

# 1

## Introdução

Em uma comparação internacional, o Brasil se encontra hoje como um dos países com maior taxa de homicídios, estando atrás apenas de países como a Colômbia, África do Sul, Jamaica e Venezuela. Um dos indicadores mais consistentes do aumento da criminalidade violenta no Brasil nas últimas décadas é a evolução da incidência de homicídios, que passou de 11 para 23 ocorrências por 100 mil habitantes entre 1980 e 2003. Com referência a roubos, o Brasil se encontra hoje como um dos países com maior taxa de roubos, estando atrás apenas de países como o Chile e África do Sul (SENASP, 2004; SENASP, 2007). O Brasil apresenta uma das mais baixas taxas do mundo em elucidação de crimes. Segundo o ISP (2007), no Rio de Janeiro, verifica-se que o percentual de êxito em uma amostragem de 381 casos de homicídios analisados no ano de 2002 é muito baixo (1,31%), onde apenas cinco casos foram relatados à Justiça com autoria.

Os crimes praticados nas metrópoles brasileiras consomem muito tempo e vultuosos recursos policiais para investigação. O policiamento ostensivo apenas captura criminosos em ação flagrante, concentrando-se em delitos considerados de maior visibilidade pública, como distribuição de drogas no varejo, violações de trânsito ou desordens urbanas.

As atividades de investigação policial utilizam históricos, relatos e boletins de ocorrências, residentes em bases de informações estruturadas ou textuais, de onde são extraídos conhecimentos para elucidação de autoria, participações, circunstâncias e objetos envolvidos nas ocorrências criminais.

Os registros policiais integram atores, objetos, eventos e relacionamentos, captados das fontes de informações, e sobre os quais são sintetizados desenhos e simuladores de cenários, conhecidos no âmbito policial como mapas criminais, mapas de investigação ou Mapas de Inteligência.

Devido ao caráter crítico do problema relacionado ao combate da criminalidade, que associa a complexidade inerente à desestruturação das fontes de informações policiais com restrições de tempo e recursos para análise das autorias e elucidação dos crimes, a extração automática de conhecimentos em

bases de dados criminais assume importância relevante na geração de Mapas de Inteligência e atividades de investigação.

A motivação desta pesquisa concentra-se na descoberta de ferramentas que possam oferecer melhor velocidade e qualidade na investigação criminal, aumentando os indicadores de eficiência policial e as taxas de elucidação de crimes como contribuição tecnológica para o segmento da Segurança Pública brasileira.

## 1.1

### O Problema

Um dos problemas fundamentais encontrados nas investigações de autorias dos crimes é descobrir vestígios que transportem o criminoso para o local do crime ou relacionem vínculos de suas atividades para consecução do delito cometido. Os estudos investigativos para análise de autorias baseiam-se em históricos policiais, que compreendem os boletins de ocorrência preenchidos, levando-se em conta, as narrativas textuais de testemunhas, participantes e policiais.

Em um registro criminal são transcritos os fatos, pessoas, circunstâncias e relacionamentos que caracterizaram a ocorrência policial. Devido à necessidade da captura das circunstâncias com maior exatidão da realidade, os documentos de transcrição, geralmente denominados de Boletins de Ocorrência, procuram oferecer livre espaço para comunicação através de narrativas, muitas vezes desenvolvidas por pessoas de diferentes funcionalidades: comunicante, testemunhas, escrivão, policial de plantão, patrulheiro externo e policial responsável pelo registro.

Os históricos policiais são muito ricos em informações das quais são extraídas redes de relacionamento e fornecem evidências preciosas e padrões criminais analisados em investigações. Os resultados extraídos podem ser representados em Mapas de Inteligência, que são representações topológicas do cenário do crime, apresentando possíveis conexões entre atores e eventos potencialmente envolvidos e que integram todas as peças da investigação (Xu & Chen, 2004).

Dados textuais em aplicações policiais têm sido utilizados para reconhecimento de padrões e verificação de similaridade entre crimes (Cortes et al., 2002). Muitas vezes os históricos constituem bases de informações de difícil análise, ocultando importantes evidências e indícios presentes no cenário do crime. O sucesso da elucidação da ocorrência criminal depende essencialmente do montante investido em pessoal, precisão dos dados coletados, recursos para investigação e prazo para início da investigação. Scott (1996) cita que pesquisas aprofundadas exigem tempo e pessoal qualificado para uma análise histórica útil e conclusiva.

Diante da relação de compromisso, entre praticar policiamento preventivo e investir em policiamento investigativo, Sparrow (1991) cita que o tempo dedicado à investigação pela polícia seja tão breve quanto eficiente, empregando para isso, de forma regular, ferramentas eficazes para atividades de inteligência.

O emprego de ferramentas de investigação que promovam a automação dos Mapas de Inteligência e a descoberta de vínculos entre atores envolvidos na investigação reduz o emprego e tempo necessários para extração do conhecimento em históricos criminais, aumentando a eficiência da atividade policial.

## 1.2

### **Objetivos da Tese**

O objetivo principal desta Tese é propor metodologias e ferramentas para criação e análise de Mapas de Inteligência.

O método proposto contempla a extração de conhecimentos em documentos criminais, utilizando mineração de dados em textos estruturados e não estruturados. Os resultados da extração serão convertidos em uma representação topológica, sobre a qual são aplicados algoritmos para descoberta de associações criminais, identificação dos mais fortes relacionamentos entre as entidades extraídas e identificação de clusters criminais.

Esta pesquisa possui três objetivos associados, compreendendo:

- Preparação de dicionários especialistas;

- Estruturação e modelagem de redes semânticas extraídas de documentos criminais (Modelagem de associações criminais) usando dicionários especialistas;
- Métodos de pesquisa, análise e identificação de relacionamentos topológicos dos elementos extraídos para a rede semântica, usando teoria dos grafos e redes neurais.

### 1.3

#### O Método

A descoberta de conhecimentos em históricos policiais é realizada através de conjunto de métodos e algoritmos que se integram para extração de conhecimento. A descoberta de conhecimento é desenvolvida em três módulos que foram, para efeitos de apresentação desta Tese, seqüenciados em três etapas, conforme abaixo especificado:

- Etapa 1 - Construção de dicionários de apoio ou dicionários especializados
- Etapa 2 - Extração de entidades úteis de históricos e modelagem de Mapas de Inteligência
- Etapa 3 - Análise de conexões e descoberta de conhecimentos

### 1.4

#### Resultados Esperados

Como resultados, esta pesquisa deverá produzir métodos para:

Extração de radicais em língua portuguesa que servirão de apoio à extração de dicionários temáticos especializados na linguagem policial.

Construção de dicionários temáticos especializados, baseados em documentos textuais estruturados e não estruturados, extraídos de históricos policiais aplicando estruturas de radicais treinados.

Extração de entidades úteis de históricos policiais, modelando uma estrutura topológica de relacionamentos entre as entidades em formato de grafos, representando um Mapa de Inteligência criminal.

Desenvolvimento de redes semânticas valoradas integradas por entidades extraídas de históricos policiais.

Identificação de organizações criminais e análise de concentrações de entidades relacionadas extraídas dos Mapas de Inteligência.

Identificação de interseções entre entidades e subgrupos criminais extraídos de Mapas de Inteligência

Identificação de centralidades em subgrupos criminais extraídos de Mapas de Inteligência.

Identificação das mais fortes conexões entre entidades extraídas de Mapas de Inteligência.

## 1.5

### Organização da Tese

Este documento encontra-se organizado da forma seguinte:

O **capítulo 1** - motivação e objetivos da pesquisa, descrevendo uma síntese do problema, uma visão geral do método e resumo dos resultados esperados com a pesquisa.

O **capítulo 2** - conceitos de Mapa de Inteligência e seus componentes, relacionamentos entre os elementos topológicos extraídos para modelagem de associações criminais e apresentação de um modelo exemplo.

O **capítulo 3** - construção de dicionários especialistas; descreve o referencial teórico utilizado para pesquisa e método aplicado para desenvolvimento de uma estrutura de radicais de apoio que servirão para construção de dicionários especializados. São descritos os algoritmos e sistemas computacionais de apoio utilizados na geração dos dicionários de radicais e entidades..

O **capítulo 4** - construção do Mapa de Inteligência e referencial teórico utilizado; descreve o método para pesquisa de vínculos entre entidades, algoritmos e sistemas computacionais de apoio utilizados na modelagem e construção das redes semânticas

O **capítulo 5** - análise do Mapa de Inteligência extraído de históricos policiais; descreve o referencial teórico utilizado e o método desenvolvido para identificação das associações entre entidades e subgrupos criminais gerados como

produto das extrações, algoritmos e sistemas computacionais de apoio utilizados para análise e classificação das associações.

O **capítulo 6** - estudos de casos criminais extraídos dos históricos tratados demonstrando os métodos propostos nesta Tese.

O **capítulo 7** - conclusões dos estudos desenvolvidos e recomendações para futuras pesquisas sobre o método proposto.

O **capítulo 8** - Bibliografia utilizada para desenvolvimento da Tese.

**Apêndices anexos.**