

Um Ambiente de Apoio à Decisão baseado em *Data Warehouse* para a Área de Segurança Pública do Estado do Rio de Janeiro

Wagner Santos¹, Daniel de Oliveira²

¹Polícia Militar do Estado do Rio de Janeiro (PMERJ)

²Instituto de Computação – Universidade Federal Fluminense (IC/UFF)

pesquisa_ei@pmerj.rj.gov.br, danielcmo@ic.uff.br

Resumo. *As novas abordagens propostas para cidades inteligentes possuem o potencial de melhorar a vida dos cidadãos em diversos aspectos. Um desses aspectos é o aprimoramento das políticas de segurança pública, em especial nos grandes centros metropolitanos. A análise integrada de dados históricos de ocorrências de crimes ajuda, sem dúvida, na tomada de decisão, na previsão e até mesmo na prevenção de ocorrências. No entanto, o acesso a esses dados históricos de forma integrada pode não ser trivial, pois se faz necessário integrar bases de dados de diferentes organizações (e.g., bombeiros, polícia civil, polícia militar, etc). Neste artigo, apresentamos uma abordagem que visa integrar dados das diferentes organizações associadas à segurança pública de forma a prover um ambiente de apoio a decisão que possa de fato auxiliar especialistas em segurança a delinear suas ações. O arcabouço proposto é denominado CHORD (Criminal dasHbOaRd Decision making). Nesse artigo, instanciamos o CHORD para o cenário do estado do Rio de Janeiro, que vem sofrendo nos últimos anos com um aumento significativo nos índices de criminalidade. A abordagem proposta é uma ferramenta promissora para auxiliar os departamentos de polícia na prevenção de crimes.*

1. Introdução

O conceito de *Cidades Inteligentes* (i.e., *Smart Cities*) tem ganhado bastante relevância na comunidade acadêmica nos últimos anos. Diversas iniciativas de sucesso podem ser encontradas tanto no Brasil quanto fora dele, como por exemplo o programa MIT *City Science*¹. As Cidades Inteligentes podem ser definidas como sistemas complexos que envolvem pessoas com uma variedade de *expertises*, interagindo e utilizando um conjunto de serviços e ambientes para melhorar o desenvolvimento da cidade e, conseqüentemente, a qualidade de vida dos seus cidadãos [Shapiro 2006, Allwinkle and Cruickshank 2011].

De acordo com a Fundação Getúlio Vargas (FGV)², podemos caracterizar o nível de inteligência de uma cidade de acordo com nove dimensões, a saber: governança, administração pública, planejamento urbano, tecnologia, meio ambiente, conexões internacionais, coesão social, capital humano e economia. Em especial, nesse artigo, nos concentramos em uma dimensão específica: a administração pública, e mais detalhadamente, em políticas de segurança na administração pública.

Cidades inteligentes são, *a priori*, cidades seguras. Dessa forma, fornecer soluções para reduzir a criminalidade urbana é uma das principais prioridades. A criminalidade urbana é uma questão antiga, ainda que seja um problema em aberto, em muitos países, em especial o Brasil [Baldwin 1975, Sociales 2001, Gribanova et al. 2017]. O problema do aumento da criminalidade tem afetado praticamente todas as cidades brasileiras não distinguindo

¹<https://www.media.mit.edu/groups/city-science/overview/>

²<http://fgvprojetos.fgv.br/noticias/o-que-e-uma-cidade-Inteligente>

etnia, sexo ou classe social. A crescente emigração para as grandes cidades em busca de oportunidades de emprego, alavancada pela crise econômica nos últimos anos, tem gerado um aumento significativo dos índices criminais nos grandes centros urbanos. De acordo com o Instituto de Segurança Pública do estado do Rio de Janeiro³, os roubos de rua têm seguido uma trajetória inconstante, mas ascendente, no período entre janeiro de 2010 e dezembro de 2017. O mesmo vêm acontecendo com homicídios no mesmo período⁴, o que corrobora com o estudo realizado pelo Conselho Cidadão para a Segurança Pública e Justiça, o Brasil ocupa o décimo lugar das cidades mais violentas do mundo⁵.

Tais fatos têm levado diversas autoridades da segurança pública – nos diversos níveis governamentais – e a sociedade civil a refletirem sobre o desenvolvimento de ferramentas para criação de políticas públicas eficazes com o propósito de reduzir a violência. Uma das soluções propostas pela Polícia Militar do Estado do Rio de Janeiro foi a criação do Escritório de Gestão da Qualidade (EGQ). O EGQ possui o intuito de estudar as ocorrências de crimes a fim de desenvolver análises complexas capazes de oferecer subsídios aos gestores no momento da tomada de decisão. Entretanto, o EGQ ainda não possui um ambiente analítico que seja capaz de integrar dados das diversas esferas governamentais de forma a apoiar o processo de tomada de decisão para elaboração de políticas públicas.

Nas últimas décadas, uma abordagem capaz de prover tal capacidade analítica, chamada de *Data Warehouse* (DW), tem sido amplamente utilizada em diversos domínios [Golfarelli and Rizzi 2009], sejam eles acadêmicos ou comerciais [Inmon 1992]. DWs são bases de dados multidimensionais que integram informações de diversas fontes a fim de facilitar a análise de dados, reunindo e consolidando informações de diversos *Data Marts* (DM) [Inmon 1992]. Estes são um subconjunto dos DWs que possuem um objetivo específico, orientado por assunto, variante no tempo, e não volátil. Uma das maiores vantagens dos DMs frente aos bancos de dados transacionais é que eles possuem dados previamente sumarizados, *e.g.*, dados agregados por mês ou ano [Inmon 1992], o que acelera e facilita o processo de análise dos dados.

Dessa forma, o objetivo desse artigo é propor um ambiente de apoio à tomada de decisão baseado em *Data Warehouse* denominado CHORD (Criminal dasHbOaRd Decision making) que seja capaz de integrar informações de várias esferas governamentais e que possibilite a sumarização dos indicadores estratégicos criminais para análise pelos membros do EGQ. Além disso, tal ambiente tem como objetivo estar alinhado com o Sistema Integrado de Metas do estado do Rio de Janeiro, sendo capaz de estabelecer uma comparação entre os valores reais dos indicadores de crimes e suas metas pré-estabelecidas pela Secretaria de Segurança Pública.

Esse artigo se encontra organizado em 5 seções além da Introdução. A Seção 2 apresenta o referencial teórico. A Seção 3 apresenta a abordagem proposta chamada CHORD. A Seção 4 apresenta a avaliação experimental. A Seção 5 discute trabalhos relacionados, e, finalmente, a Seção 6 conclui o artigo e discute trabalhos futuros.

2. Referencial Teórico

Esta seção apresenta os principais conceitos associados a abordagem proposta tanto no que tange conceitos de segurança pública quanto de *Data Warehouse*.

³<http://www.isp.rj.gov.br>

⁴<https://www.ispgeo.rj.gov.br>

⁵<http://www.seguridadjusticiaypaz.org.mx/biblioteca/prensa/send/6-prensa/239-las-50-ciudades-mas-violentas-del-mundo-201>

2.1. O Sistema Integrado de Metas do Estado do Rio de Janeiro

O Sistema Integrado de Metas (SIM) foi desenvolvido pela Subsecretaria de Planejamento e Integração Operacional (SSPIO), órgão subordinado a Secretária Estadual de Segurança Pública do estado do Rio de Janeiro - SESEG/RJ, sendo publicado no Decreto Estadual nº 41.931 de 26 de novembro de 2009.

Este sistema tem como objetivo desencadear ações integradas entre as Polícias Civil e Militar com o propósito de prevenir e controlar a ocorrência de crimes do estado do Rio de Janeiro por meio do controle de indicadores estratégicos criminais. Esses indicadores agrupam delitos em categorias, são elas:

- *Letalidade Violenta*: composto pelos crimes de Homicídio Doloso, Latrocínio (roubos seguidos de morte), Lesão Corporal Seguida de Morte e Homicídio Decorrente de Oposição à Intervenção Policial;
- *Roubo de Veículo*: Composto pelos crimes de roubo de veículo e roubo de moto;
- *Roubo de Rua*: composto pelos crimes de Roubo a Transeunte, Roubo em Coletivo e Roubo de Aparelho Celular;

Ainda segundo o decreto, para melhorar o controle dos indicadores estratégicos criminais, o estado foi dividido em 137 Circunscrições Integradas de Segurança Pública (CISP), com numerações não sequenciais agrupadas em 39 Áreas Integradas de Segurança Pública (AISP), que por sua vez estão agrupadas dentro de sete Regiões Integradas de Segurança Pública (RISP). A Figura 1 apresenta um exemplo da organização do estado do Rio de Janeiro para controle dos indicadores estratégicos.

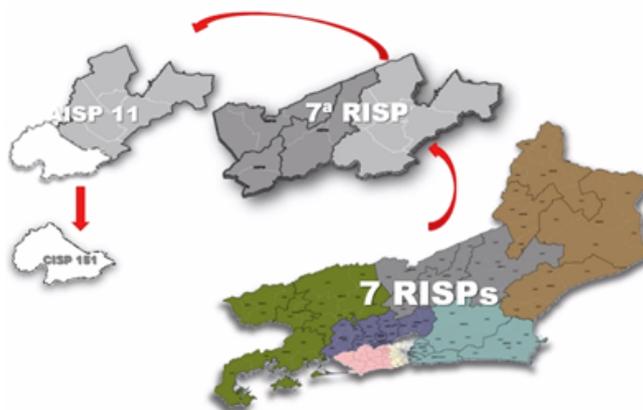


Figura 1. Divisão Territorial do Estado do Rio de Janeiro em RISP, AISP e CISP - Fonte: www.sistemademetas.rj.gov.br

Dessa forma, as polícias Civil e Militar adaptaram seus departamentos e comandos a fim de criar equivalências em suas estruturas especiais e a estrutura especial da Secretaria de Segurança Pública. No que tange a Polícia Militar do estado do Rio de Janeiro, esta se organizou da seguinte forma: (i) Criou sete Comandos de Policiamento de Área (CPA), cada um sendo equivalente a uma RISP; (ii) Criando 39 Batalhão de Polícia Militar agrupados em CPA, sendo cada batalhão equivalente uma AISP; e (iii) Cada batalhão possui em sua área de policiamento uma quantidade de delegacias, sendo que cada delegacia equivalente a uma CISP.

2.2. Uma Breve Introdução à Análise Criminal

Com os indicadores estratégicos, o Escritório de Gestão da Qualidade realiza o monitoramento utilizando técnicas oriundas de diversas áreas de conhecimento, denominada

Análise Criminal (AC). A AC serve como fonte para a confecção dos relatórios gerenciais que oferecem apoio às tomadas de decisão. De acordo com [Magalhães 2007], a AC possui três vertentes (Estratégica – Tática – Administrativa), sendo dessa forma o maior vetor de produção de conhecimento específico para a Gestão da Segurança Pública, possibilitando revelar com clareza as características do crime a partir de eventos desconexos.

A Análise Criminal Administrativa (ACD) refere-se à aplicações de técnicas estatísticas descritivas que mostram o comportamento dos indicadores ao longo de um determinado período (*i.e.* série histórica). A realização de comparação dentre períodos sazonais tem como objetivo a verificação da variação dos indicadores, além da produção do conhecimento voltada a diversos públicos (Cidadãos, Gestores Públicos, Instituições Pública, Organismos Internacionais, Organizações Não-Governamentais, etc.) devidamente selecionados pelos gestores.

A Análise Criminal Tática (ATC) produz conhecimento a médio prazo que são utilizados nas atividades de polícia ostensiva e investigativa. Na atividade de polícia ostensiva são produzidos relatórios que indicarão as áreas de maior incidência dos delitos e, por vezes, indicando os *modi operandi* dos autores dos delitos, bem como o perfil dos envolvidos (vítima e autor) e a correlação entre eles. No que tange a atividade investigativa, o mesmo conhecimento agregado de mais informações sobre os envolvidos (vítima e autor) possibilita a busca da solução dos casos investigados.

A Análise Criminal Estratégica (ACE) diz respeito à produção do conhecimento tendo como direcionamento o estudo dos fenômenos criminais a longo prazo. Nesta vertente, os relatórios servem como base para o planejamento e desenvolvimento de soluções capazes de produzir políticas públicas por meio da interação entre instituições que possuem influência na segurança pública.

Em todas as vertentes da análise criminal, o analista realiza inicialmente o processo de extração de dados oriundo de um ambiente transacional (OLTP), onde os usuários executam várias tarefas, tais como: inclusões, alterações, exclusões e pesquisas, além de dados oriundos de planilhas eletrônicas, textos, mapas entre outros tipos de arquivos.

Em um terceiro momento, o analista realiza a inferência fundamentada na aplicação de equações estatísticas e probabilísticas, bem como o uso de técnicas de geoprocessamento a fim de responder as perguntas ("O que?", "Onde?", "Quando?", "Quem?" e "Como?") necessárias ao encontro da solução para os eventos criminosos.

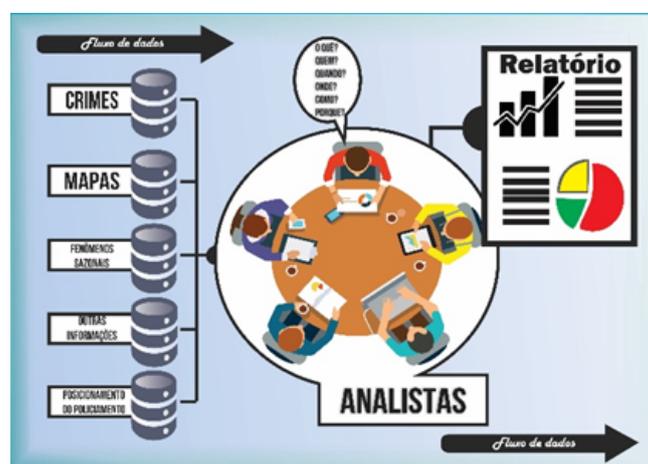


Figura 2. Fluxo de Dados no Processo de Análise Criminal

A última etapa é a construção dos relatórios e/ou *dashboards*. Neles são expostos os resultados encontrados nas análises realizadas pelos analistas utilizando técnicas básicas de visualização de dados (*i.e.* exposição de gráficos com métricas estatísticas como: contagens, totais, diferenças absolutas, diferenças percentuais, frequências acumuladas e percentuais acumulados). A Figura 2 apresenta o fluxo dos dados ao longo do processo de análise.

2.3. Data Warehousing e Modelagem Dimensional

Segundo [Kimball and Ross 2002], a modelagem dimensional é uma técnica de projeto de bancos de dados que visa apoiar consultas analíticas. Faz-se uso de redundâncias planejadas dos dados para melhorar o desempenho das consultas [Kimball and Ross 2002, Inmon 1992]. O modelo dimensional de um banco de dados é composto pelas tabelas *Fato* com suas respectivas *Dimensões*. As dimensões podem ser compartilhadas por tabelas fato diferentes. Existem dois modelos de implementação e um banco de dados dimensional: o Modelo Estrela [Kimball and Ross 2002] e o Modelo Floco de Neve [Inmon 1992]. O Modelo Estrela possui a tabela fato centralizada com as suas respectivas dimensões no seu entorno. Nesse modelo, a tabela fato possui chaves estrangeiras para todas as suas dimensões, sendo um modelo desnormalizado. O Modelo Floco de Neve é uma variação do Modelo Estrela, no qual todas as dimensões são normalizadas, fazendo com que sejam geradas quebras na tabela original ao longo de hierarquias existentes em seus atributos.

Um Data Warehouse (DW) é constituído pela união dos Data Marts (DMs). Assim, como nos DMs, um DW preferencialmente deve ser modelado de forma dimensional, pois em comparação com um banco de dados transacional e normalizado, a modelagem dimensional produz modelos mais previsíveis e compreensíveis, facilitando a utilização e assimilação pelos usuários finais (no contexto desse artigo, os analistas de segurança), além de possibilitar consultas com alto desempenho [Kimball and Ross 2002]. Portanto, a modelagem dimensional possui uma estrutura simplificada, mais próxima da visão que o físico tem do seu domínio, facilitando assim a compreensão, de forma que os próprios físicos possam criar suas consultas. Apesar de terem uma estrutura diferente de bancos de dados transacionais, os bancos de dados dimensionais podem ser modelados sobre Sistemas de Gerência de Bancos de Dados (SGBDs) relacionais como o MySQL ou o PostgreSQL.

Segundo [Raslan and Calazans 2014] antes que os dados sejam sumarizados e inseridos no DW, eles passam por um processo denominado ETL (*Extract, Transform, Load*). A extração envolve a leitura e compreensão dos dados de origem e cópia destes para a *staging area* para serem manipulados posteriormente. Na fase de transformação há o processo de filtragem dos dados, correção de possíveis erros de digitação, tratamento de conteúdos ausentes de atributos, a combinação de dados de diversas origens, exclusão de atributos que não fazem parte do domínio, além da criação de novas chaves (*surrogate keys*). A carga de dados (*load*) é o processo de inserção dos dados no DW após o processo de transformação, sendo portanto, um processo interativo e incremental ao longo do tempo. O ETL apresenta os mesmos processos realizados manualmente pelos analistas do EGQ atualmente.

3. Abordagem Proposta: CHORD

O CHORD (Criminal dasHbOaRd Decision making) é um ambiente de apoio a decisão baseado em DW para o Escritório de Gestão da Qualidade (EGQ). Utilizando um ambiente que integra dados de diferentes esferas governamentais, os especialistas em segurança do EGQ são capazes de delinear políticas de segurança mais eficazes. A arquitetura do CHORD é apresentada na Figura 3. O CHORD é composto por cinco componentes principais: (i) as fontes de dados, (ii) o componente ETL, (iii) o ODS (*Operational Data Store*), (iv) o *Data*

Warehouse CHORD-DW, que é composto por vários *Data Marts*, e (v) o *Dashboard*. A seguir discutiremos cada um dos componentes do CHORD.

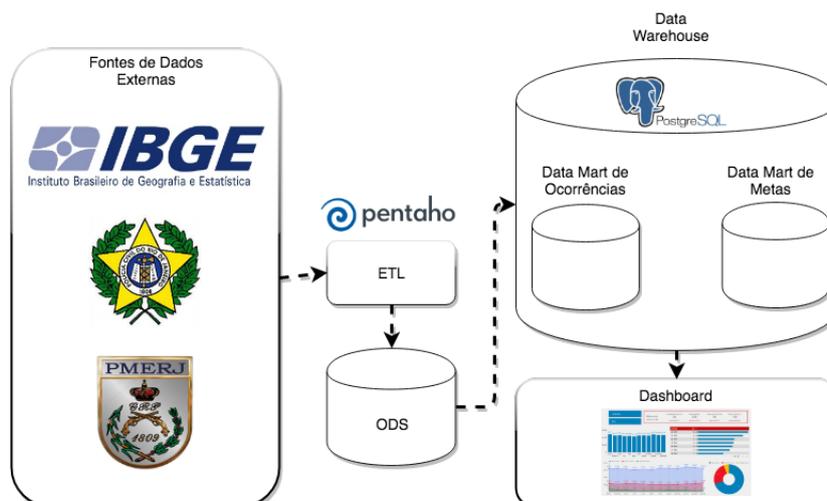


Figura 3. Arquitetura Conceitual do CHORD

3.1. Fontes de Dados

As fontes de dados representam bases de dados de diversas organizações que são utilizadas para alimentar o CHORD-DW. A principal fonte de informações para o CHORD-DW deriva do Sistema Gerencial *Web* (Gweb). O Gweb é um ambiente transacional que armazena diariamente as informações dos registros nas delegacias de Polícia Civil, sendo possível extrair do sistema as principais informações de crimes que compõem o Sistema Integrado de Metas além de todas as variantes de ocorrências intituladas como roubo. No Gweb há a opção de extração de dados para arquivos no formato .CSV com estrutura predefinida. Além dos dados do Gweb, o CHORD também importa dados das metas que devem ser atingidas dentro de um período de tempo por todos os gestores. As metas são publicadas semestralmente no diário oficial do estado do Rio de Janeiro, além de estarem disponíveis no sítio do Instituto de Segurança Pública/RJ. Outra fonte importante de dados é a base de dados do IBGE de onde são importadas informações sobre a população residente, a população flutuante da região, etc.

3.2. O Componente ETL

O componente ETL é o responsável por executar todo o processo de extração de informações e carga no CHORD-DW. O processo se inicia identificando as origens dos dados que servirão como base para compor o *Operational Data Store* (ODS), tais como: a relação de comandantes das Áreas Integradas de Segurança Pública, a relação dos indicadores que serão monitorados, os títulos de crimes que os compõem, a origem das ocorrências de crimes e o mapa da área em que as ocorrências serão monitoradas. Essas informações chegam até os analistas por meio de planilhas eletrônicas extraídas do Gweb. A ferramenta escolhida para realizar o processo de ETL foi o Pentaho, uma vez que a suíte Pentaho é uma ferramenta livre que está em expansão na comunidade de *Business Intelligence*. O Pentaho verifica se existem atualizações ou novas informações a serem importadas das múltiplas fontes de dados, e, caso haja, o mesmo realiza o processo de carga no ODS. O ODS é uma base de dados intermediária que contém dados das várias fontes. Diferentemente do CHORD-DW, o ODS não tem o objetivo de armazenar o histórico de dados e de servir como base para consultas mais complexas.

3.3. Construção dos *Data Marts* do CHORD-DW

O CHORD-DW é um componente-chave do CHORD, uma vez que todo o processo analítico depende dos dados que são integrados. Nessa sub-seção apresentamos como os DMs que compõem o CHORD-DW foram modelados. [Monteiro et al. 2013] afirma que os DWs não necessitam ser construídos de uma vez, pois a complexidade se torna bastante elevada. Em vez disso, sugere-se abordar processos de negócios de maneira incremental, e, assim, conceber o DW em partes. Neste sentido, optou-se em construir o DW utilizando a metodologia *bottom-up*, onde a construção do CHORD-DW começa a partir de uma pequena parte (assunto), *i.e.* DMs. A seguir são descritos os processos de construção dos DMs para o acompanhamento dos crimes ocorridos no estado do Rio de Janeiro, bem como a comparação com as metas estabelecidas para os indicadores.

A Figura 4 apresenta o Modelo CHORD-DW. Foi utilizada a modelagem dimensional estrela [Kimball and Ross 2002], sendo a tabela fato de ocorrências de crimes representada pela tabela *fatoOcorrencia*, responsável por armazenar os fatos, que no contexto desse artigo são as ocorrências de crimes, armazenados, respectivamente, nos atributos *valorLetalidadeViolenta*, *valorRouboVeiculo* e *valorRouboRua*. Ainda na tabela *fatoOcorrencia*, os atributos *idArea*, *idTempo* e *idTitulo* são chaves estrangeiras para as tabelas que representam as dimensões tempo, área e título, respectivamente.

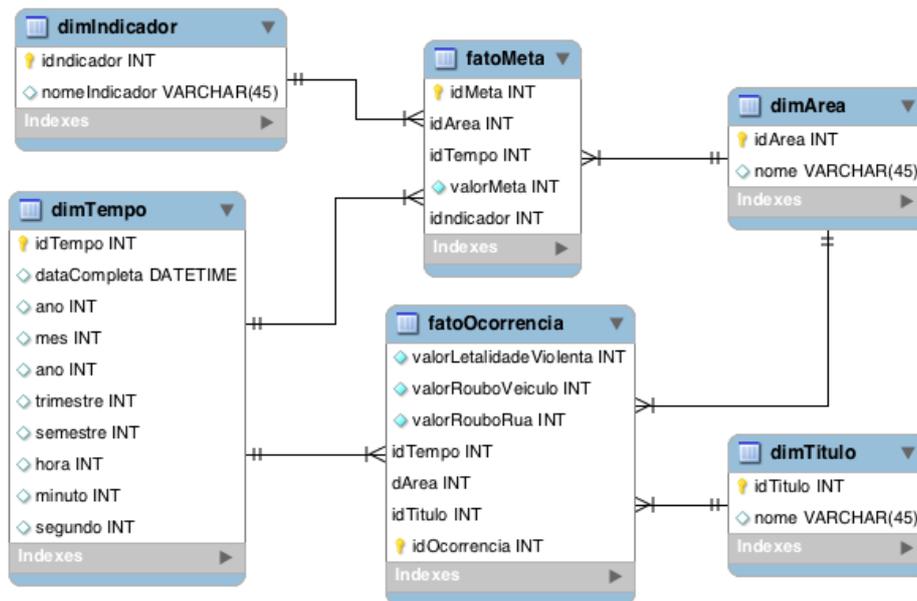


Figura 4. Modelagem do *Data Warehouse* do CHORD

A tabela fato que corresponde às metas que devem ser cumpridas por cada área se encontra representada pela tabela *fatoMeta*, responsável por armazenar as metas de cada área ao longo do tempo. Os valores das metas são armazenados no atributo *valorMeta*. Ainda na tabela *fatoMeta*, os atributos *idArea*, *idIndicador* e *idTempo* são chaves estrangeiras para as tabelas que representam as dimensões tempo, indicador e área, respectivamente.

A dimensão área *dimArea* agrupa todas as divisões e subdivisões do estado do Rio de Janeiro que estão previstas no Decreto Estadual nº 41.931 de 26 de novembro de 2009. Com exceção da Circunscrição Integrada de Segurança Pública, todas as divisões possuem um comandante da Polícia Militar. A tabela *dimTempo* representa a dimensão tempo, sendo responsável por armazenar todas as possíveis granularidades de tempo para um determinado

fato. O atributo *dataCompleta* representa a data completa (até milissegundos). Os atributos *ano*, *mes*, *dia*, *trimestre*, *semestre*, *hora*, *minuto* e *segundo* são as representações numéricas de partes da data completa. O atributo *idTempo* é a chave primária da tabela. A tabela *dimIndicador* associa a qual indicador um determinado crime está associado e possui o atributo *nomeIndicador*. A tabela *dimTitulo* representa os títulos de crimes que estão impactando diretamente os números do indicador em que pertence.

4. Avaliação Experimental

De forma a avaliar o CHORD, realizamos uma avaliação experimental com uma amostra de ocorrências de crimes ocorridos no estado do Rio de Janeiro durante os anos 2015, 2016 e 2017. A tabela *fatoOcorrencia* possui aproximadamente 500.000 registros. A Tabela 1 exhibe três consultas e seus tempos de execução associados (por limitações de espaço apresentamos apenas três consultas). Tal medição tem como objetivo demonstrar a eficiência adquirida em relação ao banco de dados transacional existente no EGQ. Para cada consulta apresentamos seu tempo de execução em milissegundos obtido por meio da sua execução via *Dashboard* do CHORD, apresentado na Figura 5.

Tabela 1. Desempenho das Consultas no CHORD

Consulta	Tempo de Execução
Consultar o somatório dos valores dos índices de criminalidade durante todos os anos	7.678 ms
Consultar o somatório dos valores dos índices de criminalidade durante o ano de 2017	3.144 ms
Consultar o somatório dos valores dos índices de criminalidade durante o ano de 2016 no 18º BPM	2.432 ms

Na primeira consulta foi explorada a agregação dos valores dos campos *valorLetalidadeViolenta*, *valorRouboVeiculo* e *valorRouboRua* em todos os anos. A segunda consulta explora agregação dos valores dos campos *valorLetalidadeViolenta*, *valorRouboVeiculo* e *valorRouboRua* durante o ano de 2017. A terceira consulta explora agregação dos valores dos campos *valorLetalidadeViolenta*, *valorRouboVeiculo* e *valorRouboRua* durante o ano de 2016 no 18º Batalhão de Polícia, que compreende os bairros de Jacarepaguá, Pechincha, Freguesia, Tanque, Vila Valqueire, Taquara, Curicica, Cidade de Deus, Anil e Gardênia Azul. Todas as consultas se beneficiam de agregações pré-calculadas presentes no CHORD-DW e, por isso, apresentam um tempo de execução bastante reduzido. É importante ressaltar que a amostra contém apenas três anos, e os tempos de execução tendem a aumentar quando todos os dados estiverem carregados no CHORD-DW. Este experimento reforça a necessidade de se utilizar agregações pré-calculadas neste cenário.

5. Trabalhos Relacionados

A segurança pública é uma prioridade absoluta tanto para a indústria quanto para a comunidade científica, principalmente devido aos benefícios potenciais que as abordagens propostas nessa área podem proporcionar à sociedade e aos seus cidadãos.

A abordagem mais proeminente nesse sentido é o *PredPol*⁶, um *software* comercial usado pelo Departamento de Polícia de Atlanta. O *PredPol* considera a história dos crimes para prever a incidência de novos crimes na cidade. Apesar de seu aparente sucesso, já que

⁶<http://www.predpol.com/>



Figura 5. Um fragmento do *Dashboard* do CHORD

a utilização do sistema foi capaz de ajudar na redução de 32% dos arrombamentos naquela cidade, não está claro como o sistema armazena seus dados ou se integra diferentes bases de dados.

Outras iniciativas mais simples usadas pelos departamentos de polícia em todo o mundo incluem os mapas de crime fornecidos pela Polícia do Reino Unido⁷ e do departamento de Polícia de Nova York⁸. No Brasil, podemos citar o portal "Onde fui Roubado"[Roubado 2013] que é uma abordagem *crowdsourcing* para coleta e visualização de ocorrências de crimes na *Web*. No entanto, tais abordagens apresentam apenas ocorrências criminais em mapas visuais, sem que seja realizado o cruzamento de tais dados com outras bases de dados.

6. Conclusões e Trabalhos Futuros

Infelizmente, nos últimos anos, os cidadãos dos grandes centros urbanos, em especial do estado do Rio de Janeiro, sofreram com ondas de violência significativamente maiores. Ambientes e soluções que fornecem capacidade analítica aos gestores têm o potencial de fornecer soluções para aliviar este problema.

Neste artigo, demos um passo nessa direção, propondo o ambiente CHORD (Criminal dasHbOaRd Decision making). O CHORD é baseado no conceito de DW que integra dados de bancos de dados de diversas esferas governamentais. O CHORD-DW, que é o componente-chave do ambiente CHORD, segue uma modelagem dimensional do tipo estrela [Kimball and Ross 2002] que permite que dados sejam armazenados de forma pré-calculada no CHORD-DW, acelerando assim consultas que eram bastante lentas anteriormente. Foi realizada uma avaliação experimental do CHORD utilizando-se um sub-conjunto dos dados

⁷<https://www.police.uk>

⁸(<https://maps.nyc.gov/crime/>)

de ocorrências de crimes nos anos de 2015, 2016 e 2017 (aproximadamente 500.000 registros), e constatou-se que o tempo de execução de todas as consultas é aceitável.

Como sugestão de trabalho futuro, podemos considerar a identificação de padrões nos dados do CHORD, utilizando-se de técnicas de mineração de dados e aprendizado de máquina. A aplicação de técnicas de mineração de dados possibilita a indução de padrões interpretáveis que podem auxiliar os departamentos de polícia na prevenção de crimes.

Agradecimentos

Os autores agradecem a CNPq, CAPES e FAPERJ por financiarem parcialmente o trabalho aqui apresentado.

Referências

- Allwinkle, S. and Cruickshank, P. (2011). Creating smart-er cities: An overview. *Journal of urban technology*, 18(2):1–16.
- Baldwin, J. (1975). Urban criminality and the ‘problem’ estate. *Local Government Studies*, 1(4):12–20.
- Golfarelli, M. and Rizzi, S. (2009). *Data Warehouse Design: Modern Principles and Methodologies*. McGraw-Hill, Inc., New York, NY, USA, 1 edition.
- Gribanova, G., Vulfovich, R., et al. (2017). Modern city safety as a complex problem. *Public administration issues*, (5):83–100.
- Inmon, W. H. (1992). *Building the Data Warehouse*. John Wiley & Sons, Inc., New York, NY, USA.
- Kimball, R. and Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, Inc., New York, NY, USA, 2nd edition.
- Magalhães, L. C. (2007). Análise criminal e mapeamento da criminalidade–gis. *Anais do Fórum Internacional de Gabinetes de Gestão Integrada. São Luís, Maranhão*.
- Monteiro, A. V. G., Pinto, M. P. O., and da Costa, R. M. E. M. (2013). Uma aplicação de data warehouse para apoiar negócios. *Cadernos do IME-Série Informática*, 16:48–58.
- Raslan, D. A. and Calazans, A. T. S. (2014). Data warehouse: conceitos e aplicações. *Universitas: Gestão e TI*, 4(1).
- Roubado, O. F. (2013). Disponível em: <http://ondefuiroubado.com.br>. *Acessado em, 27*.
- Shapiro, J. M. (2006). Smart cities: quality of life, productivity, and the growth effects of human capital. *The review of economics and statistics*, 88(2):324–335.
- Sociales, P. (2001). Crime as a social cost of poverty and inequality: a review focusing on developing countries. *Facets of Globalization*, page 171.