

Evento: XXII Jornada de Pesquisa

MODELAGEM MATEMÁTICA DA PRECIFICAÇÃO DOS RECURSOS COMPUTACIONAIS UTILIZADOS EM PLANOS DE PROVEDORES DE COMPUTAÇÃO EM NUVEM USANDO O MÉTODO HEDÔNICO¹
MATHEMATICAL MODELING OF COMPUTATIONAL RESOURCES PRECISIONING USED IN PLANS OF COMPUTER PROVIDERS IN CLOUD USING THE HEDONIC METHOD

Jussiano Regis Pacheco², Sandro Sawicki³

¹ Pesquisa desenvolvida no Programa de Pós Graduação em Modelagem Matemática, pertencente ao Grupo de Pesquisa em Computação Aplicada (GCA)

² Aluno do Curso de Mestrado em Modelagem Matemática da UNIJUI,

³ Professor Doutor do Departamento de Ciências Exatas e Engenharias e Orientador.

Resumo:

A computação em nuvem se apresenta como uma possibilidade eficiente de maximizar e flexibilizar os recursos computacionais das empresa. O modelo IaaS (Infrastructure as a Service) permite que o usuário contrate máquinas virtuais compostas por diferentes configurações de recursos computacionais a um determinado valor. A grande diversidade de planos ofertados, somado as estratégias de precificação adotadas pelas empresas do ramo, torna complexa a escolha por planos e provedores. O objetivo deste trabalho é modelar e analisar os planos de precificação utilizadas pelos principais provedores de computação em nuvem no modelo Infrastructure as a Service (IaaS) através do método hedônico, visando compreender a influência de cada um dos recursos computacionais na formação do preço final. Para a construção do modelo utilizou-se de estimativas obtidas do provedor Amazon EC2, considerando as características CPU, Memória, Armazenamento, Sistemas Operacionais e localização geográfica dos servidores. O estudo analisou uma amostra de 2.622 planos, os resultados obtidos indicam que o modelo testado é estatisticamente satisfatório, isto é, explica a realidade de forma consistente.

Abstract:

A cloud computing presents itself as an efficient solution to maximize and flex the computing resources of the company. The IaaS (Infrastructure as a Service) model allows the user to compete against different composite virtual machines by different ones. The great diversity of flat plans, added as a strategy of pricing adopted by the companies of the branch, makes complex the choice by plans and providers. The objective of this work is to model and analyze the pricing plans for the main providers of cloud computing without model Infrastructure as a service (IaaS) through the hedonic method, aiming at an influence of each of the computational resources in the final price formation. For a construction of the model, estimates obtained from the Amazon EC2 supplier are used, considering CPU, Memory, Storage, Operating Systems and geographic location of the servers. The study analyzed a sample of 2,622 plans, the results obtained indicate that the model tested is statistically satisfactory, that is, it explains the reality in a consistent way.

Evento: XXII Jornada de Pesquisa

Palavras-chave: Computação em Nuvem; Infraestrutura como Serviço; Modelo de Preços; Regressão Linear; método hedônico.

Keywords: Cloud computing; Infrastructure as a Service; Pricing Model; Linear regression; Hedonic method.

1. INTRODUÇÃO

O cenário atual de negócios exige das empresas maior velocidade, flexibilidade e consistência na implantação de suas estratégias para sobrevivência e crescimento. A computação em nuvem se apresenta como uma possibilidade eficiente de maximizar e flexibilizar os recursos computacionais de uma empresa, possibilitando o armazenamento e o processamento através da internet, consumo sob demanda e a redução de custos em infraestrutura de Tecnologias de Informação (TI).

Define-se Computação em nuvem, segundo NIST (National Institute of Standards and Technology), como um modelo que possibilita acesso, de modo conveniente e sob demanda, a um conjunto de recursos computacionais configuráveis (por exemplo, redes, servidores, armazenamento, aplicações e serviços) que podem ser rapidamente adquiridos e liberados com mínimo esforço gerencial ou interação com o provedor de serviços.

No modelo SaaS (Software as a Service), a aplicação encontra-se em um ambiente virtual, sem a necessidade de instalação no computador local, necessitando somente o acesso a Internet. O PaaS (Platform as a Service) disponibiliza ao cliente uma plataforma, incluindo sistemas e linguagens de programação, para o desenvolvimento de aplicações. Já o serviço de IaaS (Infrastructure as a Service) oferece o acesso a uma infraestrutura computacional, como servidores, rede, armazenamento e outros recursos de computação sobre os quais pode ser instalado e executado qualquer tipo de software.

O avanço da computação em nuvem, em especial o modelo IaaS (Infrastructure as a Service), potencializou aumento da migração de aplicações dos setores de Tecnologia da Informação das empresas para infraestruturas de hardware na nuvem (Rold & Maurer, 2016). Este serviço permite que as empresas utilizem recursos de computação de forma virtual, sem a necessidade de infraestrutura de hardware local. É altamente escalável e o custo varia de acordo com o uso, ou pode ser definido previamente mediante contrato. Segundo a consultoria Gartner, o IaaS crescerá 36,8% em 2017, alcançando o valor de 34,6 bilhões de dólares.

Dentre as vantagens dos serviços de computação em nuvem, destacam-se: o consumo sob demanda, onde o cliente tem disponibilidade para alocar recursos computacionais na quantidade que considerar necessário. A aquisição do serviço ocorre rapidamente, o que dá maior agilidade à empresa. Essa elasticidade também gera uma redução nos custos da empresa, pois permite contratar um serviço moldado as suas necessidades, pagando somente pelos recursos utilizados, o que dilui o custo associado à ociosidade. Dessa forma evita o investimento em uma infraestrutura, incluindo a aquisição de máquinas e equipamentos, com capacidade suficiente para dar conta de sua demanda, considerando os períodos de maiores picos (Rodamilans et. al, 2014).

Evento: XXII Jornada de Pesquisa

Um plano de computação em nuvem é composto por um conjunto de recursos computacionais, tais como CPU, memória, armazenamento, localização geográfica, sistema operacional, I/O (entrada e saída), entre outras características. Cada provedor oferece uma série de planos com diferentes combinações desses recursos e diferentes preços. Essa grande quantidade de opções disponíveis possibilita encontrar o plano mais adequado para atender as necessidades de cada cliente.

Ao mesmo tempo em que a oferta de planos permite o acesso de consumidores com diferentes necessidades e capacidades de pagamentos, essa vasta quantidade de opções torna complexa a tomada de decisão. Diferentes aspectos precisam ser observados pelo cliente, como desempenho, custo, segurança, para chegar à melhor decisão.

A concorrência entre as empresas do setor para aumentar sua participação no mercado faz com que promovam diferentes estratégias de precificação. Uma das maneiras conhecidas é o desagrupamento de preços, este sistema cobra por cada característica de computação separadamente, os clientes escolhem a quantidade dos itens e um preço específico por unidade é cobrado. De outro modo há também provedores que definem o preço a partir do seu conjunto, dessa forma não há um valor específico para cada um dos componentes (Kihal *et al.*, 2012; Mitropoulou *et al.*, 2016). Há também diferenças no formato dos planos ofertados, onde cada provedor oferece seus planos com conjuntos de recursos computacionais predefinidos.

Diante da complexidade deste serviço faz-se importante compreender as estratégias de precificação adotadas pelos provedores na formação de seus planos de venda, buscando encontrar a relação entre os recursos oferecidos e o seu preço.

O modelo hedônico é um dos métodos relevantes para análise de preços utilizados na literatura em geral. Esse modelo consiste que o preço de um produto é uma função dos atributos que o compõe, possibilitando identificar a contribuição de cada componente no preço final de um bem ou serviço.

Neste trabalho objetiva-se realizar um estudo preliminar de análise dos planos de precificação utilizados por provedores de computação em nuvem no modelo IaaS - Infrastructure as a Service através do método hedônico, visando compreender a influência de cada um dos recursos computacionais na formação do preço final.

Este artigo está organizado da seguinte forma: Na seção 2 apresenta-se a metodologia de trabalho adotada. Na terceira seção são apresentados resultados preliminares do modelo aplicado. Por fim, na Seção 4 são apresentadas as conclusões e algumas perspectivas para trabalhos futuros.

2. METODOLOGIA

Para a construção da base de dados sobre os preços dos planos de computação em nuvem, foram simuladas contratações dos planos do Amazon em todas as suas localizações geográficas, bem como as informações referentes aos sistemas operacionais. O modelo de cobrança adotado foi o On-demand, este modelo não exige a contratação do serviço por longo prazo, possibilitando ao usuário cancelá-lo a qualquer momento.

Evento: XXII Jornada de Pesquisa

Tabela 1: Critérios da amostra de planos de computação em nuvem selecionados.

Provedor	Amazon
Sistema Operacional da Máquina Virtual	Linux Windows, Windows e Web SQL Server, SUSE Linux Enterprise Server Red Hat Enterprise Linux e Linux;
Localização Geográfica do Provedor	EUA (Virginia), EUA (Oregon), Europa (Irlanda), Europa (Frankfurt), Asia (Cingapura), Asia (Japão), America do Sul (São Paulo), AWS gov cloud (EUA);
Modelo de Cobrança	On-demand;

Diante da grande quantidade de dados obtidos, retirou-se uma amostra composta pelos planos disponíveis considerando os locais dos servidores e os sistemas operacionais que apresentavam a maior quantidade de opções aos clientes. Desta forma foram definidas oito localizações geográficas e cinco sistemas operacionais, conforme apresentados na tabela 1

Os planos da Amazon são formados pelos recursos computacionais CPU, memória e armazenamento. A cobrança é feita por 744 horas de utilização por mês e os dados são referentes ao mês de maio de 2017.

Para a análise dos dados será utilizado o método de precificação hedônico (Rosen, 1974). Refere-se a modelos de regressão, onde os preços de um produto ou serviço possuem relação com suas características objetivas. Desta forma, pressupõe-se que a soma dos preços atribuídos as características de um determinado bem, define o preço deste para o mercado.

Um método hedônico obtém estimativas do valor de cada componente, através da decomposição do item principal. Com isso, é possível obter a importância relativa que cada atributo tem na determinação do preço total do objeto estudado.

Uma função hedônica relaciona um número de características do produto com o preço correspondente:

$$P=f(X)$$

Onde P é o preço do produto considerado e X é um vetor de características associadas ao bem. Esse modelo pode ser expresso por uma função:

Evento: XXII Jornada de Pesquisa

$$P = a_0 + a_1.X_1 + a_2.X_2 + \dots + a_n.X_n + b_1.Y_1 + b_2.Y_2 + \dots + b_n.Y_n + e$$

Onde

- P é a variável dependente preço do produto.
- a_0 é a constante do modelo.
- X_1, X_2, \dots, X_n , são as variáveis dummies independentes.
- Y_1, Y_2, \dots, Y_n , são as variáveis independentes (exceto dummies).
- a_1, a_2, \dots, a_n são os coeficientes que determinam o preço implícito de cada variável *dummy* independente.
- b_1, b_2, \dots, b_n são os coeficientes que determinam o preço implícito das variáveis independentes (exceto *dummies*).
- e é o residual de regressão da forma funcional assumida.

Neste trabalho os componentes de cada um dos planos, como a quantidade de RAM, o número de CPUs, a quantidade de memória, local do servidor e demais características que afetam o preço dos produtos, corresponderão ao vetor de características. A modelagem, baseada em variáveis quantitativas e qualitativas, são um diferencial desta abordagem.

O modelo a ser aplicado utilizará a análise de regressão linear múltipla para estimá-lo, usando como base os componentes dos planos de computação em nuvem. As variáveis do modelo são divididas em variáveis quantitativas, como CPU, memória e armazenamento e variáveis qualitativas, como sistema operacional e localização geográfica. Estas últimas por não possuir informações quantificáveis serão tratadas como variáveis *dummy*, que utiliza o código binário para sua definição, onde o valor 1 indica a presença de um determinado atributo e o valor 0 a sua ausência (Wonnacott, 1990)

3. RESULTADOS E DISCUSSÃO

O principal objetivo da retirada de uma amostra de dados foi analisar o preço das instâncias mediante variação dos principais recursos computacionais. Para isso, nesta seção são apresentados o modelo aplicado e as respectivas análises estatísticas para verificar a consistência do modelo.

$$P = \beta_0 + \beta_1.CPU + \beta_2.Memoria + \beta_3.Armazenamento + \beta_4.Local1 + \beta_5.Local2 + \beta_6.Local3 + \beta_7.Local4 + \beta_8.Local5 + \beta_9.Local6 + \beta_{10}.Local7 + \beta_{11}.S_operac1 + \beta_{12}.S_operac2 + \beta_{13}.S_operac3 + \beta_{14}.S_operac4 + e$$

Tabela 2: DEFINIÇÃO DAS VARIÁVEIS HEDÔNICAS

Variável	Definição da Variável
CPU	Variável quantitativa, medida em Unidades.

Evento: XXII Jornada de Pesquisa

Memoria	Variável quantitativa, medida em Gigabytes
Armazenamento	Variável quantitativa, medida em Gigabytes
Localização geográfica do servidor	<p>Variáveis Qualitativas</p> <p>EUA (Oregon): $Local1=0, Local2=0, Local3=0, Local4=0, Local5=0, Local6=0, Local7=0, Local8=0$.</p> <p>EUA (Virginia): $Local1=1, Local2=0, Local3=0, Local4=0, Local5=0, Local6=0, Local7=0, Local8=0$.</p> <p>Europa (Irlanda): $Local1=0, Local2=1, Local3=0, Local4=0, Local5=0, Local6=0, Local7=0, Local8=0$.</p> <p>Europa (Frankfurt): $Local1=0, Local2=0, Local3=0, Local4=1, Local5=0, Local6=0, Local7=0, Local8=0$.</p> <p>Ásia (Cingapura): $Local1=0, Local2=0, Local3=0, Local4=0, Local5=1, Local6=0, Local7=0, Local8=0$.</p> <p>Ásia (Japão): $Local1=0, Local2=0, Local3=0, Local4=0, Local5=0, Local6=1, Local7=0, Local8=0$.</p> <p>América do Sul (São Paulo): $Local1=0, Local2=0, Local3=0, Local4=0, Local5=0, Local6=0, Local7=1, Local8=0$.</p> <p>AWS gov cloud (EUA): $Local1=0, Local2=0, Local3=0, Local4=0, Local5=0, Local6=0, Local7=0, Local8=1$.</p>
Sistema Operacional	<p>Variáveis Qualitativas</p> <p>Linux: $S_Operac1=0, S_Operac2=0, S_Operac3=0, S_Operac4=0$.</p> <p>Windows: $S_Operac1=1, S_Operac2=0, S_Operac3=0, S_Operac4=0$.</p> <p>Windows e Web SQL Server: $S_Operac1=0, S_Operac2=1, S_Operac3=0, S_Operac4=0$.</p> <p>SUSE Linux Enterprise Server: $S_Operac1=0, S_Operac2=0, S_Operac3=1, S_Operac4=0$.</p> <p>Red Hat Enterprise Linux: $S_Operac1=0, S_Operac2=0, S_Operac3=0, S_Operac4=1$.</p>

Os resultados das regressões para o modelo de regressão múltipla proposto para este estudo são apresentados nesta seção. Pela análise das estimativas hedônicas podemos determinar se o modelo é adequado e verificar, além disso, se os dados coletados fornecem informações suficientes para a elaboração das estimativas necessárias. Por fim, os coeficientes obtidos serão analisados para que se possa determinar quais são os atributos que mais impactam o preço dos planos de computação em nuvem desta amostra. Os resultados das regressões são apresentados a tabela 3.

O coeficiente de determinação (R^2) calculado alcançou 86.26% para o modelo linear, o que traduz um poder explicativo satisfatório. Este índice serve para avaliar a qualidade da correlação geral do modelo, sendo que quanto mais próximo de 100%, maior é a consistência explicativa do modelo

Evento: XXII Jornada de Pesquisa

(Field, 2009).

Tabela 3: Estatística de regressão

R múltiplo	86,2661%
R-quadrado ajustado	86,1923%
Observações	2622
F de significação	0

Outro ponto observado é a estatística “f”, através do F de significação, este índice avalia se existe combinação linear entre as variáveis independentes, de forma que elas exerçam influência sobre a variável de estudo e possam ser consideradas determinantes do seu comportamento. O F de significação alcançado foi zero, sendo que para ser relevante deve apresentar uma significância estatística menor que 0,05 para um nível de confiança de 95% (FÁVERO et al 2009).

A análise das estatísticas dos testes “t” apresentadas na Tabela 4, permite afirmar a significância dos parâmetros (β) utilizados no estudo. Para que a modelagem pesquisada seja considerada válida ela também deve apresentar uma significância estatística maior que zero e menor que 0,05, “valor-p” < 0,05 (FÁVERO et al 2009).

Tabela 4: Coeficientes estimados e informações sobre a significância destes coeficientes.

	Coeficientes
Interseção	-233,3802***
CPU	49,6831***
MEMÓRIA	4,2407***
ARMAZENAMENTO	0,0891***
Local1: EUA (Virginia)	9,3220
Local2: Europa (Irlanda)	123,7669*
Local3: Europa (Frankfurt)	131,0978*
Local4: Asia (Cingapura)	188,4406***
Local5: Asia (Japão)	188,8932***
Local6: America do Sul (São Paulo)	299,7192***
Local7: AWS gov cloud (EUA)	219,0999***
S_Operac1: Windows	461,1352***
S_Operac2: Red Hat Enterprise Linux	72,8283
S_Operac3: SUSE Linux Enterprise Server	68,8516
S_Operac4: Windows e Web SQL Server	620,6186***

Evento: XXII Jornada de Pesquisa

***p < 0.01, **p < 0.05, *p < 0.1.

A partir dos coeficientes apresentados na tabela 4, é possível observar individualmente os componentes, bem como sua influência em relação ao preço dos planos. Em relação as variáveis quantitativas, CPU, memória e armazenamento, todas alcançaram um nível de confiança bem expressivo, com valor-p menor que 1%. Outro ponto relevante são os diferentes efeitos que cada componente realiza no preço final. A contratação de uma unidade de CPU gera uma elevação do valor do plano em US\$ 49,68 dólares, já a quantidade memória, a cada Gigabyte adicionado, faz o preço se elevar em US\$ 4,24. E no caso do armazenamento, cada unidade acresce em R\$ 0,09/GB.

Em relação as variáveis qualitativas, quando observado a localização geográfica dos servidores, destacam-se os Norte Americanos como os planos com menor impacto no preço, seguido dos Europeus, Asiáticos e por fim a América do Sul. EUA (Oregon) apresenta-se como o menor entre todos os itens pesquisados, o que gera uma influência negativa na formação dos preços (-233,38), considerando todos os demais fatores constantes. Já nos sistemas operacionais, Linux é o que apresenta o menor incremento no valor dos planos de computação. Em segundo lugar, com uma diferença de 461,13 o sistema Windows e por último Windows e Web SQL Server com valor de 620,61.

As variáveis com índice valor-p acima de 10%, foram considerados de significância estatística baixa. Desta forma, a localização geográfica EUA (Virginia) e os Sistemas operacionais Red Hat Enterprise Linux e SUSE Linux Enterprise Server não foram incluídos na análise.

4. CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentado um estudo preliminar para a elaboração do modelo de preços dos planos praticados por provedores de computação em nuvem a nível de IaaS utilizando um modelo de regressão linear múltipla.

Foram obtidos resultados estatísticos satisfatórios sobre a consistência do modelo apresentado. As análises estatísticas permitiram ampliar a compreensão sobre a formação de preços dos planos, possibilitando mensurar o poder de influência de cada componente no valor final do serviço.

Estes resultados mostram ser um importante passo para tornar a política de preços dos provedores mais transparente, contribuindo na tomada de decisão dos usuários em relação ao plano mais adequado a sua necessidade.

Para trabalhos futuros, espera-se realizar a análise de preços para os dois provedores restantes, Azure e Google, e implementar um modelo de regressão para cada um. Por se tratar de um método estatístico, pretende-se ampliar a análise considerando novas variáveis ao modelo, bem como melhorar o nível de significância.

Evento: XXII Jornada de Pesquisa

5. REFERÊNCIAS

AMAZON WEB SERVICES. <https://aws.amazon.com/>. Acesso em: 21/05/2017

ANDERSON E, ESCHINGER C, WURSTER L, DE SILVA F, CONTU R, LIU V, BISCOTTI F, PETRI G, ZHANG J, YEATES M (2013) Forecast overview: public cloud services, worldwide, 2011-2016, 4Q12 Update, Market analysis and statistics. Gartner Group, Inc., Connecticut, USA

FÁVERO, L. P. L. et al. Análise de dados: modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

FIELD, A. Descobrimo a estatística usando SPSS. Tradução de Lorí Viali. 2. ed. Porto Alegre: Artmed, 2009.

CHUN, S. e CHOI, B., Service models and pricing schemes for cloud computing. Cluster Computing, 17(2):529- 535, 2014.

GOOGLE CLOUD PLATAFORM. <https://cloud.google.com/>. Acesso em: 10/04/2017.

HUANG, J., KAUFFMAN, R. J. e MA, D.. Pricing strategy for cloud computing: A damaged services perspective. Decision Support Systems, 78:80-92, 2015.

KIHAL, S.E., SCHLERETH, C., E SKIERA, B. 2012. Price Comparison for Infraestructure as a Service. Frankfurt, Germany; Goethe University, pp. 1-12.

LANCASTER, K. (1972). Operationally Relevant Characteristics- Essays in Honour of Lord Robbins, London.

MICROSOFT AZURE. <https://azure.com/>. Acesso em: 07/06/2017

MENZEL, M. e RANJAN, R., Cloudgenius: Decision support for web server cloud migration. In Proceedings of the WWW12, pages 979-988, 2012.

MITROPOULOU P., FILIOPOULOU E., TSAROUCHA S., MICHALAKELIS C. E NIKOLAIDOU M. (2015). A Hedonic Price Index for Cloud Computing Services. In Proceedings of the 5th International Conference on Cloud Computing and Services Science, pages 499-505.

MURTHY M. K. M., SANJAY H. A. e PADMANABHA, A. J., Pricing models and pricing schemes of IaaS providers: a comparison study. In Proceedings of the ICACCI12, pages 143-147, 2012.

ROLD, C; MAURER W. Hybrid IT Infraestructure Management and Cloud Migrations: Top Providers' Strengths and Weaknesses, Gartner Inc. 2016

RONDAMILANS, C.B., Uma Metodologia para Caracterização de Aplicações e de Intancias de

Evento: XXII Jornada de Pesquisa

Maquinas Virtuais no Ambiente de Computação em Nuvem. 2014. Tese (Doutorado em Engenharia Elétrica) - Universidade de São Paulo.

ROSEN, S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, v. 82, p. 34-55, 1974.

SAMIMI, P. e PATEL, A., Review of pricing models for grid & cloud computing. In *Proceedings of the ISCI 2011*, pages 634-639, 2011.

SANCHEZ, O. P. e CAPPELLOZZA, A, Antecedentes da adoção da computação em nuvem: efeitos da infraestrutura, investimento e porte. *Rev. adm. contemp.* [online]. 2012, vol.16, n.5, pp.646-663.

TOMKOVICK, C., e K. E. DOBIE (1995). Applying Hedonic Pricing Models and Factorial Surveys at Parker Pen to Enhance New Product Success. *Journal of Product Innovation Management*, 12(4), 334-345.

WONNACOTT, T. H. e WONNACOTT, R. J. *Introductory Statistics for Business and Economics*. John Wiley & Sons, 1990.