

PROCESSOS DE AQUISIÇÃO MAIS INTELIGENTES PELA ANÁLISE E AVALIAÇÃO DE PADRÕES DE CONSUMO

MARCELO FABIANO NASCIMENTO DE OLIVEIRA*
Capitão de Corveta (T)

SUMÁRIO

Introdução
Revisão de Literatura
Metodologia
Processos de aquisição por análise e avaliação de consumo
Mundo real: reflexo sobre as licitações
Limitações e trabalhos futuros
Conclusões

INTRODUÇÃO

O caso analisado transcorre no âmbito de uma Força Armada que conta com estrutura interna de postos de venda de vestuário, utilizando um processo de controle e indicação de reposição de estoque baseado em séries históricas realizadas item a item, ou seja, de forma isolada. O presente artigo pretende introduzir a ideia de busca de padrões sequenciais entre os itens de

vestuário, objetivando reduzir o esforço de produção de uma licitação, tanto em quantidade como em recurso financeiro utilizado, além de permitir aumento de conhecimento sobre a dinâmica interna de aquisição de produtos. Para tanto, este trabalho focou no estudo de identificação de padrões no tempo, porém não enveredou pela mineração de hierarquias (LU *et al.*, 1998) ou mineração de séries temporais (ALENCAR, 2007).

* Graduado em Ciência da Computação pela Universidade Federal do Rio de Janeiro (UFRJ). Pós-graduado em Gerenciamento de Projetos pela Escola de Engenharia da Universidade Federal Fluminense (UFF) e mestre em Engenharia de Sistemas e Computação pela Coppe/UFRJ (Msc). Serve na Diretoria de Finanças da Marinha (DFM).

Este artigo foi dividido da seguinte forma: Revisão de Literatura, Metodologia, Processos de aquisição por análise e avaliação de consumo, Mundo real: reflexo sobre as licitações, Limitações e trabalhos futuros, Conclusões e Referências Bibliográficas.

REVISÃO DE LITERATURA

Gerenciamento da Cadeia de Suprimentos

Segundo Bertaglia (2000), a Cadeia de Suprimentos corresponde ao conjunto de processos requeridos para obter recursos materiais, agregando-lhes valor, de acordo com a concepção dos clientes e consumidores, e disponibilizando-os no lugar (onde) e na data (quando) que os clientes e consumidores os desejam. Ainda segundo o autor, o trabalho com a cadeia requer o domínio de várias competências, como o conhecimento de processos e de padrões de mercado, influenciando sobremaneira o processo como um todo e suas componentes, como a logística (BALLOU, 2001).

Descoberta de Conhecimento

A descoberta de Conhecimento em Bases de Dados (do inglês Knowledge Discovery in Databases – KDD) surgiu com o propósito de fazer uso inteligente e eficiente de dados acumulados pelas

organizações ao longo do tempo. A Figura 1 descreve o processo:

Sua prioridade é a descoberta de padrões que possam ser utilizados positivamente. O processo consiste em Pré-Processamento (incluindo os movimentos de preparação dos dados), Mineração dos Dados e Pós-Processamento (depuração e/ou síntese dos padrões descobertos (CALIL, 2008).

Mineração de Dados

A Mineração de Dados é o ramo da Computação, com início nos anos 80, em que foram demonstradas as primeiras preocupações das organizações com as imensas massas de dados geradas e, continuamente, armazenadas, demandando novos espaços em disco a cada dia, porém sem utilidade aparente como diferencial competitivo para essas mesmas organizações (DEVEZA, 2011). A literatura, historicamente, atribui à Mineração de Dados a convergência de três importantes áreas de conhecimento: Inteligência Artificial, Banco de Dados e Estatística.

Mineração de Padrões Sequenciais (MPS)

Atribui-se a Agrawal *et. al* (1995), originalmente, a formulação do problema de Mineração de Padrões Sequenciais, considerada uma extensão da análise de Regras de Associação, porém avaliando não apenas as relações entre atributos, mas



Figura 1 – Etapas da Descoberta de Conhecimento



Figura 2 – Os quatro passos desenvolvidos na solução

também a ordem temporal da ocorrência dos eventos. Como ponto de partida, a técnica pressupõe a existência de um banco de dados *D*, contendo registros de transações de compras identificados por algum atributo e possuindo uma lista de itens e uma data correspondendo ao momento da execução da ação.

Ferramentas de Mineração de Padrões Sequenciais – Generalized Sequential Pattern (GSP)

Proposto também por Agrawal *et al.* (1995), o algoritmo GSP foi o introdutor da busca pela mineração dos chamados padrões sequenciais, baseando-se na abordagem de geração e teste de candidatos e sendo dividida em duas etapas: geração de candidatos e etapa de cálculo do suporte. Por sua vez, a etapa de geração de candidatos divide-se em junção e poda. O conceito de *K*-sequência é o número de vezes, em diferentes elementos da sequência, em que este padrão é observado. *Lk* representa o conjunto de padrões sequenciais frequentes. A pesquisa fez uso de outras ferramentas, entre elas o Sequential Pattern Discovery Using Equivalent Classes (Spade) (ZAKI, 2001) e o Sequential Pattern Mining (Spam).

METODOLOGIA

A metodologia aplicada está apoiada em quatro passos bem distintos:

Passo 1: Tratamento e Formatação dos dados originais. Aproximadamente, 4 milhões de registros e 656 mil compras foram utilizados para realizar o estudo, referentes aos anos de 2006 a 2016;

Passo 2: Mineração de padrões sequenciais. Aplicação dos algoritmos de mineração de padrões sequenciais;

Passo 3: Interpretação dos resultados. Desenvolvimento de cenários de curto prazo, ou seja, menos de 30 dias; e

Passo 4: Formulação da conclusão. Apresentação da melhor forma de montar a licitação (aquisição).

PROCESSOS DE AQUISIÇÃO POR ANÁLISE E AVALIAÇÃO DE CONSUMO

Nesta pesquisa, os dados foram organizados em três atributos principais: Sequence ID (SID); o identificador da compra (no caso, o número da nota fiscal); e, por fim, o conjunto de compras (chamado de *itemset*) realizado pelo indivíduo, respeitando a ordem dos acontecimentos, representado na Tabela 1.

SID	Identificador	Conjunto de Compras
1	100.001	(CALCA, SAPATO), (CINTO, BOLSA)
2	203.103	(CAMISA, BOLSA, CINTO)
3	410.345	(GORRO, BOTA, MEIA)

Tabela 1 – Representação das abordagens Ano ou por Ciclo

Abordagem: Ano ou por Ciclo

O propósito é investigar e identificar os melhores padrões recorrentes espalhados entre os dez anos de dados disponíveis, olhando individualmente cada ano ou num bloco único de todos os anos.



Figura 3 – Representação das abordagens Ano ou por Ciclo

Algumas medidas e conceitos foram desenvolvidos para facilitar o processo de análise. A Recorrência avalia quantas vezes um padrão específico foi observado, em um dado período. Assim:

Recorrência (x) = n , onde x é o padrão escolhido, n é o número de vezes que o padrão p aparece no período.

A Abrangência é uma medida espacial e/ou geográfica relacionada com a observação do padrão em uma região específica do Brasil e/ou navio. Lembrando que a

pesquisa foi realizada dentro do contexto das Forças Armadas. Este trabalho relacionou 15 regiões mais navio, compondo as 16 localidades possíveis.

Abrangência (x) = m , onde x é o padrão escolhido, m o número de localidades em que o padrão p surge.

Padrões Ordinários – conceito, concebido para este trabalho, que diz respeito ao padrão constituído de partes pertencentes ao mesmo tipo de uniforme; por exemplo, calça camuflada e coturno.

Padrões Extraordinários – conceito, também concebido para este trabalho, que trata dos padrões em que as partes constituídas não pertencem ao mesmo tipo de uniforme, como calça cinza e blusa cáqui.

Seleção do algoritmo para geração dos resultados

A aplicação do critério de escolha dos algoritmos permitiu a identificação da melhor ferramenta a ser utilizada para um cenário específico, em cada abordagem, sendo escolhidos aqueles com maiores Recorrência e Abrangência no cenário desejado.

Assim, os dados foram minerados por meio de todos os algoritmos mencionados neste trabalho. Para aqueles que superaram as metas de Recorrência e Abrangência num cenário específico,

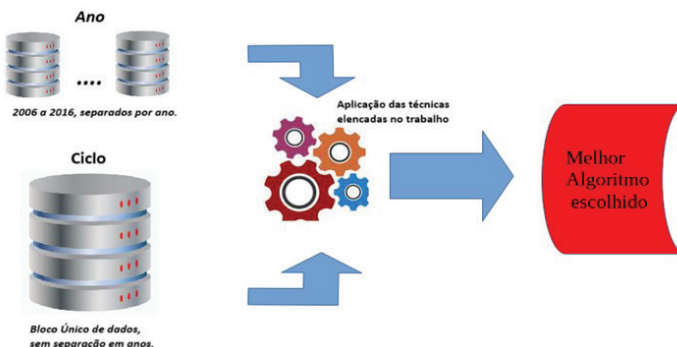


Figura 4 – Seleção do algoritmo para a geração de resultados

Médias para recorrência e abrangência para cada algoritmo							
Algoritmos	GSP	SPADE	SPAM	LAPIN	PREFIXPAN	CMSPADE	CMSPAM
Padrões extraordinários							
Recorrência	69,01	56,36	58,18	67,27	54,55	58,18	65,45
Abrangência	96,25	63,75	71,25	78,75	97,5	92,5	77,5
Padrões ordinários							
Recorrência	43,63	56,36	54,54	61,82	52,72	65,45	60,00
Abrangência	88,75	90	86,25	93,75	98,75	91,25	96,25

Tabela 2 – Seleção de algoritmos para Abordagem Ano

Médias para recorrência e abrangência para cada algoritmo							
Algoritmos	GSP	SPADE	SPAM	LAPIN	PREFIXPAN	CMSPADE	CMSPAM
Padrões extraordinários							
Abrangência	93,75	96,75	93,75	93,75	75	93,75	93,75
Confiança	100,00	97,87	100	100	100	35,84	19,45
Padrões ordinários							
Abrangência	87,5	100,00	93,75	93,75	87,5	93,75	93,75
Confiança	100,00	100,00	66,3	73,07	100	100	19,45

Tabela 3 – Seleção de algoritmos para Abordagem por Ciclo

este algoritmo foi usado para realizar a mineração e buscar padrões sequenciais.

Após a execução do processo de seleção, é possível verificar que o algoritmo GSP apresenta os melhores resultados para os padrões extraordinários e o CM-SPADE, para os ordinários.

Para a abordagem por Ciclo, verificamos que, para os padrões extraordinários, novamente, o algoritmo GSP apresenta os melhores resultados. Já para os padrões ordinários, o Spade se sobressai.

Análise de Padrões de Consumo

A seguir, serão apresentados os padrões de consumo obtidos segundo a seleção de algoritmos mencionados anteriormente.

Abordagem Ano

– Padrões Ordinários

A análise dos dados permitiu a obtenção de oito padrões considerados ordinários, relacionados na Tabela 4.

ORDEM	PADRÕES ORDINÁRIOS
1	{SAPATO PRETO 46, BLUSA AZUL 40}
2	{CALÇÃO AZUL P, BLUSA AZUL 52}
3	{SAPATO PRETO 38, CALÇA AZUL 38}
4	{SAPATO PRETO 46, CALÇA AZUL 54}
5	{SAPATO PRETO 38, CALÇA AZUL 54}
6	{GANDOLA PRETO 38, SAPATO PRETO 38}
7	{SAPATO PRETO 38, CALÇÃO AZUL P, CALÇA AZUL 38}
8	{SAPATO PRETO 38, CALÇÃO AZUL P, CALÇA AZUL 54}

Tabela 4 – Padrões encontrados com o algoritmo CM-SPADE

– Padrões Extraordinários

Com o GSP, seis padrões de consumo foram encontrados e estão na Tabela 5.

ORDEM	PADRÕES EXTRAORDINÁRIOS
1	{CALÇA AZUL 54, BOTA 40}
2	{CALÇA AZUL 54, BOTA 43}
3	{SAPATO PRETO 38, BOTA 40}
4	{CALÇA AZUL 54, DOLMÃ 46}
5	{BOTA 40, CAXAGÁ 59}
6	{CALÇÃO AZUL P, BOTA 40}

Tabela 5 – Padrões encontrados com o algoritmo GSP

Abordagem por Ciclo

– Padrões Ordinários

Com a mudança da abordagem, o Spade apresentou os cinco padrões considerados ordinários.

ORDEM	PADRÕES ORDINÁRIOS
1	{CALÇA AZUL 38, CAMISA CINZA 3}
2	{SAPATO PRETO 38, CALCAO AZUL P}
3	{CAMISA CINZA 3, SAPATO PRETO 38}
4	{CALCAO AZUL P, CALÇA AZUL 54}
5	{SAPATO PRETO 38, CALCAO AZUL P, CALÇA AZUL 54}

Tabela 6 – Padrões encontrados com o algoritmo Spade

– Padrões Extraordinários

Com o GSP, quatro padrões de consumo foram encontrados e exibidos na Tabela 7.

ORDEM	PADRÕES ORDINÁRIOS
1	{CALÇA CAMUFLADA 42, CALÇA AZUL 38}
2	{CALCAO AZUL P, BOTA 40}
3	{BOTA 40, DISTINTIVO GOLA}
4	{BOTA DE CONVÉS 40, GORRO BEGE 57}

Tabela 7 – Padrões encontrados com o algoritmo GSP

MUNDO REAL: REFLEXO SOBRE AS LICITAÇÕES

De posse dos padrões de consumo, foi possível avaliar o fator de impacto para redução do número de licitações realizadas pela identificação dos padrões revelados pela mineração de dados. Esta comparação se propõe a verificar o ganho em termos de

redução do número de licitações, caso fosse adotada a compra baseada em padrões de consumo identificados, buscando-se determinar um período de ocorrência relevante de licitações ocorridas. Assim, buscaram-se dados de licitações no período compreendido entre 2014 e 2016, quando foram observados 18 processos ocorridos em 2014, 12 em 2015 e 15 em 2016.

Um fato relevante diz respeito ao modo de execução escolhido para os diversos processos licitatórios, nos quais cada item, normalmente, é separado por categoria. Ou seja, há uma licitação para sapatos e outra para calças, por exemplo. O propósito, aqui, é evidenciar o impacto no número de licitações com a utilização do planejamento de compras baseado na identificação de padrões de consumo. Para checar o impacto causado, parte-se da seguinte situação: um padrão específico P, composto dos itens (a) e (b), (notação $P\{(a), (b)\}$) e licitações L1, composta do item (a), e L2, composta dos itens (b) e (c).

O passo seguinte é verificar quais licitações possuem o item (a) e quais possuem o item (b), lembrando que a organização pública realiza, normalmente, a compra de itens em separado, como mencionado no parágrafo anterior. Para a situação proposta, percebe-se que L1 possui o item (a) e a licitação L2 possui o item (b), além do (c), sugerindo a possibilidade de junção de L1 e L2 em uma única licitação, neste caso, L3. Dessa forma, a licitação L3 substituiria L1 e L2, contendo os itens (a), (b) e (c) e permitindo a realização de apenas um processo licitatório, podendo consumir menor tempo e, com certeza, menor quantidade de recursos públicos. Em resumo, as licitações L1 e L2 saem de cena, surgindo a licitação L3.



Figura 5 – Exemplo de Licitação Reduzida

A Figura 5 exibe a junção das licitações L1 e L2 formando a licitação L3, contendo todos os itens necessários em um único processo. A seguir, serão apre-

sentados os impactos nas licitações dos anos de 2014, 2015 e 2016.

Reflexos da Abordagem Ano no mundo real

Os resultados para o cenário de padrões extraordinários, para os anos de 2014, 2015 e 2016, são apresentados na Figura 6, na qual foi realizada uma análise da licitação e de seus itens, verificando a possibilidade de encaixe de um padrão minerado anteriormente. Dessa forma, em 2014, foram realizadas 18 licitações. Com a aplicação dos padrões minerados, essa realidade mudaria para a necessidade de realização de, apenas, 12 processos, reduzindo os esforços em cerca de 33,34%. A mesma análise foi realizada para os anos de 2015 e 2016, obtendo resultados importantes.



Figura 6 – Padrões extraordinários, anos 2014, 2015 e 2016

Trabalhando com os dados de 2015, sob as mesmas bases de comparação e aplicação dos padrões minerados, a simulação sugere a obtenção dos melhores resultados, com uma redução de 41,67% dos esforços para a realização de processos licitatórios, saindo de 12 para sete licitações. Em 2016, obtivemos 33,34%, saindo de 15 para a necessidade de dez processos.

Os padrões ordinários, também, foram beneficiados pela introdução dos padrões minerados, exibindo sensíveis reduções de esforços, os quais podem ser entendidos como recursos públicos, conforme Figura 7.



Figura 7 – Padrões Extraordinários, anos 2014, 2015 e 2016

Assim, em termos percentuais, em 2014, o fator de redução foi de 44,45%; em 2015, 33,34%; e, em 2016, teve 40% de redução no número de licitações efetivamente realizadas.

Reflexos da Abordagem por Ciclo no mundo real

Já a abordagem por ciclo, para os anos de 2014, 2015 e 2016 apresentou os respectivos fatores de redução do número final de licitações a serem realizadas: 38,9%, 41,67% e 60%.



Figura 8 – Padrões Extraordinários, anos 2014, 2015 e 2016

Em termos de padrões ordinários, os resultados foram:

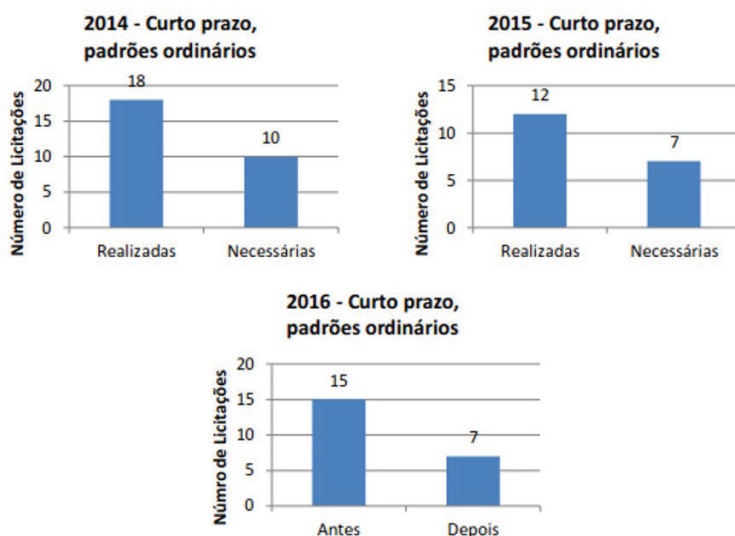


Figura 9 – Padrões Ordinários, anos 2014, 2015 e 2016

Resumindo, os anos de 2014, 2015 e 2016 apresentam os seguintes fatores de redução do número final de licitações a serem realizadas: 44,45%, 41,67% e 53,34%. Este e outros resultados podem ser conferidos na Tabela 8.

ABORDAGEM/PADRÃO	ANO	LICITAÇÕES ANTES	LICITAÇÕES DEPOIS	% REDUÇÃO
ANO/EXT	2014	18	12	33,34
ANO/EXT	2015	12	7	41,67
ANO/EXT	2016	15	10	33,34
ANO/ORD	2014	18	10	44,45
ANO/ORD	2015	12	8	33,34
ANO/ORD	2016	15	6	40
CICLO/EXT	2014	18	11	38,9
CICLO/EXT	2015	12	7	41,67
CICLO/EXT	2016	15	6	60
CICLO/ORD	2014	18	10	44,45
CICLO/ORD	2015	12	7	41,67
CICLO/ORD	2016	15	7	53,34

Tabela 8 – Resultados da Mineração no mundo real

Legenda:

EXT –Padrões Extraordinários

ORD – Padrões Ordinários

LIMITAÇÕES E TRABALHOS FUTUROS

O trabalho teve como limitação a concentração da análise entre os anos de 2014 e 2016, verificando a possibilidade de redução do número de licitações realmente realizadas. Foi investigado o universo de itens de fardamento, não admitindo outro contexto, e, por fim, utilizou-se apenas um método não supervisionado, conhecido como regras de associação. É preciso dizer que este artigo apresentou a forma como os itens de fardamento estão relacionados, mas não a proporção aplicada a eles.

Como trabalhos futuros, seria interessante e desejável utilizar outras ferramentas dos métodos supervisionado e não supervisionado, como forma de investigar outras possibilidades de geração de informação.

CONCLUSÕES

A proposta deste trabalho foi sugerir uma alternativa à geração de previsão de demanda tradicional, por meio do alcance

de alguns resultados importantes, possibilitando o desenvolvimento de uma solução que apresenta resultados relevantes. Para isso, foi utilizada a busca de padrões sequenciais, uma vez que essa abordagem traz consigo inúmeras oportunidades de aplicação, como, por exemplo, conhecer momentos em que um padrão é mais demandado que outro e em que proporção (não desenvolvido). Os padrões foram divididos em ordinários (quando as relações são esperadas) e extraordinários (quando as relações não são esperadas nem conhecidas).

O passo seguinte foi a confrontação dos padrões recuperados, em relação aos dados reais de compras por licitações realizadas pela organização pública. Os resultados favoráveis mostram que a abordagem proposta neste artigo pode contribuir efetivamente para alcançar menores gastos de dinheiro público com processos administrativos ligados à previsão de demanda, além de poupar tempo pela redução do número de licitações de fato necessárias.

 CLASSIFICAÇÃO PARA ÍNDICE REMISSIVO:

<ADMINISTRAÇÃO>; Aquisição; Gerenciamento; Licitação; Obtenção; Processo; Uniforme;

REFERÊNCIAS BIBLIOGRÁFICAS

- AGRAWAL, Rakesh; SRIKANT, Ramakrishnan. "Mining sequential patterns". In: *DataEngineering*, 1995. Proceedings of the Eleventh International Conference on IEEE, pp. 3-14, 1995.
- ALENCAR, Aretha Barbosa, 2007. "Mineração e visualização de coleções de séries temporais". Tese de D.Sc., Universidade de São Paulo, São Paulo, SP, Brasil.
- BALLOU, R.H. *Logística Empresarial*. São Paulo: Atlas, 1995.
- BERTAGLIA, Paulo Roberto. *Logística e gerenciamento da cadeia de abastecimento*. Editora Saraiva, 2000.

- CALIL, Leonardo Aparecido de Almeida *et al.* “Mineração de dados e pós-processamento em padrões descobertos”. 2008.
- DEVEZA, CECÍLIA HENRIQUES. “Minerando Padrões Sequenciais para Base de Dados de Lojas Virtuais”. Monografia (Curso de Bacharelado em Ciência da Computação), Universidade Federal de Ouro Preto (Ufop), Minas Gerais, MG, 2011.
- LU, Yijun *et al.* “Concept hierarchy in data mining: specification, generation and implementation”. Simon Fraser University, 1998.
- MAYER-SCHONBERGER, Viktor; CUKIER, Kenneth. *Big data: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana*. Elsevier Brasil, 2014.
- MANNILA, Heikki; TOIVONEN, Hannu; VERKAMO, A. Inkeri. “Discovering frequent episodes in sequences extended abstract”. In: 1st Conference on Knowledge Discovery and Data Mining, 1995.
- MARCONI, M. A. & LAKATOS, E. M. *Metodologia científica*. São Paulo: Atlas, 2000.
- MARTINEZ, A. *et al.* “Las categorias o facetas fundamentales: una metodologia para el diseño de taxonomias corporativas de sitios Web argentinos”. *Ciência da Informação*, v. 33, n. 2, pp. 106-111, 2004.
- SILVA, Marcelino Pereira dos Santos. “Mineração de dados: Conceitos, aplicações e experimentos com weka”. *Sociedade Brasileira de Computação*, v. 1, 2004.
- SRIKANT, Ramakrishnan; AGRAWAL, Rakesh. “Mining sequential patterns: Generalizations and performance improvements”. In: International Conference on Extending Database Technology. Springer Berlin Heidelberg, 1996, pp. 1-17.
- STANK, Theodore P.; KELLER, Scott B.; DAUGHERTY, Patricia J. “Supply chain collaboration and logistical service performance”. *Journal of Business logistics*, v. 22, n. 1, pp. 29-48, 2001.
- TERRA, José Cláudio Cyrineu. “Gestão do conhecimento: o grande desafio empresarial”, 2005.
- TRUJILLO FERRARI, A. *Metodologia da ciência*. 3ª ed. Rio de Janeiro: Kennedy, 1974.
- ZAKI, Mohammed J. “Sequence mining in categorical domains: Algorithms and applications”. In: Sequence Learning. Springer Berlin Heidelberg, 2000, pp. 162-187.
- ZAKI, Mohammed J. “Spade: An efficient algorithm for mining frequent sequences”. *Machine learning*, v. 42, n. 1-2, pp. 31-60, 2001.
- ZELENY, Milan. “Management support systems: towards integrated knowledge management”. *Human systems management*, v. 7, n. 1, pp. 59-70, 1987.
- VIANA, Reinaldo. “Mineração de dados: Introdução e Aplicações”. *SQL Magazine*, Ed. 10, Ano, v. 1, 2004. Simple Data Mining. Disponível em: <http://simpledatamining.blogspot.com.br/2015/03/generalized-sequential-pattern-gsp.html>. Acesso em: 17 nov. 2017.
- VASCONCELOS, Fernanda. “Licitação pública: análise dos aspectos relevantes do Pregão”. *Prima Facie – Direito, História e Política*, v. 4, n. 7, pp. 151-163, 2005.