

Predição de Ocorrências Policiais em Natal: Uma Abordagem em Análise de Séries Temporais

Adelson Dias de Araújo Júnior
Departamento de Engenharia de Computação e
Automação
UFRN
Natal, Brasil
adelsondias@live.com

Allan de Medeiros Martins
Departamento de Engenharia Elétrica
UFRN
Natal, Brasil
allan@dec.ufrn.br

Rémi Verdier
Génie Mathématique et Modélisation
INSA
Toulouse, France
rverdier@etud.insa-toulouse.fr

Nélio Alessandro Azevedo Cacho
Departamento de Informática e Matemática Aplicada
UFRN
Natal, Brasil
neliocacho@dimap.ufrn.br

Resumo—Integrar inteligência aos processos diários de gestão de uma cidade é uma tarefa de grande importância. Por isso, na área de segurança pública, tratar de policiamento preditivo é uma urgência quando se quer construir uma cidade inteligente. Neste trabalho, é apresentado uma predição quantitativa das ocorrências policiais da cidade do Natal, utilizando a abordagem de análise de séries temporais. Levando-se em conta regiões estratégicas adotadas pela polícia, este trabalho, então exemplifica a construção de um modelo ARIMA para uma dessas regiões a fim de descrever o comportamento da frequência de ocorrências observadas. (*Abstract*)

Palavras-chave: policiamento preditivo; séries temporais; ARIMA.

I. INTRODUCTION

É bastante intuitiva a ideia de que as cidades passem a se integrar com soluções de tecnologia a medida em que o tempo passa. E com toda a necessidade de integração de modelos científicos e tecnológicos criados pela humanidade em um aspecto maior de convivência, o conceito de cidades inteligentes passou a tomar maior atenção nas ciências de computação, de comunicações, urbanística, antropológica entre várias outras. Segundo [1], o conceito de cidade inteligente ainda continua bastante nebuloso, mas é possível se dizer que uma das suas principais características é a capacidade de prestar atenção ao papel da sociedade e do capital relacionado ao desenvolvimento urbano. Ainda, [1] se refere ao poder de usar o capital tecnológico em seu benefício, para fazer a comunidade aprender sobre si mesmo, adaptar-se mais facilmente e, ainda, inovar.

Nessa perspectiva, a área da segurança pública possui total interseção na esfera ação de uma cidade inteligente, e trazer inteligência à segurança de uma cidade provavelmente vem a ser uma missão mais antiga do que o conceito discutido anteriormente. Um desses esforços antigos é a busca pelo policiamento preditivo [2], que consiste em desenvolver estratégias eficientes que previnam crimes ou investigações

desnecessárias. Não é como numa bola de cristal, em que se sabe onde e quando ocorrerão eventos que necessitem de intervenção policial, entretanto, serão produzidos resultados que podem guiar chefes de polícia em torno da gestão dos recursos limitados da segurança pública da cidade.

Nessa perspectiva, diversos trabalhos apresentam soluções para o policiamento preditivo em si. Um deles [3] aborda diversas técnicas para implementações de análise preditiva de crimes em cidades pequenas, discutindo aquisição da informação e análises de regressão linear em dados georreferenciados. Entretanto, neste trabalho, o foco consiste em se determinar modelos preditivos para regiões estratégicas, definidas pela polícia (ditas AISPs) da cidade do Natal, no Brasil, que ajustem os dados quantitativos das ocorrências policiais em torno de séries temporais únicas para cada uma dessas regiões as quais descreverão o comportamento da frequência dessas ocorrências ao longo do tempo.

Para alcançar esses modelos preditivos, a metodologia de análise de séries temporais utilizada é a Box-Jenkins [4], amplamente utilizada na literatura da área. Essa metodologia consiste em um processo iterativo de identificação, estimação e diagnóstico do modelo auto-regressivo integrado de médias móveis (ARIMA) correspondente da série temporal, de forma que a predição em um tempo futuro possua menor erro possível em torno das métricas de diagnóstico, ajustando, assim, o comportamento do crime em Natal.

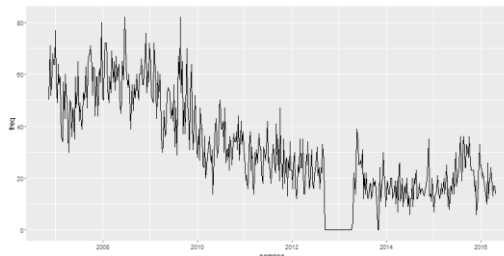


Figure 1. Série temporal para AISP 12 somente com o grupo “desordem”.

II. PRÉ-PROCESSAMENTO

Antes de tudo, foram necessários alguns esforços para enriquecer o experimento feito. A proposta é selecionar um modelo de série temporal para cada região da cidade, chamadas AISP, e para isso, é necessário incluir aos dados a informação sobre a AISP correspondente ao ocorrido. Através da utilização de consultas de cruzamentos espaciais em sistemas de gerenciamento de banco de dados com extensão espacial, foi possível se adicionar a coluna referente à AISP nos dados das ocorrências. Após isso, também foi necessário classificar os tipos de ocorrência em um número pequeno de grupos para que o modelo correspondente a AISP possua um grau maior de semântica, já que será tratado por grupos de ocorrências específicos, sem levar em consideração ocorrências domésticas, por exemplo. Os grupos foram feitos manualmente agrupando tipos de ocorrência semelhantes entre si.

III. ARIMA PARA UMA AISP E UM GRUPO

De maneira sintética, a predição pode ser resumida ao se explicitar o processo para estimar um modelo ARIMA para somente uma AISP e um grupo de ocorrência, que no caso é o grupo de “desordem”. A série temporal correspondente é ilustrada na figura 1, que possui granularidade temporal de semanas. Na série, é bastante visível que existe uma lacuna no período de Setembro de 2012 a Abril de 2013. Neste caso, a solução para não desvirtuar a análise foi utilizar somente os dados pós-lacuna, que correspondem a 107 semanas de ocorrências. Para a formação do modelo, foram utilizadas 70 semanas chamadas de amostras de aprendizado, e as 37 restantes foram utilizadas para validação da predição.

Na etapa de identificação da série, que corresponde a encontrar o grau de diferenciação (d), quantos parâmetros auto-regressivos (p) e quantas médias-móveis serão utilizados, são necessários testes de estacionariedade da série para cada diferenciação d e funções de auto correlação e auto correlação parcial para identificar os parâmetros p e q do modelo ARIMA. No caso, foi identificado o modelo ARIMA(3,1,0). A partir dessa conclusão, pôde-se partir para estimação, que resultou no seguinte modelo regressivo:

$$(I - (-0.6181)B - (-0.3830)B^2 - (-0.2869)B^3)(I - B) X_w = a_w$$

Em que B representa o operador lag e I a matriz identidade.

Para o diagnóstico, a métricas do MAPE (em português, erro percentual absoluto médio) foi utilizada em conjunto com os testes de hipótese comumente utilizados, como Shapiro-Wilk para normalidade dos resíduos e Ljung-Box para independência dos resíduos. O valor do MAPE encontrado foi de 0.3420, o p-valor do teste Shapiro-Wilk foi de 0.181 e o p-valor do teste Ljung-Box foi de 0.659. Isso significa que

para um nível de significância de 5%, o modelo escolhido não é rejeitado pelos testes feitos.

Por fim, os plots da predição são mostrados a seguir. Primeiro, na figura 2 é mostrado a comparação entre a predição e os valores observados para a amostra de aprendizado. Na figura 3, da mesma forma, podemos comparar a predição com a observação para a amostra de validação. Analisando as curvas, é possível enxergar que a predição, de fato, representa bem a série temporal, mas ela aparenta apresentar uma espécie de atraso em ambas as situações. Esse fato pode ser utilizado na inferência de amostras futuras desconhecidas para atrasar levemente a curva de predição realizada.

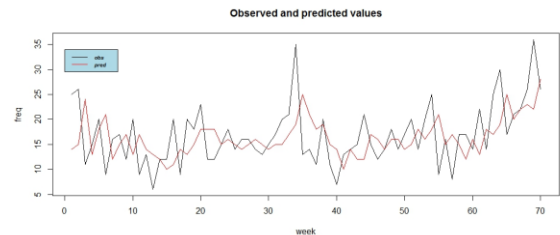


Figure 2. Predição para a amostra de aprendizado (70 semanas)

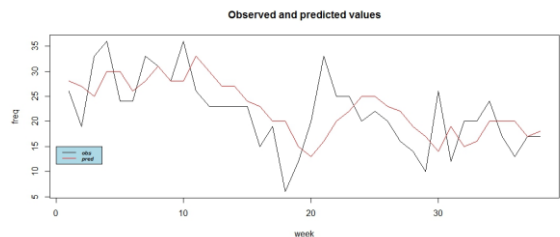


Figure 3. Predição para a amostra de validação (37 semanas)

Sendo assim, a avaliação das ocorrências policiais em torno de uma metodologia de predição pode ser realizada com técnicas de análise de série temporal, as quais buscam refletir como pode ser o comportamento futuro dessas variáveis. Esse esforço de policiamento produtivo pode incitar diferentes estratégias na alocação de recursos de segurança pública, fazendo com que a eficácia das estratégias montadas comprometa previamente a ação dos malfeitores.

REFERENCES

- [1] Caragliu, A., Del Bo, C., Nijkamp, P. (2009), “Smart Cities in Europe” In: 3rd Central European Conference in Regional Science – CERS.
- [2] Perry, Walt L. Predictive policing: The role of crime forecasting in law enforcement operations. Rand Corporation, 2013.
- [3] Lamas, João Paulo Campolina. “Predição de crimes e otimização de ações de segurança pública para cidades de pequeno porte utilizando geotecnologias.” (2013).
- [4] Box, George EP, et al. Time series analysis: forecasting and control. John Wiley & Sons, 2015.