

Gestão de Conhecimento e Integração de Bases: A Experiência do Disque Denúncia

A experiência do Disque Denúncia (DD), no tratamento de questões relativas à segurança pública em seus vinte três anos de operações, dando suporte às ações policiais através da difusão de informações relevantes em um cenário complexo, como o crime organizado no Rio de Janeiro. Com características multifacetadas – denunciante, polícia e imprensa – apoia a criação de um conjunto de tecnologia de gestão do conhecimento, em denúncias anônimas e na integração de bases entre as agências e órgãos de segurança pública.

Tudo isto irá se refletir na transferência e aplicação de conhecimento de informações de dados abertos, não estruturados e estruturados, contribuindo na aplicação e transferência de conhecimento aplicável gerando incremento de valor e informação na integração de bases, como por exemplo, Instituto de Segurança Pública (ISP), Secretaria de Segurança (SESEG), Ministério Público (MPRJ), agências de inteligência, aplicativos de segurança - Fogo Cruzado, entre outros. O presente estudo reflete a visão virtual que está sendo implementada junto ao DD. O que vem a seguir, descreve nossa experiência no Rio de Janeiro, no trato com os dados referidos.

Introdução

Um setor que mundialmente vem buscando tecnologias mais eficientes para dar suporte às suas operações é o de segurança pública. Praticamente todas as instituições que atuam nessa área possuem, em sua estrutura, uma ou mais unidades de inteligência, dedicadas ao trato de grandes quantidades de dados. Atualmente, as atividades de investigação, denúncias e ocorrências envolvem receber, armazenar e processar dados das mais diversas fontes (estruturadas e não estruturadas), como perfis em redes sociais, dados cadastrais de pessoa física e jurídica, informações sobre pessoas, fotos, veículos, áudios obtidos com processos de interceptação legal, informações via web e aplicativos móveis e outras tantas fontes disponíveis.

Segurança pública é inegavelmente um tema que nos últimos anos tem adquirido cada vez mais importância em diferentes contextos e falar dela, quando se trata do estado e cidade do Rio de Janeiro, torna-se ainda mais interessante e desafiador ao considerar-se os aspectos socioeconômicos e geográficos do problema da violência. Esses aspectos permitem a aproximação e interação entre favelas e o “asfalto”, e com isso o convívio cotidiano entre diferentes classes sociais e conseqüentemente todos os conflitos oriundos desse contato, onde se destaca como mais marcante a violência que atinge a todos no cotidiano. Historicamente, desde década de 90 observa-se uma ascensão da criminalidade no município e a crescente crise das instituições de segurança pública, que vêm se somar à ligação entre tráfico de drogas e armas, contexto esse que caracterizou uma profunda mudança na dinâmica da criminalidade e da violência na cidade.

No Rio de Janeiro, o DD atua como um canal entre a população e os agentes de segurança pública no combate à criminalidade e solução de problemas que atingem comunidades, bairros e cidades. Ele recebe as denúncias e coleta informações de importância fundamental na apuração de ações criminosas e conseqüente redução da criminalidade.

Entretanto, quando se trata de análises mais profundas e armazenamento e tratamento de dados, surgem dificuldades de natureza operacional derivadas da limitação tecnológica desde a coleta de informações e dados até a sua análise, passando pelo tratamento, armazenamento e seus filtros. Desse modo, faz-se necessária a busca por inovações tecnológicas na área de inteligência artificial e tratamento de dados que acompanhem o próprio progresso tecnológico da sociedade e sua crescente complexidade, para melhor analisá-los e estudá-los.

Diante do panorama atual no qual os setores de segurança pública dos governos ao redor do mundo vêm buscando tecnologias mais eficientes para dar suporte às suas operações na área de investigação e inteligência, destaca-se o caso especial do estado e da cidade do Rio de Janeiro, cujas especificidades de geografia física, humana e urbana representam grandes desafios para os agentes envolvidos no contexto da segurança pública do estado. Neste sentido, levando em consideração que pretendemos trabalhar com o armazenamento, tratamento e associação em grandes bases de dados, devemos considerar que esses dados precisam ser compreendidos a partir da natureza de suas fontes e as implicações, tanto sociais quanto as da perspectiva do processo de análise desses dados.

Justificativa

O acesso da sociedade a novas tecnologias tem alterado a rotina e o processo de captura de denúncias dentro do DD. O principal canal de comunicação da organização com a sociedade tem sido a linha telefônica, em um digitronco de número 2253-1177 que já recebeu mais de dois milhões de denúncias ao longo dos últimos 20 anos. Mais recentemente, registra-se o incremento de denúncias realizadas através de mídias sociais como Facebook e WhatsApp, entre outras, e que já correspondem a cerca de 15% do número de denúncias recebidas. As informações recebidas por estes canais têm como característica a forma não estruturada de dados, o que demanda tratamento, armazenamento e análise diferenciados. Sem dúvida, o surgimento desses novos canais digitais com a sociedade abre um número infinito de possibilidades para a contribuição do DD com a política pública de segurança do Estado do Rio de Janeiro.

Assim sendo, nosso parecer é uma iniciativa que se baseia na pressão imposta pela tecnologia sobre a sociedade e a necessidade de adaptação do DD a essa nova realidade para melhor contribuir com as políticas públicas de combate à criminalidade. O acesso às informações e a geração de conhecimentos regionais e setoriais é o propósito fundamental que orienta este projeto. A iniciativa é pioneira e está voltada ao desenvolvimento, capacitação e implantação de novos processos tanto no que tange tanto ao processo de coleta da denúncia quanto da entrega de novas respostas à sociedade. São incorporados todos os temas já analisados pela instituição, como por exemplo, violência doméstica, meio ambiente, entre outros. Entretanto, a estratégia é preparar-se para atender às novas demandas que se façam presentes na sociedade e dessa maneira melhor colaborar com a política pública estadual.

É importante considerar que uma denúncia anônima, mesmo que seja uma única denúncia, é um instrumento poderoso. Ela alcança todo seu potencial quando chega a tempo às mãos certas do agente certo, a quem compete fiscalizar ou investigar os fatos que ela aponta. A experiência de mais de duas décadas de atuação do DD mostra que o correto

tratamento e utilização da informação favorecem o melhor desempenho das autoridades. Quanto maior a adesão da sociedade a esse canal de informação, aliada a uma maior e mais inteligente capacidade de gestão de dados, maior se torna a eficácia dos serviços públicos.

Descrição do Programa

Objetivos Gerais

Baseada no interesse em trabalhar com Sistemas de Informação e Gestão de Sistemas Complexos com dados de Segurança Pública, conjuntamente ao tratamento e associação em grandes bases de dados, esta proposta tem por objetivo elaborar uma estratégia de inteligência para combate e prevenção à criminalidade, trabalhando com inteligência artificial nas bases de dados do Disque Denuncia e do Instituto de Segurança Pública (ISP), do estado do Rio de Janeiro.

Mais especificamente, no contexto dessa base de dados é necessário encontrar os vínculos entre esses dados, entender como estão geograficamente distribuídos, correlacionar a abordagem de textos diferentes que possuam o mesmo sentido e garimpar informações relevantes entre centenas ou milhares desses textos, dados, fotos, vídeos e objetos. Para isso, partimos de uma abordagem quantitativa e qualitativa que possibilite a compreensão desse processo como captação de dados, que geram informações (valores), e posteriormente conhecimento. Para dar conta dessa missão e obter resultados dinâmicos e eficientes, as unidades de inteligência do setor de segurança pública precisam focar na utilização de novas tecnologias para dotar investigadores, agentes e ONGs de ferramentas poderosas que permitam a realização de todas essas etapas de análises. Consideramos ainda que essa nova abordagem pode ser útil a potenciais pesquisadores interessados no resultado desse trabalho, como fonte de informação para a criação de novas políticas públicas e estudos concernentes as áreas afins.

Podemos citar como exemplo existente o sistema utilizado pela Secretaria de Estado da Segurança Pública de Santa Catarina, que adquiriu um produto de tratamento de Big Data para apoio à inteligência chamado “Guardião”. O sistema é parte de uma plataforma que aplica vários conceitos com a capacidade de tratar de maneira eficiente um número muito maior de dados, onde agentes de segurança podem responder de forma mais efetiva à resolução e investigação de denúncias e crimes e, principalmente, também preveni-los e (ou) antecipá-los. Observado pela perspectiva da segurança pública, como no exemplo, fica fácil observar que Big Data não é algo que pode ser considerado uma moda momentânea das empresas de tecnologia, mas sim um conceito que já está sendo aplicado em benefício de toda a sociedade. O uso inteligente da informação pode auxiliar o trabalho das autoridades policiais na diminuição dos índices de violência urbana.

Objetivos Estudados

O que propomos tecnicamente é a geração de uma camada de inteligência artificial para melhorar o processamento e encaminhamento das denúncias recebidas pela central de atendimento do DD, gerando uma base de dados que poderá ser compartilhada e transmitida

aos órgãos de segurança pública e também com as das autoridades necessárias. Para tanto é necessário classificar e relacionar a grande quantidade de informação de maneira que seus sentidos e valores sejam claros e úteis às autoridades competentes, desse modo aumentando a eficiência das informações recebidas pelos operadores do DD, um dos escopos fundamentais desta proposta. Todavia, os seguintes objetivos específicos são visados por esta proposta e conduzirão ao alcance do objetivo geral:

- Aumento da qualidade na captação de informações;
 - Aumento do número de denúncias por atendentes;
 - Aumento da capacidade de armazenamento das informações;
 - Aumento da capacidade de tratamento das informações, produção de análises customizáveis e relatórios;
-
- Aumento do número de horas de treinamento, capacitação e difusão de todo o pessoal;
 - Aumento da capacidade de otimização e envio de informações a Secretaria de Segurança, agentes pré-estabelecidos, secretarias municipais, estaduais, parceiros, governo, entre outros.

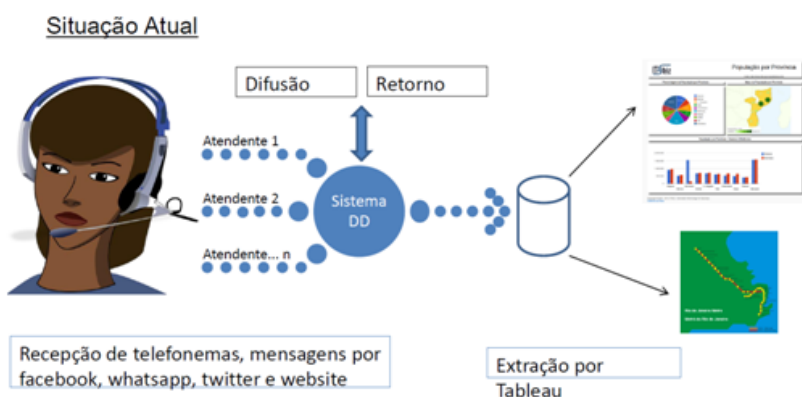
Metodologia

Esta proposta pretende seguir as etapas e análises descritas a seguir para a consecução de seus objetivos de estudo. O novo tipo de abordagem no tratamento dos dados do DD, bem como seus conhecimentos e valores serão obtidos não somente do rigor metodológico no tratamento dos dados apurados, mas também da observação, da leitura e troca de ideias com os principais atores envolvidos. O anonimato dos denunciante será mantido, pois é um dos principais pilares estratégicos que sustentam a existência do DD.

Para oferece maior sinergia entre processos de interação com os demandadores de solução, as equipes de apoio e a "ponta" do processo, que oferece a solução da denúncia, a equipe terá papel ativo no processo de inovação, na coordenação e busca de informações e nas planificações temáticas e regionais. Será um importante canal capaz de, por um lado, descrever o "estado de arte" e acompanhar tendências na sociedade prevendo cenários, e por outro, conhecer e divulgar a realidade regional. Não basta estruturar um sistema de informações através do Big Data, há a necessidade de criar um ambiente tecnológico propício às interações que favoreçam a inovação, acompanhem, validem e monitorem os dados recebidos.

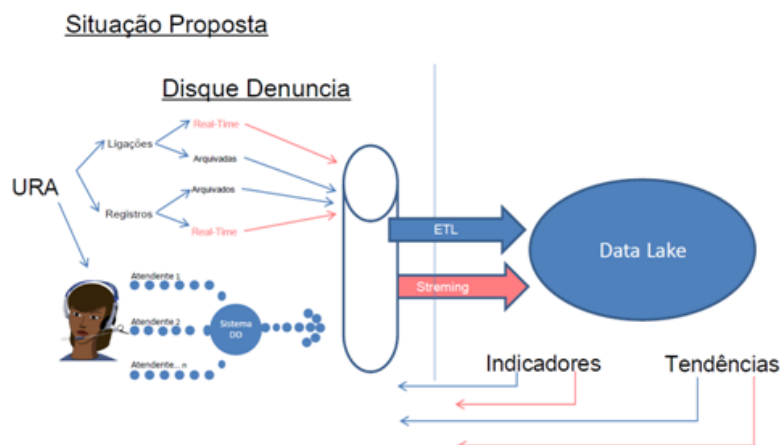
CENÁRIOS:

CENÁRIO ATUAL Figura 1



No cenário atual, as informações passadas à Central DD através de ligações anônimas ao sistema são filtradas e classificadas por analistas de informações, visando uma melhor compreensão da mecânica criminal. Especificamente, no caso dos principais crimes cometidos, é possível perceber que há relações e padrões que podem ser observados e analisados. Conforme a Figura 1, hoje as informações recebidas de todas as fontes são registradas de forma individual (telefone separado de mídias sociais) em software in-house para tratamento e, posteriormente, através de um software proprietário ser extraídas para difusão junto às instituições responsáveis pela apuração. Como não há softwares de descoberta de informações em grandes bases de dados, mineração de dados, e/ou outros de desempenho e análises correlatas, a tecnologia existente limita as possibilidades de interpretação das informações coletadas e, conseqüentemente, a capacidade de atuação das forças de segurança que recebem essas informações.

CENÁRIO PROPOSTO Figura 2



Nesse novo cenário é alocado um serviço de distribuição de chamadas (URA) para garantir agilidade no processo de denúncias e facilitar o treinamento dos atendentes na condução do script de recepção. As informações recebidas permaneceram registradas no software in-house, entretanto a utilização de um software open source altamente customizável permitirá que todas as informações sejam depositadas em um lago de dados. Trata-se de um grande repositório de armazenamento baseado em objetos que armazena dados em seu formato original até que seja necessário usá-los para serem tratadas por tecnologias de mineração as quais poderão ler uma grande quantidade de dados a fim de correlacioná-los na produção de tendências e indicadores. Utilizando dessa forma como princípio, não somente o conjunto de dados armazenados como também a lógica de análise usada para obter melhores resultados é preciso que o conjunto de dados seja extraído, transformado e carregado, utilizando novas tecnologias de armazenamentos capazes de receberem e lidarem com dados estruturados e não estruturados de diferentes naturezas e origens.

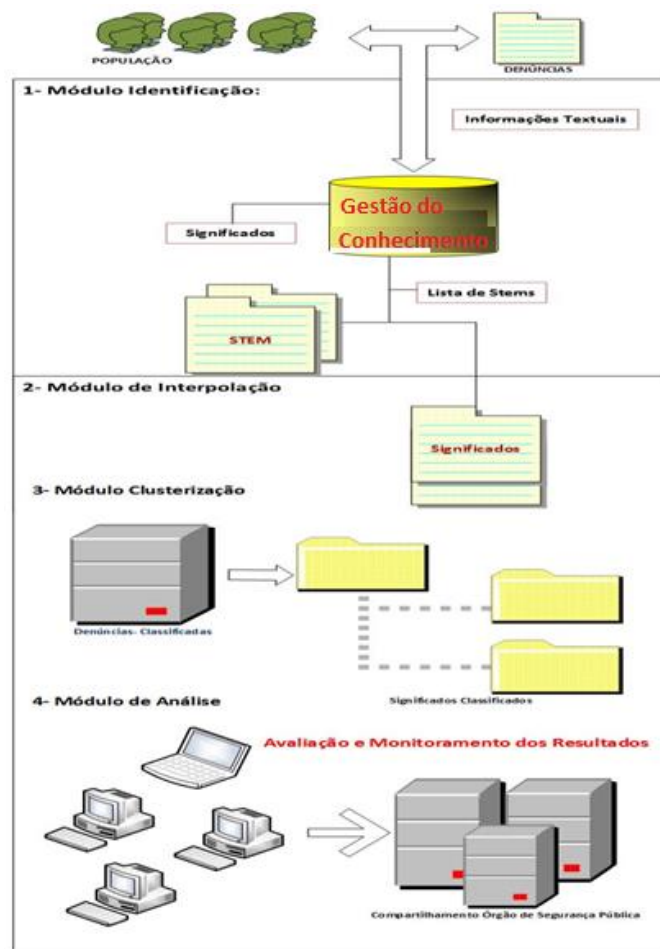
Conforme descrevemos nos objetivos específicos dessa proposta, no que tange à melhora do processamento e encaminhamento das denúncias recebidas pela central de atendimento do DD, para tanto faz-se necessária a aplicação de uma camada de tecnologia, que consiste na utilização de técnicas de inteligência artificial, propondo técnicas de mineração de textos nas etapas de extração, classificação, análise das informações recebidas, processamento e associação dessas informações às respectivas denúncias que podem ligar a possíveis crimes, recuperando informações relevantes. Aqui temos a descrição desse modelo proposto:

1. Módulo de Identificação: Através do texto fonte será produzido uma lista de palavras-chaves (stems chaves), aplica-se um algoritmo de stemming que transformará cada palavra no seu respectivo stem ou forma raiz. Algoritmos de Stemming, ou stemmers tem o objetivo de reduzir uma palavra ao seu stem, à sua forma raiz. As palavras-chave de uma consulta ou de um documento são mais bem representadas por stems do que pelas palavras originais. Com isso reduziremos o tamanho da ontologia, e com isso o número de termos distintos necessários para representar uma lista de documentos.

2. Módulo de Interpolação: A partir da lista de stems de cada denúncia busca-se a lista de significados mapeados com suas respectivas estatísticas de frequência do banco.

3. Clustering: As denúncias serão clusterizadas a partir dos significados extraídos de cada uma delas. Primeiro agrupa-se as denúncias similares e adiciona-se informações estatísticas aos temas classificados. Cria-se um sumário executivo para posterior validação.

4. Módulo Análise: A partir dos tópicos classificados e clusterização, utiliza-se a base de conhecimento em que estará contida uma lista de medidas que servirá para o direcionamento dos tópicos extraídos e para a realização de inferências do usuário.



- Trata-se de um estudo de doutorado em Informática na área de sistemas complexos adaptativos na UFRJ.