

MapReduce aplicado na detecção de fracionamento irregular de compras públicas

Lucas de Souza Siqueira¹; Leandro da Silva Foly²

¹ Estudante no Instituto Federal Fluminense; ² Professor no Instituto Federal Fluminense

lucassouza.ti@gmail.com

Resumo

Apesar do avanço do Brasil nas legislações sobre Transparência Pública, casos de corrupção ainda acontecem com frequência, lesionando a Administração Pública. A legislação sobre compras e contratações públicas dita que o processo por meio de licitação deva ser a regra, salvo exceções previstas em lei. Porém ocorre que em alguns casos, gestores públicos recorrem a estes subterfúgios para driblar o certame legal. A análise de dados dos portais de transparência poderia auxiliar na detecção destes casos, porém muitas vezes eles são publicados pelos entes públicos em formato não estruturado e em grandes volumes de dados, dificultando que sejam analisados de formas manuais. Neste trabalho implementou-se técnicas do *Big Data*, a fim de atestar a eficiência de sua aplicabilidade na detecção de fracionamento irregular de compras públicas. Utilizou-se como apoio um portal em que os dados dos entes públicos podem ser disponibilizados através de interfaces de programação. O resultado se mostrou exitoso, ainda que seja necessário por ora, a intervenção humana na análise de dados.

Palavras-Chave: Big Data. Transparência. Fiscalização Pública.

Introdução

O Brasil atualmente possui uma legislação robusta no âmbito da Transparência Pública. Em 2000 já se tinha a Lei Complementar nº 101 (Lei da Responsabilidade Fiscal), em que se definia alguns pontos de transparência do setor público. A partir da Lei Complementar nº 131 de 2011, estes pontos passaram a ter mais clareza, pois a partir desta lei especificavam-se os dados que deveriam estar disponíveis através dos sites institucionais dos órgãos públicos.

Desta forma, entende-se do que se trata a Transparência Ativa, pois o governo deve ativamente divulgar esses dados dentro de prazos legais, sem a necessidade que alguém os solicite. Dentre esses dados estão as informações referentes às compras, licitações e contratos públicos, essenciais para que a sociedade fiscalize e comprove como os gastos públicos estão sendo executados.

Por mais robusta que seja a legislação da Transparência Pública brasileira, casos de corrupção continuam acontecendo, mesmo nas compras públicas, onde os certames processuais devem ser ditados nos princípios da isonomia da eficiência da Administração Pública. Entre casos de fraudes, pode-se citar as que acontecem em processos licitatórios, onde empresas pagam propina para pessoas envolvidas na elaboração do processo, a fim de facilitar que empresas com preços superfaturados ganhem a licitação.

No entanto, mesmo com estes casos de fraude, a licitação deve ser a regra que a Administração Pública deve seguir na efetuação de compras e contratações, ao menos nas exceções previstas em lei, segundo o inciso XXI do art. 37 da Constituição Federal, regulamentado pela Lei 8.666 de 21 de Junho de 1993. Pois ela tem mecanismos que visam garantir a competitividade e a publicidade do certame. Então se com a licitação existem casos de fraude, sem ela a corrupção fica ainda mais fácil.

O que costumeiramente vem a acontecer, seja por desonestidade ou por falta de planejamento de algum gestor público, é o fracionamento irregular de compras ou contratações pela administração pública (BARRETO, 2017). A lei prevê como exceção e situação de dispensa de licitação quando o valor total do processo não ultrapassa determinado valor. Porém, não é incomum que no mesmo exercício fiscal, o mesmo órgão público compre o mesmo item ou contrate o mesmo serviço, através de um processo de dispensa de licitação, onde a soma destes processos ultrapassa o limite legal. Neste caso, caracteriza-se o fracionamento irregular de serviços ou compras.

Mas ao tentar acessar um portal de transparência de algum órgão público a fim de tentar identificar essa situação, muita das vezes depara-se com um emaranhado de dados não estruturados e em grandes volumes. Fazer essa análise por meios manuais resultaria num processo dispendioso e ineficaz. Dessa forma, o auxílio de técnicas computacionais se torna conveniente, como as técnicas da área denominada *Big Data*.

Portanto o objetivo deste trabalho foi avaliar a aplicabilidade de técnicas do *Big Data* como auxílio na fiscalização pública através de dados públicos divulgados em portais de transparência, através de um estudo de caso.

Metodologia ou Materiais e Métodos

Neste trabalho, utilizou-se uma área da computação chamada *Big Data* onde se manipula grandes volumes de dados que precisam de cuidados específicos. Através de técnicas específicas desta área realizou-se a análise de um conjunto de dados.

A princípio, constatou-se de forma preditiva um número incomum de dispensas de licitação em um determinado portal da transparência de um ente público. Neste caso, as dispensas são listadas como “contratos de dispensa de licitação”, onde todas elas individualmente não ultrapassavam o valor limite previsto em lei.

Os dados foram exportados deste portal da transparência, e posteriormente carregados em um portal online de apoio a fiscalização pública, chamado SLACS – Serviço Livre de Apoio ao Controle Social (SIQUEIRA, 2017). Lá eles ficaram disponíveis para serem consumidos através de uma API (*Application Interface Programming*). Então um algoritmo na linguagem Python foi desenvolvido a fim de reconhecer com mais facilidade possíveis casos de fracionamento irregular de compras públicas, através do consumo destes dados.

Este reconhecimento é feito através de uma análise da semelhança das descrições dos objetos das compras públicas. O cálculo é feito através de um comparativo entre a frequência de termos nos respectivos descritivos.

Resultados e discussão

Primeiramente, implementou-se um algoritmo de contagem de palavras baseada no modelo de programação *MapReduce*. Este modelo foi desenhado a fim de processar grandes volumes de dados em paralelo, dividindo o trabalho em um conjunto de tarefas independentes. Sua primeira implementação foi proposta pela empresa Google (DEAN & GHEMAWA, 2004).

Conforme pode ser visto na Figura 1, a função *mapReduce* recebe como parâmetro a variável *inputs*, que neste caso são descrições de contratos de dispensa recebidos pela API do SLACS.

A função *mapper* é responsável por aplicar técnicas de Processamento de Linguagem Natural através da biblioteca NLTK (*Natural Language Toolkit*) e depois são contabilizadas as ocorrências das palavras.

Logo após, a função *reducer* é responsável por somar todas as ocorrências de uma determinada chave, neste caso as palavras encontradas. Então a função *mapReduce* retorna um vetor com a palavra e todas as suas ocorrências.

```
def map_reduce(inputs):
    collector = defaultdict(list)

    for input in inputs:
        for key, value in mapper(input):
            collector[key].append(value)

    return [output
            for key, values in collector.iteritems()
            for output in reducer(key, values)]

def mapper(document):
    stopwords = nltk.corpus.stopwords.words('portuguese')
    punctuations = ['!', '"', '#', '$', '%', '&', '"', '(', ')', '*',
                    '!', '"', '#', '$', '%', '&', '"', '(', ')', '*',
                    '!', '"', '#', '$', '%', '&', '"', '(', ')', '*']
    for word in nltk.word_tokenize(document):
        if word not in stopwords and word not in punctuations:
            yield (word, 1)

def reducer(word, counts):
    yield (word, sum(counts))
```

Figura 1. Função *MapReduce* implementada.

Fonte: Autor.

Através da API de consumo do Portal SLACS, foi possível carregar no algoritmo todos os contratos do ente público escolhido. A função *loads* da biblioteca *Json* do Python, carrega um arquivo em formato JSON para um dicionário de dados Python.

Então foi implementada uma estrutura de repetição, onde foi atribuído um vetor de frequência de palavras para cada descrição dos contratos importados, formando uma nova lista de contratos com o atributo 'frequencia_palavras'.

Após esta contabilização, foi realizada uma comparação entre os contratos do ente público. Então, foi calculada uma porcentagem de semelhança ps_i de um contrato i , baseada na divisão do total de termos similares ts_i com outro contrato j , pelo total de palavras $tot_{palavras_i}$ da descrição do contrato em questão. O ts_i é definido pela a frequência de um termo ft_i que se repete no contrato j pela frequência do mesmo ft_j . Então pode se definir o valor de ps_i por:

$$ps_i = \frac{ts_i}{tot_{palavras_i}} \text{ onde } ts_i = \begin{cases} ts_i + 1 & \text{se } ft_i = ft_j \\ 0 & \end{cases}$$

O resultado disso foi um novo dicionário de dados, em que eram especificados dois contratos de dispensa de licitação e o percentual de semelhança dos mesmos. Estes dados foram tratados e exportados para uma planilha eletrônica do LibreOffice, para uma melhor análise visual. Foi aplicado um filtro para contratos com percentual de semelhança acima de 50%, e depois uma análise visual na qual se encontrou itens que possivelmente atenderam à hipótese do fracionamento irregular, como pode ser visto na Tabela 1.

Tabela 1. Contratos de dispensa com portencagem de semelhança superior a 50% selecionados.

Contrato	Descrição	Valor (R\$)
19 / 2016	Aquisição de material para atender as necessidades do Departamento de Tecnologia. Recurso Próprio.	448,00
44 / 2016	Aquisição de material para atender aos setores de licitação e tecnologia. Recurso Próprio.	7850,00

A soma do valor destes dois contratos de dispensa de licitação resulta em mais de R\$ 8.000, cujo valor até a data em que este estudo foi realizado era o máximo permitido para os casos em que não era necessário a realização de procedimento licitatório pelo valor da compra.

Conclusão

Este trabalho buscou avaliar a aplicabilidade de técnicas de *Big Data* como apoio à fiscalização pública, tendo em vista o grande volume de dados disponíveis em um portal de transparência.

Através do estudo de caso, conseguiu-se encontrar um possível caso de fracionamento irregular de compras públicas. Portanto, é possível concluir a eficiência do método para o respectivo fim.

Vale ressaltar que a técnica não trouxe o resultado de imediato e de forma totalmente automatizada. Ela auxiliou em fases, nas quais houve intervenção humana para conseguir-se chegar ao resultado e comprovar a hipótese.

Referências

- BRASIL. Lei 8.666, de 21 de Junho de 1993. Regulamenta o art. 37 inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8666cons.htm>. Acesso em: 22 set. de 2019.
- BRASIL. Lei 12.527, de 18 de Novembro de 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 27 jan. de 2017.
- BRASIL. Lei Complementar 101, de 10 de Maio de 2000. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/LCP/Lcp101.htm>. Acesso em: 22 set. de 2017.
- BARRETTO, Carlos R. O fracionamento de despesas na Lei n. 8666/93 e a imprevisibilidade, 2017. Disponível em: <<https://jus.com.br/artigos/58847/o-fracionamento-de-despesas-na-lei-n-8-666-93-e-a-imprevisibilidade>>. Acesso em: 22 set. 2019.
- BIG Data. Disponível em: <<http://www.gartner.com/it-glossary/big-data/>>. Acesso em: 22 set. 2019.

DEAN, J.; GHEMAWT, S. MapReduce: Simplified Data Processing on Large Clusters. 2004. Disponível em: <<https://static.googleusercontent.com/media/research.google.com/pt-BR/archive/mapreduce-osdi04.pdf>>. Acesso em: 22 set. de 2019.

SIQUEIRA, Lucas de Souza. Serviço Livre de Apoio ao Controle Social. 2017. 103 f. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação)-Instituto Federal Fluminense Campus Itaperuna, Itaperuna, 2017.