clusters* e sejam  $m_1, m_2, \dots, m_k$  elementos escolhidos em cada um dos *clusters*, representando o centro dos mesmos. Define-se o erro quadrático da partição como sendo  $Erro(C) = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2$ .

O método *K-means* constrói uma partição *C* contendo *k clusters*, para o qual o erro quadrático é mínimo. A seguir apresenta-se as etapas do método *k-means*:

- (1) Escolhe-se arbitrariamente *k* objetos  $\{p_1, \dots, p_k\}$  do banco de dados. Estes objetos serão os centros *k clusters*, cada *clusters*  $C_i$  formado somente pelo objeto  $p_i$ .
- (2) Os outros objetos do banco de dados são colocados nos *clusters* da seguinte maneira: para cada objeto *O* diferente de cada um dos  $p_i$ 's, considera-se a distância entre *O* e cada um

dos  $p_i$ 's (dado pela matriz dissimilaridade). Considera-se aquele  $p_i$  para o qual a distância é mínima. O Objeto  $O$  passa a integrar o *clusters* representado por  $p_i$ .

(3) Calcula-se a média dos elementos de cada *clusters*, isto é, o seu centro de gravidade. Este ponto será o novo representante do *clusters*.

Volta para o passo 2: varre-se o banco de dados e para cada objeto  $O$  calcula-se a distância entre este objeto  $O$  e os novos centros dos *clusters*. O objeto  $O$  será realocado para o *clusters*  $C$  tal que a distância entre eles seja a menor possível.

(4) Quando todos os objetos foram devidamente realocados entre os *clusters*, calculou-se os novos centros dos *clusters*.

(5) O processo se repete até que nenhuma mudança ocorra, isto é, os *clusters* se estabilizam.

A vantagem do método é sua eficiência em tratar grandes conjuntos de dados e suas desvantagens são: o fato do usuário ter que fornecer o número  $k$  de *clusters*, o fato de não descobrir *clusters* de formatos não convexos e o fato de ser sensível a ruídos. Sendo assim, objetos com valores altos podem causar uma grande alteração no centro da gravidade dos *clusters* e distorcer a distribuição dos dados.

A seguir são apresentadas estas etapas aplicadas na descoberta de conhecimento na base de dados.

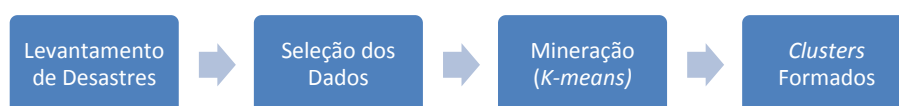
## 5. METODOLOGIA PROPOSTA E EXPERIMENTOS COMPUTACIONAIS.

Nesta sessão do artigo, será apresentada a metodologia proposta para a formação dos *clusters*, bem como o experimento computacional para validação e verificação do modelo apresentado, utilizando para tal dados reais sobre desastres naturais do estado de Santa Catarina.

### 5.1. Metodologia

Foi utilizado o gerenciador de banco de dados Postgre SQL 8.2 para organização da base de dados, juntamente com a linguagem de programação Java 1.6; os mapas apresentados são os mapas georeferenciados pelo IBGE (IBGE, 2011) e para a formação dos *clusters* utilizou-se a ferramenta computacional RapidMiner 5 (RAPIDMINER 5,2011).

A figura 5 esquematiza as etapas da metodologia proposta neste trabalho.



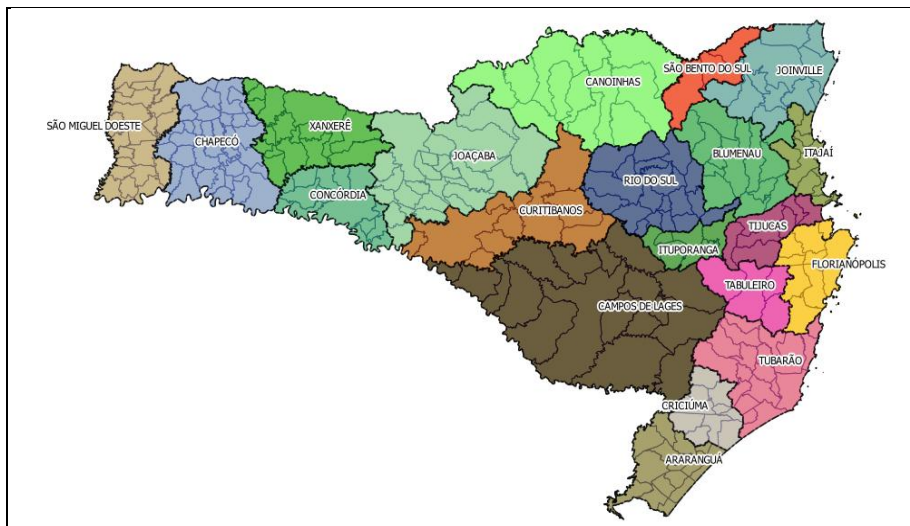
**Figura 5:** Esquema dos passos da metodologia proposta.

### 5.2. Experimento Computacional

Para testes da metodologia proposta, foram utilizados dados reais de desastres ocorridos no período entre os anos de 2000 e 2010 no estado de Santa Catarina. Estes dados foram levantados através da base de dados do CEPED – Centro Universitário de Estudos e Pesquisas sobre Desastres da UFSC. Esta base de dados utilizou documentos oficiais como AVADANS- Formulário de Avaliação de Danos; NOPRED- Notificação Preliminar de Desastre; DECRETO- Decreto Municipal; portarias; relatórios de danos e outros considerados não oficiais como jornal (DEFESA CIVIL,2011).

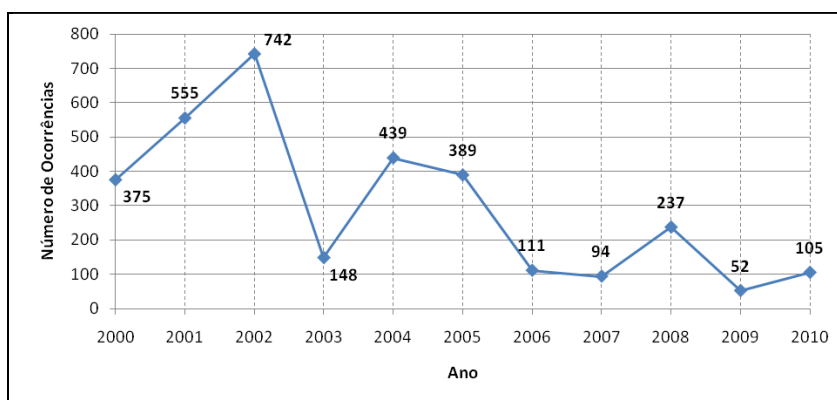
Inicialmente as ocorrências de desastres naturais foram separadas de acordo com as microrregiões do estado. O mapa político de Santa Catarina e suas microrregiões podem ser observados na Figura 6.

Em seguida, como os dados do CEPED incluíam vários tipos de desastres, foram filtrados da bases de dados somente as ocorrências de desastres naturais, resultando assim em 3.247 ocorrências selecionadas. Consideraram-se os tipos de desastres naturais conforme classifica a Secretaria de Defesa Civil Nacional (DEFESA CIVIL, 2011).



**Figura 6:** Microrregiões do estado de Santa Catarina (IBGE, 2011).

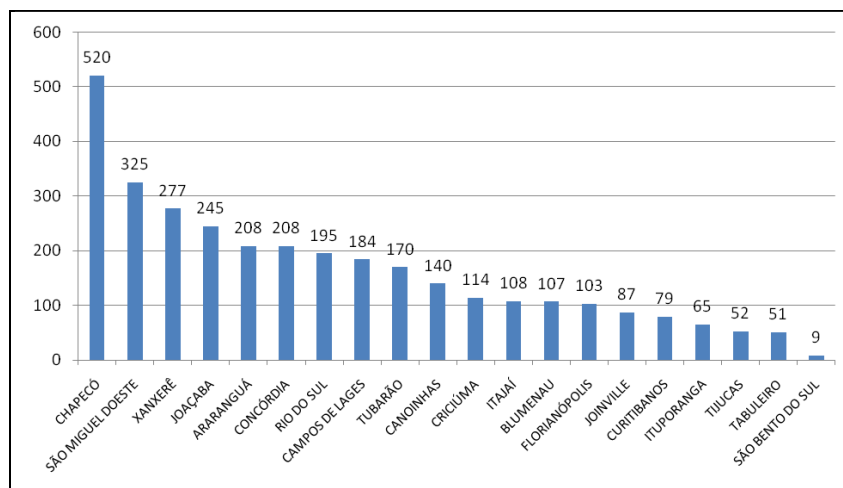
Pôde-se, com isso, constatar a distribuição anual dos municípios impactados por desastres naturais (Figura 7), ocorrências por microrregião (Figura 8) e por tipo de desastre (Figura 9).



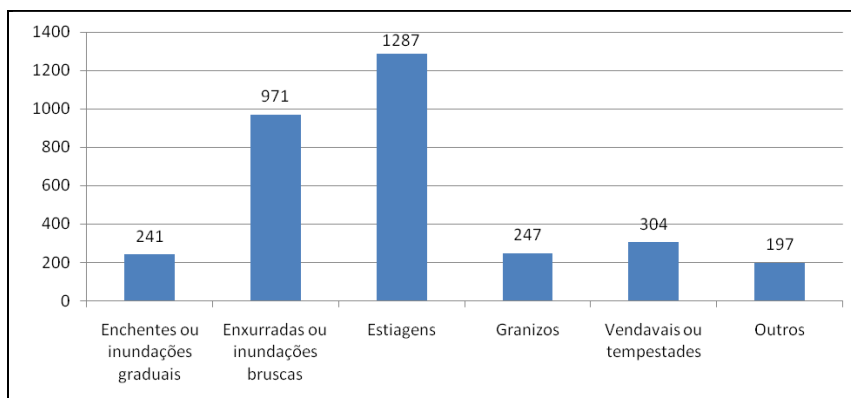
**Figura 7:** Distribuição anual de municípios impactados em Santa Catarina.

Como os dados do CEPED já estavam organizados, não foi necessário pré-processamento ou uma limpeza dos dados antes de proceder a base de dados à mineração de dados, como a remoção de dados incompletos ou eliminação da repetição dos dados. No entanto, para o objetivo deste trabalho, era necessária uma adequação e, para tal, os dados foram transformados numa matriz contendo a frequência de cada tipo de desastre natural por microrregião.

Esta matriz transformou-se, então, na base de dados a ser explorada pelo algoritmo *K-means* objetivando a formação de *clusters* com base na quantidade de municípios impactados por desastres naturais em Santa Catarina por microrregião.



**Figura 8:** Soma absoluta dos municípios impactados por Microrregião.



**Figura 9:** Soma absoluta dos municípios impactados por Tipo de Desastre.

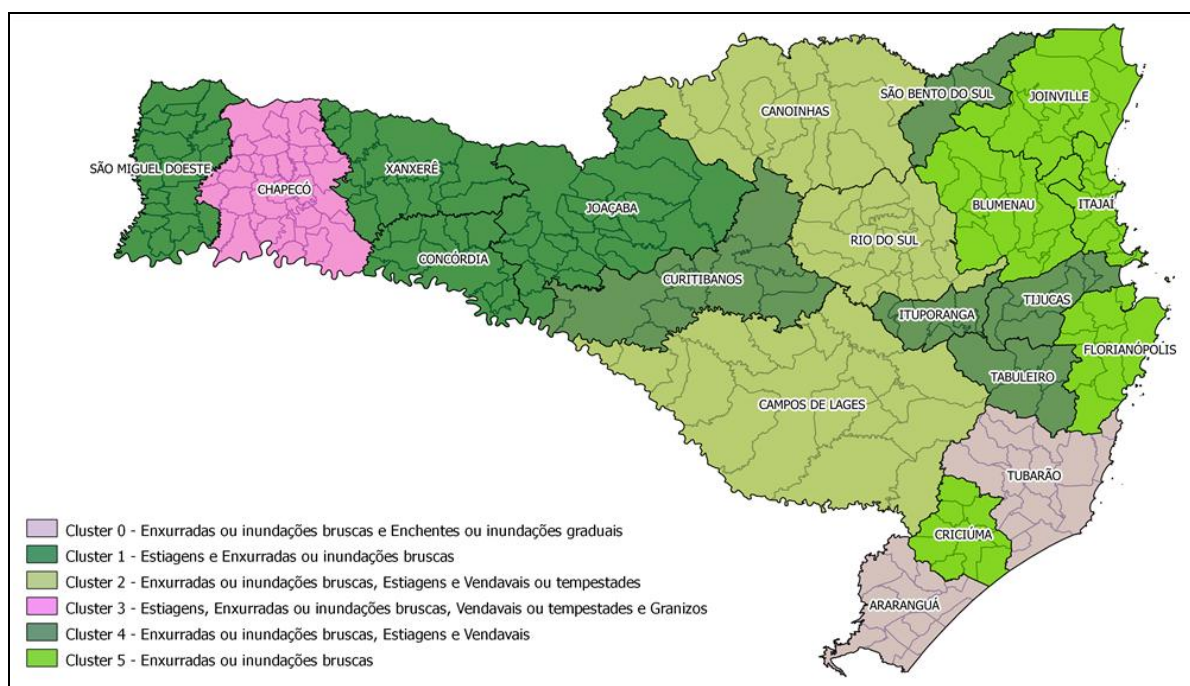
### 5.2.1. Resultados Obtidos

Verificou-se junto ao corpo de especialistas do CEPED que a quantidade de 6 *clusters* responde ao teste computacional esperado neste artigo. Logo, executando o algoritmo *K-means* com  $k=6$  chegou-se ao resultado sumarizado na Tabela 1 e demonstrado graficamente na Figura 10.

**Tabela 1:** Resultado do algoritmo *k-mean* com  $k=6$ .

Cluster	Microrregiões	Predominância	
		Menor Representatividade	Maior Representatividade
0	Araranguá e Tubarão	<ul style="list-style-type: none"> <li>• Vendavais ou tempestades;</li> <li>• Granizos;</li> <li>• Enchentes ou inundações graduais.</li> </ul>	<ul style="list-style-type: none"> <li>• Enxurradas ou inundações bruscas.</li> </ul>
1	São Miguel D'Oeste, Xanxerê, Concórdia e Joaçaba	<ul style="list-style-type: none"> <li>• Vendavais ou tempestades;</li> <li>• Granizos;</li> <li>• Enxurradas ou inundações bruscas.</li> </ul>	<ul style="list-style-type: none"> <li>• Estiagens.</li> </ul>
2	Campos de Lages, Canoinhas e Rio do Sul	<ul style="list-style-type: none"> <li>• Vendavais ou tempestades;</li> <li>• Granizos;</li> <li>• Enchentes ou inundações graduais.</li> </ul>	<ul style="list-style-type: none"> <li>• Enxurradas ou inundações bruscas;</li> <li>• Estiagens.</li> </ul>
3	Chapecó	<ul style="list-style-type: none"> <li>• Vendavais ou tempestades;</li> <li>• Granizos;</li> <li>• Enchentes ou inundações graduais;</li> <li>• Enxurradas ou inundações bruscas.</li> </ul>	<ul style="list-style-type: none"> <li>• Estiagens.</li> </ul>

4	Curitibanos, São Bento do Sul, Ituporanga, Tabuleiro e Tijucas	<ul style="list-style-type: none"> <li>• Estiagens.</li> </ul>	<ul style="list-style-type: none"> <li>• Enxurradas ou inundações bruscas.</li> </ul>
5	Blumenau, Criciúma, Florianópolis, Itajaí e Joinville	<ul style="list-style-type: none"> <li>• Vendavais ou tempestades;</li> <li>• Enchentes ou inundações graduais.</li> </ul>	<ul style="list-style-type: none"> <li>• Enxurradas ou inundações bruscas.</li> </ul>



**Figura 10:** *Clusters* resultantes do experimento.

## 6. CONSIDERAÇÕES FINAIS

Neste artigo foram apresentados alguns conceitos envolvidos na gestão em desastres com o objetivo de localizar o foco da pesquisa. Foram abordados conceitos e características de um sistema *clusters* tendo como intuito verificar a aplicação do sistema de *clusters* na Logística Humanitária.

Pôde-se verificar que o método apresentado neste trabalho, corresponde às expectativas esperadas e traz resultados satisfatórios, construindo grupos homogêneos de municípios impactados por desastres naturais. Isto que mostra a validade do método apresentado e sua diversidade de aplicações.

Constata-se que os *clusters* formados em algumas microrregiões apresentam uma grande aproximação espacial, além do mesmo perfil de desastres nos municípios impactados. Isso possibilita às organizações responsáveis criar parcerias para investir no tipo de infra-estrutura necessária, facilitar o compartilhamento das experiências, visando cooperação, espelhar soluções e aproximar interesses, gerando, com isso, mútuo aprendizado a custos mais baixos. Tais concentrações podem beneficiar, também, cada município os quais passam a ter uma maior dimensão, sem que sua flexibilidade seja afetada.

Pretende-se, para trabalhos futuros, dar continuidade a esta pesquisa direcionando o foco para formação de *clusters* por municípios e compará-los com os resultados obtidos neste artigo, a fim de identificar possíveis ruídos que podem ter surgido em função da divisão política das

microrregiões, que nem sempre refletem a realidade geográfica a ser considerada para cada município impactado.

#### AGRADECIMENTOS

Os autores agradecem ao corpo de especialistas do CEPED/UFSC e ao apoio financeiro do CNPq, através do projeto nº 470774/2010-0.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, S.J. e Zarate, L.H. (2007) Identificação de Grupos de Estados Brasileiros Similares em Relação a Índices de Mortalidade. *III Workshop em Algoritmos e Aplicações de Mineração de Dados. WAAMD*.
- Balcik, B.; Beamon, B.M.; Krejci, C; Muramatsu, K.M e Ramirez, M. (2010) Coordination in humanitarian relief chains: Practices, challenges and opportunities. *International Journal Production Economics. Science Direct*, vol.126 issue 1, pp.22-34, 2010.
- CRED ( 2011) disponível em [http://cred.be/sites/default/files/Disaster\\_numbers\\_presentation\\_2010.pdf](http://cred.be/sites/default/files/Disaster_numbers_presentation_2010.pdf) acesso em 20/06/2011.
- CREA ( 2011) Revista do Conselho Regional de Engenharia, Arquitetura e Agronomia de Santa Catarina. Ano 6- n° 9 maio.
- DEFESA CIVIL (2011). Disponível em [www.defesacivil.gov.br](http://www.defesacivil.gov.br). Acesso em: 20 junho 2011.
- Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (1996). Advances in Knowledge Discovery and Data Mining. *AAAI Press, Menlo Park, CA*.
- Han, J. e Kamber, M. (2000). Data mining: concepts and techniques. *Morgan Kaufmann Publishers Inc., San Francisco, CA, USA*.
- IBGE (2011) disponível em <http://www.ibge.gov.br>. Acesso em 27 de maio de 2011
- Jahre, M. e Jensen, M. (2010) Coordination in humanitarian logistics through clusters. *International Journal of Physical Distribution & Logistics Management* Vol. 40 No. 8/9, pp. 657-674.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics and Probability*, 1:281-296.
- Nogueira, C. W.; Gonçalves, M. B. (2010) O Desenvolvimento e aplicação de um modelo para a localização de uma central de inteligência e suporte para recebimento, controle e distribuição de recursos em situações emergenciais com foco na logística humanitária. *Anais do XXIV Congresso de Pesquisa e Ensino em Transportes. Salvador, Bahia*.
- OCHA (2007), Appeal for Building Global Humanitarian Response Capacity, *Office for the Coordination of Humanitarian Affairs, New York, NY*, available at: <http://ochaonline.un.org/HUMANITARIANAPPEAL/webpage.asp?Page=1566>
- ONERESPONSE (2011). Disponível em <http://onerresponse.info/Coordination/ClusterApproach/Pages/Cluster%20Approach.aspx> acesso em 21/06/2011.
- ONERESPONSE (2011a). Disponível em <http://onerresponse.info/GlobalClusters/Pages/default.aspx> acesso em 21/06/2011.
- Pyle, D. (1999). *Data Preparation for Data Mining*. Morgan Kaufmann Publishers, Inc., California.
- Porter, M.E. (1998) Clusters and the new economics of competition, *Harvard Business Review*, Vol. 76 No.6, pp.77-90.
- RAPIDMINER 5 (2011). Disponível em <http://www.rapid-i> acesso em 31 de março de 2011
- Sapir, G. D. (2011) Disasters in Numbers 2010. *CRED, Catholic University of Louvain, Brussels (Belgium) January 24, Geneva*.
- Schultz, S.F. (2008) Disaster Relief Logistics: Benefits of and Impediments to Horizontal Cooperation between Humanitarian Organizations. *Tese. Technischen Universität Berlin*.
- Souza, F.T. (2004) Predição de escorregamentos das encostas do município do rio de janeiro através de técnicas de mineração de dados. *tese (doutorado). Programa de pós- graduação em engenharia, universidade federal do rio de janeiro, Rio de Janeiro, 2004*.
- Tufinkgi, P. (2006) Logistik im kontext internationaler katastrophenhilfe: Entwicklung eines logistischen referenzmodells für katastrophenfälle. *Ed. Haupt Verlag. Bern. Stuttgart Wien*