



Análise Fatorial e Análise de Conglomerados para Determinação de Clusters em Empresas de Internet no Brasil

Luiz Paulo Lopes Fávero (Administração FEA-USP) - lpfavelo@usp.br

Patrícia Prado Belfiore (Engenharia de Produção POLI-USP) -
patricia.belfiore@poli.usp.br

Claudio Felisoni de Angelo (Administração FEA-USP) - cfa@usp.br

Sérgio Renato Carmo Brejon (Engenharia Naval e Oceânica – POLI-USP)

Este trabalho busca a determinação de comportamentos semelhantes de atuação operacional e financeira em empresas de internet no Brasil, em função da utilização de indicadores econômico-financeiros. Para isto, aplica-se o método multivariado de análise fatorial e análise de conglomerados para a determinação de eventuais clusters existentes. Para a realização deste estudo, utilizou-se como base de dados os indicadores econômico-financeiros das 24 maiores empresas de internet do Brasil, segundo a revista Balanço Anual da Gazeta Mercantil de 2004. As empresas estão classificadas em um ranking pela sua receita operacional e o tratamento dos dados se deu por meio do software estatístico SPSS (Statistical Package for Social Sciences). O objetivo deste estudo consiste na análise dos clusters formados, com o intuito de identificar características comuns de cada agrupamento de empresas.

Palavras-chave: indicadores econômico-financeiros; tecnologia da informação; internet; análise fatorial; análise de clusters

1. INTRODUÇÃO

Na atual conjuntura econômica de competitividade em oferta de produtos e serviços, o setor terciário apresenta-se como sendo o mais suscetível a mudanças de gestão de indicadores econômicos e financeiros. Entre os vários serviços ofertados por esse setor, destaca-se o de tecnologia da informação que, segundo FREEMAN e SOETE (1987), representa um conjunto integrado de inovações em computação eletrônica, engenharia de *software*, sistemas de controle, circuitos integrados e telecomunicações que reduzem os custos de armazenagem, processamento, comunicação e disseminação da informação. Portanto, trata-se de uma atividade integrada cuja relevância não decorre apenas de sua enorme expressão econômica, mas também por ser o elo entre todos os *players* de uma cadeia de informação, desde a geração da informação por uma fonte emissora até a aceitação da informação pela entidade receptora.

Durante muitas décadas os aspectos relacionados à administração da tecnologia da informação foram negligenciados e as vastas e rentáveis possibilidades oferecidas pelo mercado financeiro permitiam que as empresas do setor compensassem com grande

facilidade seus problemas operacionais (BLECHER, 2001, p. 32). Por outro lado, a política de redução drástica da inflação, introduzida em 1994 com o Plano Real, alterou significativamente a forma de atuação das empresas e reestruturou de imediato o setor de tecnologia da informação.

Com os avanços tecnológicos, a abordagem sistêmica da informação começou a ser uma tendência e uma necessidade nas organizações. Novas ferramentas e metodologias vêm surgindo desde então, na busca de aprimoramento dos sistemas de informação, que começaram a ganhar cada vez mais importância tanto na área operacional como estratégica das corporações. Abriu-se espaço para questionamentos e aprimoramento do seu gerenciamento de indicadores financeiros, outrora pouco avaliados pela condição mais propícia ao incremento dos mesmos.

Com o intuito de tornar possível a determinação de comportamentos semelhantes de atuação em empresas do setor de tecnologia da informação no Brasil, em especial empresas de internet, em função da utilização de indicadores econômico-financeiros, este artigo utiliza técnicas multivariadas de análise fatorial e análise de conglomerados para a determinação respectivamente, de fatores combinados de variáveis correspondentes a índices financeiros e de eventuais *clusters* existentes entre os *players*.

Para a realização deste estudo, portanto, utilizou-se como base de dados os indicadores econômico-financeiros das 24 maiores empresas de internet do Brasil, segundo a REVISTA BALANÇO ANUAL DA GAZETA MERCANTIL de 2004. As empresas estão classificadas em um ranking pela sua receita líquida e o tratamento dos dados se deu por meio do software estatístico SPSS (*Statistical Package for Social Sciences*). O objetivo deste estudo consiste na análise dos *clusters* formados, com o intuito de identificar características comuns de cada agrupamento de empresas, em função das respectivas atuações regionais.

Embora essa reflexão possa ser aplicada a todos os segmentos que compõem os setores produtivos e de prestadores de serviços no Brasil, não há dúvidas que o ramo composto pelas empresas de internet mereça atenção especial. O crescimento anual deste setor está situado em torno de 30%, segundo a REVISTA BALANÇO ANUAL DA GAZETA MERCANTIL de 2004 e este número mostra a possibilidade de expansão do faturamento pela otimização da operação e pelo gerenciamento da informação, plausível de ser obtida por métodos estatísticos e econométricos.

2. OBJETIVO DO ESTUDO

O objetivo deste trabalho é avaliar, por meio das técnicas de análise fatorial e análise de conglomerados, como se comportam as 24 maiores empresas de internet do Brasil, em relação à combinação de doze variáveis financeiras e de performance, com o objetivo de verificar a existência de algum padrão de atuação geográfica que possa indicar conglomerados específicos. Para tanto, com o intuito de verificação, serão utilizados dados provenientes da REVISTA BALANÇO ANUAL DA GAZETA MERCANTIL de 2004.

Nesse tipo de estudo econométrico, poder-se-ia considerar outros indicadores financeiros, porém, para manter um critério mais parcimonioso, optou-se pela utilização de variáveis referentes apenas a condições econômico-financeiras.

3. PANORAMA DO SETOR DE TECNOLOGIA DA INFORMAÇÃO NO BRASIL

A tecnologia da informação vem contribuindo para a mudança dos cenários organizacionais, mais notadamente no que se refere ao aumento da capacidade de processamento, da estrutura e dos fluxos de informação. Do ponto de vista interno, as novas tecnologias que afetam a organização, o processamento e o uso da informação têm se refletido na eficiência organizacional. Quanto ao cenário externo, a forma de fazer negócios também está sendo afetada, as bases da competição estão se transformando, a relação entre antigos rivais se configura, com a formação cada vez mais constante de redes e parcerias de negócios.

Todos os segmentos são afetados: varejo, serviços financeiros, manufatura e educação. Estes segmentos estão cada vez mais próximos uns dos outros, e suas relações parecem ser facilitadas pelo fluxo de informação existente entre eles, potencializado pelas tecnologias de informação.

O aumento da velocidade das mudanças, devido à aplicação das Tecnologias da Informação nos vários segmentos da sociedade, é um fato incontestável. Embora não se tenham referências precisas sobre o alcance dessas mudanças, bem como sobre suas implicações, seus reflexos repercutem em todos os lugares e causam inquietação em todos os setores da sociedade (ANDRADE, 2002, p. 50).

Com as tecnologias da informação incorporadas, os fluxos da informação na empresa tendem a tornar-se mais efetivos em virtude da elevação da capacidade de coletar, estocar, processar e transferir informações, o que torna possível a obtenção de uma maior velocidade na comunicação, redução no prazo das respostas às variações dos ambientes interno e externo, melhoria na tomada de decisão, enfim, um aumento da eficiência organizacional em vários aspectos (NORONHA, 2000, p. 5).

Para a obtenção de ganhos de produtividade, as empresas buscam cada vez mais a sua eficiência, e a tecnologia tem sido uma grande aliada neste processo. A tecnologia da informação está inserida neste contexto, vindo a contribuir com as empresas na busca da elevação de sua eficiência através da melhoria dos fluxos de informação. AMOR (2000) afirma que a Tecnologia da Informação está modificando a forma de fazer negócios, entretanto, o negócio vem vagarosamente se adaptando às novas possibilidades. A convivência da empresa dentro da nova configuração da sociedade necessita de um novo referencial, que será alicerçado pela Tecnologia da Informação.

Para avaliar a situação atual deve-se considerar, pelo menos, três processos desenvolvidos durante a última década e que se realimentam, tornando o mundo diferente do que era. Os dois primeiros são a globalização e a estabilidade econômica, que tornaram o mundo menor, mais misturado e mais dinâmico. O terceiro é a informatização, que introduzida no mundo dos negócios na década de cinquenta, ganhou força depois da invenção do microprocessador, na década de setenta (CASTELLS, 1999, p. 58).

Os computadores da década de cinquenta eram grandes e difíceis de manejar. Suas falhas implicavam reparos frequentes, dificultados por seu tamanho e sua construção. Além disso, esses computadores só eram acessíveis em linguagem de máquina. Quanto ao hardware, era impossível transferir um programa de aplicação de uma máquina para outra.

Nas décadas seguintes, a evolução dos componentes eletrônicos foi, sem dúvida, a mais carregada das conseqüências. Sua miniaturização, assim como seu custo reduzido, permitiu sua disseminação e utilização em massa, principalmente com a expansão dos PC's.

Ao mesmo tempo, a característica de rigidez da geração anterior desapareceu. As informações só eram acessíveis de forma seqüencial. Surgiram procedimentos particulares, que permitiam a obtenção da informação de acordo com o interesse do usuário. Surgiram

também os bancos de dados, que avaliam a vantagem de uma capacidade de armazenamento maior com as possibilidades de fácil acesso. A máquina passou a gerar, de forma cada vez mais eficaz, seu próprio funcionamento, otimizando as seqüências de trabalho e organizando melhores os períodos de tratamento, de obtenção de dados e de impressão.

Surgiram as redes em tempo real. Os meios de transmissão, anteriormente separados, aproximaram-se. Os pontos de acesso multiplicaram-se onde os terminais, cada vez mais numerosos, dialogavam entre si e com os computadores centrais.

Nos anos oitenta, havia uma suposição de que a informatização revolucionaria sozinha todo o mundo dos negócios, mais o que se viu foi que, sem transformações organizacionais, a tecnologia agravaria os problemas de burocracia e rigidez das empresas, em vez de solucioná-los.

As inovações na tecnologia da informação permitem oferecer produtos mais variados a usuários cada vez mais numerosos. A inter-relação crescente da tecnologia da informação com seus usuários, dentro ou fora da empresa propriamente dita, transforma os modelos organizacionais, pois, a partir da possibilidade de criação de redes, as empresas ganham novos contornos do ponto de vista organizacional. Segundo SHAPIRO e VARIAN (1999, p. 25) formar uma rede envolve mais do que fabricar um produto. Encontrar parceiros, formar alianças estratégicas e saber como liderar o movimento de apoio pode ser tão importante quanto as habilidades de engenharia. As regras tradicionais de estratégia competitiva focalizam os concorrentes, fornecedores e clientes. Na economia da informação, as empresas que vendem componentes complementares, ou complementadores, são igualmente importantes (SHAPIRO; VARIAN, 1999, p. 23).

O lugar em que esses avanços socioeconômicos estão atingindo seu ponto mais alto é a internet. Segundo a Revista Balanço Anual de 2004, o crescimento do setor de internet no último ano foi de 30%. É aí que a tecnologia de informação entra em contato com a globalização e os processos organizacionais de uma forma jamais vista. Este contato proporciona várias transformações nas organizações, nas comunicações internas, no relacionamento com clientes e fornecedores, na própria natureza dos produtos e na configuração das empresas. E mais uma vez, como em toda a história, novas tecnologias de informação irão possibilitar e estimular novas formas organizacionais (TAPSCOTT, 1997, p. 17).

Percebe-se, portanto, uma nova fase de crescimento do setor de tecnologia da informação no Brasil, em especial a internet. Investimentos na melhoria dos processos de logística e tecnologia da informação evidenciam o foco na melhoria da eficiência operacional. Assim, uma possibilidade de expansão do faturamento é pela otimização da operação e pelo gerenciamento da informação, plausível de ser obtida por métodos estatísticos e econométricos.

A tabela abaixo apresenta o ranking das 26 maiores empresas de internet no Brasil, classificadas em ordem decrescente de receita líquida, segundo a REVISTA BALANÇO ANUAL 2004.

Tabela 1: Ranking das empresas de Internet em função da Receita Líquida

1. Terra Networks – RS	10. Onda Provedor – PR	19. Mercador/RS – RS
2. Americanas.Com-SP	11. Webmotors – SP	20. Estrutura.Net – SP
3. Submarino – SP	12. Investnews – SP	21. Klick Net – SP

4. BrT – DF	13. Real Estate Web-SP	22. TCO IP – DF
5. Ibest – RJ	14. Pakprint – SP	23. EPC Am. Latina-SP
6. Blah! Servs –RJ	15. Fulano.Com – SP	24. Netcash – SP
7. Net Site – SP	16. MLS Wireless – RJ	25. Casa Interativa – RS
8. Escelsanet – ES	17. Procurement – RJ	26. Hicorp - SP
9. Merc. Eletrônico-SP	18. Conectt - RS	

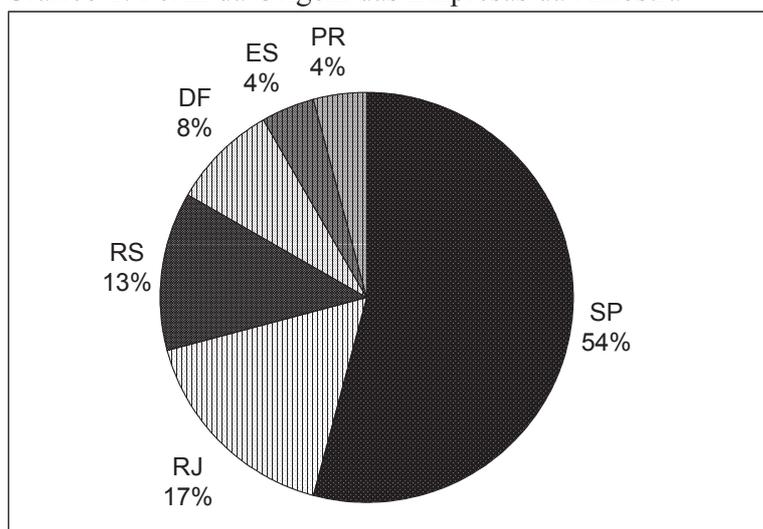
Fonte: Revista Balanço Anual da Gazeta Mercantil 2004.

4. MÉTODO EMPREGADO NO ESTUDO E REVISÃO TEÓRICA DOS CONCEITOS

4.1. Composição da amostra

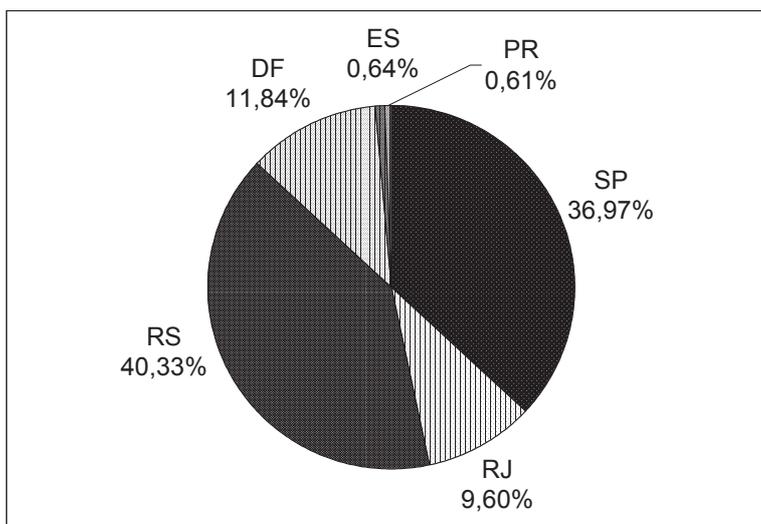
A amostra de empresas de internet coletada para este estudo compõe uma receita líquida anual acima de R\$ 1 bilhão, representando um ativo total de aproximadamente R\$ 1,5 bilhões. Estas 26 empresas são provenientes de seis estados do Brasil (São Paulo, Paraná, Espírito Santo, Rio de Janeiro, Distrito Federal e Rio Grande do Sul), com a seguinte composição geográfica:

Gráfico 1: Perfil da Origem das Empresas da Amostra



Em termos de receita líquida, o perfil das empresas está no gráfico a seguir:

Gráfico 2: Perfil de Receita Líquida das Empresas da Amostra



Com a análise dos gráficos 1 e 2, pode-se notar que as empresas do RS apresentam uma melhor performance financeira em relação à média, uma vez que seus percentuais incrementaram mais do que proporcionalmente, quando a análise partiu do ponto de vista de número de empresas para o ponto de vista de receita líquida, o que pode oferecer uma eventual formação de um *cluster* para as empresas deste estado.

A empresa Terra Networks do Rio Grande do Sul representa 40% da receita líquida total (99% da receita do RS), enquanto que as empresas Americanas.Com e Submarino de São Paulo representam 32,7% da receita líquida total (88% da receita de SP), possibilitando a formação de eventuais *clusters* destas empresas.

Para o desenvolvimento deste trabalho, foi realizada uma pesquisa quantitativa, com o levantamento de dados secundários provenientes do ranking da REVISTA BALANÇO ANUAL DA GAZETA MERCANTIL de 2004, por meio da qual possibilitou-se a coleta de indicadores econômico-financeiros das maiores empresas de internet do Brasil.

As variáveis coletadas encontram-se na tabela 2 a seguir:

Tabela 2: Variáveis utilizadas no estudo

CÓDIGO	VARIÁVEL
RL	RECEITA LÍQUIDA
LO	LUCRO OPERACIONAL
LL	LUCRO LÍQUIDO
ATIVO	ATIVO
PL	PATRIMÔNIO LÍQUIDO
EBITD	EBITDA
ENDIV	ENDIVIDAMENTO
CG	CAPITAL DE GIRO
NCG	NECESSIDADE DE CAPITAL DE GIRO

Fonte: Revista Balanço Anual da Gazeta Mercantil 2004.

4.2. Revisão Teórica de Análise Fatorial

Desta forma, foi elaborado primeiramente o método de análise fatorial para a determinação de fatores que compõem as variáveis apresentadas. Segundo HAIR, ANDERSON, TATHAM e BLACK (1998, p. 89), um fator representa uma combinação linear de variáveis originais. Segundo AAKER, KUMAR e DAY (1998, p. 582), a análise fatorial serve para a combinação de variáveis que criam novos fatores. No âmbito da análise fatorial, o que se pretende é a identificação de possíveis associações entre as variáveis observacionais, de modo que se defina a existência de um fator comum entre elas. Assim, pode-se dizer que a análise fatorial, ou análise de fator comum tem como objetivo a identificação de fatores ou *constructos* subjacentes às variáveis observacionais, o que, sem dúvida, contribui para facilitar sobremaneira a interpretação dos dados (RODRIGUES, 2002, p. 76).

Segundo JOHNSON e WICKERN (1992, p. 396), na análise do fator comum, as variáveis são agrupadas em função de suas correlações. Isto significa que variáveis que compõem um determinado fator devem ser altamente correlacionadas entre si e fracamente correlacionadas com as variáveis que entram na composição do outro fator.

O modelo de análise fatorial estima os fatores e as variâncias, de modo que as covariâncias ou as correlações previstas pelo estejam o mais próximo possível dos valores observados. Para tal os métodos de estimação ou extração mais usados são o das componentes principais e o da máxima verosimilhança. Neste trabalho será usado o método das componentes principais.

Antes de iniciar a análise fatorial deve-se explorar individualmente cada variável em termos de *outliers* e do enviesamento da distribuição. Embora a normalidade não seja um pressuposto necessário de análise fatorial, tanto as distribuições muito enviesadas como os *outliers* podem distorcer os resultados, uma vez que alteram as estimativas das médias e dos desvios padrões, modificando as estimativas das covariâncias e das correlações. Deve-se igualmente analisar o número de não respostas, pois quando são elevadas (correspondem pelo menos a 20% dados) e não são aleatórias, podem distorcer os resultados da análise fatorial.

Será apresentado um roteiro básico que, segundo PESTANA E GAGEIRO (2000, p. 389 a 427), é fundamental para se avaliar a adequação da análise fatorial.

4.2.1. Matriz de correlações

Para se poder aplicar o modelo de análise fatorial deve haver correlação entre as variáveis. Se essas correlações forem pequenas é pouco provável que estas variáveis partilhem fatores comuns. Assim, um valor de correlação pode ser considerado aceitável se estiver acima de 0,4.

4.2.2. KMO e Teste de Esfericidade de Bartlett

O segundo passo é o teste KMO (Kaiser-Meyer-Olkin) e o teste de esfericidade de Bartlett. O teste de esfericidade de Bartlett pode ser usado para testar a hipótese da matriz das correlações ser a matriz identidade com determinante igual a 1. Este teste requer que os dados provenham de uma população normal multivariada. Caso o nível de significância do teste seja inferior a 5% para uma significância definida de 5%, deve-se rejeitar a hipótese da matriz das correlações entre as variáveis ser a identidade, mostrando, portanto, que

existe correlação entre as variáveis. Caso tal não se verificasse dever-se-ia reconsiderar a utilização deste modelo fatorial.

O teste KMO, que varia entre 0 e 1, testa a adequação da amostra quanto ao grau de correlação parcial entre as variáveis, que deve ser pequeno. O KMO perto de 1 indica coeficientes de correlação parciais pequenos, enquanto valores próximos de zero indica que a análise fatorial pode não ser adequada, pois existe uma correlação fraca entre as variáveis. A análise dos valores do KMO está apresentada no quadro abaixo.

Quadro 1: Teste KMO (Kaiser-Meyer-Olkin)

KMO	Análise Fatorial
1 – 0,9	Muito boa
0,8 – 0,9	Boa
0,7 – 0,8	Média
0,6 – 0,7	Razoável
0,5 – 0,6	Má
< 0,5	Inaceitável

4.2.3. Matriz Anti-Imagem

A matriz antiimagem é uma medida da adequação amostral de cada variável para uso da análise fatorial, onde pequenos valores na diagonal levam a considerar a eliminação da variável.

As medidas da adequação da amostra (MAS) são os valores da diagonal principal da matriz, e quanto maiores melhor será a análise fatorial. Os valores fora da diagonal representam o simétrico da matriz de correlações, e deverão ser pequenos para a aplicação do modelo das componentes principais.

4.2.4. Comunalidades

A proporção de variância de cada variável explicada pelas componentes principais retidas designa-se por comunalidade. As comunalidades exibem o valor inicial e após a extração do número desejado de fatores. As comunalidades iniciais são iguais a 1, e após a extração variam entre 0 e 1, sendo 0 quando os fatores comuns não explicam nenhuma variância da variável e 1 quando explicam toda a sua variância.

4.2.5. Fatores retidos e variância total explicada

Pelo critério de Kaiser escolhe-se o número de fatores a reter, em função do número de valores próprios maiores que 1. Os valores próprios são ordenados por tamanho. No método de extração de componentes principais, a soma dos valores próprios iguala o número de variáveis.

4.2.6. Matriz de componentes antes e após rotação

Conforme dito anteriormente, o modelo de análise fatorial estima os fatores *loadings* e as variâncias, de modo a que as covariâncias ou as correlações previstas pelo modelo estejam o mais próximo possível dos valores observados, em geral pelo método de extração das componentes principais ou pelo método de estimação da máxima verosimilhança. A matriz de componentes inicial apresenta os coeficientes ou pesos (*loadings*) que correlacionam as variáveis com os fatores antes da rotação.

Na matriz de componentes, a soma (em coluna) dos quadrados dos *loadings* das variáveis para cada fator é o valor próprio dos componentes. Caso não haja nenhum *loading* elevado em ambos os fatores, é necessário o uso de rotações.

Os métodos de rotação podem ser ortogonais ou oblíquos. O primeiro produz fatores que não estão correlacionados entre si chamados fatores ortogonais, os quais são interpretados a partir dos seus *loadings*. Na rotação oblíqua, os fatores estão correlacionados, e para a interpretação da solução torna-se necessário considerar simultaneamente a matriz das correlações e os *loadings*. As rotações Varimax, Quartmax e Equamax são exemplos de rotações ortogonais, enquanto que o Direct Oblimin e o Promax são exemplos de rotações oblíquas. O Varimax, por exemplo, minimiza o número de variáveis com elevados *loadings* num fator. Em geral consideram-se significativos os *loadings* maiores ou iguais a 0,5 por serem pelo menos responsáveis por 25% da variância.

A matriz de componentes após rotação ortogonal tem como objetivo extremar os valores dos *loadings*, de modo que cada variável se associe apenas a um fator. Quanto menos iterações forem necessárias, melhor os dados se adequam ao modelo.

Na análise fatorial as variáveis com baixos *loadings* devem ser eliminadas, de forma que sejam utilizadas apenas as variáveis com elevados *loadings*.

4.3. Revisão Teórica de Análise de Conglomerados

Com a determinação dos fatores obtidos na análise fatorial, parte-se para o método multivariado de análise de conglomerados. A análise de conglomerados, ou *clusters*, visa o agrupamento não mais de variáveis, mais de indivíduos, ou unidades observacionais, segundo determinados critérios de distância entre os respectivos vetores de dados. Como afirmam JOHNSON e WICHERN (1992, p. 573), a análise de *clusters* corresponde também a uma importante técnica exploratória que busca identificar uma estrutura de agrupamentos com o intuito de avaliar a dimensionalidade dos dados, identificar *outliers* e fornecer interessantes hipóteses acerca de associações.

A análise de conglomerados tem como objetivo principal neste estudo determinar em qual *cluster* se insere cada uma das empresas de internet, permitindo avaliar algumas características semelhantes de atuação, em função de aspectos regionais.

5. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Primeiramente, apresenta-se a matriz de correlação entre as variáveis, que pode oferecer algum subsídio à determinação dos fatores:

Quadro 2: Matriz de Correlação das Variáveis

	RL	LO	LL	ATIVO	PL	EBITD	ENDI V	CG	NCG
RL	1,00 0								
LO	,364	1,000							
LL	,304	,910	1,000						
ATIVO	,932	,333	,161	1,000					
PL	,821	,333	,108	,970	1,000				
EBITD	,181	,798	,794	,023	-,045	1,000			
ENDIV	,972	,261	,240	,850	,696	,166	1,000		
CG	-,661	,221	,220	-,713	-,681	,432	-,629	1,000	
NCG	,894	,210	,234	,709	,529	,299	,947	-,535	1,000

Por meio da análise do quadro 2, percebe-se que as variáveis oferecem correlações altas entre si. Além disso, o teste KMO sugere que há a adequação da amostra quanto ao grau de correlação parcial entre as variáveis, uma vez que ofereceu um valor de 0,696. Por fim, o teste de esfericidade de Bartlett permite rejeitar a hipótese nula de que o modelo não é adequado em razão das associações verificadas, uma vez que o nível de significância é 0 (menor do que 5%).

No quadro 3 são apresentados os resultados da matriz antiimagem. Para todas as variáveis, os valores na diagonal são elevados mostrando que estas variáveis são adequadas para o uso da análise fatorial.

Quadro 3: Matriz Anti-Imagem

	RL	LO	LL	ATIVO	PL	EBIT	ENDIV	CG	NCG
RL	,728 ^a								
LO	,355	,620 ^a							
LL	-,774	-,771	,479 ^a						
ATIVO	-,137	,005	,100	,641 ^a					
PL	,121	-,016	-,085	-1,000	,591 ^a				
EBIT	,373	-,444	-,171	-,125	,122	,561 ^a			
ENDIV	,104	-,026	-,076	-,999	,999	,133	,637 ^a		
CG	-,427	-,522	,676	,087	-,076	-,331	-,078	,658 ^a	
NCG	-,631	,142	,448	,121	-,111	-,857	-,121	,438	,651 ^a

a Medidas de Adequação da Amostra (MSA).

O método de extração ou estimação utilizado foi o das componentes principais. Os resultados das comunalidades são apresentados a seguir:

Quadro 4: Comunalidades

	Inicial	Extração
RL	1,000	,985
LO	1,000	,899
LL	1,000	,881
ATIVO	1,000	,930
PL	1,000	,772
EBIT	1,000	,882
ENDIV	1,000	,901
CG	1,000	,834
NCG	1,000	,746

Método de Extração: Análise de Componentes Principais.

Pode-se notar que todas as variáveis têm uma forte relação com os fatores retidos, por terem comunalidades elevadas conforme a coluna Extração.

O próximo passo consiste em analisar os fatores retidos e a variância total explicada. Os resultados encontram-se no quadro 5.

Quadro 5: Variância Total Explicada

Comp o n e n t e s	Valores Próprios Iniciais			Extração Soma do Quadrado dos <i>Loadings</i>			Rotação Soma dos Quadrados dos <i>Loadings</i>		
	Total	% de Variân cia	Cumula tiva %	Total	% de Variân cia	Cumul ativa %	Total	% de Variân cia	Cumulati va %
1	5,055	56,163	56,163	5,055	56,163	56,163	4,907	54,517	54,517
2	2,775	30,834	86,997	2,775	30,834	86,997	2,923	32,480	86,997
3	,731	8,119	95,116						
4	,279	3,096	98,212						
5	,136	1,516	99,728						
6	,016	,173	99,901						
7	,007	,080	99,981						
8	,002	,019	100,00 0						
9	,000	,000	100,00 0						

Método de Extração: Análise dos Componentes Principais.

Na coluna Total do quadro 5, os valores próprios ordenam-se por tamanho. Na situação inicial, a soma dos valores próprios iguala o número de variáveis, que neste caso são 9. Como existem dois valores próprios maiores do que 1 o número de fatores retidos é 2.

Assim sendo, ao invés de se trabalhar com nove variáveis, pode-se trabalhar com duas para a análise de conglomerados, já que estes dois fatores conseguem explicar 87% da associação total entre os dados. A seguir, apresenta-se a matriz dos fatores rotacionados pelo método Varimax:

Quadro 6: Matriz dos Fatores Rotacionados – Método Varimax

	FATORES	
	1	2
RL	,969	,212
ATIVO	,961	,078
ENDIV	,934	,166
PL	,878	,029
NCG	,837	,213
CG	-,804	,433
EBITD	-,017	,939
LL	,110	,932
LO	,198	,927

Método de Extração: Análise dos Componentes Principais.
Rotações convergidas após 3 iterações.

O quadro 6 confirma a existência de dois fatores. O fator 1 é formado pela Receita Líquida, Ativo, Endividamento, Patrimônio Líquido, Necessidade de Capital de Giro, Capital de Giro. Já o fator 2 compõe as variáveis Ebitda, Lucro Líquido e Lucro Operacional.

O próximo passo é a determinação dos *clusters* para cada empresa de internet, em função da utilização de duas variáveis, que agora correspondem aos fatores 1 e 2 determinados anteriormente.

Aplicando a análise de conglomerados em função dos dois fatores obtidos, verifica-se que três *clusters* foram formados. Um dos *clusters* é formado apenas pela empresa Terra Networks do RS, o segundo *cluster* é formado pelas empresas Americanas.Com e Submarino de SP, ficando as empresas restantes no terceiro *cluster*.

Conforme os resultados obtidos nos gráficos 1 e 2, o estado do Rio Grande do Sul apresentou melhor performance, representando uma receita líquida total de 40,33%. A empresa Terra Networks do RS representa 99% da receita total do RS, o que justifica a formação do primeiro *cluster*. Já as empresas Americanas.Com e Submarino de SP representam 32,7% da receita líquida total (88% da receita de SP), e por apresentarem indicadores semelhantes, possibilitou-se a formação do segundo *cluster*.

Este resultado pode oferecer um entendimento de que empresas que se situam em localidades semelhantes, com mesmas incidências de tributação estadual, acabam por apresentar indicadores econômico-financeiros semelhantes e, por consequência, situam-se em mesmos *clusters*.

6. CONCLUSÕES E PERSPECTIVAS

Primeiramente, é importante ressaltar o poder explicativo que as variáveis relacionadas a indicadores econômico-financeiros têm sobre o comportamento da atuação de empresas de

internet no Brasil. É claro que não se pode ignorar o fato de que muitos destes indicadores representam um resultado da própria operação organizacional, porém o incremento dos mesmos ao longo do tempo oferece subsídios a esta análise.

Além disso, a melhor performance percebida pelas empresas Terra Networks do RS e Americanas.Com e Submarino de SP, além de características semelhantes das empresas do estado de São Paulo, oferecem uma explicação para a formação de dois *clusters* específicos.

Estes resultados, embora restritos a um número limitado de empresas e estados, pode oferecer algum subsídio em termos de alocação de recursos e gerenciamento dos aspectos que compõem os indicadores financeiros. Para ser generalizado a todos os estados do Brasil, deve-se coletar dados sobre outras empresas de internet que atuam em diversas regiões geográficas do país.

Por fim, percebe-se uma nova fase do setor de internet no Brasil, com um incremento exacerbado de competitividade, o que faz com que os *players* ofertantes de produtos e serviços apresentem diferenciação no nível de serviços, quer seja pelo gerenciamento de estoques, quer seja pela automação ou pela melhora dos indicadores de atendimento e prazo de entrega. Independentemente de qual seja o principal atributo de diferenciação, é importante ressaltar que os índices financeiros representam o sucesso da organização no médio prazo e, portanto, devem ser modelados e avaliados para que sejam determinados padrões e tendências de comportamento.

Sugestões para novos estudos podem se relacionar com a continuidade da análise dos impactos dos indicadores financeiros para outras regiões ou outros estados do Brasil. Além disso, outras variáveis também podem ser incluídas na modelagem, como indicadores sociais e capacidade de consumo da região de atuação. Assim, podem ser fornecidas respostas diferentes do comportamento destas variáveis sobre a formação de *clusters* de empresas, permitindo o alcance de resultados mais profundos sobre a atuação da internet no Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

AAKER, D.; KUMAR, V.; DAY, G. **Marketing research**. 6. ed. New York: John Wiley & Sons Inc., 1998.

AMOR, D. **A (r)evolução do e-business: vivendo e trabalhando em um mundo interconectado**. São Paulo: Makron Books, 2000.

ANDRADE, A. R. Comportamento e estratégias de organizações em tempos de mudança sob a perspectiva da tecnologia da informação. **Caderno de Pesquisas em Administração**, São Paulo, v. 9, n.2, p. 49-58, abril/junho 2002.

BLECHER, N. Quer preço ou quer marca? **Revista Exame**, ano 35, n. 12, p. 32, 2001.

CASTELLS, M. **A sociedade em rede**. São Paulo: Paz e Terra, 1999 (A Era da Informação, v. 1).

FREEMAN, C; SOETE, L. **Technical change and full employment**. São Paulo, ano 27, junho 2003.

GAZETA MERCANTIL. Revista Balanço Anual da Gazeta Mercantil. São Paulo, n.28, agosto 2004.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Multivariate data analysis**. 5. ed. New Jersey: Prentice Hall, 1998.

JOHNSON, R.; WICHERN, D. **Applied multivariate statistical analysis**. 3 ed. New Jersey: Prentice Hall, 1992.

NORONHA, J.M. **A empresa e o processo de mudança**. Notas de aula, Rio de Janeiro, 2000.

PESTANA, M. H.; GAGEIRO, J.N. **Análise de dados para ciências sociais: a complementaridade do SPSS**. Edições Silabo, Ltda, 2000.

RODRIGUES, M. C. P. Potencial de desenvolvimento dos municípios Fluminenses: uma metodologia alternativa ao IQM, com base na análise fatorial exploratória e na análise de *clusters*. **Caderno de Pesquisas em Administração**, São Paulo, v. 9, n. 1, p. 75-89, jan - mar 2002.

SHAPIRO, C; VARIAN, H. R. **A economia da informação**: como os princípios econômicos se aplicam à era da *Internet*. Rio de Janeiro: Campus, 1999.

TAPSCOTT, D. **Economia digital**: promessa e perigo na era da inteligência em rede. São Paulo: Makron Books, 1997.