ESTRUTURA E DINÂMICA DAS REDES AÉREAS

Alunos: Antonio Rossano e Eduardo Henrique Filizzola Colombo
Orientadora: Celia Anteneodo

Introdução

Muitos aspectos dos sistemas complexos podem ser modelados mediante redes onde os elementos e as suas interações estão representados por vértices e arestas, respectivamente. A maioria das redes biológicas e sociais apresenta uma topologia não trivial em que o padrão de conexões entre seus elementos não é puramente regular, nem aleatório, tendo padrões estatísticos únicos, como distribuição de graus com caudas longas, menor caminho entre arestas que cresce lentamente com o tamanho do sistema e alto coeficiente de aglomeração [1,2]. Por inúmeros fatores, mas, principalmente por estar imersa em meio complexo, sofrendo influência econômica e social, as redes aéreas são redes complexas. Um exemplo ilustrativo da estrutura de uma rede aérea é apresentado na Fig. 1, em que os vértices (aeroportos) aparecem unidos por arestas se existem voos entre eles. Nota-se desde a existência de aeroportos com uma única conexão até aeroportos com alta conectividade (*hubs*).

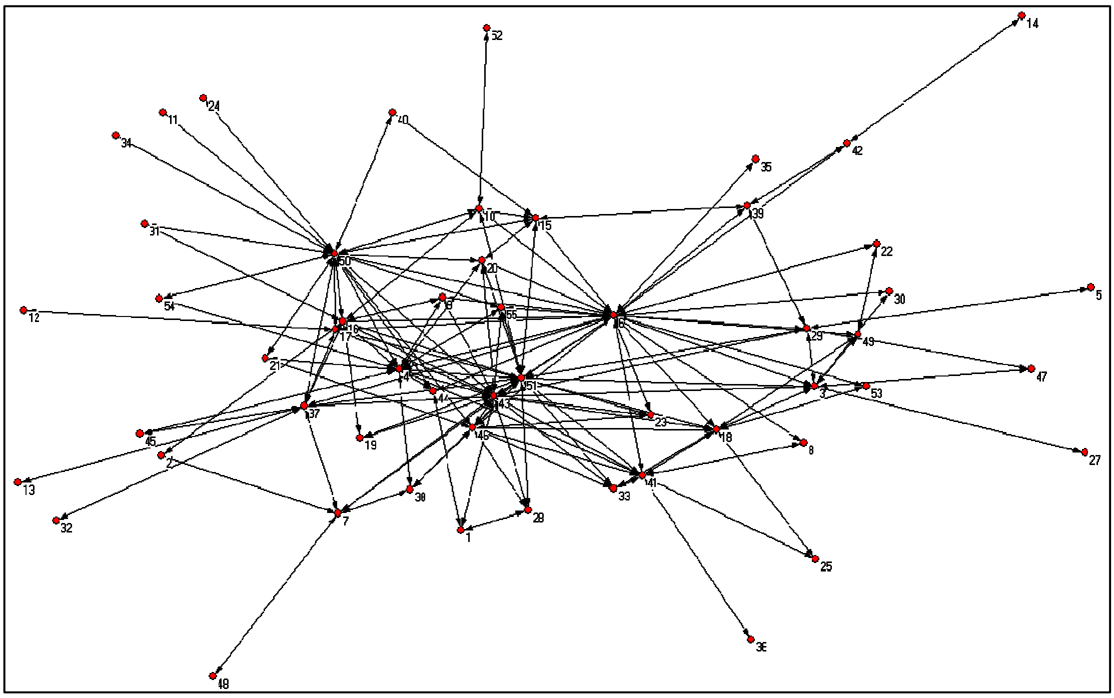

Fig. 1: ***Rede da Companhia Aérea Gol (dados obtidos do site da companhia, 2009).***

Objetivos

Estudaremos algumas das características estruturais e dinâmicas das redes aéreas, como primeiro passo visando diminuir a sua vulnerabilidade e otimizar a eficiência dos fluxos, o que motiva o presente trabalho.

Serão adquiridos e analisados dados de voos domésticos de EUA, de diversas companhias, fornecidos pela *Research and Innovative Technology Administration* do departamento de transporte dos Estados.

Metodologia

Como, em muitos casos reais, não conhecemos, e inclusive possa ser impossível determinar teoricamente por primeiros princípios, as leis de organização e evolução de uma rede, o estudo empírico de redes reais é um dos meios principais para abordar o entendimento de redes complexas e inclusive dos aspectos dos sistemas complexos por elas descritos.

Com efeito, existem inúmeras tentativas de modelar a evolução e crescimento de grafos até atingir as características de redes complexas reais a partir de regras simples que estabelecem as conexões com certas tendências de concentração em nós específicos (ligação preferencial), e outras mais sofisticadas que tentam reproduzir as propriedades estatísticas que emergem nas redes reais. Assim a caracterização de uma rede a partir dos dados empíricos representa uma base para o posterior modelado teórico da mesma.

A estrutura das conexões de uma rede escrita na forma matricial permite expressar não somente a informação da existência, ou ausência, de conexão entre dois vértices e o seu direcionamento (matriz de adjacência), mas também, a intensidade das interações (matriz de pesos). De modo a caracterizar a rede são usados parâmetros pré-definidos, como a conectividade, k , que mede o número de conexões que partem de um dado vértice, e o menor caminho médio, L, entre dois vértices. Para definir e analisar essas e outras características são úteis noções de teoria de grafos, álgebra linear e física estatística.

Dados analisados

A partir do banco de dados disponível no site (http://www.rita.dot.gov/) do departamento de transporte dos Estados Unidos, foram obtidas e analisadas informações sobre os voos domésticos de todas as companhias aéreas ali cadastradas (um total de 194). Entre os anos de 1990 e 2010 foi observada a evolução de um conjunto de parâmetros globais para toda a rede de fluxo de voos e para alguns anos particulares foi feita uma análise mais detalhada em relação à estrutura de cada camada de voos de companhias, observando a distribuição de probabilidade de graus e dos menores caminhos entres vértices. Foi também objetivo deste trabalho entender como o tamanho da companhia influencia no uso da malha já predefinida pelos aeroportos existentes, propondo de modo qualitativo a diferença de estruturas em função do número de aeroportos operantes.

Resultados e discussão

A definição de uma rede ou grafo pode ser feita por meio da matriz de adjacência.

*Seja **A** a é matriz de adjacência de um grafo de N vértices, o elemento A_{ij} é igual a um, se existir ligação i → j , caso contrário, A_{ij} é nulo para todo* i,j=1,2, *N*.

Note-se que a matriz ***A*** não é necessariamente simétrica. Em particular no caso de redes aéreas vôos podem ocorrer num determinado sentido somente.

Definimos, então, conectividade.

A conectividade $k_{OUT,i}$ mede o número de conexões que partem do i-ésimo vértice, sendo

$$k_{OUT,i} = \sum_j A_{ij}.$$

Analogamente podemos definir $k_{IN,i}$, ou também k_i a partir da matriz de adjacência simétrica (ou não-direcionada, com elementos unitários se existe ligação entre *i* e *j* sem importar o sentido). A conectividade é uma característica particular de cada vértice, que informa certo grau de hierarquia em relação aos pólos de conexões, como no caso de grandes aglomerados, e servidores centrais.

Considerando a conectividade de todos os vértices da rede, pode ser definida *a distribuição de probabilidade p(k) de um vértice escolhido aleatoriamente ter k graus, ou conectividade k*. Teoricamente, pode ser visto, que em certos casos, a construção de todas as características de uma rede parte da distribuição *p(k)*. No caso de redes aleatórias, um desenvolvimento em funções geradoras permite mostrar, de forma geral, para qualquer distribuição de graus arbitrária, a relação entre *p(k)* e outras diversas quantidades que medem propriedades da rede [3]. Porém, a obtenção de *p(k)* não é suficiente para caracterizar por completo redes complexas.

Outra medida relevante sobre uma rede é baseada na proximidade dos nós. Está claro que no contexto de redes, a proximidade não está vinculada a uma distância geográfica, mas à distância percorrida sobre arestas da rede. A distância entre dois vértices pode ser obtida de varias maneiras. Uma delas é fazendo uso do algoritmo de Dijkstra, outra, aplicando o seguinte teorema:

Seja $\boldsymbol{A}$ a matriz de adjacência de um grafo. O elemento $A^{(n)}{}_{ij}$da n-ésima potência de $\boldsymbol{A}$, determina o número de caminhos de n passos de i para j.

A matriz $\boldsymbol{S}$ que determina todos os menores caminhos entre vértices pode ser obtida fazendo $S_{ij} =$ m, com *m*> 0, tal que $A^{(m)}{}_{ij} = 1$ e $A^{(n)}{}_{ij} = 0$ para *n*<*m*. Desta forma pode-se simplesmente aplicar a força bruta no cálculo desse problema, já que elevar, em uma ordem típica para redes aéreas, à décima potência, uma matriz de 200x200 não é grande problema computacional. Este caminho é válido quando ***A*** é simétrica (grafo não-direcionado) e também quando $\boldsymbol{A}$ tem forma arbitrária, caso direcionado. A partir de ***S*** pode ser obtida a distribuição $g(L)$ dos menores caminhos, descrevendo a probabilidade de um par de vértices quaisquer estarem distantes de L passos.

O gráfico da Fig. 2 resume a evolução de diversas medidas computadas para a rede aérea norteamericana (nacional) como um todo ao longo do período 1990-2000. Para definir a existência de arestas e determinar os pesos foram considerados voos de passageiros, sem escalas, entre aeroportos. O gráfico mostra o número de aeroportos operantes *N*, a conectividade k, tanto direcionada quanto simétrica. Outra medida que consta é a força média *s* (*strength*),

$$s_{OUT,i} = \sum_j W_{ij},$$

onde, ***W*** a matriz de pesos, que neste caso, é construída levando em conta o número de voos de passageiros entre *i* e *j* (sem escalas). Igual que no caso da matriz de adjacência, pode interessar s_i, sem direcionamento. Poderíamos avaliar o peso das conexões de outras maneiras, como pelo fluxo de passageiros ou de carga, que podem ser importantes em diversas aplicações. Também foi calculado o diâmetro *D* da rede que é definido pela maior das distância entre vértices, o coeficiente de reciprocidade *R* que mede o grau de simetria da rede (sendo 1 para a rede completamente simétrica), o coeficiente de aglomeração *cc* que é a média dos coeficientes de cada vértice

$$cc_i = \frac{1}{k_i(k_i-1)} \sum_{j,k} a_{ij} a_{jk} a_{ki}$$

e a assortatividade

$$r = \frac{\frac{1}{M}\sum k_i k_j a_{ij} - \left[\frac{1}{2M}\sum (k_i + k_j) a_{ij}\right]^2}{\frac{1}{2M}\sum [(k_i)^2 + (k_j)^2] a_{ij} - \left[\frac{1}{2M}\sum (k_i + k_j) a_{ij}\right]^2},$$

onde *M=Nk/2,* que mede a ligação preferencial entre vértices de similar conectividade [4].

Dentro do período de tempo estudado, observa-se (ver Fig. 2) um suave aumento no número de aeroportos operantes, com um pulo em 1995 que volta ao nível normal em 1997. Existem diferenças entre os valores computados mensalmente (linhas) e os acumulados a cada ano (símbolos). Além do aumento no número de aeroportos, houve um leve crescimento no número de voos por aeroporto por mês (*s*). Contudo as medidas que caracterizam a topologia da rede são praticamente invariantes exceto por pequenas flutuações e mudanças sazonais, indicando uma rede com estrutura consolidada. Uma das medidas que apresenta maior variação é o coeficiente *r* definido acima, cuja diminuição indica a tendência crescente a ligar vértices com diferente grau de conectividade. Cabe observar que o curto caminho médio (entre 2 e 3 arestas) indica a característica de "mundo pequeno".

O período 2000-2010 também disponível na base de dados foi analisado, mostrando características de estabilidade semelhantes. Existe uma descontinuidade em 2001-2002 devida aparentemente à inclusão de aeroportos menores na base de dados, mas que não representa uma mudança estrutural da rede.

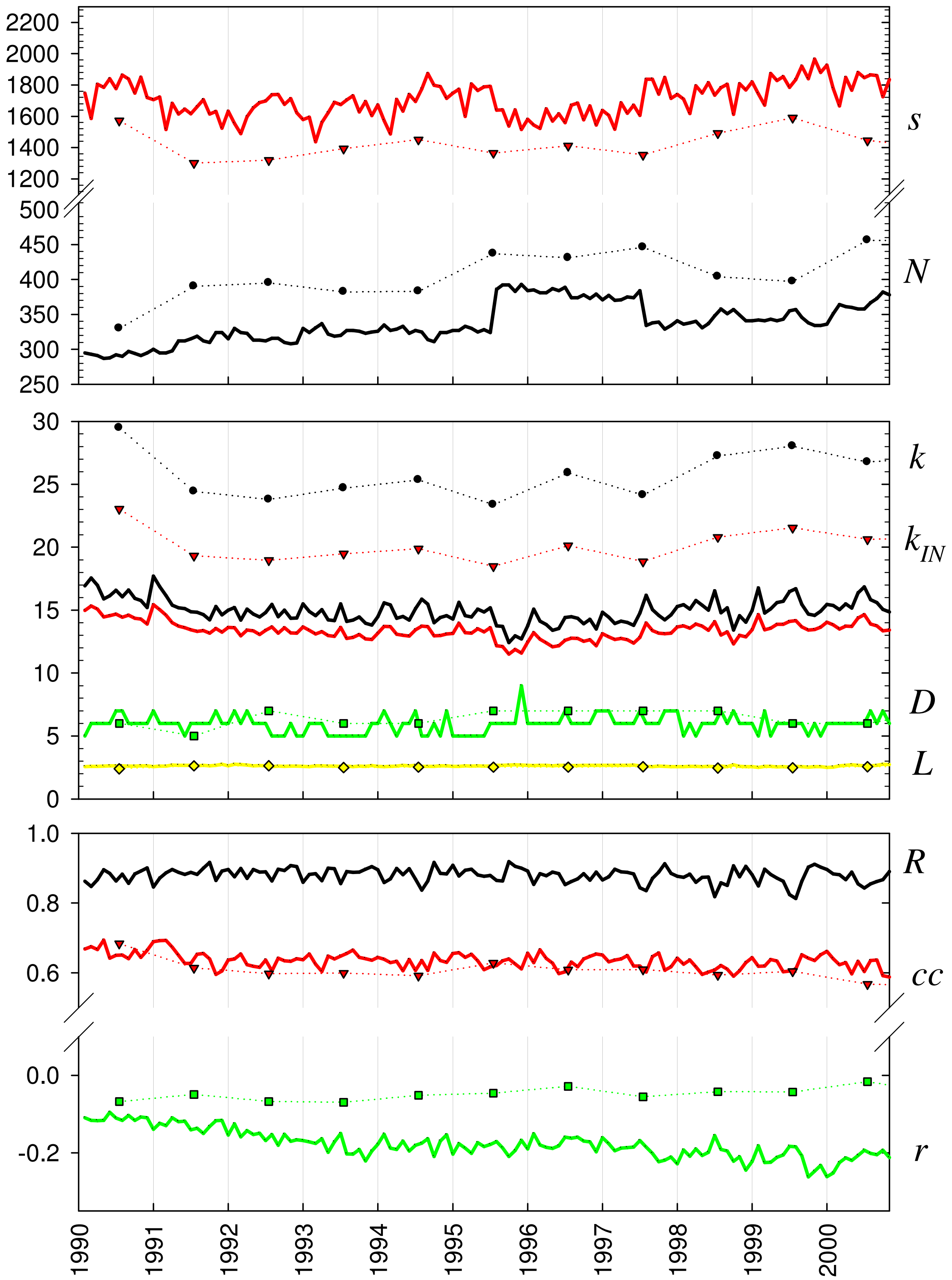

Fig. 2: ***Diversas medidas da rede aérea norteamericana (malha nacional) como função do tempo. Linhas continuas representam valores mensais, símbolos valores acumulados anualmente.***

Medidas da conectividade e menor distância entre vértices são extremamente importantes para o entendimento da estrutura de conexões. A Fig. 3 mostra a distribuição das conectividades direcionadas para a rede completa. A linearidade do gráfico nas escalas linear-logarítmica indica uma dependência *P(k)* logarítmica. Porem, acumular as informações de todas as companhias pode distorcer a estatística de cada companhia individual. Assim analisamos as companhias separadamente. A Fig. 4 mostra a distribuição da conectividade para as 5 maiores companhias em 2010. Em quanto algumas companhias (as 3 maiores) apresentam uma distribuição exponencial, outras (4a e 5a) possuem conectividade mais próxima de uma lei de potência, se aproximando da característica "livre de escala".

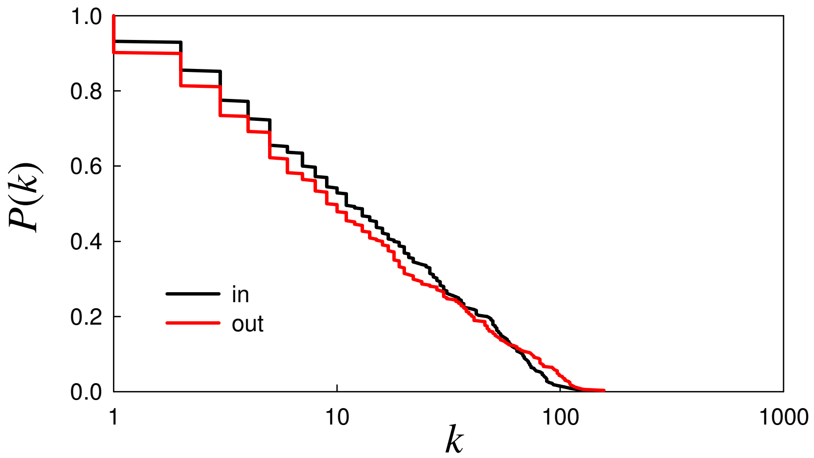

Fig. 3: ***Distribuição acumulada (complementar) da conectividade direcionada da rede completa (1990).***

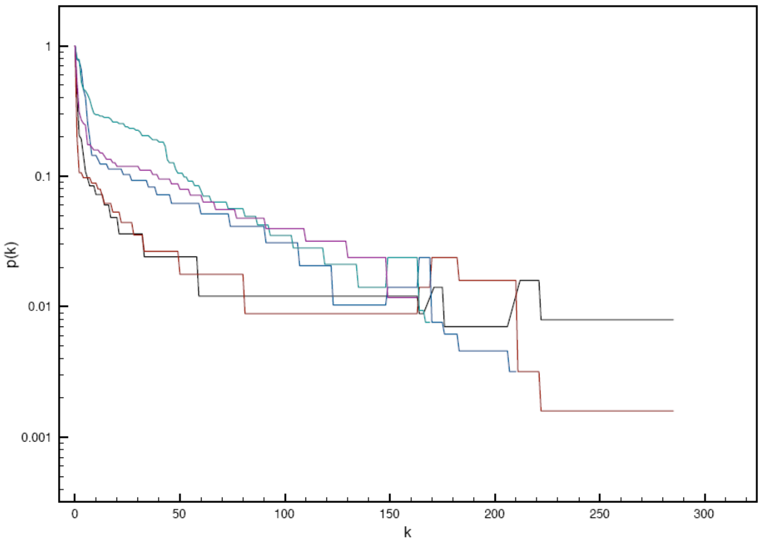

Fig. 4: ***Distribuição acumulada (complementar) da conectividade k de cada uma das cinco maiores companhias operando na rede nacional norteamericana (2010).***

A distribuição de menores caminhos também foi obtida para cada companhia. Em todos os casos o histograma apresenta um máximo bem definido, exceto por alguns casos de companhias extremamente pequenas, 2-5 aeroportos operantes, em que a estatística é insuficiente.

Para cada companhia foi analisado a correlação entre diversas grandezas. O gráfico da Fig. 5 mostra que a conectividade, o menor caminho médio e o caminho típico entre aeroportos crescem com o número de aeroportos operantes da companhia, indicando que redes maiores (neste caso relacionamos diretamente tamanho ao número de aeroportos) apresentam uma estrutura que fornece a possibilidade da criação de grandes aglomerados (*clusters*), onde em geral voos até o destino final efetuam uma única escala (duas arestas de distância).

Com devido cuidado podemos interpretar o tamanho das companhias como tempo de evolução o que permite extrair informações sobre a dinâmica.

Resultados para outras redes aéreas podem ser encontrados na literatura, e.g., a rede nacional Brasileira [5], a da China [6] ou a mundial [7].

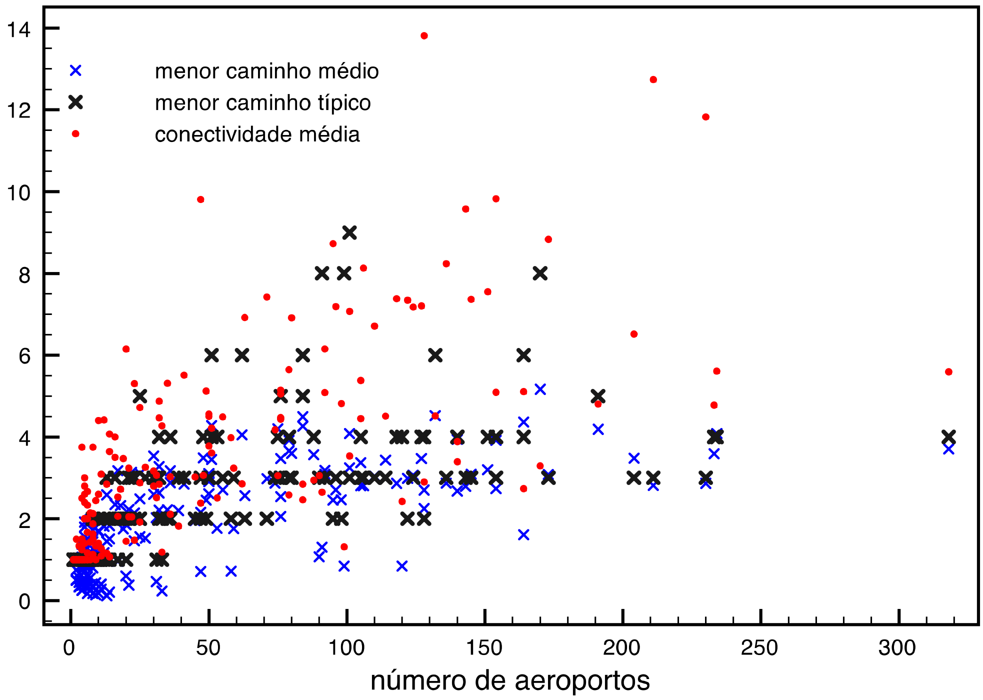

Fig. 5: ***Conectividade, menor caminho médio e caminho típico entre aeroportos vs. número de aeroportos operados, para cada companhia (1990).***

Conclusões

O estudo empírico das redes aéreas norte-americanas trouxe informações de como cada companhia organiza sua sub-rede de voos dentro da estrutura espacial pré-definida de aeroportos. Pode ser visto, mesmo que qualitativamente, que empresas de diferentes tamanhos tem estruturas distintas. O aumento do tamanho de uma companhia tem como consequência o aumento do menor caminho médio e da conectividade. Em termos gerais, pode ser dito que companhias maiores criam centros de conexão com certo detrimento da coesão da rede.

Os dados disponíveis durante o período de 1990-2010 não manifestam mudanças estruturais, indicando uma rede já estabilizada.

Entretanto, como perspectiva futura, pode-se buscar casos de companhias criadas recentemente esclarecer a dinâmica de crescimento das redes aéreas reais.

Referências

1 - WATTS. D. J.. STROGATZ S. H.. Collective dynamics of small world networks.**Nature**. 393. 440–442. 1998.

2 - BARABASI. A.-L.. ALBERT R. Emergence of scaling in random networks. **Science**. v.286. 509-512. oct. 1999.

3 – M. E. J Newman; S. H. Strogatz; D. J. Watts. Random graph with arbitrary degree distribution and their aplications. **Physical Review** 64. 026118. 2001

4 - RODRIGUES. F.A. **Caracterização, classificação e análise de redes complexas**. Tese de doutorado - Instituto de Física de São Carlos (USP). 2007.

5 - ROCHA. L. Structural evolution of the Brazilian airport network. **J. Stat. Mech.**. v.04. P04020. 2009.

6 – LI, W. CAI, X. Statistical analysis of airport network of China**, Physical Review** 69. 046106. 2004

7 – GUIMERA R. et al, The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles, **PNAS** v. 102. n. 22. 7795 2005.