Vanessa Helena Pereira

Redes complexas em presença de falhas induzidas

Dissertação apresentada à Faculdade de Tecnologia da Universidade Estadual de Campinas, como parte dos requisitos para a obtenção do grau de Mestre em Tecnologia.

Área de Concentração: Tecnologia e Inovação

Orientador: Prof. Dr. Varese Salvador Timóteo

Limeira
2010

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

Pereira, Vanessa Helena
P414r Redes complexas em presença de falhas induzidas / Vanessa Helena Pereira. --Limeira, SP: [s.n.], 2010.

Orientador: Varese Salvador Timóteo.
Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Tecnologia.

1. Redes complexas. 2. Teoria dos grafos. 3. Falha de sistema. I. Timóteo, Varese Salvador. II. Universidade Estadual de Campinas. Faculdade de Tecnologia. III. Título.

Título em Inglês: Complex networks in presence of induced failures
Palavras-chave em Inglês: Complex networks, Graph theory, System failure
Área de concentração: Tecnologia e Inovação
Titulação: Mestre em Tecnologia
Banca examinadora: Sérgio Szpigel, Vitor Rafael Coluci
Data da defesa: 17/12/2010
Programa de Pós Graduação: Tecnologia

DISSERTAÇÃO DE MESTRADO ACADÊMICO

Redes Complexas em presença de falhas induzidas

Autor: Vanessa Helena Pereira
Orientador: Prof. Dr. Varese Salvador Timóteo

A Banca Examinadora composta pelos membros abaixo aprovou esta Dissertação:

Prof. Dr. Varese Salvador Timóteo
FT/UNICAMP

Prof. Dr. Vitor Rafael Coluci
FT/UNICAMP

Prof. Dr. Sérgio Szpigel
CCH/ UPM

Dedico este trabalho
ao único Senhor e Salvador:
Jesus Cristo.

Agradecimentos

Sobretudo agradeço ao único que é digno de receber a honra, a glória, a força e o poder: o Senhor Deus, maravilhoso, tremendo, justo e fiel.

Agradeço o professor Dr. Ted Lewis, do *Center for Homeland Defense and Security, Naval Postgraduate School,* pelo exemplo de pessoa, apoio, incentivo, amizade, e valiosos ensinamentos.

A minha família, pela compreensão e persistente apoio em todos os momentos.

Ao meu professor orientador Dr. Varese Salvador Timóteo, pelos ensinamentos e grande incentivo.

A todos os professores, colegas e funcionários da que ajudaram de forma direta ou indireta neste trabalho.

A CAPES, pelo auxílio financeiro.

"Mas a sabedoria que do alto vem é,
primeiramente pura, depois pacífica, moderada,
tratável, cheia de misericórdia e de bons frutos,
sem parcialidade e sem hipocrisia."

Tiago 3:17, Bíblia Sagrada

Resumo

A necessidade da operação intermitente de redes complexas leva ao estudo das falhas nas redes de topologia livre de escala (*Scale-Free*) de Barabási-Albert. Neste trabalho introduzem-se as teorias fundamentais ao estudo das redes complexas, além da revisão de vários trabalhos científicos relacionados às falhas e aos mecanismos de contenção destas. Utilizando o software *Attacker-Defender*, são construídas várias redes complexas *Scale-Free* de diferentes tamanhos, representadas por grafos. Estas redes são utilizadas para simular dois tipos de falhas mais frequentes: falhas aleatórias e falhas direcionadas aos *hubs* (nós com maior número de arestas incidentes) em duas etapas. Na primeira etapa, em dez tamanhos de redes são testadas quatro situações distintas. Na segunda etapa, em sete tamanhos de redes, são testadas dez diferentes vulnerabilidades. A partir da análise dos resultados da primeira etapa, observa-se qual dos quatro cenários analisados é o mais vantajoso para contenção de falhas nas redes. A análise da segunda etapa permite definir a descrição matemática do comportamento dos nós sobreviventes e atingidos no pós-falha, em cada uma das redes, para cada tipo de falha, através de métodos e funções específicas encontradas.

Palavras Chave: Redes complexas; Teoria dos grafos; Falha de sistema.

Abstract

The need for the intermittent operation of complex networks leads to the study of failures in these networks topology called Scale-Free, Barabási-Albert. In this work we introduce the fundamental theories to the study of complex networks, in addition to reviewing various scientific studies related to the failures and the mechanisms against cascade failures. Using the Attacker-Defender software, Scale-Free complex networks of different sizes are built, represented as graphs. These networks are used to simulate the two most common types of failures: random failures and attacks to hubs (nodes with the largest number of incident edges) in two steps. In the first step in ten sizes of networks are tested in distinct four cases. In the second step, in seven sizes of networks are tested ten different vulnerabilities. From the analysis of the results of the first step, it is observed which cases are best for the networks. The analysis of the second step provides the mathematical description of the behavior of the survivors and the affected nodes, after the failure in each network for each type of failure, through methods and specific functions was found.

Key Words: Complex networks, Graph theory, System failure.

Lista de Ilustrações

Lista de Tabelas

SUMÁRIO

1 INTRODUÇÃO

O que são sistemas complexos? Muitas vezes entendem-se sistemas complexos como aqueles complicados na compreensão e na dinâmica de seu funcionamento. Porém, é necessário esclarecer a principal distinção entre o que é considerado complicado e o que é complexo. Um carro pode ser considerado um sistema complicado, mas não necessariamente complexo. Isto porque seu comportamento não é complexo: se a direção é girada para a esquerda, esse carro fará uma trajetória para a esquerda e não para cima ou para baixo. Já um sistema complexo tem a característica da imprevisibilidade no seu comportamento e, portanto seu estudo é ainda mais desafiador. O tráfego de veículos, diferente do sistema do veículo em si, é um sistema complexo, devido ao seu comportamento complexo.

Nos últimos anos, pesquisadores passaram a dar grande importância aos sistemas complexos, uma vez que tais sistemas podem ser modelados como redes de conexões ou redes complexas. Percebeu-se que estruturas complexas dos mais diversos campos da ciência podem ser organizadas e modeladas como redes. Estas estruturas estão presentes em nosso cotidiano, a exemplo das redes de distribuição elétrica, redes sociais, redes rodoviárias, redes de computadores, redes de neurônios, etc. Tais redes são passíveis de modelagem, o que possibilita o estudo e análise de estratégias que possam aperfeiçoar tanto seu projeto quanto sua dinâmica.

Cidades, regiões, bem como a sociedade fazem parte de uma rede complexa de infraestrutura interdependente, como sistemas de energia elétrica, distribuição, comunicação, transporte, economias, mercados e até mesmo decisões humanas formam sistemas complexos. A descoberta de leis que governam tais sistemas, as correlações entre eles, a ocorrência de *blackouts*, etc., sugerem a necessidade crescente de modelos dinâmicos que revelem comportamento, pontos críticos, estado dos sistemas e causas que disparam ocorrências imprevisíveis (BARABÁSI, 2003).

Modelos que capturem todas as características são necessários para investigação do *design* e de técnicas para melhorar a resiliência e robustez desses sistemas. Mesmo as infraestruturas aparentemente individuais são muitas vezes conectadas. Nesta relação simbiótica, uma falha inicial de uma delas pode causar ou aumentar a probabilidade de ocorrência de falha em outra. O estudo das correlações de tempo entre as ocorrências de falhas ou eventos em um sistema podem

revelar se um evento ocorrido em uma semana poderá ocorrer na próxima, em meses ou no próximo ano (BANKS e CARLEY, 1996).

Surge então a necessidade de compreensão de quais são as implicações para a vulnerabilidade e risco de falhas em cada sistema. Outra importante distinção a se fazer é a relacionada à resiliência e robustez de um sistema. Um sistema resiliente é aquele capaz de se recuperar mais rapidamente quando da ocorrência de instabilidades ou falhas. Já um sistema robusto é aquele cuja frequência de falhas que o fazem colapsar é menor. Um sistema complexo possui a maioria das partes interagindo de maneira não linear. O comportamento total (dinâmica) destes sistemas não é igual à soma dos comportamentos de seus componentes individuais.

Os sistemas evoluem com o tempo e isto também apresenta consequências no seu comportamento. A universalidade do estudo dos sistemas complexos também possibilita predições estatísticas de comportamentos. Uma árvore pode ser considerada complicada. Já uma floresta é complexa. O estudo de sistemas complexos permite a utilização de modelos com menos detalhes individuais (e complicados, como árvore, carro, etc.) para existir a possibilidade de investigar o comportamento complexo do conjunto e extrair características universais dos mesmos (pontos críticos, medidas, etc.). Além dos sistemas complexos já citados existe a possibilidade de estudo de sistemas acoplados: a mistura das inter-relações de dois ou mais sistemas complexos e a descoberta das implicações desta mistura.

Muitas vezes a frequência de uma falha pode ser pequena em termos estatísticos, mas suas consequências podem ser grandes e os custos altos. Correlações de tempo entre as ocorrências de problemas em sistemas são muitas vezes ignoradas. O fato do uso de políticas de segurança terem diminuído a ocorrência das falhas/problemas excluiu a analise temporal, mas a memória do sistema e sua vulnerabilidade permanecem e continuam a impactá-lo. Um exemplo é a persistência de *blackouts* em redes de energia elétrica. Apesar de ter havido redução na frequência das falhas, o grande problema é que estas ainda persistem. Pesquisadores afirmam que o *blackout* ocorrido no nordeste dos Estados Unidos e Canadá em 2003 teve as mesmas causas de outras falhas ocorridas recentemente. Se a vulnerabilidade tivesse sido sanada, novas falhas não ocorreriam (CARRERAS et al., 2002).

Implementações que utilizam o método de Monte Carlo muitas vezes não funcionam nesta situação, pois tratam de problemas de ocorrências aleatórias, sem previsão do que há tempos vem influenciando o sistema. Isto está relacionado à criticalidade auto-organizada dos sistemas

complexos. Também é observado que muitos sistemas operam próximos à sua capacidade total, devido ao fato de a demanda sempre crescer em uma velocidade maior do que a geração e transmissão da rede.

A área da matemática denominada teoria dos grafos possibilitou a evolução dos estudos de sistemas modelados por redes. Em meados de 1959 a escassez de dados experimentais destes sistemas levou ao surgimento de uma abordagem aleatória que explicaria a formação de uma estrutura de rede, conhecida por *random-graphs networks*. Seus autores, Paul Erdös (*1913- † 1996) e Alfred Rényi (*1921- † 1970), introduziram as primeiras análises de estruturas em redes. O modelo proposto consistia de nós interconectados entre si com probabilidade *p*. Através desse tipo de consideração uma rede aleatória segue uma distribuição de Poisson, fazendo com que seja raro encontrar nós com concentração de conexões ou muito grande ou muito pequena (ERDÖS e RÉNYI, 1959).

Redes complexas motivaram vários trabalhos, porém, sempre houve a dificuldade de coletar dados e de comprovar seus métodos e teorias. Surgem então perguntas: como verificar quais modelos são adequados para redes de conexões aéreas? Como obter a distribuição de conexões de neurônios? Como simular comportamento e reações das redes quando elas falham? Anteriormente, os estudos se limitavam a pequenas redes, com poucos nós e arestas.

Com a difusão da Internet, crescimento do ambiente web, com suas páginas que apontam outras páginas que por sua vez se conectam a muitas, têm-se um exemplo claro de uma rede complexa, de grande dimensão em constante crescimento. Além disso, a Internet possibilitou aquisição acessível de dados para estudos diversificados.

Esta facilidade providenciada pela Internet permitiu que teorias, modelos e técnicas fossem desenvolvidas e testadas. A Internet e outras redes se tornaram o objeto de estudo para pesquisa na subárea de Sistemas Complexos da Física Estatística. A extensão e alcance dos benefícios vão para diversas redes, muitas vezes de difícil aquisição de dados, como rede de neurônios. (COSTA e DIAMBRA, 2004).

Uma das principais motivações para o estudo das redes complexas é o seu rápido crescimento, devido à utilização e dependência cada vez maior pela sociedade nacional e mundial.

Duncan Watts (*1971 - Presente) e Steven Henry Strogatz (*1959 - Presente), em 1998, desenvolveram um modelo de rede denominado Pequeno Mundo ou *Small World*, graças ao

aumento do poder computacional e a evolução tecnológica. Quando se tem uma rede conexa e se adiciona um número muito pequeno de arestas aleatoriamente, o diâmetro desta rede tende a cair drasticamente. Este fenômeno é conhecido como *Small World*. O trabalho de Watts-Strogatz se tornou uma clássica referência na área de redes complexas cuja figura principal passou a ser citada em vários trabalhos posteriores (WATTS e STROGATZ, 1998).

Em 1999, descobriu-se que a grande maioria das redes do mundo real foram formadas diferentemente do que a teoria das redes aleatórias previa. As redes passaram a ser chamadas de livres de escala, ou *Scale-Free*, graças aos estudos de Albert Lasló Barabási (*1967 - Presente) e Réka Albert, em 1999. Os princípios que caracterizam este modelo são: anexação preferencial (na qual novos nós tendem a se conectar a nós que já apresentam grande número de conexões - *richs-gets-richers*) e distribuição de graus seguindo uma lei de potência ou *power-law*. Tais redes são observadas em uma infinidade de sistemas de naturezas diversas, por possuírem um grande número de elementos interconectados e por serem estudadas pela área de sistemas complexos. Somente nos últimos anos, com o avanço tecnológico dos sistemas de aquisição de dados e o aumento do poder computacional, pesquisas nesta área puderam se desenvolver. Assim, o estudo dos princípios que governam essas redes é de fundamental importância para a compreensão das diversas redes do mundo real (BARABÁSI e ALBERT, 1999).

Dada a importância, expansão e alta dependência das redes complexas, mostra-se fundamental o seu estudo, compreensão e operabilidade eficiente. Chama a atenção a dinâmica das falhas em redes, que prejudica sua operação e afeta todos aqueles que dependem do seu pleno funcionamento.

Logo, o foco deste trabalho é utilizar as técnicas disponibilizadas pela teoria das redes complexas, para caracterização e modelagem de redes de topologia *Scale-Free*, visando compreensão da ocorrência de falhas nas mesmas. Assim torna-se possível fornecer uma base para desenvolvimento de novas estratégias de proteção das redes complexas quando sujeitas a falhas, compreensão do comportamento destas falhas, novas metodologias para conectividade e alocação de recursos para que se possa estimular o desenvolvimento de novas técnicas para construção ou remodelagem de redes que atinjam operação satisfatória.

Introduzido o tema, ainda neste capítulo são colocadas as motivações dos estudos e os objetivos deste trabalho. No Capítulo 2, são expostos os conceitos relacionados à Teoria dos Grafos, que possibilita modelagem das redes. No Capítulo 3 colocam-se os conceitos de redes

complexas, os modelos existentes, além de medidas fundamentais. No Capítulo 4 são revisados mecanismos de contenção de falhas disponíveis na literatura recente, também os métodos e materiais utilizados para execução das simulações das falhas em redes. No Capítulo 5 abordam-se os resultados e discussões. No Capítulo 6 colocam-se as conclusões e sugestões para trabalhos futuros e ao final as referências.

1.1 Motivações

Três de julho de 2008: o estado mais populoso do Brasil, com mais de 40 milhões de habitantes fica sem Internet. Serviços essenciais, como bancos, agências da Previdência, prefeituras de 407 municípios e demais usuários do *Speedy* totalmente paralisados. Tudo isso por causa de 6 roteadores cujas falhas se espalharam através da rede. O chamado "apagão tecnológico" resultou em significativas perdas.

Podem-se citar exemplos de quedas em grandes redes de comunicação, tanto no Brasil quanto no exterior, que causaram prejuízos em larga escala. Conforme Zhao et al. (2005), *blackouts* no nordeste dos Estados Unidos e no Canadá possuem características de falhas em redes.

Sistemas como estes e outros de grande importância econômica, social e tecnológica, a exemplo das redes de distribuição de água, de malhas viárias, de rotas aéreas e de relacionamentos comerciais, são modelados por Redes Complexas (KURRANT e THIRAN, 2006).

Esta área de investigação, Redes Complexas, está estabelecida através da sua estreita relação com outros domínios como a Teoria de Grafos e a Mecânica Estatística, formando um quadro teórico geral e poderoso para representar e modelar sistemas complexos (GOH et al., 2001) (WANG e CHEN, 2003).

Redes complexas são um campo de pesquisa relevante e recentemente aberto, que vai ao encontro dos "Desafios em Computação no Brasil", os quais incluem a Modelagem computacional de sistemas complexos artificiais, naturais e socioculturais (MEDEIROS, 2006; CÂMARA, 2006)

Enquanto ferramenta de estudo de sistemas, as redes complexas abrangem a conectividade e também a dinâmica das estruturas complexas, conforme se constata pelo grande número de artigos em revistas atuais, a exemplo dos trabalhos de Barrat e Weigt (2000), Almaas et al. (2002), Albert et al. (1999), DeMarco (2001) e Carreras et al. (2002).

Estes sistemas, nas suas diversas formas, estão sujeitos a falhas em seus diferentes componentes. Estas falhas podem se propagar de um componente a outro podendo causar colapso da rede (ADAMIC e HUBERNA, 2000). Falhas causadas intencionalmente são denominadas ataques ou falhas induzidas. Uma vez que se espalhem pela estrutura são denominadas "falhas em cascata" e podem ocorrer inclusive na rede formada pelo relacionamento entre várias Redes Complexas (WANG e CHEN, 2003).

O controle destas falhas é de fundamental importância, tanto em uma situação em que é extremamente necessário contê-las (como no caso de uma rede de distribuição de energia) quanto em uma situação em que o objetivo é justamente espalhá-la, para que a rede entre em colapso (visando, por exemplo, evitar que vírus de computadores se disseminem).

Compreender a intrincada dinâmica das Redes Complexas com o objetivo de conter propagação de falhas é um desafio que precisa ser superado para possibilitar a evolução dos sistemas.

Sun et al. (2008) propõem em seu trabalho a alocação estratégica de capacidade nos nós de uma rede, para que ocorra ganho de robustez e redução do tamanho de uma falha em cascata na rede. Zhao e Xu (2009) mostraram como aumentar a robustez de uma rede pela adição de novos *links* entre os nós de baixo grau (aqueles que apresentam poucas conexões). Eles notaram que graças ao ciclo formado entre esses nós, a rede mantém o seu funcionamento mesmo após a falha de um nó de alto grau (denominado nó *hub*).

A combinação destas duas estratégias é estudada neste trabalho, para análise do comportamento destas duas ações de contenção de cascatas em conjunto.

Espera-se contribuições na eficiência dos mecanismos, na redução de custos de construção, manutenção, ampliação e controle de Redes de Complexas.

1.2 Objetivos deste trabalho

Este trabalho tem como objetivo estudar o comportamento de Redes Complexas em presença de falhas induzidas e estudar a ocorrência de falhas em cascata. Para alcançar este objetivo consideram-se as Redes Livres de Escala ou *Scale-Free*, modeladas como grafos direcionados computacionalmente.

O foco deste trabalho são as falhas em cascata induzidas em redes. Assume-se, por hipótese, que uma cascata pode ser contida a partir de estratégias específicas conhecidas. Todavia, estas são restritas aos seus cenários.

Este trabalho consiste no estudo e teste de dois mecanismos descritos na literatura e investiga o resultado da combinação de ambos, além de analisar matematicamente o comportamento das falhas em cascata apresentadas após a simulação das falhas. A avaliação de mecanismos de contenção é feita através de simulações de ataques intencionais às redes, considerando a propagação da falha entre nós e conexões. Os ataques simulados devem ser de dois tipos: a) aleatórios; b) direcionados segundo características topológicas da rede. Na primeira situação, investiga-se o resultado de uma ação casual e, na segunda, de uma ação planejada com conhecimento dos alvos potenciais.

As simulações da primeira e segunda etapas produzem dados sobre a extensão dos danos causados pelas falhas em cascata induzidas e avalia a eficiência dos mecanismos de contenção. A extensão dos danos é medida a partir da contabilização dos componentes em operação e atingidos, da análise da interconectividade, da quantidade de nós afetados. A eficiência dos esquemas de contenção é avaliada pela dimensão da cascata, contabilizando a parcela atingida e sobrevivente da rede. Espera-se contribuir com avanços teóricos sobre o tema e analisar o comportamento de mecanismos de contenção das falhas em cascata.

2 TEORIA DOS GRAFOS

A Teoria dos Grafos teve seu início com o artigo de Leonhard Euler (*1707-†1783), publicado em 1736, denominado as Sete pontes de Königsberg. Este problema deu início à história da matemática e a partir dele se desenvolveu o fundamento inicial da teoria dos grafos.

O problema é baseado na cidade de Königsberg, que até 1945 era território da Prússia, e hoje denomina-se Kaliningrado, pertencente a Rússia. Em Konigsberg o posicionamento do rio Prególia formava duas grandes ilhas. Juntas, tais ilhas formavam um complexo com sete pontes.

A possibilidade de atravessar todas as pontes sem repetir nenhuma foi então levantada. Havia-se tornado uma lenda popular tal façanha, até que Euler, em 1736, provou que não existia caminho que possibilitasse tais restrições.

Euler representou os caminhos como retas e suas intersecções como vértices fazendo o que deve ter sido o primeiro grafo da história. Notou que poderia atravessar todo o caminho passando uma única vez em cada ponte, somente se houvesse nenhum ou dois pontos dos quais saísse um número ímpar de caminhos.

Deveria existir um número par de caminhos em cada um dos nós, pois eram necessárias uma entrada e uma saída. Dois nós com caminhos ímpares referiam-se ao início e ao fim do trajeto. Se não existissem nós com número ímpar de caminhos, seria possível iniciar e terminar o trajeto no mesmo nó.

Das sete pontes originais (Figura 1), uma foi demolida e reconstruída em 1935, duas foram destruídas durante a Segunda Guerra Mundial e outras duas foram demolidas para dar lugar a uma única via expressa. Atualmente apenas duas pontes são da época de Euler (Figura 2).

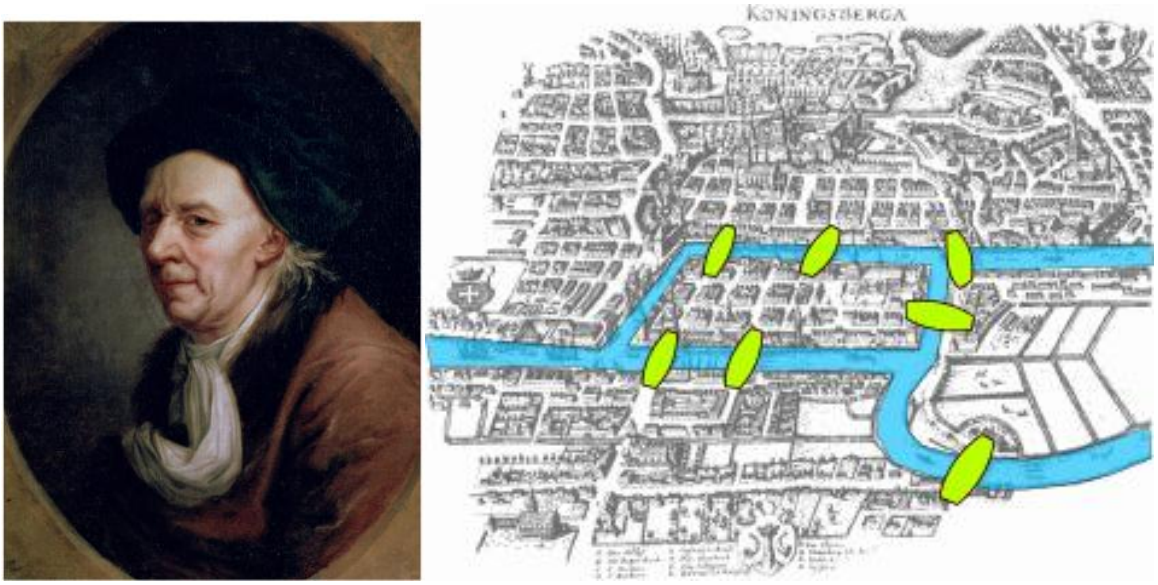

Figura 1: Leonard Euler (à esquerda) e as pontes de Konigsberg (à direita).[1]

Figura 2: Konigsberg e as pontes atualmente.[2]

[1] Fonte: 9TTP://feedblog.org/2009/07/14/the-konigsberg-bridge-problem-or-where-graph-theory-was-invented/

[2] Fonte: GoogleMaps

O trabalho de Euler é considerado o primeiro resultado da teoria dos grafos. É também considerado um dos primeiros resultados topológicos na geometria; isto é, não dependente de quaisquer medidas. Isso ilustra a profunda conexão entre a teoria dos grafos e topologia.

Colocam-se neste item as principais características dos grafos. Um grafo *G(V, A)* é definido pelo par de conjuntos *V* e *A*, onde:

V é o conjunto finito e não vazio de vértices ou nós do grafo;

A é o conjunto de pares ordenados *a=(v,w)*, onde *v* e *w* são elementos distintos e pertencem a *V*, formando as arestas do grafo.

Seja, por exemplo, o grafo *G(V, A)* dado por:

$V = \{c \mid c$ é um computador$\}$,

$A = \{(c_1, c_2) \mid < c_1$ é conectado a $c_2 >\}$.

No exemplo acima têm-se um grafo formado por um conjunto de computadores e as conexões existentes entre eles. Pode-se definir que este grafo é uma representação de uma rede de computadores interconectados.

Diz-se que duas arestas são adjacentes quando possuem um extremo comum. O conceito de grau de um vértice ou nó é definido pela somatória do número de arestas incidentes a este nó. A teoria dos grafos também diz que um laço é uma aresta cujo nó origem e nó destino são os mesmos.

Um nó é isolado quando ele não possui arestas e por consequência seu grau é zero. A sequência de vértices necessária para se partir de um vértice origem a um vértice destino é denominada caminho. Um caminho de *k*-vértices é formado por *k-1* arestas (v_1,v_2), (v_2,v_3) ... (v_{k-1},v_k), e o valor *k-1* é o comprimento do caminho. Caminho simples ou elementar é quando todos os vértices do caminho são distintos. Um caminho $v_1, v_2 \ldots v_k, v_{k+1}$, onde $v_1 = v_{k+1}$ e $k>=3$ é denominado ciclo. Um grafo que não apresenta ciclos é denominado acíclico.

Um Grafo Direcionado ou Dígrafo é representado por *D(V, A),* onde *D* é um conjunto finito não vazio de *V* (vértices) e um conjunto *A* (arestas) de pares ordenados de vértices distintos. Cada aresta *(v,w)* possui uma única direção de *v* para *w*. *(v,w)* é divergente de *v* e convergente a *w*.

Nos grafos direcionados tem-se os conceitos de grau de entrada de *v*, que é o número de arestas convergentes a *v*, também o grau de saída de *v*, que é o número de arestas divergentes de *v* ainda o sumidouro que é um vértice com grau de saída nulo.

A Matriz de Adjacências tem a seguinte definição: dado um grafo *G*, a matriz de adjacências $R = (r_{ij})$ é uma matriz m x m tal que:

m = número de vértices,

$r_{ij} = 1$ se (v_i, v_j) pertence a A,

$r_{ij} = 0$ cc ou seja, $r_{ij} = 1$ quando os vértices v_i, v_j forem adjacentes e $r_{ij} = 0$ caso contrário.

A matriz de adjacências representa um grafo sem ambiguidade. A matriz R é simétrica para um grafo não direcionado.

Uma matriz de adjacências caracteriza unicamente um grafo, contudo a um mesmo grafo *G* podem corresponder várias matrizes diferentes.

A Matriz de Adjacências para Dígrafos é dada por: tomando $r_{ij}=1$ se (v_i,v_j) for aresta divergente de v_i e convergente a v_j; e $r_{ij}=0$, caso contrário a matriz não é mais (necessariamente) simétrica e o número de 1's é exatamente igual a m (número de arestas).

Uma Matriz de Incidências é dada pela seguinte definição: Seja $G(V,E)$ um grafo, define-se a matriz de incidências $B=(b_{ij})$ uma matriz $n \times m$ tal que:

$b_{ij} = 1$ se o vértice v_i e a aresta e_j forem incidentes,

$b_{ij} = 0$ caso contrário.

A matriz de adjacência representa univocamente um grafo, mas este último pode ser representado, em geral, por várias matrizes de incidências diferentes. Note que cada coluna de *B* tem exatamente dois números um (GUIMARÃES, 2005).

Um Grafo é dito simples se é não direcionado, sem laços e existe no máximo uma aresta entre quaisquer dois vértices (sem arestas paralelas). Um Grafo é dito completo se for um grafo simples em que, para cada vértice do grafo, existe uma aresta conectando este vértice a cada um dos demais. Ou seja, todos os vértices do grafo possuem mesmo grau. O grafo completo de n vértices é frequentemente denotado por K_n. Ele tem $n(n-1)/2$ arestas (correspondendo a todas as possíveis escolhas de pares de vértices).

Um Grafo nulo é o grafo cujo conjunto de vértices é vazio. Um Grafo vazio é o grafo cujo conjunto de arestas é vazio. Um Grafo trivial é o grafo que possui apenas um vértice e nenhuma aresta. Um Grafo regular é um grafo em que todos os vértices tem o mesmo grau. Um multigrafo é um grafo que permite múltiplas arestas ligando os mesmos vértices (arestas paralelas). Pseudografo é um grafo que contém arestas paralelas e laços. Um Ponto de articulação ou Vértice

de corte é um vértice cuja remoção desliga um grafo. Uma ponte é uma aresta cuja remoção desliga um grafo.

Um componente biconectado é um conjunto máximo de arestas tal que qualquer par de arestas do conjunto fazem parte de um ciclo simples comum. O contorno de um grafo é o comprimento do ciclo simples mais curto no grafo. O contorno de um grafo acíclico é, por definição, infinito. Uma Árvore é um grafo simples acíclico e conexo. Às vezes, um vértice da árvore é distinto e chamado de *raiz*. Árvores são comumente usadas como estruturas de dados em computação. Floresta é um conjunto de árvores; equivalentemente a uma floresta, em algum grafo acíclico.

Um subgrafo de um grafo *G* é um grafo cujo conjunto dos vértices é um subconjunto do conjunto de vértices *G*, cujo conjunto de arestas é um subconjunto do conjunto de arestas de *G*, e cuja função *w* é uma restrição da função de *G*. Um subgrafo Gerador é aquele obtido pela remoção de uma ou mais arestas de um outro grafo, dizemos então que este novo grafo obtido é gerador do primeiro. Um Subgrafo Induzido é obtido pela remoção de vértices e consequente das arestas relacionadas com ele de um outro grafo. Dizemos que este novo grafo é um grafo induzido do original. Grafo parcial de um grafo *G* é um subgrafo com o mesmo conjunto de vértices que *G* (GUIMARÃES, 2005).

Uma árvore parcial é um grafo parcial que é árvore. Todo grafo tem pelo menos uma árvore parcial. Clique em um grafo é um subgrafo que também é um grafo completo. Conjunto independente em um grafo é um conjunto de vértices não adjacentes entre si. Grafo planar é aquele que pode ser representado em um plano sem qualquer intersecção entre arestas. Caminho é uma sequência de vértices tal que de cada um dos vértices existe uma aresta para o vértice seguinte. Um caminho é chamado simples se nenhum dos vértices no caminho se repete. O comprimento do caminho é o número de arestas que o caminho usa, contando-se arestas múltiplas vezes.

O custo de um caminho num grafo balanceado é a soma dos custos das arestas atravessadas. Dois caminhos são independentes se não tiverem nenhum vértice em comum, exceto o primeiro e o último. Caminho euleriano em um grafo é o caminho que usa cada aresta exatamente uma vez. Se tal caminho existir, o grafo é chamado traversável. Um ciclo euleriano é um ciclo que usa cada aresta exatamente uma vez. Caminho hamiltoniano em um grafo é o caminho que visita cada

vértice exatamente uma vez. Um ciclo hamiltoniano é um ciclo que visita cada vértice uma só vez.

Enquanto determinar se um dado grafo contém um caminho ou ciclo euleriano é trivial, o mesmo problema para caminhos e ciclos hamiltonianos é extremamente árduo. O lema do aperto de mãos diz que se os convidados de uma festa apertarem as mãos quando se encontrarem pela primeira vez, o número de convidados que apertam a mão um número ímpar de vezes é par. Também em grafos não direcionados a soma dos graus de todos os vértices é igual ao dobro do número de arestas.

Grafo bipartido é o grafo cujos vértices podem ser divididos em dois conjuntos, nos quais não há arestas entre vértices de um mesmo conjunto. Para um grafo ser bipartido ele não pode conter circuitos de comprimento ímpar.

Um grafo G é considerado bipartido se todo o circuito de *G* possui comprimento par**.** Sejam V_1 e V_2 os dois conjuntos em que, de acordo com a definição de grafo bipartido, se particiona *V(G).* Toda a aresta de *G* conecta um vértice em V_1 com outro em V_2. Assim sendo, se *X* for um vértice de V_1, para "voltar" a esse vértice terá de se ir a V_2 e voltar a V_1 um número indeterminado de vezes, e de cada vez serão percorridas duas arestas, uma de um vértice em V_1 para um vértice em V_2 e outra de um vértice em V_2 para um vértice em V_1. Logo, o número de arestas a percorrer será par, ou seja, o comprimento do circuito é par.

Grafo bipartido completo é o grafo bipartido, cujo qualquer vértice do primeiro conjunto é adjacente a todos vértices do segundo conjunto. Grafo *k*-partido ou grafo de *k***-**coloração é um grafo cujos vértices podem ser particionados em *k* conjuntos disjuntos, nos quais não há arestas entre vértices de um mesmo conjunto. Um grafo 2-partido é o mesmo que grafo bipartido. Emparelhamento de grafos consiste em partir o grafo em conjuntos de vértices a qual não compartilham nenhuma aresta entre eles.

Teorema das quatro cores é baseado no problema das cores necessárias para se colorir um mapa sem que os países vizinhos compartilhem da mesma cor. Transformando o mapa em um grafo pode-se provar que pode-se representar qualquer mapa (um grafo planar) com apenas 4 cores (4 partições).

Percurso Árvores: Percorrimento sistemático em todos os vértices e arestas do grafo. O grafo pode ser dirigido ou não. O percurso em árvores é o processo de visitar cada nó da árvore exatamente uma vez. O percurso pode ser interpretado como colocar todos os nós em uma linha,

não existindo uma ordem para ser seguida. Existem n percursos diferentes, quase todos caóticos. Os básicos são percurso em profundidade e percurso em largura. A Fila trata da busca em largura. A Pilha, da busca em profundidade (DIESTEL, 2000).

Busca em Extensão ou Largura (*Breadth-First Search* ou BFS): propriedade especial que está relacionada ao fato de a árvore não possuir ciclos: dados dois vértices quaisquer, existe exatamente um caminho entre eles. Um percurso em extensão é visitar cada nó começando do menor nível e movendo-se para os níveis mais altos nível após nível, visitando cada nó da esquerda para a direita. Sua implementação é direta quando uma fila é utilizada. Depois que um nó é visitado, seus filhos, se houver algum, são colocados no final da fila e o nó no início da fila é visitado. Assim, os nós do nível *n+1* serão visitados somente depois de ter visitados todos os nós do nível n. Computa a menor distância para todos os vértices alcançáveis. O subgrafo contendo os caminhos percorridos é chamado de *breadth-first tree*.

Busca em profundidade (*Depth-first search* ou DFS): um algoritmo de busca em profundidade realiza uma busca não-informada que progride através da expansão do primeiro nó filho da árvore de busca, e se aprofunda cada vez mais, até que o alvo da busca seja encontrado ou até que ele se depare com um nó que não possui filhos (nó folha). Então a busca retrocede (*backtrack*) e começa no próximo nó. Em uma implementação não-recursiva, todos os nós expandidos recentemente são adicionados a uma pilha, para realizar a exploração (DIESTEL, 2000).

A complexidade espacial de um algoritmo de busca em profundidade é muito menor que a de um algoritmo de busca em largura. A complexidade temporal de ambos os algoritmos é proporcional ao número de vértices somados ao número de arestas dos grafos aos quais eles atravessam. Quando ocorrem buscas em grafos muito grandes, que não podem ser armazenadas completamente na memória, a busca em profundidade não termina, em casos onde o comprimento de um caminho numa árvore de busca é infinito. O simples artifício de "lembrar quais nós já foram visitados" não funciona, porque pode não haver memória suficiente. Isso pode ser resolvido estabelecendo-se um limite de aumento na profundidade da árvore.

Grafos têm sido estudados por centenas de anos e suas propriedades são entendidas e provadas. No contexto das redes, a teoria dos grafos e grande parte dos seus conceitos são fundamentais para estudos aprofundados. Isso ocorre em especial pelo fato de redes poderem ser reduzidas a grafos matemáticos. A teoria das redes fornece uma maneira de modelar sistemas

críticos como grafos abstratos. As redes contêm nós, *links* ou arestas, além de da distribuição dos graus, que nos diz quais nós estão conectados em quais. Um entendimento fundamental da teoria das redes é que esta provê a base da visão global de um sistema, pois a maioria dos sistemas críticos podem ser modelados em alguma forma de rede. O primeiro passo é compreender teoria dos grafos e a teoria das redes. O segundo passo é aprender como aplicar a teoria dos grafos nas redes (LEWIS, 2009).

3 MODELOS PARA REDES COMPLEXAS

3.1 Redes Complexas

Possibilidades como a modelagem proporcionada pelos grafos enriqueceram a teoria das redes e possibilitaram a representação complexa associada à coleta de medidas fundamentais para o entendimento da dinâmica e estrutura das redes. Muitos pesquisadores, ao longo dos anos, apresentaram contribuições cruciais que possibilitaram a evolução da teoria das redes. Stanley Milgram (*1933 - †1984), um psicólogo norte-americano graduado da Universidade de Yale, realizou a experiência chamada de "pequeno mundo", que comprovou a existência de seis graus de separação em redes sociais. Milgram calculou a quantidade de graus de separação existentes entre uma pessoa e qualquer outra. No experimento, correspondências eram enviadas a pessoas conhecidas de um grupo específico até que outra pessoa, desconhecida das iniciais, a recebesse. A conclusão foi que cada pessoa dista de qualquer outra em *seis graus de separação*. (MILGRAM, 1963).

Em 1959, surgiu o estudo de outros pesquisadores, são eles: Paul Erdös e Alfréd Rényi, que propuseram um modelo de geração de grafos visando a compreensão da forma pela qual os grafos cresciam. O modelo denominado *Random Graphs* incluía o conceito de que pares de nós de um grafo se conectavam com probabilidade equivalente, de maneira aleatória, independente de outras ligações. Desta forma, o resultado era um grafo com nós com poucas conexões e praticamente com o mesmo número de conexões.

Quase quatro décadas após a descoberta de Erdos e Rényi, surgiu o trabalho de Watts e Strogatz de 1998. Inspirados pelo experimento de Milgram e graças ao aumento do poder computacional, desenvolveram o modelo *Small-World,* que se tornou uma clássica referência na área de redes complexas cuja figura principal passou a ser citada em vários trabalhos posteriores.

Em 1999, Barabási e Albert estudaram sistemas de origem bem distintas, naturais como redes genéticas e criados pelo homem como o ambiente *World Wide Web.* Neste trabalho, pela primeira vez relacionado com a Internet, surgiu o conceito de S*cale-Free* network, redes sem escala ou livre de escala, quando perceberam que o grau de conectividade *k* dos nós ou vértices da rede seguiam uma distribuição em lei de potência ou *power-law*. Este comportamento era devido a dois principais mecanismos: (*i*) crescimento contínuo da rede pela adição de novos nós e (*ii*) novos nós se conectam preferencialmente a outros nós já bem conectados. Baseado nestes dois princípios, um modelo de crescimento pode reproduzir as distribuições *Scale-Free* encontradas nos sistemas reais e isto indica que o desenvolvimento de grandes redes pode ser governado por um fenômeno auto-organizado e auto-consistente que predomina sobre as características pontuais de cada elemento do sistema.

Ainda em 1999, em outro trabalho, os irmãos Faloutsos, apresentaram um estudo da topologia da Internet mundial e revelaram sua surpresa diante do comportamento da distribuição ser do tipo lei de potência. Suas observações foram feitas entre Sistemas Autônomos - ASs, por meio de dados coletados de roteadores públicos mantidos pelo *National Laboratory for Applied Network Research* - NLANR. Devido à sua área de atuação, os autores perceberam a utilidade desta área da ciência no desenho e desempenho de análise de protocolos de comunicação para Internet.

Em 2001, Barabási submete a Internet, sob o ponto de vista da Web, aos recentes conceitos de redes *Scale-Free* da área de Redes Complexas, traçando um paralelo com a teoria clássica de redes aleatórias de Erdös e Rényi. Também neste ano, Pastor-Satorras et al., seguindo esta mesma linha de estudo, apresentaram as propriedades dinâmicas e as correlações de vizinhança da Internet em um estudo que considerou a evolução da topologia durante três anos. Os resultados obtidos apresentam um comportamento que também segue uma lei de potência, com coeficiente angular do ajuste linear de dados em escala log – log, para a análise ordenada do número de conexões e para a distribuição de conexões de Sistemas Autônomos. Os estudos apresentados por

Barabási e Pastor-Satorras consideram a Internet como um processo complexo e dinâmico, resultado da topologia observada.

Em 2003, Barabási e Bonabeau transmitiram ao público menos acadêmico as particularidades das redes *Scale-Free* encontradas na natureza e no cotidiano da sociedade (BARABÁSI e BONABEAU, 2003).

Maslov, Sneppen e Zaliznyak (2004) apresentaram um sistema de análise de padrões topológicos para redes complexas de grande porte. Neste sistema, a rede estudada é comparada com sua versão totalmente aleatória chamada de modelo nulo. O desvio das propriedades deste modelo reflete seu desenho e/ou sua história de crescimento. A Internet foi usada como exemplo ao quantificarem as correlações entre os coeficientes de conexão. Verificaram também, que existem diferenças entre a Internet e redes moleculares estudadas anteriormente. O grande mérito deste trabalho está na introdução do conceito de padrões de correlação que permitem quantificar as diferenças entre redes complexas que possuam a mesma distribuição de graus de conectividade (LEWIS, 2009).

Neste breve histórico, onde foram destacados cronologicamente alguns dos principais estudos acerca da área de redes, alguns dos aspectos que serviram de base para este trabalho foram citados. Podem-se citar também, diversas outras contribuições na área das redes complexas ao longo dos anos. Isto é mostrado na Tabela 1, que apresenta em sequência cronológica os principais estudos fundamentais na teoria das redes complexas com os respectivos autores.

Tabela 1: Histórico de estudos em redes complexas com seus respectivos autores (LEWIS, 2009).

Ano	Autor	Contribuição
1736	Euler	Pontes de Konigsberg
1925	G. Yule	Anexação Preferencial, Distribuição Yule–Simon
1927	Kermack, McKendrick	Primeiro modelo epidêmico
1951	Solomonoff, Rappaport	Espalhamento de Infecção em redes aleatórias
1955	Simon	Lei de Potência observada em análise de palavras
1959	Gilbert	Primeiro procedimento gerativo para grafos aleatórios
1960	Erdos, Renyi	Grafos Aleatórios

1967	Milgram	Experimento Mundo-Pequeno
1969	Bass	Difusão da Inovação em populações: modelo não-rede
1971	Fisher, Pry	Difusão pela substituição por produto, modelo não-rede
1972	Bollobas	Grafos Complexos
1972	Bonacich	Idéia da influência em redes sociais levando a diagramas de influência
1973	Granovetter	Redes de busca por emprego formam clusters com ligações fracas entre si
1978	Pool, Kochen	Primeira prova teórica de mundos-pequenos
1984	Kuramoto	Sincronização de Sistemas Lineares
1985	Bollobas	Publica livro sobre "grafos aleatórios"
1988	Waxman	Primeiro modelo em grafo da Internet
1989	Bristor, Ryan	Redes de Compras aplicação da ciência das redes para modelar sistemas econômicos
1990	Guare	Criada a frase: “seis graus de separação” nos nomes das suas peças na Broadway
1995	Molloy, Reed	Geração de redes com arbitrárias sequências de distribuição dos graus
1996	Kretschmar, Morris	Início de aplicação da ciência das redes para disseminação de contágio de doenças infecciosas impulsionado pela maior componente conexa
1998	Holland	Introdução de emergência em sistemas complexos adaptativos
1998	Watts, Strogatz, Faloutsos	Renovam interesse no trabalho original de Milgram em mundo pequeno, exemplos de clustering; primeiro procedimento gerativo para mundo pequeno
1999	Faloutsos	Lei de Potencia observada na Internet
1999	Albert, Jeong, Barabasi	Lei de Potência observada na WWW
1999	Dorogovtsev, Mendes	Propriedades Mundo Pequeno
1999	Barabasi, Albert	Modelo *Scale-Free*
1999	Dorogovtsev, Mendes, Samukhim, Krapivsky Redner	Solução exata para sequência de graus em redes *Scale-Free*
1999	Watts	Explicação do "dilema mundo pequeno": alto clustering, pequeno caminho
1999	Adamic	Distância entre sites .edu mostrou-se mundo pequeno
1999	Kleinberg, Kumar, Raghavan, Rajagopalan Tomkins	Formalizaram modelos de WWW como “Grafo Web”

1999	Walsh	Dificuldade de pesquisa em mundo pequeno usando propriedades locais
2000	Marchiori, Latora	Distância harmônica substituiu tamanho do caminho: funciona para redes desconectadas
2000	Broder, Kumar, Maghoul, Raghavan, Rajagopalan Stata, Tomkins, Wiener	Mapa completo em Grafo Web da WWW
2000	Kleinberg	Mostra a busca O(n) em mundo pequeno usando " Distância Manhattan"
2000	Albert, Jeong, Barabasi	Redes *Scale-Free* são resilientes se hubs estão protegidos ("Calcanhar de Aquiles" da Internet)
2001	Yung	Taxonomia de aplicações da teoria do mundo pequeno para: SNA, colaboração, internet, negócios, ciências da vida
2001	Pastor-Satorras, Vespignani	Alegação do limiar não epidêmico em redes *Scale-Free*; Internet suscetível a vírus SIS
2001	Tadic, Adamic	Uso da informação local pode acelerar busca em redes *Scale-Free*
2002	Levene, Fenner, Loizou, Wheeldon	Melhoria do modelo Grafo Web que concluiu estrutura da WWW não poderia ser explicado pela anexação preferencial sozinho
2002	Kleinfeld, Claims	Experimento de Milgram não bem fundamentado: rede social mundo pequeno é um mito urbano
2002	Wang, Chen, Barahona, Pecora, Liu, Hong, Choi Kim, Jost, Joy	Sincronização em mundo pequeno é equivalente a estabilidade em sistemas acoplados
2003	Wang, Chakrabarti, Wang, Faloutsos	Mostram espalhamento de epidemia determinado pelo raio espectral, maior componente conexa da matriz
2003	Virtanen	Levantamento completo dos resultados da ciência das redes até 2003
2003	Strogatz	Sincronização dos grilos, dos batimentos cardíacos
2005	NRC	Definição da Ciência das Redes
2006	Atay	Sincronização em redes com seqüência de distribuição dos graus - aplicações para redes
2007	Gabbay	Consenso em redes de influência —modelos linear e não linear

Mas, o que é uma rede? Em termos simples, uma rede passível de estudo é representada por um conjunto de nós e conexões (*links*). Estes últimos conectam pares de nós. Nós e *links* são

conceitos abstratos. Um nó pode representar cidades, um comutador de internet ou uma pessoa. Um *link* pode representar uma rodovia que conecta duas cidades ou um cabo de fibra óptica que conecta os switches de internet ou um relacionamento entre duas pessoas na sociedade.

A Teoria geral das redes pode ser usada para modelar uma variedade de coisas do mundo real. Abstrações de redes podem ser aplicadas em diferentes níveis de detalhes. Nós e *links* podem representar muitos sistemas. Mas, uma vez que se têm uma infraestrutura abstraída em rede, técnicas podem ser aplicadas e trazer importantes resultados e ideias (LEWIS, 2009).

Utiliza-se para construção da rede uma abstração matemática chamada grafo. Na verdade, utiliza-se o termo "rede" e "grafo" de forma equivalente. Um grafo é um formalismo matemático para a teoria das redes. Um caminho é uma sequência de conexões de um nó líder para outro nó na rede. O comprimento de um caminho é igual ao número de saltos (*links*) ao longo do caminho. Assim, se um nó A está ligado ao nó C através de uma ligação de A para B, então B para C, o caminho é de comprimento dois. Leva dois saltos para ir do nó A para C. Se nenhum caminho existe entre uma parte da rede e outra parte, diz-se que a rede é dividida em segmentos ou componentes.

Um componente é uma rede independente ou uma ilha no grafo que representa uma rede. Uma rede de energia é separada em grafos com componentes ou ilhas quando ocorre uma queda de energia maciça, como aconteceu com a Rede de Energia do Nordeste dos Estados Unidos em agosto de 2003.

Aqui estão alguns termos que serão usados: Os grafos são coleções de nós e *links*, além de possuírem uma função de mapeamento que define como os nós e os *links* estão conectados. Esta função de mapeamento funciona como uma tabela que lista os pares de nós que são conectados por um *link*. A preocupação aqui é só com grafos simples - isto é, grafos em que cada par de nós conectados tem um *link*. Por exemplo, se duas cidades estão ligadas por duas estradas paralelas, utiliza-se apenas um *link* para representar esta conexão. Se dois oleodutos correm lado a lado considera-se este *link* único. Grafos simples são igualmente desprovidos de laços – que ocorrem quando um nó está ligado a si mesmo. Por exemplo, se uma rede de computador conecta um computador para si próprio, ignora-se esta ligação. Somente *links* que conectam os nós são considerados distintos em um grafo simples. Sempre que cada nó em um grafo pode ser alcançado por qualquer outro nó no grafo, pelo menos, um caminho, diz-se que o grafo é conexo. Por exemplo, cerca de 20% da Internet é um grafo fortemente conexo, porque nós dentro deste

conjunto de usuários podem enviar e-mails para todos os outros. A parcela que sempre conecta dois usuários de e-mail para um outro é chamada de componente fortemente conexo.

Se é possível percorrer alguns dos nós de um grafo e retornar ao seu ponto de partida, é dito que o grafo têm um ciclo. Por exemplo, se as estradas dentro de seu estado permitem viajar de uma cidade, por meio de outras cidades, e retornar ao ponto de partida, sem recuo, em seguida, o mapa das estradas do seu estado é um grafo com um ciclo. Grafos que não têm ciclos são chamados acíclicos, por razões óbvias. Nos grafos acíclicos é impossível completar um ciclo e retornar a um nó previamente visitado. Visando esclarecimento acerca dos conceitos relacionados a teoria das redes, na seção seguinte aborda-se a teoria dos grafos.

3.2 Modelo de Erdös-Rényi

Em meados de 1959 surgiu uma proposta de organização topológica de uma determinada rede, desenvolvida por Paul Erdös e Alfred Rényi. Os autores propuseram um modelo considerado simples para geração de grafos aleatórios. Tal modelo foi bastante discutido e estudado por matemáticos até que surgissem outros modelos, quase meio século mais tarde.

O modelo de rede aleatória, primeiramente, propõe a definição de um conjunto de vértices V conectados aos pares com probabilidade p. Com $p = 0$, obtém-se uma rede dita fragmentada. Com $p = 1$, a rede fica completamente conectada e seu coeficiente de agrupamento tem o valor máximo, $ca = 1$. Também se pode construir uma rede aleatória a partir da definição do número máximo de arestas, e conectar pares de vértices aleatoriamente, até que se alcance esse número máximo.

Este modelo, independente do esquema de formação, resulta em uma típica distribuição de conexões, dada pelo grau médio da rede $\langle k \rangle = p(N-1)$, onde N é o número de nós. A distribuição é definida como a de *Poisson* que é

$$\mathrm{P}(\mathrm{k}) = \frac{k^k e^{-k}}{k!}. \qquad (1)$$

A Figura 3 mostra a distribuição de *Poisson*, típica desta topologia complexa.

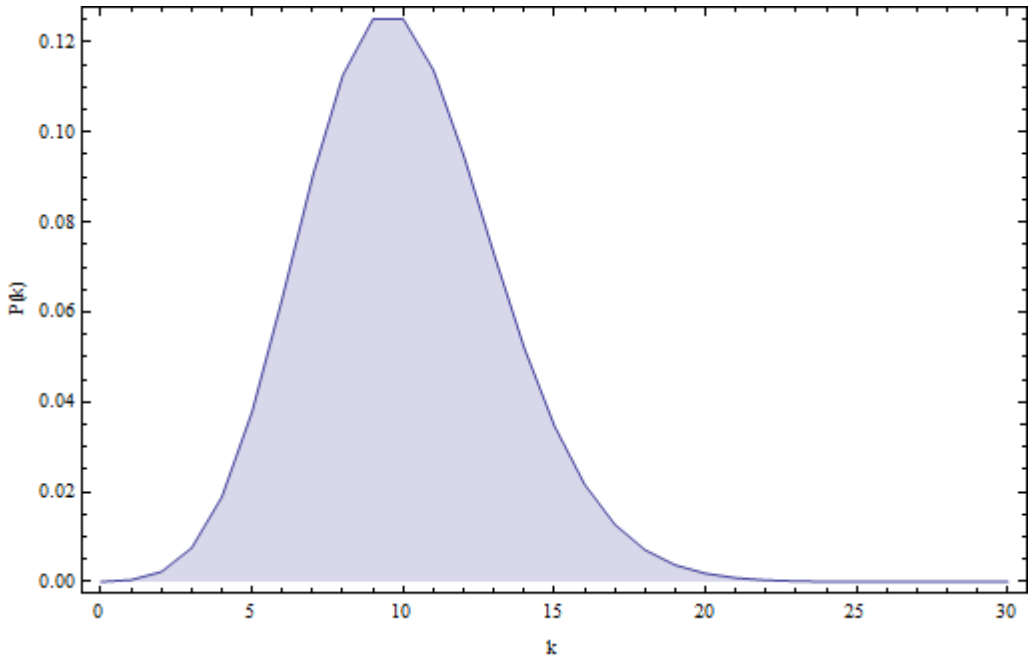

Figura 3: Distribuição de Poisson.

O resultado deste esquema de construção de uma rede resulta em vértices com poucas arestas, isto é, vizinhança pouco conectada quando probabilidade p é baixa. Neste caso, o coeficiente de agrupamento médio é $ca = \langle k \rangle / N$ é baixo e a rede é dita esparsa.

O fato de as conexões serem aleatórias faz com que o caminho médio entre dois vértices quaisquer seja muito pequeno. Este fato leva a consideração proposta por Duncan Watts e Strogatz, em 1998, que uma rede pode possuir um efeito denominado mundo pequeno, inicialmente testado em uma rede de contatos sociais em 1967 pelo sociólogo Stanley Milgram (MILGRAM, 1968).

Uma rede aleatória é caracterizada por um histograma de frequência que se encaixa em uma distribuição de Poisson.

3.3 Modelo de Watts-Strogatz

Um esquema de organização denominado de retornos crescentes é capaz de converter redes aleatórias em redes *Small-World*, obtendo então o modelo descoberto por Watts-Strogatz. Ocorre que os nós organizados inicialmente de forma aleatória vão sendo agrupados em vizinhanças até atingir estrutura denominada *Small-World* (NEWMAN, 2003).

O surgimento do modelo pequeno mundo a partir de uma rede aleatória, foi descoberto por Duncan Watts e Strogatz em 1998. Neste modelo, os nós se movimentam aleatoriamente e a cada instante um *link* é selecionado e substituído. A rede se torna cada vez mais agrupada (WATTS e STROGATZ, 1998).

O número de *links* e nós permanece constante e a rede continua a ser esparsa e totalmente conectada. O que se pode ver é a existência de grupos de nós. Estes grupos podem se formar e dissipar conforme os nós realizam a sua caminhada aleatória em torno de uma área retangular.

Há autores que comparam o procedimento como as nuvens de chuva num dia de verão ameno. Assim como as nuvens, estes agrupamentos ou *clusters* constroem-se, dispersam-se e formam-se novamente.

A observação do histograma deste tipo de rede não é uma lei de potência e não é estritamente uma distribuição de Poisson. Na verdade, o histograma de distribuição de *links* é mais próximo de uma rede aleatória do que uma rede *Scale-Free*.

Redes mundiais formam pequenas vizinhanças de nós críticos, em vez de *hubs* únicos. O modelo pequeno mundo representa redes do mundo real nas situações em que seus nós são orientados espacialmente.

Em uma rede de energia elétrica os nós de geração de energia, subestações e linhas de força são espacialmente fixos. Centrais não se movem. Mas a economia de disseminação de energia favorece ligações de mais curtas distâncias (menores *links*) ao invés de ligações de longas distâncias. Assim, a rede de energia tem evoluído de acordo com o princípio organizador do mundo pequeno. Apesar de extremamente simples, este princípio explica porque a maioria das redes fixas espacialmente são frequentemente redes mundo pequeno.

Duncan Watts, um pioneiro da pesquisa do mundo pequeno, define uma rede de pequeno mundo, em termos mais rigorosos como uma rede que é:

i. Grande - da ordem de bilhões de nós,
ii. Esparsa - a maioria da rede está vazia, com poucos *links*,
iii. Não há *hub* dominante - em vez disso, há conjuntos de nós,
iv. Agrupamento de vizinhanças - caminhos de um nó a outro são curtos,
v. Seis graus de separação - cada nó é conectado a cada outro nó em um pequeno número de etapas, por exemplo, não mais de 6 *links* separam cada par de nós.

O histograma de uma rede pequeno mundo não apresenta um *hub*, essa distribuição mostra *clusters*. Também não é definitivamente uma lei de potência. As redes de pequeno mundo são diferentes das redes *Scale-Free*, porque têm *clusters* de nós em vez de *hubs* simples. Em vez de proteger um *hub* simples, esta estrutura exige proteção de todo o *cluster* de nós, se a idéia é aperfeiçoar a estratégia de proteção de infraestruturas críticas que apresentem esse esquema de organização.

3.4 Modelo de Barabási-Albert

Apesar da evolução promovida pelas descobertas do modelo aleatório e do efeito mundo pequeno, observou-se que algumas redes, como as sociais, apresentavam as propriedades já descritas, mas com diferenças no mecanismo de construção e na distribuição das conexões. Isto levou a tentativa de buscar uma nova definição para as novas observações.

Em 1973, Price sugeriu um modelo de crescimento das redes de citações entre autores, no qual autores já citados tinham predisposição a atrair mais e mais citações. Todavia, apenas em 1999 Albert Lasló Barabási e sua então aluna Réka Albert propuseram uma modelo de construção de rede para explicar a estrutura de ponteiros entre paginas da web. A descoberta do crescimento das conexões preferenciais rompeu com a ideia de aleatoriedade na formação das redes. (WANG e CHEN, 2003)

Neste modelo, uma rede é construída sobre um conjunto de vértices, completamente conectados no inicio. A chegada de novos vértices traz um número fixo de arestas, conectadas preferencialmente aos vértices com maior número de conexões do conjunto inicial. Este fenômeno é denominado *preferential attachment* (anexação preferencial) ou ainda *richs gets richers* (ricos ficam mais ricos). O resultado estrutural é a distribuição dos graus denominada *power-law* ou lei de potência na qual vértices mais antigos (minoria da rede) concentram a maior parte das conexões e vértices recentes (maioria da rede) possuem poucas conexões, mostrado na Figura 4, com o valor do grau dos nós no eixo horizontal e a quantidade de nós no eixo vertical.

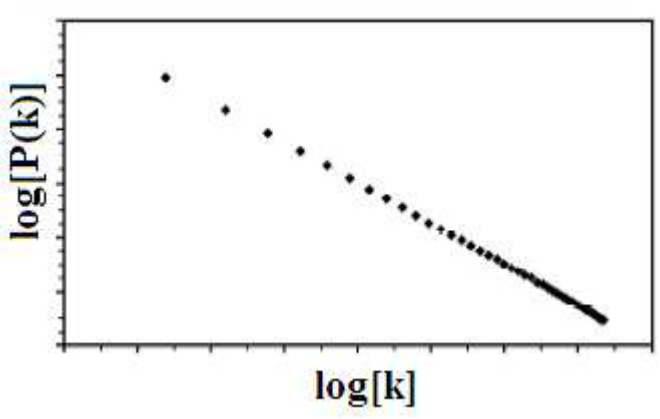

Figura 4: Distribuição *power-law*.

Esta rede resultante não possui uma escala especifica (como a rede aleatória), daí a sua denominação *Scale-Free* (Livre de escala). A presença de vértices concentradores, mais conectados, denominados *hubs*, faz o caminho médio entre quaisquer dois vértices ser diminuído. A conexão preferencial resulta no baixo índice de conectividade na vizinhança do vértice, mas com valor superior ao da rede aleatória.

A classificação de uma rede arbitrária de acordo com sua topologia é dada a partir do histograma de frequências dos graus dos nós. O histograma funciona como a "assinatura" da rede e representa a estrutura interna desta.

Todas as redes têm estrutura específica e esta estrutura é representada por um histograma. Portanto, o histograma é uma "fotografia" da estrutura interna da rede, definida pelo número de

ligações associadas a cada nó. O grau de um nó é definido como o número de ligações do nó para o resto da rede.

Desta forma, histogramas de diferentes formas correspondem a estruturas de rede diferentes. Mas, o que é a estrutura de uma rede? É a disposição dos *links* para cada um dos nós. Por exemplo, observa-se uma rede com 8 nós, numerados de 1 a 8, para cada nó é etiquetado seu grau. Se o nó 1 tem dois *links*, o seu grau é dois. Se o nó 2 tem 3 *links*, seu grau é 3 e assim por diante.

Estas contagens são registradas e contam-se quantos nós existem com grau igual a um, quantos com grau igual a dois e assim por diante.

Computados o número de nós com um grau *d*, converte-se cada uma dessas contagens em uma frequência dividindo cada contagem ao número total de nós. O resultado é o histograma de frequências.

A função matemática que começa com ajuste da curva apresentada pelo histograma pode ser classificada - e essa classificação é o nome que se dá para a rede.

Quando as barras do histograma de frequência diminuem rapidamente de acordo com a lei de potencia, a maioria dos nós tem grau baixo, poucos têm grau alto. Chama-se isso lei de potência, e as redes com essa estrutura são chamadas de *Scale-Free*.

Pensar em redes abrange, portanto, um espectro de "estruturas" - em um extremo, a rede é perfeitamente aleatória e obedece a distribuição de Poisson. No outro extremo a rede é sem escala, e obedece a uma lei de potência. Todo o resto é entre elas. Por exemplo, uma rede pequeno mundo exibe características de ambas - a aleatoriedade e uma certa quantidade de estrutura de lei de potência. Em outras palavras, uma rede de pequeno mundo não é aleatória nem *Scale-Free*. Ela cai em algum lugar no meio. Normalmente, o histograma de frequência de uma rede de mundo pequeno pode ser moldado como uma montanha russa.

A estrutura de rede é uma propriedade que se explora para determinar o que é crítico em uma infraestrutura crítica. A maioria das infraestruturas críticas são estruturadas, então se pode explorar a estrutura para reduzir o problema a um tamanho administrável.

Não se pode dar ao luxo de proteger tudo em todos os setores de infraestruturas críticas, assim deve-se escolher o que proteger. A classificação dos pontos críticos mais importantes torna-se, então, fundamental.

Para classificá-los, explora-se a estrutura das redes não-aleatórias. Isso significa alocar mais recursos para os nós críticos e *links* que os outros.

No mundo das infra-estruturas críticas, os nós mais importantes são aqueles que são raros, porque eles têm mais *links* do que quaisquer outros nós.

3.5 Centralidade

Na Teoria dos Grafos e na análise de redes existem várias medidas relacionadas ao conceito de centralidade de um vértice em um grafo. Tais medidas determinam a importância relativa de um vértice no grafo em que este se apresenta. Para a análise de redes existem quatro principais medidas de centralidade: grau de centralidade, *betweenness* e proximidade e autovetor de centralidade. (OPSAHL et al., 2010)

O primeiro e mais simples é o grau de centralidade, definido como o número de arestas incidentes em um nó. O grau é interpretado em termos do risco imediato de capturar quaisquer fluxos suportados pela rede, como um vírus ou dados. Se a rede é direcionada, isto é, as arestas possuem direção, usualmente se definem duas medidas para esse tipo de centralidade, denominadas grau de entrada e de saída.

O grau de entrada é o número de arestas direcionadas ao nó e o grau de saída é o número de arestas direcionadas a outros nós. Para relacionamentos de amizade ou conselhos, normalmente interpreta-se grau de entrada como uma forma de popularidade e o grau de saída como os agregados. Para um grafo $G = (V,E)$ com n vértices, o grau de centralidade $C_D(v)$ para o vértice v é dado por

$$C_D(v) = \frac{\deg(v)}{n-1}. \qquad (2)$$

Calculando o grau da centralidade para todos os vértices *V* em um grafo leva em uma densa matriz de adjacência a representação do grafo e para as arestas E em um grafo leva a uma representação de matriz esparsa[3].

A definição de centralidade pode ser estendida a grafos. Considera-se v sendo o nó com maior centralidade em *G*. Seja *X*: = (*Y*,*Z*) o nó n conectado ao grafo que maximiza a quantidade

$$H = \sum_{j=1}^{|Y|} C_D(y*) - C_D(y_i). \tag{3}$$

então o grau de centralidade do grafo *G* é definido como

$$C_D(G) = \frac{\sum_{i=1}^{|V|} |C_D(v*) - C_D(v_i)|}{H}. \tag{4}$$

H é maximizado quando o grafo *X* contém um nó que é conectado a todos os outros nós e todos os outros nó são conectados somente a este nó central (grafo estrela). Neste caso,

$$H = (n-1)\left(1 - \frac{1}{n-1}\right) = n - 2\,. \tag{5}$$

Então o grau de centralidade de *G* se reduz a:

[3] Uma matriz é esparsa quando a maioria de seus elementos são iguais a zero. A elaboração de uma representação para matrizes esparsas leva em consideração a facilidade de acesso tanto para linhas como para colunas e a eficiência nas operações (soma, multiplicação, escalonamento, etc.)

$$C_D(G) = \frac{\sum_{i=1}^{|V|} |C_D(v*) - C_D(v_i)|}{n-2} . \quad (6)$$

Betweenness é a centralidade de intermediação ou entrocamento. É mensurada pelo vértice no interior de um grafo. Existe também a *betweenness* de arestas. Vértices em que ocorrem muitos menores caminhos entre outros vértices tem grande *betweenness* do que aqueles que não ocorrem.

Para um grafo G: = (V,E) com n vértices o *betwenness* $C_B(v)$ para um vértice v é calculado como segue:

1. Para cada par de vértices *(s,t)* calcule todos os menores caminhos entre eles.

2. Para cada par de vértices *(s,t)* determinar a fração de menores caminhos que passam através do vértice em questão, aqui, v.

3. Somar esta fração sobre todos os pares de vértices *(s,t).*

Matematicamente têm-se:

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}, \quad (7)$$

onde σ_{st} é o numero de menores caminhos de s para t e $\sigma_{st}(v)$ é o numero de menores caminhos de s para t que passa pelo vértice v. Este pode ser normalizado dividindo-se o numero de pares de vértices que não incluem v, que são $(n - 1)(n - 2)$ para grafos direcionados e $(n - 1)(n - 2) / 2$ para grafos simples. Por exemplo, em um grafo estrela simples, o vértice central, contido em cada possível menor caminho, teria a *betweenness* $(n - 1)(n - 2) / 2$ igual a 1, se normalizada,

enquanto as folhas, que não estão contidas em nenhum menor caminho teriam a *betweenness* igual a 0.

Calcular a *betweenness* e a centralidade de proximidade de todos os vértices do grafo envolve o calculo os menores caminhos de todos os pares de vértices no grafo. Isto leva ao algoritmo de Floyd-Warshall, modificado não apenas para encontrar os menores caminhos como também para somá-los. Em um grafo esparso, o algoritmo Johnsons´s pode ser até mais eficiente. Em grafos não ponderados, o calculo da centralidade do *betweenness* tempo distinto do utilizado pelo algoritmo Brandes (BRANDES, 2001).

Em matemática, a proximidade é um dos conceitos básicos de topologias e espaços. Diz-se que dois conjuntos são próximos se estão arbitrariamente com pouco espaço entre si. A noção de distancia entre elementos do espaço é definida, mas também generalizada para espaços topológicos onde não existem formas de mensurar distancias.

Na teoria dos grafos a proximidade é a centralidade medida de um vértice em um grafo. Vértices pouco profundos com relação aos demais, com pequenas distâncias geodésicas, tem maior proximidade. Centralidade é utilizada em análise de redes para mensurar o tamanho do menor caminho, conforme dá maiores valores para vértices mais centrais, sendo também associada a outras medias como o grau.

Na teoria das redes, proximidade é uma medida sofisticada da centralidade. Esta é definida como a distância geodésica média (menor caminho) entre um vértice v e todos os outros vértices alcançáveis a partir dele, ou seja,

$$\frac{\sum_{t \in V/v} d_G(v,t)}{n-1}, \tag{8}$$

onde $n \geq 2$ é o tamanho do componente de conectividade da rede V alcançável a partir de v. Proximidade pode ser considerada como a medida de quão longe a informação deverá percorrer para se espalhar de um vértice para todos os outros por ele alcançáveis (NEWMAN, 2003).

3.6 Coeficiente de Agrupamento ou Transitividade

Em teoria dos grafos, um coeficiente de agrupamento ou *clustering* é uma medida do grau em que nós em um gráfico tendem a se agrupar. Evidências sugerem que na maioria das redes do mundo real, e em particular as redes sociais, nós tendem a criar grupos coesos caracterizados por uma densidade relativamente alta de laços. Em redes do mundo real, este risco tende a ser maior do que a probabilidade média de um empate aleatoriamente estabelecida entre dois nós (WATTS e STROGATZ, 1998).

Existem duas versões dessa medida: o coeficiente de agrupamento global e o local. A versão global foi concebida para dar uma indicação geral do agrupamento na rede, enquanto que o local dá uma indicação da inserção de nós.

3.6.1 Coeficiente de Agrupamento Global

O coeficiente de agrupamento global é baseado em trios de nós. Uma trinca são três nós que estão ligados por dois laços (trinca aberta) ou três (trinca fechada) não direcionados. Um triângulo é composto por três trincas fechadas, centradas em cada nó. O coeficiente global de agrupamento é o número de trincas fechadas (ou 3 triângulos) sobre o número total de trincas (abertas e fechadas).

A primeira tentativa de medida foi feita por Luce e Perry (1949). Esta medida dá uma indicação do agrupamento em toda a rede (global), e pode ser aplicado tanto a simples quanto as direcionadas.

Formalmente, pode ser definido como

$$C = \frac{3X\ nº\ de\ triângulos}{nº\ de\ triplas\ de\ vértices\ conectados} = \frac{nº\ de\ triplas\ fechadas}{nº\ de\ triplas\ de\ vértices\ conectados}. \quad (9)$$

3.6.2 Coeficiente de Agrupamento Local

O coeficiente de agrupamento local de um nó de é calculado como a proporção de ligações entre os seus vizinhos que são efetivamente realizadas em comparação com o número de todas as conexões possíveis. Por exemplo, um nó que tem três vizinhos, pode ter um máximo de três conexões entre ele e os demais.

O coeficiente de agrupamento local de um vértice em um grafo quantifica o quão perto os seus vizinhos estão no grafo. Duncan J. Watts e Steven Strogatz introduziram a medida em 1998, para determinar se um grafo é uma rede de pequeno porte, com efeito "mundo pequeno".

Um grafo $G = (V, E)$ é formalmente constituído por um conjunto de vértices V e um conjunto de arestas E entre eles. Uma aresta e_{ij} conecta o vértice i com vértice j. A vizinhança N de um vértice é definida como vizinhos imediatamente conectados a ele da seguinte forma:

$$N_i = \{v_j : e_{ij} \in E, e_{ji} \in E\}. \quad (10)$$

o grau k_i de um vértice é definido como o numero de vértices, $|N_i|$ na vizinhança de N_i.

O coeficiente de agrupamento local C_i de um vértice v_i é dado pela proporção de ligações entre os vértices na sua vizinhança, dividido pelo número de ligações que poderiam existir entre eles.

Para um grafo direcionado, e_{ij} é distinta da e_{ji} e, portanto, para cada vizinho de N_i existem k_i $(k_i - 1)$ ligações que poderiam existir entre os vértices na vizinhança (k_i é o grau total (entrada + saída) do vértice). Assim, o coeficiente de agrupamento local em grafos direcionados é dado por

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i-1)} : v_j, v_k \in N_i, e_{jk} \in E. \quad (11)$$

Em um grafo simples, existe a propriedade que e_{ij} e e_{ji} são considerados idênticos. Portanto, se um vértice v_i tem k_i vizinhos, podem existir

$$\mathrm{N}_{\text{arestas}} = \frac{1}{2} k_i(k_i - 1) \quad (12)$$

arestas entre os vértices com vizinhança. Logo, o coeficiente de agrupamento local para grafos simples pode ser definido como

$$C_i = \frac{2|\{e_{jk}\}|}{k_i(k_i-1)} : v_j, v_k \in N_i, e_{jk} \in E. \quad (13)$$

Denomina-se $\lambda_G(v)$ como o número de triângulos em $v \epsilon\, V(G)$ para o grafo simples G. Isto é, $\lambda_G(v)$ é o numero de subgrafos de G com 3 arestas e 3 vértices, um dos quais é v. Define-se $\tau_G(v)$ como sendo o numero de trincas de $v \epsilon\, V(G)$, ou seja, $\tau_G(v)$ é o número de subgrafos com 2 arestas e 3 vértices, um dos quais é v e tal como v é incidente a ambas as arestas. Então, pode-se definir o coeficiente de agrupamento

$$C_i = \frac{\lambda_G(v)}{\tau_G(v)}. \quad (14)$$

Ambas as definições precedentes são as mesmas, desde que

$$\tau_G(v) = C(k_i, 2) = \frac{1}{2} k_i(k_i - 1). \quad (15)$$

Estas medidas resultam em 1 se cada vizinho conectado a v_i é também conectado a cada outro vértice da vizinhança, e é igual a 0 se nenhum vértice conectado a v_i conecta-se a qualquer outro vértice que seja conectado a v_i.

3.6.3 Coeficiente de Agrupamento Médio

O Coeficiente de agrupamento de uma rede é dado por Watts e Strogatz como a média do coeficiente de agrupamento local de todos os vértices *n*.

$$C = \frac{1}{n} \sum_{i=1}^{n} C_i \, . \quad (16)$$

Um grafo é considerado "mundo pequeno" se seu coeficiente de agrupamento médio $\bar{C}$ é significativamente maior que o de um grafo aleatório construído no mesmo conjunto de vértices e se o grafo tem aproximadamente o mesmo tamanho médio de menor caminho e se o possui grafo aleatório correspondente. Redes com grandes coeficientes de agrupamento são conhecidas por terem estrutura modular e ao mesmo tempo tem a menor possível distância media entre diferentes nós.

3.7 Análise Espectral

Definir um grafo em termos de sua matriz laplaciana auxilia na extração de informações referentes à sua topologia. A matriz Laplaciana de um grafo *G*, denominada *L(G)* é a combinação

da matriz de conexões e a matriz de graus (diagonal): $L = C\text{-}D$, onde D é a matriz diagonal e C é a matriz de conexões. A matriz D tem elementos diagonais $d_{i,i,}$ igual ao grau do nó v_i :

$$d_{i,j} = \begin{cases} \text{grau}(v_i), \text{ se } j=i; \\ 0, \text{ outros casos.} \end{cases}$$

Extrair propriedades dinâmicas a partir destas matrizes usando técnicas similares a métodos usados no entendimento do comportamento de sistemas lineares como o vibrar de um instrumento musical, estados de partículas em mecânica quântica ou vários outros sistemas eletromecânicos. O raio espectral é computado a partir da matriz de adjacência e o gap espectral é computado a partir da matriz laplaciana do grafo (LEWIS, 2009).

Uma matriz é a transformação de um sistema linear que se move no tempo e no espaço de acordo com uma equação dinâmica: $D[A] = AX$ onde D é um operador linear, A é algum tipo de transformação linear e X é o estado do sistema. Por exemplo, D pode ser o operador derivativo de tempo e A pode ser a força de amortecimento agindo em um sistema mecânico, como um braço mecânico ou um sistema de suspensão automotivo. Um típico sistema linear terá ciclos através de algum tipo de padrão. Pode-se determinar a natureza desse padrão pela decomposição da resposta do sistema ao estimulo em um grupo fundamental de modos ou bases vetoriais em processos matemáticos chamados de decomposição espectral. Os vetores-base são chamados vetores ortogonais ou autovetores. Especificamente se a diagonal da matriz é λ, então A pode ser decomposto como $A = \lambda\mathbf{I}$, onde $\mathbf{I}$ é a matriz identidade e λ é a matriz contendo autovalores ou $det[A - \lambda\mathbf{I}] = 0$.

Com o risco de simplificação, o processo de análise espectral é uma forma de encontrar modos básicos de oscilação em sistemas lineares e expressá-los em termos de constantes denominadas autovalores. O maior autovalor é tipicamente denotado λ_1 é de particular interesse por representar o modo dominante de sistemas dinâmicos. Se $\lambda_1 < 0$, a oscilação eventualmente cessa, caso contrário, aumenta e leva a instabilidade do sistema.

A Ciência das Redes encontra uso para duas medidas espectrais: o raio espectral e o *gap* espectral. O raio espectral de um grafo G é o maior autovalor não nulo da matriz de adjacência de G e o *gap* espectral é o maior autovalor não nulo da matriz Laplaciana de G. O raio espectral pode explicar o espalhamento de dados através da rede e o *gap* espectral pode explicar a estabilidade ou a falta desta em uma rede.

3.7.1 Raio Espectral

O raio espectral *r(G)* é mensurado a partir da matriz de adjacência de um grafo. Ele é o maior autovalor não trivial do $det[A(G)- \lambda I] = 0$, onde *A* é a matriz de adjacência e *I* é a matriz identidade. Autovalores são as diagonais $\lambda_1, \lambda_2, \lambda_3, ..., \lambda_n$ de λI**.**

Os autovetores de raio espectral são também chamados de autovalores característicos por caracterizarem a topologia do grafo em termos sucintos. De fato, são idênticas ou proximamente relacionadas ao grau médio λ do grafo estruturado. Relembrando, grau médio $\lambda = (2m/n)$ é o número médio de *links* conectados a um nó. Grau médio e raio espectral não são geralmente os mesmos, mas raio espectral pode ser considerado uma representação compacta da função de mapeamento de um grafo. De fato, o raio espectral é determinado completamente pela topologia (sequência de graus) de um grafo e, portanto, é mensurado pela topologia do grafo.

3.7.2 Gap Espectral

A matriz Laplaciana *L(G)* de um grafo é a matriz de adjacência deste grafo com suas diagonais definidas como: $Diagonal(i,i)=-\sum c_{ij}=-grau(v_i)$, com *j* diferente de *i.* Seja A a matriz de adjacência, *D* a diagonal contendo os graus dos nós e *L* a Laplaciana. Logo, $L = A-D$.

O *Gap* espectral de *G* é o maior não trivial autovalor de *L*. Se a matriz de adjacência for simétrica, a matriz Laplaciana também será. Há outras definições de matriz Laplaciana, como por exemplo $L = D^{-1}A-I$. Porém, definições alternativas como esta citada é fornecida para matrizes que apresentam linhas e colunas que somam zero. Por causa desta propriedade, no contexto das redes, se torna mais útil a definição ideal de matriz Laplaciana como $L = A - D$.

3.8 Redes Cascata

As implicações mundiais do funcionamento das redes *Scale-Free* é de grande interesse para os setores das infraestruturas críticas. Primeiro, as redes *Scale-Free*, são menos vulneráveis a ataques, se protegidos os seus *hubs*. Em segundo lugar, muitos dos setores de interesse são de estrutura *Scale-Free*. A Internet, a maioria dos sistemas de transporte, sistemas de água e sistemas de transmissão de energia são redes consideradas *Scale-Free*. A rede elétrica tem mostrado ser uma rede mundial complexa e crítica.

A teoria das redes tem sido usada para explicar um fenômeno semelhante à organização de muitas infraestruturas críticas e suas "arquiteturas". Entre a mais intrigante é a noção de uma rede social. Por exemplo, se os atores são os nós e os filmes são *links*, em seguida, uma rede social pode ser formada pela colocação de um *link* entre dois atores se realizaram um mesmo filme. Em um nível mais sério, as redes sociais têm sido usadas para explicar por que as epidemias se espalham rapidamente entre as pessoas. A epidemia de SARS, por exemplo, foi espalhada através de uma rede social.

Agora, uma análise mais profunda das redes levanta questões como: "Como as falhas espalham-se através de uma infraestrutura, fazendo com que todo o setor falhe como o apagão na Rede Elétrica de 2003 no Leste dos EUA?" Este é um exemplo claro de uma falha em cascata - um dos tipos mais graves de falhas em todas as infraestruturas críticas.

Falhas em cascata são falhas do sistema de rede que começam com uma falha relativamente insignificante, que se propaga ao longo de uma parte importante da infraestrutura, que termina em colapso. Normalmente, um evento pequeno, como um curto-circuito em uma linha de alimentação, progride para uma grande falha, como o desligamento de um gerador de energia, o que leva a um erro ainda maior, como desligar toda a energia de uma subestação.

Em uma rede aleatória, uma epidemia vai se extinguir se a taxa de propagação (probabilidade de que um nó *A* infectado irá infectar nó *B*, se *A* e *B* estão ligados por um *link*) é menor que um limite específico. Caso contrário, a epidemia persistirá.

Uma falha ou o contágio se propaga através de *links*, saltando de nó em nó. Quanto mais *hubs* puderem ser protegidos menores danos poderão ocorrer por uma falha em cascata.

É pouco provável que um ataque aleatório no *hub* irá ocorrer, em relação à probabilidade de que um ataque ocorra no outro nó, uma vez que os nós *hubs* são minoria na rede. Todavia,

quando um ataque ocorre, se o nó *hub* está protegido, ele não se atinge e não contribui para a propagação da falha.

Em resumo, o aumento de tolerância a falhas, em uma rede, ocorre conforme há aumento no número de nós protegidos, em especial os nós *hubs*.

4 SIMULAÇÕES

Para execução das simulações, foi utilizado o software denominado *Attacker-Defender,* versão 4.1, totalmente desenvolvido em linguagem de programação Java, pelo professor Dr. Theodore Gyle Lewis, do Centro de Defesa e Segurança da Pátria, da *Naval Postgraduate School*, localizada na cidade de Monterey, Califórnia (EUA).

As simulações são focadas na análise da ocorrência de falhas em cascata na topologia de rede denominada *Scale-Free*.

4.1 Mecanismos de Contenção de Falhas

As Redes Complexas têm aptidão natural para representar praticamente qualquer sistema discreto. Elas possibilitam a integração e unificação de vários aspectos da ciência moderna, incluindo a inter-relação entre estrutura e dinâmica dos sistemas. Uma rede é composta por um conjunto de elementos (nós ou vértices) e suas ligações (arestas ou *links*). O número de ligações diretas de um nó aos seus vizinhos é chamado grau. Uma rede é definida como complexa quando seus nós apresentam diferentes graus. Muitas redes complexas são apenas parte de grandes sistemas e as topologias coexistentes interagem e dependem umas das outras (ALBERT et al, 1999) (DeMARCO, 2001).

Uma falha induzida nas redes pode ser ocasionada por testes de sistemas, ataques aleatórios ou direcionados, dentre outros (DeMARCO, 2001). Autores ao redor do mundo têm elaborado estudos acerca das falhas em cascata. Alguns resultados dos últimos oito anos são brevemente revisados a seguir.

Zhao et al. (2005) formularam uma teoria que produz estimativas para a integridade máxima alcançável de uma rede através da remoção controlada de um pequeno conjunto de nós de grau baixo. Todavia, o estudo não trata da remoção de nós de grau alto que são alvos potenciais de ataques direcionados.

De maneira semelhante, na busca de estratégias para contenção de falhas, DeMarco (2001) estudou as falhas em cascata em sistemas dinâmicos transitórios e bi-estáveis representados por equações diferenciais não lineares (Exemplo: circuito elétrico RLC[4]). Utilizando um detalhado circuito, descreveu como as características geométricas da função construída, juntamente com as trajetórias parciais de informação a partir de simulações, pode ser utilizada para predizer mais eficientemente quais ramos da rede estão sujeitos a uma falha em um cenário específico de perturbação. Entretanto, em uma situação em que o sistema não é transitório e bi-estável, as inferências são escassas, estando em aberto pesquisas que possam mostrar quais ramos estão mais sujeitos a falhas em sistemas diferentes do analisado.

Modelando uma rede de distribuição de energia, Carreras et al. (2002) mostra um ponto crítico de comportamento quando uma carga é incrementada na rede. Ao estudar diferentes pontos de transição e propriedades características de distribuição de função dos *blackouts*, consequências das falhas em cascata, observou que se tratam do primeiro passo na concepção de uma nova estratégia dinâmica para sistemas de distribuição de energia. Da mesma forma, o estudo de redes de telecomunicações poderá possibilitar conclusões que contribuam para novas estratégias de contenção de falhas.

Para representar o efeito cascata de linhas sobrecarregadas em um sistema de energia, Rios *et al*. (2002) propuseram um método baseado na simulação de Monte Carlo e avaliaram o custo esperado de um *blackout.* O método pôde mostrar que as falhas em cascata são, dentre outras coisas, consequências do mau funcionamento do sistema de proteção, de instabilidades do sistema e das respostas da própria rede a estes fenômenos. Deste modo, torna-se fundamental a análise aprofundada da rede, pois isto possibilitaria um novo caminho para desenvolver topologias melhoradas e bem estruturadas.

Visando o tratamento da falha em cascata por sobrecarga em sistemas de transmissão de energia, Dobson *et al*. (2005) propuseram uma estratégia que captura algumas das principais

[4] Consiste de um resistor (R), um indutor (L) e um capacitor (C), conectados em sério ou paralelo, no qual qualquer tensão pode ser descrita por uma equação diferencial.

características de grandes *blackouts*. A estratégia possibilitou mostrar como a carga no sistema pode influenciar no risco de ocorrência da falha em cascata. À medida que a carga é aumentada e ultrapassa seu ponto crítico, a distribuição do número de componentes que falham se satura a ponto de ocorrer falha em cascata, atingindo todo o sistema. Se, em redes de energia, que operam abaixo da capacidade total, a distribuição de cargas é problemática a ponto de ocasionar falhas, muito mais crítica é a distribuição de carga em Redes de Telecomunicações e maior é a chance de falhas ocorrerem, porque estas operam, frequentemente, acima de suas capacidades. Por exemplo, sistemas de telefonia comumente possuem mais clientes do que a capacidade de atendimento simultâneo das centrais de comutação.

Outro estudo acerca das falhas em cascata foi apresentado por Crucitti *et al*. (2004), que mostraram uma estratégia para contenção destas falhas baseado na redistribuição dinâmica de fluxos na rede. Eles mostraram que a quebra de um simples nó é suficiente para colapsar a eficiência de todo o sistema, se o nó estiver entre os nós de maior carga. Isto é particularmente importante para redes do mundo real com distribuição altamente heterogênea de cargas, assim como a Internet e redes de distribuição elétrica. Neste caso, reafirma-se a necessidade do estudo de mecanismos de contenção, uma vez que o ataque a um único nó, se este for o de alto grau, já possibilita o colapso da rede.

Sansavini et al. (2009) estudaram a influência de uma configuração de rede na sua vulnerabilidade ao espalhamento da falha em cascata. Observam que tal falha é iniciada por distúrbios seqüenciais que impõem cargas adicionais em cada nó, o que leva a falha total. A redistribuição resultante destas cargas adicionais sobre os nós vizinhos, por sua vez, leva a uma cascata por falhas locais ao longo das conexões de rede. Observam que a extensão destas falhas na rede depende da acessibilidade de suas partes constituintes e também da configuração específica do *link*-nó, das propriedades que são quantitativamente caracterizadas pela eficiência da conexão global e de um coeficiente de variação. Os resultados mostram que os menores valores de eficiência global tornam a rede mais resiliente a cascata pelo aumento da carga considerada crítica. No entanto, uma vez que uma carga crítica é ultrapassada, a transição para completar a falha ocorre mais rapidamente. A análise fornece parâmetros que são relevantes para novos *designs* de redes.

Wang et al. (2009) estudaram a redistribuição preferencial carga após a remoção de um nó da rede, propondo um mecanismo de falha em cascata para quatro modelos típicos de redes:

Barabási-Albert com propriedade S*cale-Free*, rede pequeno mundo Watts-Strogatz e a aleatória Erdos-Rényi. Assumindo que um nó falho lida somente com a redistribuição de carga para seus vizinhos, encontraram que todas as redes alcançam menor ou maior robustez contra falhas em cascata dependendo de parâmetros específicos ajustáveis. A robustez é quantificada por um limiar crítico, no qual ocorre uma transição do estado normal para o de colapso. Para um tamanho constante de rede, discutem-se correlações entre grau médio e o limiar crítico, concluindo que o limiar tem correlação negativa com o grau médio, isto é, quanto maior o valor do grau médio menor é o limiar. Estes resultados podem ser úteis para redes reais evitarem que falhas em cascata resultem em desastres.

Ash e Newth (2007) enfatizaram que a sociedade moderna tem se tornado cada vez mais dependente de grandes redes infraestruturadas, principalmente para escoamento de recursos para a sociedade e o eficiente fluxo dos negócios. Nos últimos dez anos, existem numerosos exemplos nos quais um distúrbio local resulta em falhas globais em sistemas. Devido a isso, desenvolveram um algoritmo evolutivo para projeto de redes complexas resilientes a tais falhas. Analisam as redes e as regularidades topológicas para explicar a origem de tal resiliência. Revelam que agrupamento, modularidade e longos tamanhos de caminhos desempenham papel fundamental no *design* de grandes e robustas infraestruturas.

Wu et al. (2006) estudaram a propagação da falha em cascata em redes ponderadas heterogêneas pela adoção de uma regra de redistribuição de fluxo local, onde o peso e a tolerância de um nó é correlacionado com o grau médio $\langle k \rangle$. Assumem que um nó falho lida somente com a redistribuição de fluxo para seus nós vizinhos mais próximos. Dá-se estimativas teóricas do início da falha para diferentes valores de $\langle k \rangle$. Encontra-se que a falha em cascata surge mais dificilmente em redes com $\langle k \rangle = 1.0$, enquanto se desenvolve mais vagarosamente para grandes valores de $\langle k \rangle$. Exploram também características estatísticas do tamanho da cascata em redes pela variação de $\langle k \rangle$ e obtêm cenários dinâmicos versáteis do processo da cascata, que exibe comportamentos considerados subcríticos, críticos e super-críticos.

Wu et al. (2010) simularam as falhas em cascata em redes S*cale-Free* com estrutura de comunidade, nas quais o grau local e o global obedecem a lei de potência. Enfocam o estudo da cascata em diferentes estratégias de remoção, denominadas *inter* e *inner*, para compreender a influência da cascata em tais redes. Encontraram que diferentes estratégias de remoção de uma

única importante aresta e diferentes modularidades[5] de rede podem desencadear diferentes cascatas de sobrecarga capazes de desativar totalmente a rede. Além disso, pouca modularidade significa maior vulnerabilidade. Estes resultados sugerem que, para evitar falhas em cascata em estruturas de comunidade são necessárias grandes modularidades e coeficientes de capacidade de reserva.

Hines et al. (2010) mostram que resultados de estudos de modelos de grafos topológicos são úteis para modelar vulnerabilidade em sistemas de energia. Mensuraram a susceptibilidade de redes de energia a falhas aleatórias e ataques utilizando três medidas de vulnerabilidade: características dos tamanhos de caminhos, perda da conectividade e tamanho de *blackouts*. Os dois primeiros são medidas puramente topológicas, já o tamanho do *blackout* resulta de um modelo simplificado de falha em cascata em redes de energia. Os testes foram feitos em seções da rede de energia do leste dos EUA, indicando que a dinâmica da topologia é similar aos grafos aleatórios, que tem distribuição exponencial dos graus. Entretanto, a perda de conectividade e a cascata indicam que redes de energia se comportam mais como redes S*cale-Free*, portanto, mais vulneráveis a ataques do que a falhas aleatórias. Os resultados mostraram que é necessário cautela ao se chegar a conclusões sobre vulnerabilidade das redes por métricas simplesmente topológicas. Isso porque leis físicas governam os fluxos em redes de energia (leis de Ohm's e leis de Kirchhoff's), bem como a localização de fontes e saídas de energia. Concluem que a estrutura elétrica da rede de energia é dramaticamente diferente da sua estrutura topológica e então, acaba influindo na dinâmica das falhas nas mesmas.

Arianos et al. (2009) estudaram que redes elétricas exibem padrões de reações para interrupções similares a redes complexas. Sequências de *blackouts* segue uma lei de potência assim como sistemas complexos operando próximos de um ponto critico. A tolerância da rede elétrica a interrupções acidentais e intencionais são analisadas sob a teoria das redes complexas. A eficiência é modificada pela introdução de um novo conceito de distancia entre os nós. Como resultado, um novo parâmetro chamado de *net-ability* é proposto para avaliar o desempenho das redes elétricas. Uma comparação entre eficiência e *net-ability* é colocada pela estimativa de vulnerabilidade de uma amostra de redes em termos das duas métricas.

[5] A modularidade é um valor que representa a qualidade da distribuição dos nós em uma parcela/módulo da rede.

Wang et al. (2009) estudaram falhas em cascata, em redes complexas que se congestionam/sobrecarregam. Estas são expressas como um processo de três fases: geração do congestionamento, difusão e dissipação. Uma função de congestionamento é proposta para representar a extensão deste em um dado nó. Inspirados pelo processo de reconexão do nó introduziram o conceito de "tempo de atraso", durante o qual o nó sobrecarregado não pode receber nem retransmitir qualquer tráfego. Desta forma, a gradação entre remoção permanente e a não remoção é construída e a flexibilidade do modelo é apresentada e demonstrada. Consideram que a conectividade de uma rede antes de depois de uma falha em cascata não é destruída porque o nó sobrecarregado não é removido permanentemente. Uma função de avaliação da eficiência da rede é também proposta para medir o dano causado pela cascata. Também investigam os efeitos da estrutura e tamanho da rede, tempo de atraso, processamento e velocidade de geração do tráfego na propagação da falha. Concluem que além de o processo de cascata ter três fases, outros fatores também afetam a propagação.

Denomina-se aqui "Estratégia 1" aquela proposta por Sun *et al*. (2008), que salienta que a maioria das redes reais possui topologia *Scale-Free*, o que significa que o número de *links* dos nós seguem a distribuição chamada de *power-law:* $P(k) \sim k^{-\theta}$ onde k é o número de *links* de um nó escolhido aleatoriamente na rede e θ é um expoente de dimensionamento. Os autores apontam que a segurança destas redes é uma das maiores preocupações dos trabalhos atuais. A propriedade "robustez ainda frágil" (redes robustas a falhas aleatórias, mas frágeis a ataques intencionais) inerente à topologia destas redes, leva à consideração de que é crucial a distribuição correta de cargas entre os nós, uma vez que a falha ou ataque em um único nó pode resultar em falha em cascata.

Tanto os nós quanto os *links* são sensíveis a sobrecarga, o que facilita a ocorrência das quedas. A remoção de nós, intencional ou aleatoriamente, muda o balanceamento dos fluxos na rede, que necessita lidar com a redistribuição das cargas. A rede pode não tolerar esta redistribuição e disparar sobrecargas em cascata. O grande problema é o fato de os nós ou *links* não estarem preparados para lidar com o tráfego extra. Em *Scale-Free Networks (SFNs),* os nós chamados *hubs* são aqueles que possuem grau muito maior do que os demais e desempenham papel dominante na conservação de conexões na rede. Em *SFNs*, a tolerância às falhas em cascata tem atraído interesse nos estudos. Todavia, muitos deles se centram em propriedades estáticas da rede, mostrando que a remoção de nós ou *links* podem ter importantes conseqüências. Alguns

poucos trabalhos abordam o processo de alocação dinâmica de capacidade para resistência à falha em cascata.

É difícil tornar a rede ao mesmo tempo robusta (saída) e com o menor custo (entrada). Seria possível obter o ganho máximo, definido como o valor da entrada menos saída, através da atribuição de capacidade de cada nó de maneira razoável? A Estratégia 1 propõe a alocação de capacidade (medida da quantidade de tráfego suportada pelos nós) contra as falhas em cascata para maximizar o ganho entre a atribuição de capacidade e a eficiência da rede. Também foi investigado o papel da Estratégia 1 na resistência de falhas em cascata e verificou-se que esta pode obter maior ganho que as outras.

Nesta estratégia, assume-se que, a cada intervalo de tempo, em média, um número λ de dados é gerado e o fluxo é encaminhado ao longo do menor caminho. A centralidade B_i pode ser utilizada para caracterizar o número de menores caminhos entre os pares de nós que correm através do nó *i*. Portanto, a carga em um nó é o número total de menores caminhos que passam através desse nó. A centralidade de um nó *i* pode ser definida como:

$$B_i = \sum_{j,l \epsilon N, j \neq l} \frac{n_{jl}(i)}{n_{jl}}, \tag{17}$$

onde n_{jl} é o número de menores caminhos conectando *j* e *l*, enquanto $n_{jl}(i)$ é o número de menores caminhos conectando *j* e *l* e passando por *i*.

Assume-se a capacidade do nó como a carga máxima que o nó pode lidar e é proporcional à sua carga inicial. O mecanismo de alocação da capacidade é dado pela equação (2):

$$C_i = \left(1 + \alpha \frac{B_i}{\lambda ND + \lambda}\right) L_i, \tag{18}$$

onde C_i é a capacidade do nó *i*, L_i é a carga do nó *i,* α é o parâmetro de tolerância, λ é a taxa média de geração de fluxo, *N* é o tamanho da rede, *D* é o tamanho médio do menor caminho.
Como se sabe, em redes reais feitas pelo homem, a capacidade é extremamente limitada pelo custo, assim, se

$$\prod_i = 1 + \alpha \frac{B_i}{\lambda ND + \lambda}, \tag{19}$$

quando *N* tende a infinito, a função custo pode ser escrita como

$$e = \frac{1}{N} \sum_{i=1}^{N} \alpha \frac{B_i}{\lambda ND}. \tag{20}$$

Como

$$\sum_{i=1}^{N} B_i = N(N-1)D, \tag{21}$$

a equação pode ser simplificada

$$e \approx \frac{\alpha}{\lambda}. \tag{22}$$

O tamanho da cascata é mensurado em termos da razão $G_{max} = N'/N$, número final sobre número inicial de nós, depois e antes da cascata, respectivamente.

A função lucro é definida como

$$R = Gmax - e, \tag{23}$$

onde G_{max} é a função renda e *e* a função custo.

Considerando estas equações, podem-se obter ótimos efeitos, tanto na robustez quanto no desempenho da rede, através da distribuição de capacidade razoável com custo mínimo possível. Assim, é possível propor uma infra-estrutura de rede de um ponto de vista econômico que confirma a idéia de que uma rede com o menor caminho possível nem sempre significa ganho pela sobrecarga que a rede pode suportar.

Denomina-se “Estratégia 2” aquela proposta por Zhao e Xu (2009), que tem seu foco em como prover robustez em redes já existentes, pela adição gradual de novos *links*. Quando se adicionam *links* entre nós de baixo grau na rede, esta pode melhorar a sobrevivência a ataques sem influenciar a tolerância a erros.

A tolerância a erro é a medida do quão tolerante uma rede se apresenta em relação a erros aos quais está sujeita. Quanto mais tolerante a rede for, melhor. A vulnerabilidade a ataques é o quão suscetível uma rede se apresenta em relação a uma falha intencionalmente criada. Estas são duas propriedades importantes das Redes Complexas que são normalmente utilizadas para avaliar a robustez de uma rede. Recentemente, muitos trabalhos foram dedicados a determinação do projeto de rede com robustez ideal. No entanto, pouca atenção se deu ao problema de como melhorar a robustez das redes existentes. Apresentou-se um novo parâmetro α', chamado parâmetro de reforço, para orientar o processo de ganho de robustez das *SFNs* acrescentando gradualmente novos *links*. Quando $\alpha' < 0$ os nós com graus inferiores são selecionados, preferencialmente. Enquanto os nós com graus mais elevados serão mais provavelmente selecionados quando $\alpha' > 0$. Foi mostrado teórica e experimentalmente no trabalho de Zhao e Xu (2009) que, quando $\alpha' < 0$ a sobrevivência ao ataque da rede melhora.

Através de experimentos e extensas comparações, conclui-se que o estabelecimento de novas ligações entre os nós com grau baixo pode reforçar drasticamente a capacidade de sobrevivência a ataque de redes *Scale-Free* e têm pouco impacto sobre a tolerância a erro.

O tamanho relativo do maior cluster e a média das distâncias entre pares de nós i e j inversa, denominada L^{-1}, são usadas para caracterizar o comportamento da rede durante os ataques. Pode-se também introduzir a eficiência média da rede,

$$E = \frac{1}{N(N-1)} \sum_{i=1,j=1}^{N} \frac{1}{d_{ij}}, \tag{24}$$

onde N é o tamanho da rede e d_{ij} é o tamanho do menor caminho entre i e j. Durante os ataques aos *hubs* ou falhas aleatórias, f_c é sempre usada para caracterizar a fração crítica (mínima) de nós que precisa ser removida até que a rede entre em colapso. f_c^{rand} e f_c^{targ} foram usados como respostas a falhas aleatórias e ataques aos *hubs*, respectivamente. Em particular, f_c^{rand} pode ser

calculada da seguinte forma: o custo é o fator chave que confina a estrutura no estabelecimento da rede. Define-se o custo *C* como razão do número de novos *links* sobre o número de *links* na rede inicial, que é:

$$C = \frac{E_{new}}{E_{init}} . \tag{25}$$

Pode-se computar f^{rand} usando a seguinte expressão:

$$f_c^{rand} = 1 - \frac{1}{\frac{k^2}{k} - 1} , \tag{26}$$

onde <*k*> é o primeiro momento (valor médio) do grau e <k^2> é o segundo momento do grau. Mas para o f_c^{targ} é preciso resolver algumas funções, dependendo da distribuição de graus. Considerando

$$\kappa = \frac{k^2}{k} , \tag{27}$$

a equação (10) pode ser reescrita como

$$f_c^{rand} = 1 - \frac{1}{\kappa - 1} . \tag{28}$$

Mas para f^{targ} precisa-se checar quando a condição $k < 2$ é satisfeita, para então, a fração crítica ser f^{targ}. Pelo uso de f^{rand} e f^{targ} pode-se caracterizar a tolerância a erro e sobrevivência a ataque diretamente.

É considerado impraticável manter <*k*> constante e religar os *links* para redes reais. Por exemplo, tanto a Internet quanto redes de energia foram formadas há tanto tempo que é quase impossível restabelecer <*k*> para melhorar a robustez da rede. No entanto, uma coisa que se pode fazer é adicionar um certo número de ligações à rede para alcançar uma maior robustez.

O ataque aos *hubs* a uma rede baseia-se na heterogeneidade da distribuição de graus. Contrário à falha aleatória, primeiro ocorre remoção do nó mais importante. Supõe-se que os atacantes conhecem a topologia global da rede: assim podem localizar o nó chave e atacá-lo. De fato, isso pode acontecer em redes reais. Existe importante relação do nó com o seu grau e a importância aumenta com o grau. Zhao e Xu (2009) realizaram experimentos de ataques baseados na topologia da rede *backbone* dos Sistemas Autômonos IPv6 (IPv6 AS). Utilizou-se f^{rand} e f^{targ} para indicar respostas às falhas aleatórias e aos ataques aos *hubs*, respectivamente. A rede IPv6 AS é vulnerável sob ataques aos *hubs* com $f^{targ} \approx 0{,}14$, mas robusta para falhas aleatórias com $f^{rand} \approx 0{,}96$. Relaciona-se a importância do nó com a sua centralidade devido ao fato de que a centralidade do nó estar fortemente relacionada ao seu grau na rede *backbone* IPv6 AS.

Em aplicações práticas, espera-se que a rede possa ser tolerante a erros ou falhas ocasionais, e também possa ser robusta a ataques aos *hubs*, especialmente no campo bélico. No entanto, as experiências mostram que o atacante só precisa remover uma parte muito pequena dos nós chave para fazer toda a rede entrar em colapso. A robustez elevada significa tanto a tolerância de erro quanto capacidade de sobrevivência a ataques.

Redes *Scale-Free* podem ser representadas como um grafo não direcionado *G(V, E)*, onde *V* é o conjunto de nós (do inglês, *vertex*) e *E* é o conjunto de arestas (do inglês, *edge*). Define-se Ψ como o conjunto de todas as ligações possíveis entre os nós em *V*.

Define-se o conjunto de arestas no grafo como

$$E = \Psi - \bar{\mathrm{E}}\,, \tag{29}$$

e o grau da aresta é dado por:

$$k_e = k_s k_d\,, \tag{30}$$

onde k_s e k_d são os graus dos dois nós da aresta e. Esta definição do grau da aresta como o produto dos graus dos nós conectados se deve, principalmente, porque o produto está fortemente relacionado à centralidade da ligação, a qual pode ser usada para caracterizar a importância da aresta. Zhao e Xu (2009) propõem um novo parâmetro α chamado parâmetro de execução e definem a probabilidade de escolha de um novo *link* e_i a partir do conjunto E da seguinte forma:

$$p(k_{e_i}) = \frac{k_{ei}^{\alpha}}{\sum_{e_j \in \bar{E}} k_{ej}^{\alpha}}. \quad (31)$$

Observa-se que adição de *links* entre nós de baixo grau não só melhora a sobrevivência a ataque da rede, como mantém a alta tolerância a erro e o custo decresce.

De fato, o coeficiente de agrupamento, que é a medida do grau em que os nós de uma rede tendem a se agrupar, representa a proximidade dos nós com relação aos seus vizinhos. Uma vez que um nó *hub* é atacado, seus vizinhos com baixos graus entrariam em colapso por perder transitividade do nó central. Podem-se estabelecer novas ligações entre os nós de borda, que são os vizinhos do nó *hub* para formar um *loop* local. Devido ao *loop* local, os nós de borda podem ainda ligar-se uns aos outros, mesmo quando o nó *hub* é atacado e sai de operação. Quando novos *links* entre nós de baixo grau são adicionados à rede, ocorre aumento de graus dos nós, mas ainda assim nós de baixo grau predominam. Em contraste, nós com elevados graus mantêm-se inalterados com alguns novos *links* de conexão para eles.

As estratégias apresentadas constituem uma pequena parcela daquilo que é possível de se desenvolver na área de redes. Dada a relevância e abrangência das Estratégias 1 e 2, escolheu-se simular a aplicação de ambas em redes para avaliar se há ganho maior na contenção das falhas em cascata do que a atuação dos dois em separado.

Neste trabalho, a avaliação é feita pela extensão dos danos causados pelas falhas em cascata induzidas e pela eficiência das estratégias de contenção. A extensão dos danos deve ser medida a partir da contabilização dos componentes, da análise da interconectividade, da quantidade de nós e conexões afetadas. A eficiência dos esquemas de contenção deverá ser avaliada pela capacidade de redução da cascata e retardo da propagação.

4.2 O software *Attacker Defender*

O software *Attacker Defender* é um software desenvolvido com o objetivo de gerar várias topologias de redes de vários tamanhos representando-as como grafos para que se simulem falhas nas mesmas. Existe a possibilidade de geração de redes regulares, redes aleatórias, redes mundo pequeno e redes *Scale-Free.* Também é possível pré-definir redes com topologia qualquer no seu código fonte e fixá-la em seu menu principal. Além disso, pode-se desenhar em tempo de execução a rede de preferência, dar nome aos nós, definir a capacidade de cada um deles, o direcionamento dos *links* e também editar esta rede.

Uma vez definida a rede com que se pretende trabalhar, é possível salvá-la em um diretório e abri-la novamente no futuro. Com qualquer uma das redes geradas, pode-se simular ataques a elas, podendo ser estes aos *hubs*, aleatórios, aos nós mais centrais, a um agrupamento de nós, ou a um nó único, de maior grau. Estas mesmas possibilidades se apresentam com relação à defesa da rede, que pode proteger cada uma das opções anteriores, de forma que a falha não acometa o nó cuja opção for escolhida.

A Figura 5 apresenta a interface do referido programa.

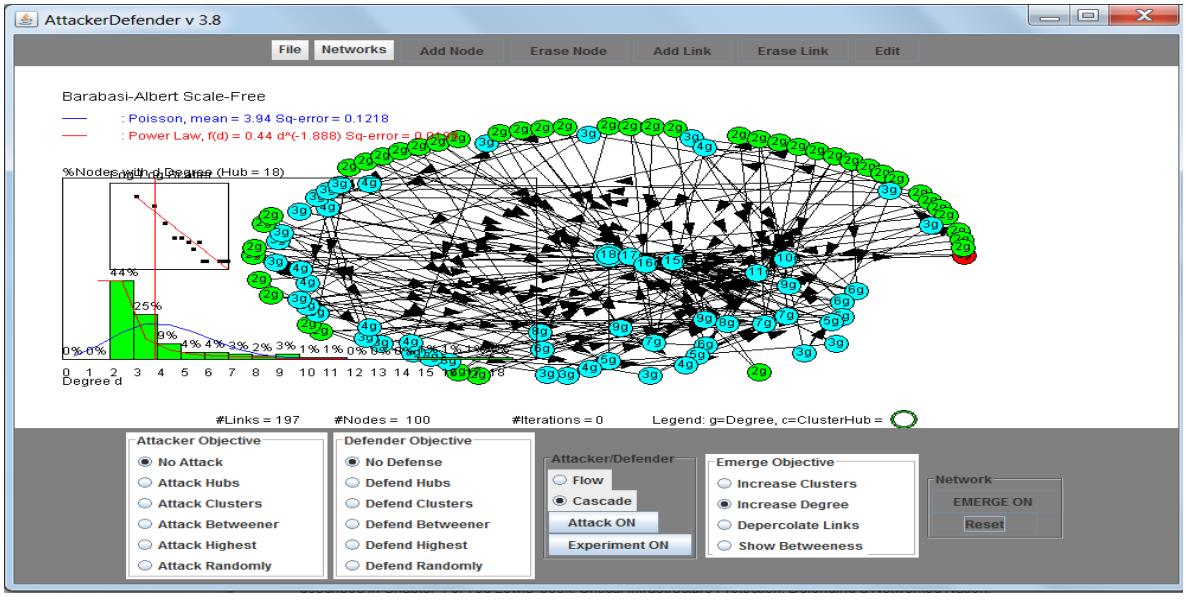

Figura 5: Inferface do software *Attacker Defender* quando da geração de uma rede de 100 nós.

Uma vez apresentada a rede, mostram-se o número de nós e *links* que esta apresenta, a distribuição de seus graus e em cada nó é mostrado o grau do mesmo. Os nós são apresentados

em 3 diferentes cores: nós verdes possuem apenas *links* de saída, nós azuis possuem *links* que entram e que saem e nós vermelhos são apenas receptores.

Neste trabalho, utiliza-se a opção de geração de rede *Scale-Free* para vários tamanhos e dentre as opções disponíveis de ataque utilizam-se o ataque aleatório e o aos *hubs* para a falha em cascata, que obedece ao modelo de infecção dos nós a uma dada taxa.

Com o objetivo de testar novas estratégias de defesa, distintas das já apresentadas pelo *software,* incluem-se no menu três novas opções de defesa: Adição de capacidade, Adição de *links* e Adição de Capacidade e *Links* ao mesmo tempo.

Cada nó possui uma capacidade *default,* padronizada na geração da rede. A implementação da adição de capacidade obedece aos requisitos da Estratégia 1, aqui já apresentada, incrementando essa capacidade inicial da rede em trinta por cento da capacidade *default* inicial. No código fonte, alteram-se: o menu, incluindo essa nova opção, e dentro da geração dessa opção o código faz com que a propriedade capacidade, presente em todos os nós, seja percorrida na rede atual exibida e então que ela seja alterada e aumentada.

Os nós também possuem informações do seu grau, e para implementação da adição de *links,* seguindo os princípios da Estratégia 2, percorre-se cada um dos nós, e em cada nó, é comparado o grau deste nó e o grau de cada um de todos os outros nós. Se o grau do nó atual for menor ou igual a 2 e ao mesmo tempo o grau de um outro nó também for menor ou igual a 2 e se já não possuírem *link* entre si, então é adicionado um novo *link* entre eles. Desta forma se adicionam *links* entre nós de baixo grau.

4.3 Etapa 1

Nesta etapa as Estratégias 1 e 2 são comparadas e combinadas via simulação computacional.

Os recursos necessários para desenvolvimento das simulações são: sistema operacional Windows, linguagem de programação Java, software de modelagem de redes e simulação de

falhas em redes *Attacker-Defender*[6] e o software de computação algébrica de *Wolfram Mathematica* para elaborar os gráficos.

As simulações consistem na observação do comportamento de redes em dois cenários distintos. O primeiro deles tem a presença de falhas/ataques direcionados aos *hubs*, ou seja, nós com maior número de arestas incidentes são afetados. O segundo cenário tem a presença de falhas/ataques aleatórios, ou seja, nós quaisquer são afetados.

A simulação da rede sujeita a falhas tem seu foco na análise de dez ordens: redes com dez, vinte, trinta, quarenta, cinquenta, sessenta, setenta, oitenta, noventa, e cem nós. A criação destas redes segue o modelo de criação *Scale-Free* em cada uma das ordens. Em cada uma das redes, para cada tipo de falha, são analisadas quatro situações: (i) a rede sem estratégias de contenção (controle), (ii) a rede com a Estratégia 1, (iii) a rede com a Estratégia 2 e (iv) a rede com as Estratégias 1 e 2 associadas.

Cada nó da rede tem uma probabilidade de espalhar a falha, variável de zero a um. Em cada simulação, considera-se a mesma probabilidade para todos os nós. Para o eficaz teste das estratégias, tal probabilidade é mantida fixa no valor meio. Implementa-se no programa *Attacker-Defender* as mudanças necessárias no seu código fonte que viabilizam tanto a adição de *links* quanto a de capacidade, bem como a adição simultânea de capacidade e *links*.

Para cada rede é fixada a capacidade equivalente a probabilidade de espalhamento, mantida em 0,5.

Para execução do ataque intencional aos *hubs*, primeiramente, é suposto que os responsáveis pelos ataques em redes conheçam sua estrutura topológica. Desta forma, os ataques são direcionados àqueles nós que causariam maior impacto em caso de falha, ou seja, os *hubs*. No primeiro cenário, os *hubs* são atacados e em seguida recolhe-se, para cada rede, a porcentagem não afetada pela falha, ou seja, a parcela "sobrevivente" de cada rede.

Para análise do impacto da adição de capacidade nas redes, a capacidade de cada nó é aumentada, seguindo os princípios da Estratégia 1. Em seguida, é simulada a falha com ataque aos *hubs* e recolhe-se a porcentagem da rede não afetada pela falha.

Para análise do impacto da adição de *links* nas redes, são colocados *links* extras na rede, seguindo os princípios da Estratégia 2, no qual os nós de baixo grau recebem novos *links* entre si.

[6] Gentilmente cedido pelo Prof. Dr. Ted Lewis.

Em seguida, é simulada a falha com ataque aos *hubs* e recolhe-se a porcentagem da rede não afetada pela falha, ou seja, parcela "sobrevivente" de cada rede.

Por fim, para a análise do impacto da associação das duas estratégias, adicionam-se *links* entre nós de baixo grau e também se incrementa a capacidade em cada uma das redes. Em seguida, é simulada a falha com ataque aos *hubs* e recolhe-se a porcentagem "sobrevivente" de cada rede.

Após estas simulações, os procedimentos relatados são repetidos, mas aí o que se simula é a falha aleatória.

A Tabela 2 mostra cada uma das redes geradas e utilizadas em cada um dos cenários para ser submetida à falha em cascata. Nesta etapa realizam-se oitenta simulações de falha em cascata.

Tabela 2: Simulações realizadas para teste das estratégias de contenção de falhas.

Número de nós	Rede	Ataque	Número da Simulação
10, 20, 30, 40, 50, 60, 70, 80, 90, 100	Normal	Hub	1ª – 80ª
		Aleatório	
	Com Estratégia 1	Hub	
		Aleatório	
	Com Estratégia 2	Hub	
		Aleatório	
	Com Est. 1 e Est. 2	Hub	
		Aleatório	

A análise leva em conta a eficiência das soluções de contenção testadas. A abrangência dos danos é mensurada através do número dos componentes da rede que foram atingidos pela falha.

4.4 Etapa 2

A segunda etapa de simulações consiste em analisar o comportamento da cascata sob diferentes vulnerabilidades para as redes, também nos cenários de ataque aleatório e aos *hubs*.

Para cada cenário de cada tamanho de rede, foram testadas as probabilidades de espalhamento da falha, variáveis de um décimo em um décimo, com zero (nenhum

espalhamento), 0.1, 0.2 e assim por diante, até um. Por exemplo, se um nó afetado pela falha tiver quatro conexões e sua probabilidade de espalhamento for 0.5, a falha será transmitida para duas conexões (50% dos nós conectados), ou seja, dois nós serão afetados.

A tabela a seguir mostra em detalhes as simulações realizadas, com os tamanhos das redes, o tipo do ataque, as vulnerabilidades testadas e o número da simulação.

Tabela 3: Simulações realizadas para teste do comportamento das falhas sob diferentes vulnerabilidades.

Número de nós	Ataque	Vulnerabilidade	Número da Simulação
10, 20, 70, 100, 200, 500, 1000	Hub	0,1	81ª – 220ª
		0,2	
		0,3	
		0,4	
		0,5	
		0,6	
		0,7	
		0,8	
		0,9	
		1,0	
	Aleatório	0,1	
		0,2	
		0,3	
		0,4	
		0,5	
		0,6	
		0,7	
		0,8	
		0,9	
		1,0	

A avaliação da robustez da rede pode ser visualizada, conforme se mudam os tamanhos da rede e as probabilidades de espalhamento da falha em cascata (vulnerabilidades).

A avaliação da rede tem seu foco na extensão dos danos causados pelas falhas. A quantidade de componentes que restam em operação auxilia no entendimento da abrangência da cascata, vulnerabilidade ou robustez da rede. A análise da interconexão pós-falha abrange a

análise de características dos nós, da distribuição dos graus, da localização e das correlações entre eles que influenciam a desintegração da rede e suas conexões.

Além disso, a submissão das redes a diferentes vulnerabilidades mostra o comportamento da cascata e se há variação em diferentes tamanhos. Também é possível definir matematicamente as curvas que representam o comportamento das falhas em cascata, tanto nos casos de ataque aos *hubs* quanto nas falhas aleatórias.

5 RESULTADOS E DISCUSSÕES

5.1 Eficiência das estratégias e da associação

Foram executadas as oitenta e uma simulações previstas na primeira etapa, visando análise e coleta dos dados relacionados a eficiência das estratégias de contenção de falhas. São recolhidas as porcentagens de nós sobreviventes, ou seja, aqueles não atingidos pela falha, para que, em cada um dos cenários e estratégias se possa mensurar o benefício ou malefício causado pela respectiva estratégia.

A Figura 6 mostra o número de sobreviventes S em função do número de nós da rede nas simulações da falha com ataque aos *hubs.*

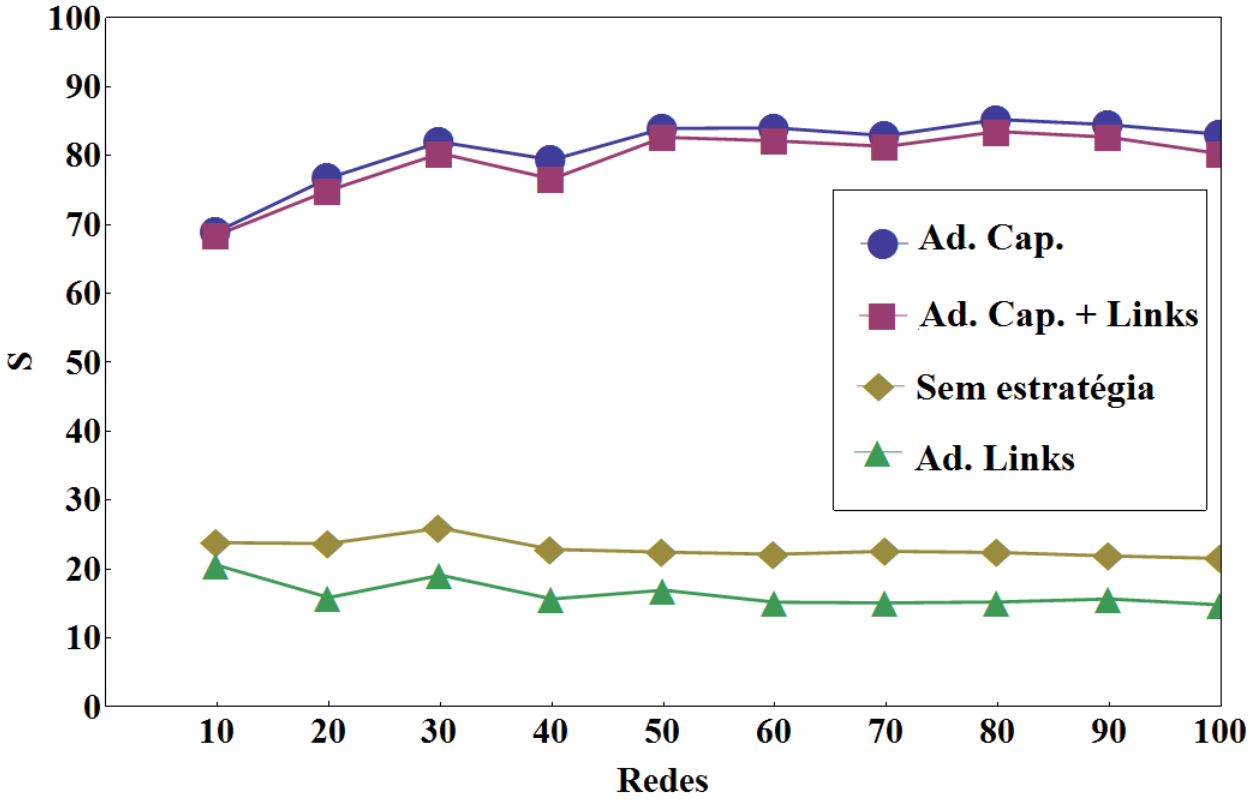

Figura 6: Número de sobreviventes para em função do número de nós para a falha aos *hubs*.

A Figura 6 mostra que a adição de *links* não se apresenta vantajosa para a rede, uma vez que reduz a parcela "sobrevivente" da rede, para todos os casos analisados, ou seja, reduz sua robustez em todos os tamanhos de rede.

A adição simultânea de *links* e capacidade nas redes incrementa a parcela de sobreviventes, porém não supera a adição isolada de capacidade, que passa a ser a melhor estratégia dentre as analisadas. A adição de capacidade na rede apresenta-se, portanto, a mais vantajosa, pois supera a parcela de sobreviventes apresentada por todos os outros cenários.

A Tabela 4 relaciona a porcentagem das redes não atingidas pela falha, nas redes sem estratégia, com cada uma das estratégias, e também com ambas as estratégias. Todos os casos se referem a falha disparada pelo ataque aos *hubs.*

Tabela 4: Redes e parcela não atingida nas simulações das estratégias quando do ataque aos *hubs*.

Nós	Ataque Hub	Links + Ataque Hub	Cap. + Ataque Hub	Cap. + Links + Ataque Hub
10	23,84%	20,56%	69,02%	68,52%
20	23,71%	15,87%	76,77%	74,97%
30	25,98%	19,12%	82,00%	80,35%
40	22,87%	15,67%	79,40%	76,63%
50	22,47%	16,95%	83,96%	82,68%
60	22,16%	15,20%	84,05%	82,16%
70	22,59%	15,10%	82,90%	81,35%
80	22,41%	15,23%	85,27%	83,51%
90	21,92%	15,67%	84,50%	82,69%
100	21,55%	14,81%	83,09%	80,26%

Na Figura 7 observa-se o comportamento das redes quando da aplicação isolada das estratégias e simultaneamente, mas agora a falha é introduzida aleatoriamente.

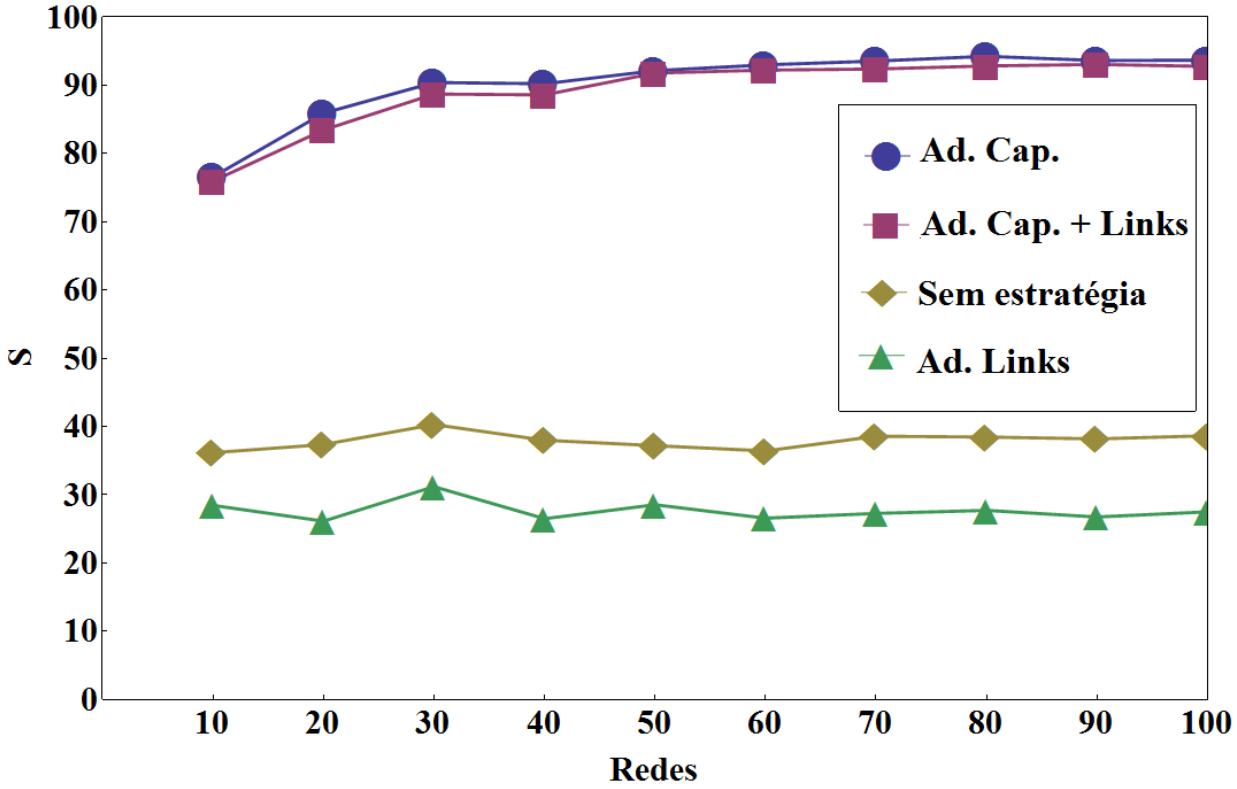

Figura 7: Distribuição de sobreviventes antes e após estratégias para a falha aleatória.

A análise das Figuras 5 e 6 confirma as conclusões do que ocorre com a rede para o cenário de ataque aos *hubs*. A diferença mais notável é o fato de a falha aleatória, para todos os casos da rede, apresenta sempre um número de sobreviventes maior do que os mesmos casos da falha aos *hubs*.

Isso se explica pelo fato de o ataque direcionado aos *hubs* ter a capacidade de atingir mais nós, uma vez que os nós *hubs* possuem mais conexões. Pode-se concluir também que as redes são mais vulneráveis aos ataques direcionados aos *hubs* e menos vulneráveis às falhas aleatórias.

A tabela a seguir relaciona a porcentagem de cada uma das redes que não foram atingidas pela falha, nas redes sem estratégia, com cada uma das estratégias, e também com ambas as estratégias. Todos os casos se referem à falha disparada pelo ataque aleatório.

Tabela 5: Redes e parcela não atingida nas simulações das estratégias quando do ataque aleatório.

Nós	Ataque Aleatório	Links + Ataque Aleatório	Cap. + Ataque Aleatório	Cap. + Links + Ataque Aleatório
10	36,17%	28,45%	76,58%	75,96%
20	37,37%	26,15%	85,91%	83,45%
30	40,29%	31,19%	90,40%	88,71%
40	38,00%	26,50%	90,22%	88,60%
50	37,20%	28,54%	92,15%	91,79%
60	36,43%	26,58%	92,99%	92,22%
70	38,58%	27,27%	93,55%	92,38%
80	38,46%	27,71%	94,24%	92,83%
90	38,19%	26,76%	93,62%	93,05%
100	38,66%	27,50%	93,69%	92,77%

Além da parcela sobrevivente, também se recolhem medidas estruturais de cada uma das redes em que foram adicionadas capacidade e *links*. A adição de capacidade não apresentou variação nestas medidas, já a adição dos *links* mostrou mudanças nas medidas das redes. Esta variação é colocada na Tabela 6.

Tabela 6: Redes, número de *links*, *betweeness*, coeficiente de agrupamento médio e raio espectral antes e após adição de *links*.

Nós	***Links***		***Betweeness***		***Cluster Coefficient***		***Spectral Radius***	
	Antes	**Depois**	**Antes**	**Depois**	**Antes**	**Depois**	**Antes**	**Depois**
10	17	18	21	18,5	0,4585	0,3966	3,91	4,04
20	37	42	37	62	0,503	0,306	4,92	5,16
30	55	61	57	104	0,282	0,1946	4,94	5,10
40	76	87	82	148	0,3284	0,2649	5,92	6,11
50	96	105	44	109,5	0,058	0,0588	5,53	5,62
60	117	135	116	345	0,1149	0,0824	6,01	6,14
70	135	152	198	310	0,1021	0,0545	6,27	6,41
80	157	176	201	435	0,098	0,07411	6,28	6,41
90	176	196	190	368	0,074	0,06015	6,61	6,73
100	197	223	200	376	0,17669	0,1312	7,1	7,25

Conforme as redes aumentam em tamanho, aumentam também o número de nós de baixo grau. Logo, o número de *links* adicionados também aumenta. A centralidade ou *betweenness* máxima é a medida daquele vértice presente em caminhos mais curtos entre quaisquer outros dois vértices. A centralidade máxima aumenta em todos os casos, devido ao aumento no número de caminhos possíveis e da ocorrência do vértice nestes caminhos, exceto na rede de dez nós, em que essa ocorrência foi reduzida devido à adição de um único *link*.

Já a medida do coeficiente de agrupamento médio é reduzida em todas as redes submetidas à adição de *links*. Essa medida mostra o quão agrupados os vizinhos de cada um dos nós da rede se apresentam. Após o cômputo destas medidas, ocorre o cálculo da média dos mesmos. Nos casos analisados, esta medida é sempre inferior a rede original, sem a adição de *links*.

Através dos experimentos, também se vê que a robustez das redes cai conforme se adicionam *links,* uma vez que os sobreviventes vão diminuindo. A comprovação deste fato é o

cálculo do Raio Espectral das redes antes e após a adição de *links*. O Raio Espectral aumenta após a adição de *links* em todos os casos, e o aumento do raio espectral em vários trabalhos significa redução da robustez da rede, e, portanto, o aumento de sua vulnerabilidade a ataques (LEWIS, 2009). Isto pode ocorrer em especial pelo fato de mais conexões significarem persistência no espalhamento da falha, que se espalha de forma semelhante a epidemias a uma dada taxa de infecção.

A Tabela 7 detalha as redes, as estratégias testadas, os cenários, os números das simulações com a porcentagem de nós sobreviventes e atingidos após a simulação das falhas em cascata.

Tabela 7: Redes, estratégias, cenários, nós sobreviventes e atingidos em cada simulação.

Número de nós	Rede	Ataque	Número da Simulação	Sobreviventes	Atingidos
10	Normal	Hub	1	23,84%	76,16%
		Aleatório	2	36,17%	63,83%
	Com Estratégia 1	Hub	3	20,56%	79,44%
		Aleatório	4	28,45%	71,55%
	Com Estratégia 2	Hub	5	69,02%	30,98%
		Aleatório	6	76,58%	23,42%
	Com Est. 1 e Est. 2	Hub	7	68,52%	31,48%
		Aleatório	8	75,96%	24,04%
20	Normal	Hub	9	23,71%	76,29%
		Aleatório	10	37,37%	62,63%
	Com Estratégia 1	Hub	11	15,87%	84,13%
		Aleatório	12	26,15%	73,85%
	Com Estratégia 2	Hub	13	76,77%	23,23%
		Aleatório	14	85,91%	14,09%
	Com Est. 1 e Est. 2	Hub	15	74,97%	25,03%
		Aleatório	16	83,45%	16,55%
30	Normal	Hub	17	25,98%	74,02%
		Aleatório	18	40,29%	59,71%
	Com Estratégia 1	Hub	19	19,12%	80,88%
		Aleatório	20	31,19%	68,81%
	Com Estratégia 2	Hub	21	82,00%	18,00%
		Aleatório	22	90,40%	9,60%
	Com Est. 1 e Est. 2	Hub	23	80,35%	19,65%
		Aleatório	24	88,71%	11,29%

40	Normal	Hub	25	22,87%	77,13%
		Aleatório	26	38,00%	62,00%
	Com Estratégia 1	Hub	27	15,67%	84,33%
		Aleatório	28	26,50%	73,50%
	Com Estratégia 2	Hub	29	79,40%	20,60%
		Aleatório	30	90,22%	9,78%
	Com Est. 1 e Est. 2	Hub	31	76,63%	23,37%
		Aleatório	32	88,60%	11,40%
50	Normal	Hub	33	22,47%	77,53%
		Aleatório	34	37,20%	62,80%
	Com Estratégia 1	Hub	35	16,95%	83,05%
		Aleatório	36	28,54%	71,46%
	Com Estratégia 2	Hub	37	83,96%	16,04%
		Aleatório	38	92,15%	7,85%
	Com Est. 1 e Est. 2	Hub	39	82,68%	17,32%
		Aleatório	40	91,79%	8,21%
60	Normal	Hub	41	22,16%	77,84%
		Aleatório	42	36,43%	63,57%
	Com Estratégia 1	Hub	43	15,20%	84,80%
		Aleatório	44	26,58%	73,42%
	Com Estratégia 2	Hub	45	84,05%	15,95%
		Aleatório	46	92,99%	7,01%
	Com Est. 1 e Est. 2	Hub	47	82,16%	17,84%
		Aleatório	48	92,22%	7,78%
70	Normal	Hub	49	22,59%	77,41%
		Aleatório	50	38,58%	61,42%
	Com Estratégia 1	Hub	51	15,10%	84,90%
		Aleatório	52	27,27%	72,73%
	Com Estratégia 2	Hub	53	82,90%	17,10%
		Aleatório	54	93,55%	6,45%
	Com Est. 1 e Est. 2	Hub	55	81,35%	18,65%
		Aleatório	56	92,38%	7,62%
80	Normal	Hub	57	22,41%	77,59%
		Aleatório	58	38,46%	61,54%
	Com Estratégia 1	Hub	59	15,23%	84,77%
		Aleatório	60	27,71%	72,29%
	Com Estratégia 2	Hub	61	85,27%	14,73%
		Aleatório	62	94,24%	5,76%
	Com Est. 1 e Est. 2	Hub	63	83,51%	16,49%
		Aleatório	64	92,83%	7,17%

90	Normal	Hub	65	21,92%	78,08%
		Aleatório	66	38,19%	61,81%
	Com Estratégia 1	Hub	67	15,67%	84,33%
		Aleatório	68	26,76%	73,24%
	Com Estratégia 2	Hub	69	84,50%	15,50%
		Aleatório	70	93,62%	6,38%
	Com Est. 1 e Est. 2	Hub	71	82,69%	17,31%
		Aleatório	72	93,05%	6,95%
100	Normal	Hub	73	21,55%	78,45%
		Aleatório	74	38,66%	61,34%
	Com Estratégia 1	Hub	75	14,81%	85,19%
		Aleatório	76	27,50%	72,50%
	Com Estratégia 2	Hub	77	83,09%	16,91%
		Aleatório	78	93,69%	6,31%
	Com Est. 1 e Est. 2	Hub	79	80,26%	19,74%
		Aleatório	80	92,77%	7,23%

5.2 Comportamento das falhas sob diferentes vulnerabilidades

As simulações para compreensão do comportamento da falha em cascata em redes *Scale-Free* é testada em sete tamanhos de redes: com dez, vinte, setenta, cem, duzentos, quinhentos e mil nós.

Para cada tamanho de rede, testam-se dois cenários: a rede sob falha aleatória e sob ataque intencional nos *hubs*. Para cada cenário, são simuladas diferentes probabilidades de espalhamento da cascata, ou vulnerabilidades, variáveis de zero até um, de valor equivalente para todos os nós da rede.

Em cada situação, coletam-se dados relativos à porcentagem da rede que sofreu falha e também a porcentagem que não foi atingida.

As Figuras 8 e 9 mostram o comportamento da cascata quando do ataque aos *hubs* em todos os tamanhos de rede, para todas as probabilidades de espalhamento e as porcentagens de nós sobreviventes (não atingidos) na Figura 8 e atingidos na Figura 9.

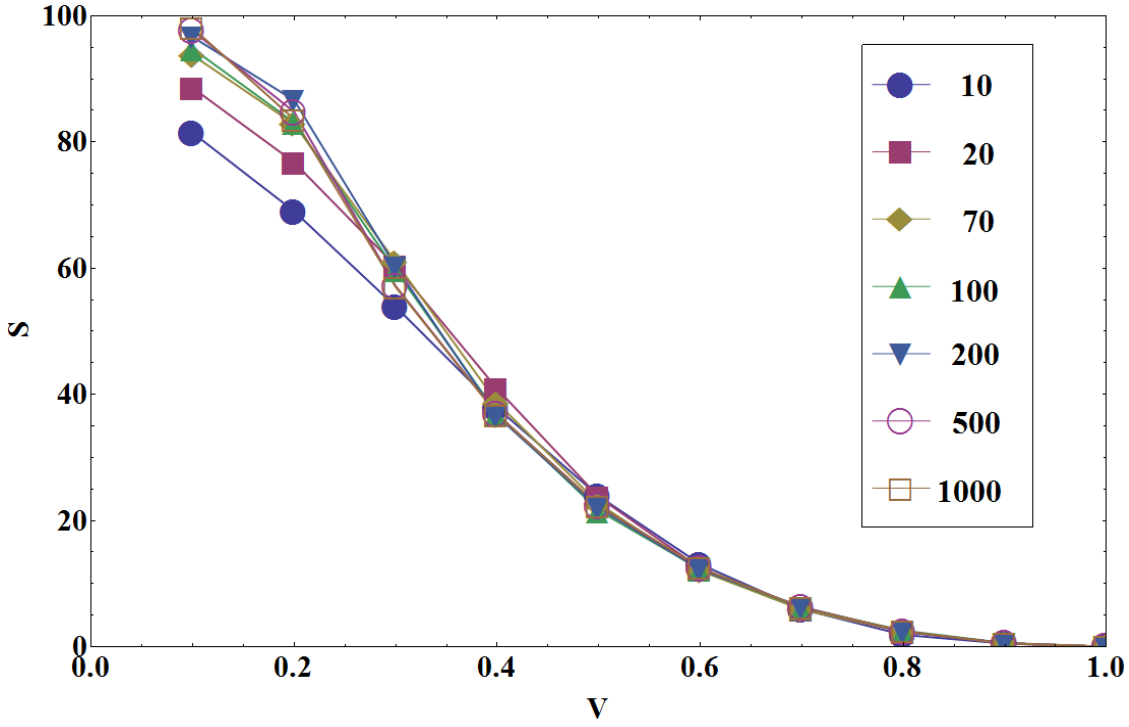

Figura 8: Distribuição de nós sobreviventes em várias redes sob diferentes vulnerabilidades quando ocorre falha nos *hubs*.

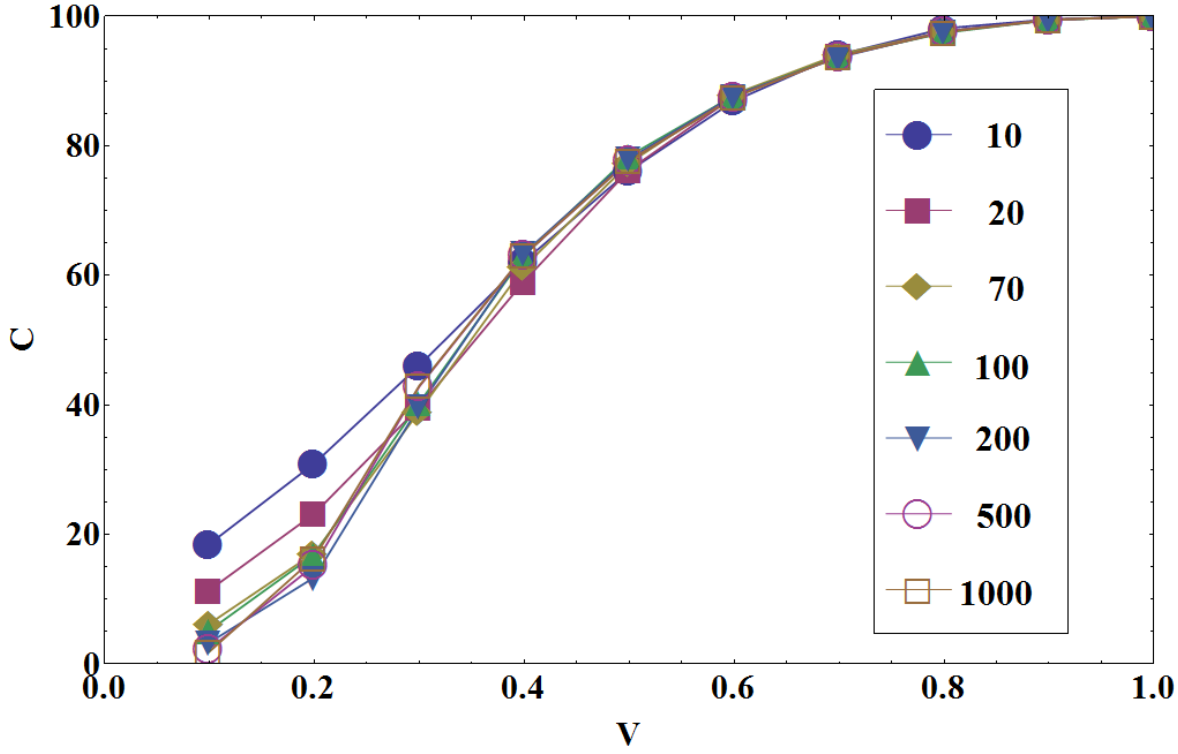

Figura 9: Distribuição de nós atingidos em várias redes sob diferentes vulnerabilidades quando ocorre a falha nos *hubs*.

A Tabela 8 mostra os dados utilizados para elaboração dos gráficos das Figuras 7 e 8.

Tabela 8: Dados coletados na segunda etapa das simulações, nós sobreviventes e atingidos quando ocorre falha nos *hubs*.

Simulações	Nós	Vulnerabilidade	Sobreviventes	Atingidos
81ª a 90ª	10	0,1	81,46%	18,54%
		0,2	69,02%	30,98%
		0,3	53,89%	46,11%
		0,4	37,81%	62,19%
		0,5	23,84%	76,16%
		0,6	13,07%	86,93%
		0,7	5,93%	94,07%
		0,8	1,87%	98,13%
		0,9	0,48%	99,52%
		1,0	0,00%	100,00%
101ª a 110ª	20	0,1	88,59%	11,41%
		0,2	76,77%	23,23%
		0,3	60,33%	39,67%
		0,4	40,86%	59,14%
		0,5	23,71%	76,29%
		0,6	12,39%	87,61%
		0,7	5,97%	94,03%
		0,8	2,37%	97,63%
		0,9	0,50%	99,50%
		1,0	0,00%	100,00%
121ª a 130ª	70	0,1	93,76%	6,24%
		0,2	82,83%	17,17%
		0,3	60,99%	39,01%
		0,4	38,66%	61,34%
		0,5	22,65%	77,35%
		0,6	12,02%	87,98%
		0,7	5,85%	94,15%
		0,8	2,37%	97,63%
		0,9	0,53%	99,47%
		1,0	0,00%	100,00%
141ª a 150ª	100	0,1	94,80%	5,20%

		0,2	83,09%	16,91%
		0,3	59,68%	40,32%
		0,4	36,90%	63,10%
		0,5	21,55%	78,45%
		0,6	12,23%	87,77%
		0,7	6,13%	93,87%
		0,8	2,55%	97,45%
		0,9	0,59%	99,41%
		1,0	0,00%	100,00%
161ª a 170ª	200	0,1	96,66%	3,34%
		0,2	86,61%	13,39%
		0,3	60,21%	39,79%
		0,4	36,54%	63,46%
		0,5	21,96%	78,04%
		0,6	12,24%	87,76%
		0,7	6,31%	93,69%
		0,8	2,38%	97,62%
		0,9	0,55%	99,45%
		1,0	0,00%	100,00%
181ª a 190ª	500	0,1	97,63%	2,37%
		0,2	84,76%	15,24%
		0,3	57,15%	42,85%
		0,4	36,83%	63,17%
		0,5	22,29%	77,71%
		0,6	12,42%	87,58%
		0,7	6,20%	93,80%
		0,8	2,39%	97,61%
		0,9	0,55%	99,45%
		1,0	0,00%	100,00%
201ª a 210ª	1000	0,1	98,08%	1,92%
		0,2	83,59%	16,41%
		0,3	57,02%	42,98%
		0,4	36,76%	63,24%
		0,5	22,27%	77,73%
		0,6	12,50%	87,50%
		0,7	6,19%	93,81%
		0,8	2,44%	97,56%
		0,9	0,55%	99,45%
		1,0	0,00%	100,00%

Observa-se que conforme a vulnerabilidade ou probabilidade de espalhamento aumenta, a porcentagem de sobreviventes tende a se igualar em todos os tamanhos de redes. Já quando a vulnerabilidade é menor, a tendência é a quantidade de nós não atingidos ser coerente com o tamanho da rede: quanto maior o tamanho, maior a porcentagem, além de próximas entre si.

As Figuras 10 e 11 mostram o comportamento da cascata quando do ataque aleatório em todos os tamanhos de rede, para todas as probabilidades de espalhamento e as porcentagens de nós não atingidos (Figura 10) e atingidos (Figura 11).

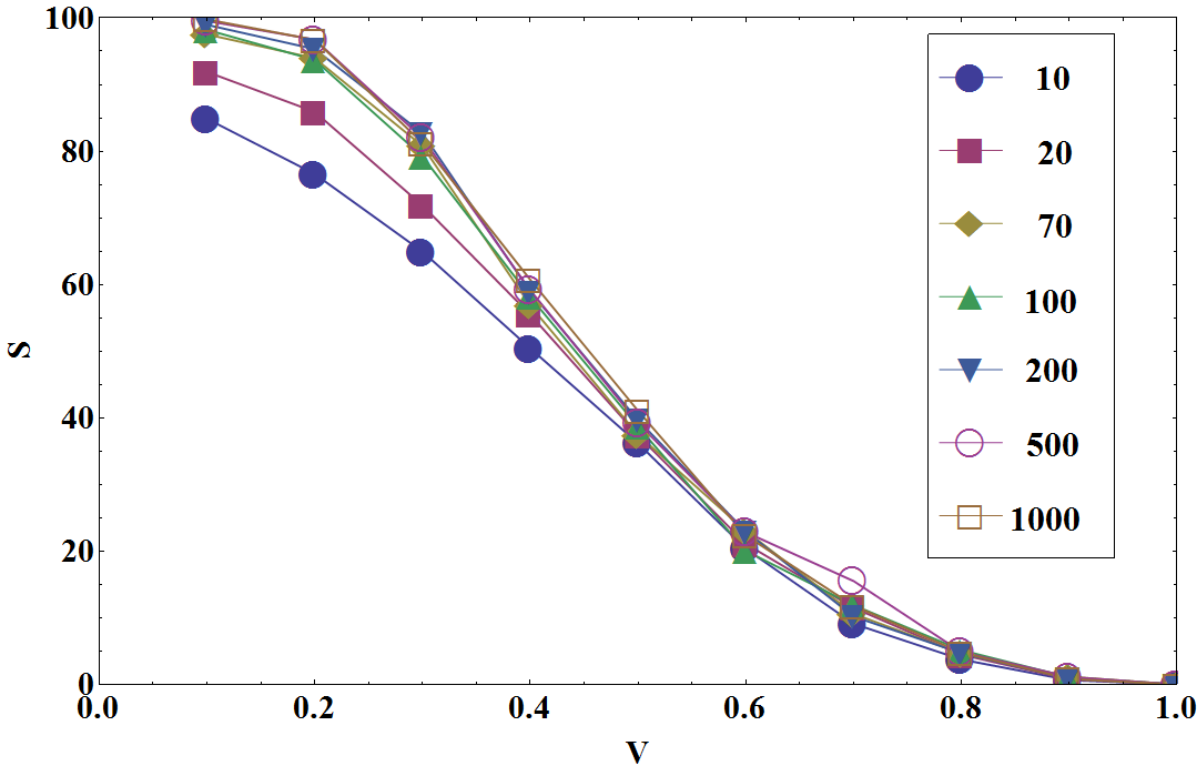

Figura 10: Distribuição de nós sobreviventes em várias redes sob diferentes vulnerabilidades quando ocorre a falha aleatória.

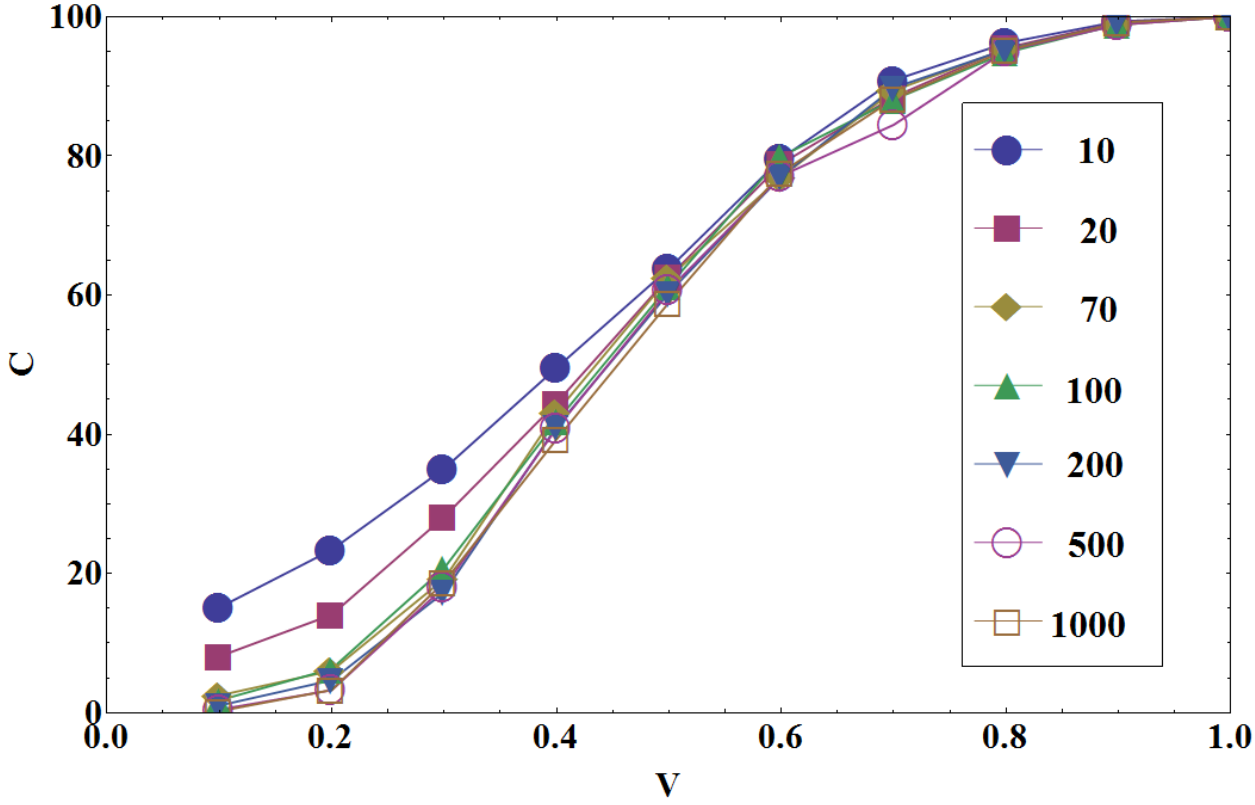

Figura 11: Distribuição de nós atingidos em várias redes sob diferentes vulnerabilidades quando ocorre a falha aleatória.

É evidente o aumento na porcentagem de nós não atingidos para todos os tamanhos, se comparada com as mesmas probabilidades de espalhamento nos casos de ataque aos *hubs*. Isso é devido ao fato de neste caso a falha ser aleatória, a qual atinge parcela menor da rede quando disparada. Observa-se também uma queda um pouco mais suave no número de nós sobreviventes, se comparada ao ataque aos *hubs* conforme se aumenta a probabilidade de espalhamento da falha. Isso pode ser consequência do tipo da falha, que neste caso é aleatória e ocasiona menor impacto global na rede.

A Tabela 9 mostra os dados das simulações no cenário aleatório para as redes sob diferentes vulnerabilidades, com a parcela de nós sobreviventes e atingidos.

Tabela 9: Dados coletados na segunda etapa das simulações, nós sobreviventes e atingidos quando ocorre falha aleatória.

Simulações	Nós	Vulnerabilidade	Sobreviventes	Atingidos
91ª a 100ª	10	0,1	84,88%	15,12%
		0,2	76,58%	23,42%
		0,3	64,93%	35,07%
		0,4	50,43%	49,57%
		0,5	36,17%	63,83%
		0,6	20,35%	79,65%
		0,7	9,12%	90,88%
		0,8	3,76%	96,24%
		0,9	0,66%	99,34%
		1,0	0,00%	100,00%
111ª a 120ª	20	0,1	91,91%	8,09%
		0,2	85,91%	14,09%
		0,3	71,86%	28,14%
		0,4	55,48%	44,52%
		0,5	37,37%	62,63%
		0,6	21,10%	78,90%
		0,7	11,51%	88,49%
		0,8	4,47%	95,53%
		0,9	0,78%	99,22%
		1,0	0,00%	100,00%
131ª a 140ª	70	0,1	97,52%	2,48%
		0,2	93,95%	6,05%
		0,3	80,78%	19,22%
		0,4	56,84%	43,16%
		0,5	37,45%	62,55%
		0,6	23,02%	76,98%
		0,7	10,59%	89,41%
		0,8	4,64%	95,36%
		0,9	1,17%	98,83%
		1,0	0,00%	100,00%
151ª a 160ª	100	0,1	98,21%	1,79%
		0,2	93,69%	6,31%
		0,3	79,20%	20,80%
		0,4	58,14%	41,86%
		0,5	38,66%	61,34%

		0,6	20,06%	79,94%
		0,7	11,90%	88,10%
		0,8	5,15%	94,85%
		0,9	1,07%	98,93%
		1,0	0,00%	100,00%
171ª a 180ª	200	0,1	98,96%	1,04%
		0,2	95,33%	4,67%
		0,3	82,68%	17,32%
		0,4	58,93%	41,07%
		0,5	39,65%	60,35%
		0,6	22,93%	77,07%
		0,7	10,20%	89,80%
		0,8	4,69%	95,31%
		0,9	0,86%	99,14%
		1,0	0,00%	100,00%
191ª a 200ª	500	0,1	99,54%	0,46%
		0,2	96,74%	3,26%
		0,3	81,96%	18,04%
		0,4	59,14%	40,86%
		0,5	39,21%	60,79%
		0,6	22,84%	77,16%
		0,7	15,52%	84,48%
		0,8	4,87%	95,13%
		0,9	1,14%	98,86%
		1,0	0,00%	100,00%
211ª a 220ª	1000	0,1	99,77%	0,23%
		0,2	96,72%	3,28%
		0,3	81,22%	18,78%
		0,4	60,72%	39,28%
		0,5	41,05%	58,95%
		0,6	22,35%	77,65%
		0,7	11,81%	88,19%
		0,8	4,80%	95,20%
		0,9	0,84%	99,16%
		1,0	0,00%	100,00%

5.3 Análise matemática do comportamento da cascata

A partir dos resultados obtidos e exibidos nos gráficos que explicitam em dois cenários os nós sobreviventes e os atingidos pelas falhas, considerou-se modificar a apresentação de tais gráficos para a escala logarítmica. É interessante observar o comportamento da cascata em todas as redes e observar que as curvas se apresentam próximas entre si.

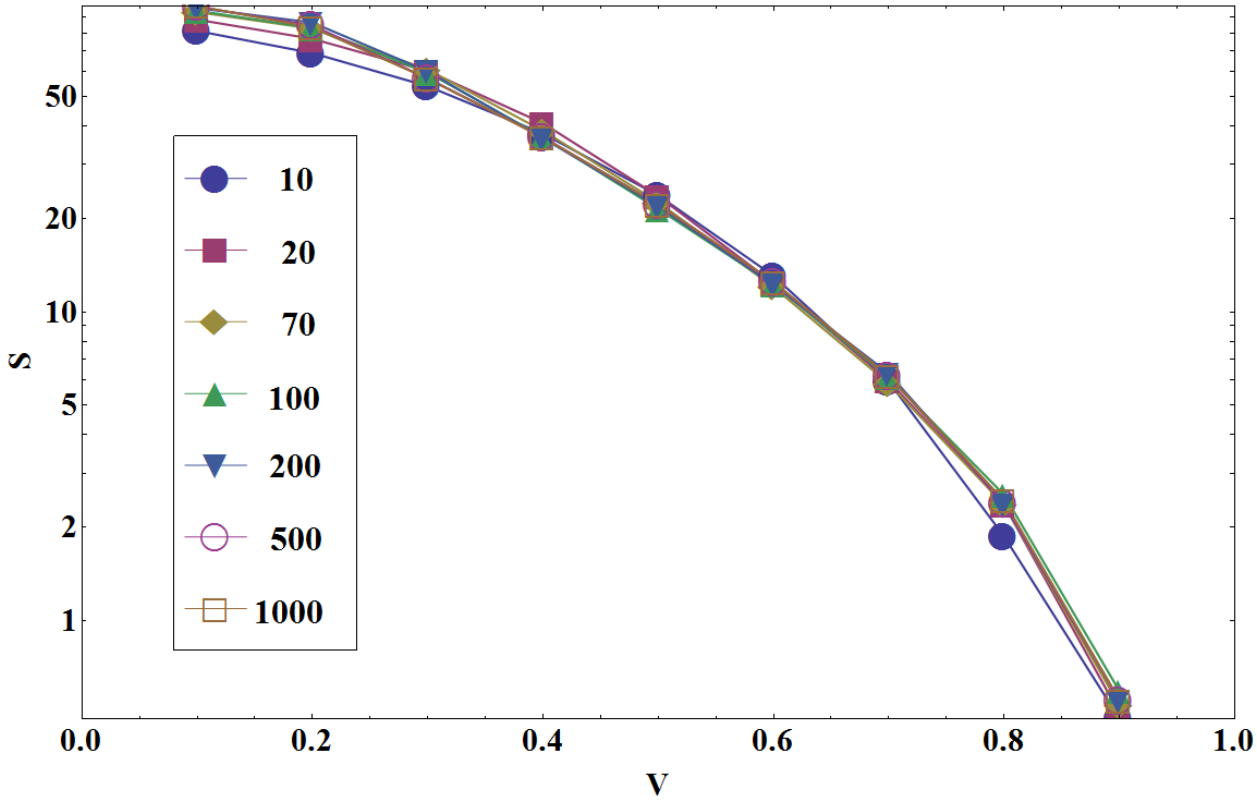

Figura 12: Distribuição de nós sobreviventes em várias redes sob diferentes vulnerabilidades quando ocorre a falha aos *hubs*, em escala logarítmica.

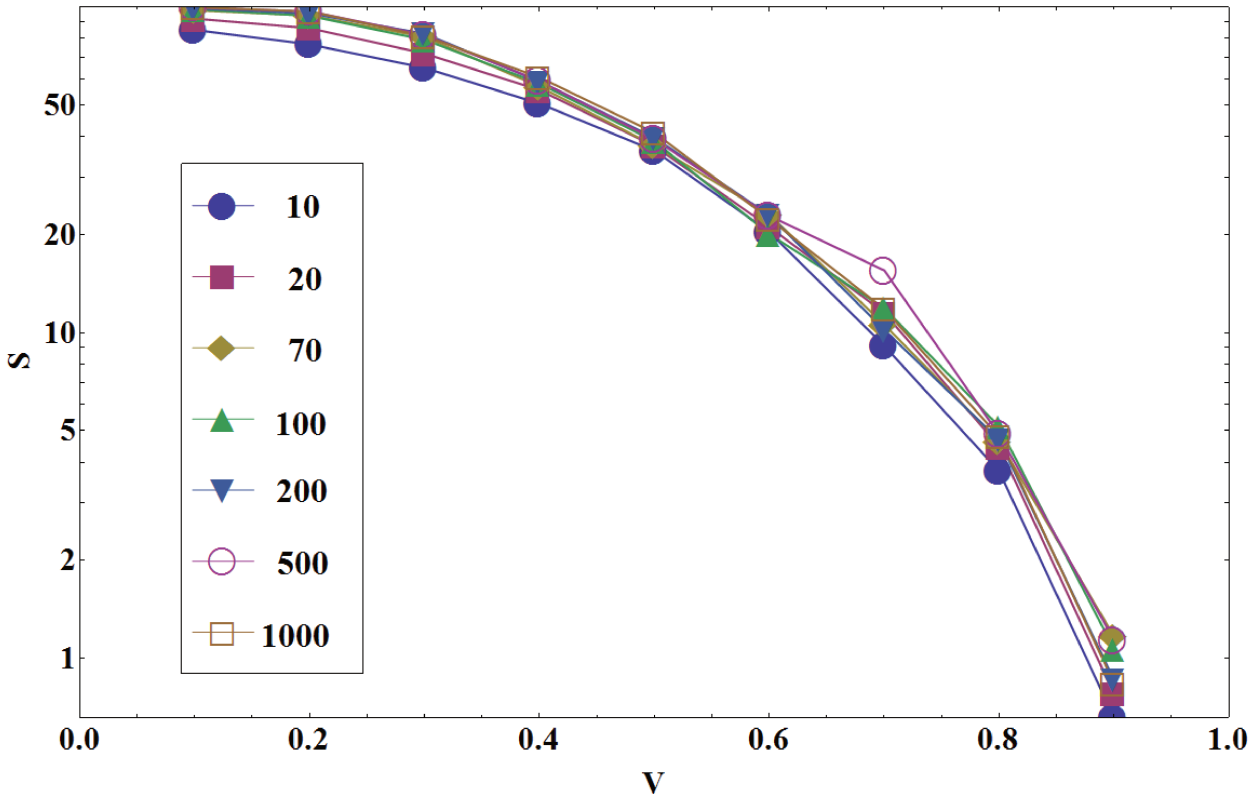

Figura 13: Distribuição de nós sobreviventes em várias redes sob diferentes vulnerabilidades quando ocorre a falha aleatória, em escala logarítmica.

A análise destas novas curvas mostradas nas Figuras 12 e 13 leva a consideração de tentar encontrar um modelo matemático capaz de descrever a trajetória destas. Considera-se então uma função exponencial.

Para o comportamento das curvas no ataque aos *hubs* da rede, considera-se o uso da equação a seguir:

$$S = e^{a(1-bv^2)} \tag{32}$$

A partir deste modelo faz-se um ajuste por mínimos quadrados para obtenção dos parâmetros *a* e *b* para cada tamanho de rede. Tal curva deve então ser resultado de uma função de variáveis que melhor se adéquem as curvas específicas dos nós sobreviventes.

Inicialmente, é encontrada uma função para a falha aos *hubs* para a rede de mil nós sob distintas vulnerabilidades. Ao comparar o modelo encontrado com os resultados das simulações, observa-se que se apresenta bem próximo da curva da falha original.

A partir das descobertas, utiliza-se o modelo proposto para os demais tamanhos de rede, contudo, o que se vê é que os parâmetros *a* e *b* devem ser modificados para cada tamanho, uma vez que a função ideal para a rede de mil nós não se apresenta ideal para as demais redes, quando da falha aos *hubs*.

Desta forma, a partir de cada curva de falha aos *hubs,* em cada tamanho de rede, utiliza-se uma função para os cálculos de parâmetros ideais de cada uma das curvas, de cada uma das redes. Aí sim são encontrados valores que se adéquam em cada caso de falha aos *hubs* em cada um dos tamanhos de rede.

Uma vez encontradas as curvas para as falhas aos *hubs* para cada uma das redes, considera-se utilizar os mesmos valores encontrados nos parâmetros para a falha aleatória. O que se observa é certo distanciamento da curva de falha aleatória em cada tamanho. Logo, vê-se a necessidade da definição de uma nova função adequada as curvas da falha aleatória.

Para o comportamento das curvas na falha aleatória, utiliza-se a equação a seguir:

$$S = e^{a(1-bv^c)} \tag{33}$$

Observa-se que na falha aos *hubs* o parâmetro *v* em todos os casos é elevado ao quadrado e sofrem mudanças os demais parâmetros. Mas para a falha aleatória, *v* elevado ao quadrado não se aproxima de maneira satisfatória, então, faz-se necessário encontrar mais um parâmetro, que é definido como *c*.

Uma vez definida a função, são determinados os valores ideais para os parâmetros *a, b* e *c* para cada um dos tamanhos de rede, para cada tipo de falha.

Colocando no mesmo gráfico a curvas da falha original e a curva da função exponencial, já com os valores de variáveis encontrados pelo *software* em cada um dos tamanhos de rede, é possível visualizar matematicamente o comportamento da cascata. Comportamento este, muito próximo ao da função exponencial parametrizada com os valores de variáveis cabíveis para cada cenário de falha, em cada tamanho de rede.

A Tabela 10, a partir dos dados dos sobreviventes à cascata em cada rede, relaciona os valores das variáveis substituíveis nas funções exponenciais citadas que descrevem as curvas apresentadas nos gráficos em escala logarítmica.

Tabela 10: Variáveis e funções que descrevem o comportamento das simulações de falhas para os nós sobreviventes.

Sobreviventes

Simulações utilizadas	Nós	Ataque	a	b	c	S, v = [0,1;1,0]
81ª a 90ª	10	Hub	4,45141	1,17071	-	$e^{a(1-bv^2)}$
101ª a 110ª	20		4,55630	1,18756		
121ª a 130ª	70		4,62180	1,27201		
141ª a 150ª	100		4,63104	1,31698		
161ª a 170ª	200		4,65923	1,33271		
181ª a 190ª	500		4,65369	1,35289		
201ª a 210ª	1000		4,65175	1,35710		
91ª a 100ª	10	Aleatório	4,44183	1,17263	2,51885	$e^{a(1-bv^c)}$
111ª a 120ª	20		4,53832	1,17263	2,52544	
131ª a 140ª	70		4,61326	1,26296	2,60895	
161ª a 160ª	100		4,61499	1,25503	2,59426	
171ª a 180ª	200		4,62226	1,31230	2,70485	
191ª a 200ª	500		4,64025	1,14506	2,51106	
211ª a 220ª	1000		4,63083	1,25573	2,65949	

Os valores relacionados na tabela 10 para as variáveis *a*, *b* e *c*, mostram pouca variação entre os vários tamanhos de rede, para cada cenário: o de ataque aleatório e aos *hubs*. A Figura 14 mostra essa pouca variação, ainda que necessária.

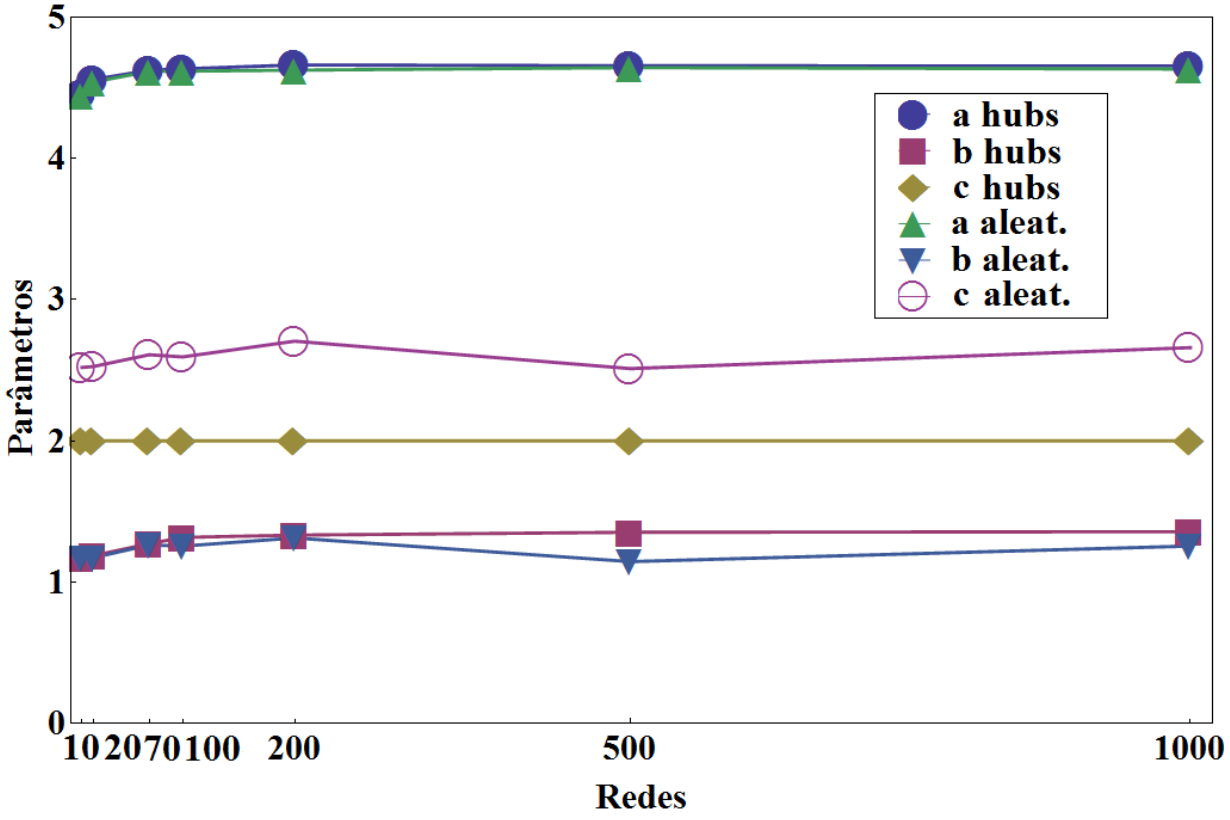

Figura 14: Variação dos parâmetros *a, b e c* nos diversos tamanhos de rede para as simulações de ataque aos *hubs* e aleatórios.

Nas Figuras 15 à 21, visualizam-se os gráficos do ataque aos *hubs* para todos os tamanhos de redes. Cada gráfico mostra o comportamento da falha em cascata disparada pelo ataque aos *hubs*. No eixo horizontal, têm-se os valores das probabilidades de espalhamento da cascata, isto é, as vulnerabilidades, e no eixo vertical, as porcentagens de sobreviventes da rede, em escala logarítmica. Os resultados são os pontos na cor azul e foram obtidos a partir das simulações no *software AttackerDefender*. Em vermelho, têm-se cada uma das funções encontradas, que mostram um comportamento muito próximo da curva original, e define matematicamente, de maneira aproximada, os resultados da falha cascata.

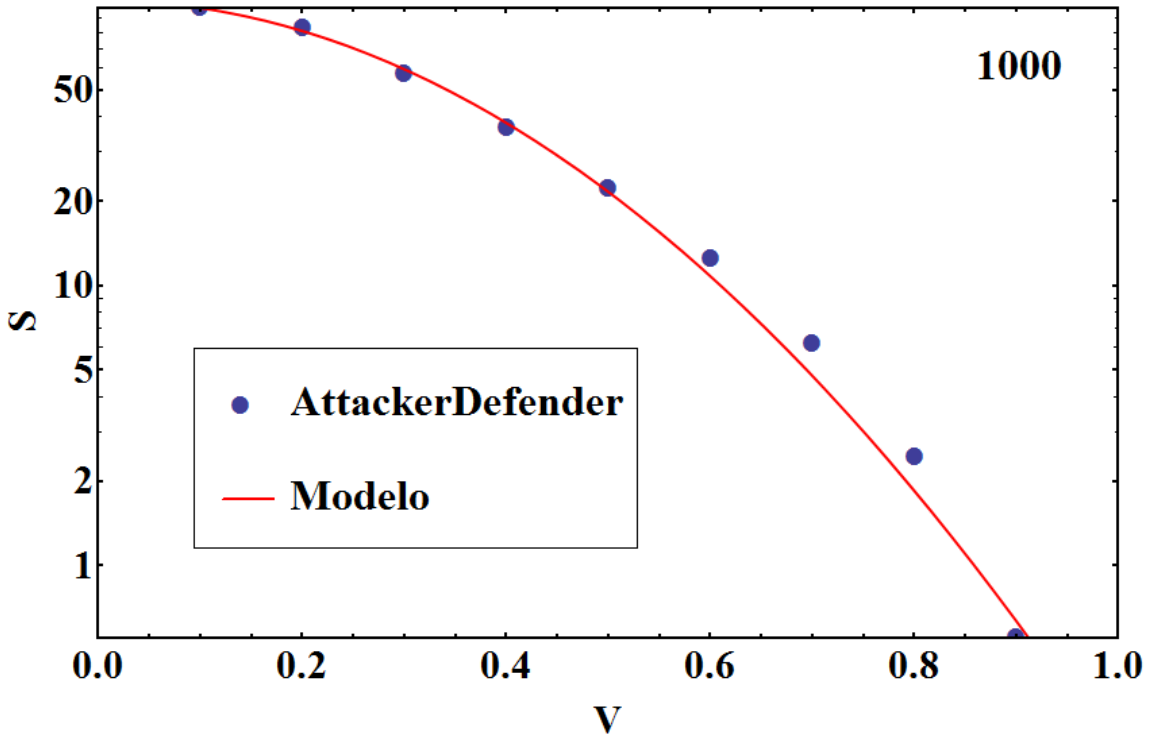

Figura 15: Sobreviventes à falha aos *hubs* na rede de mil nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

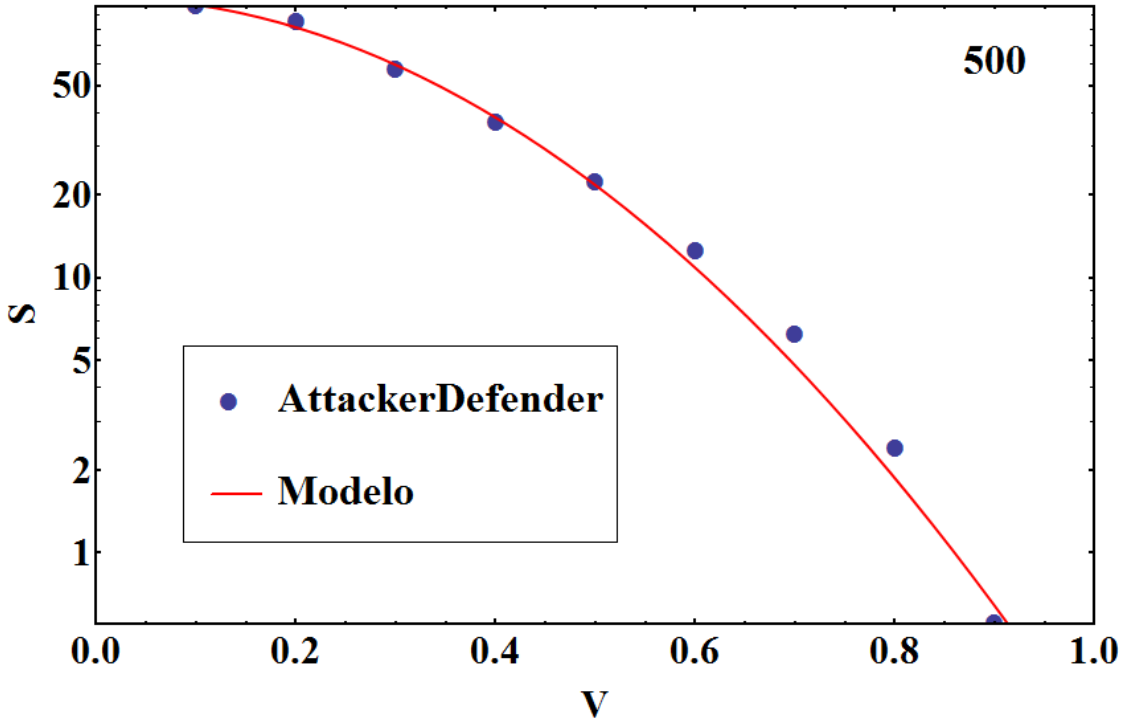

Figura 16: Sobreviventes à falha aos *hubs* na rede de quinhentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

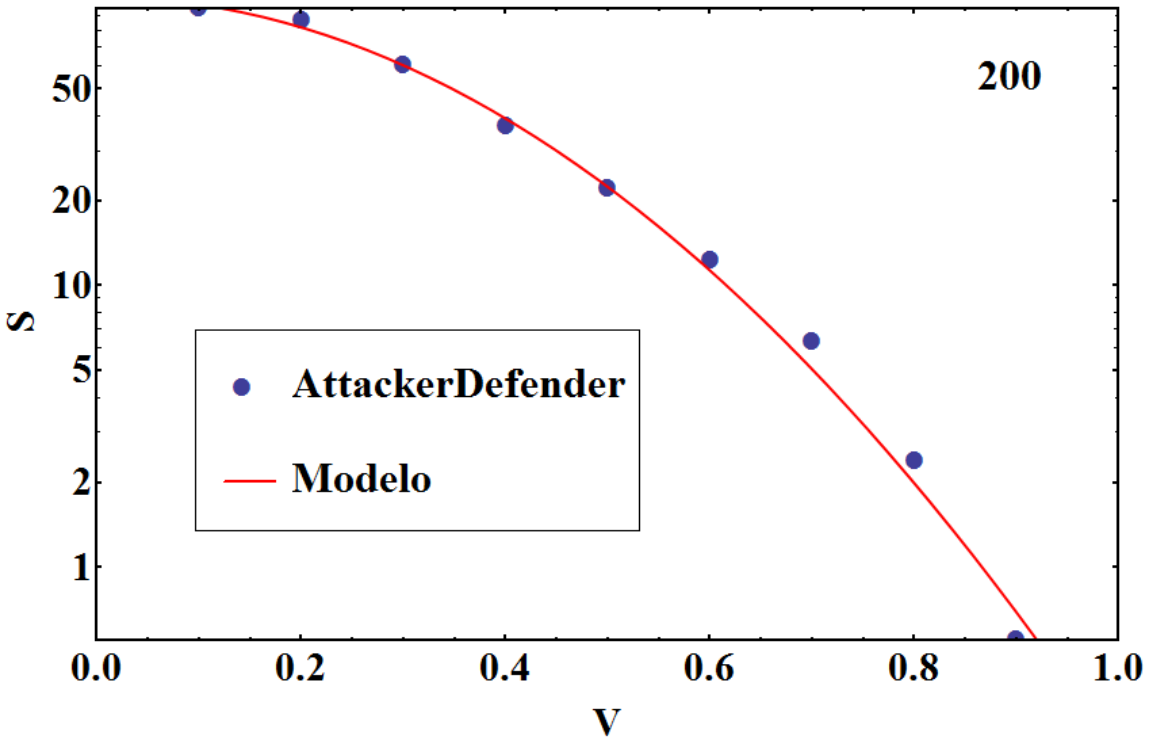

Figura 17: Sobreviventes à falha aos *hubs* na rede de duzentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

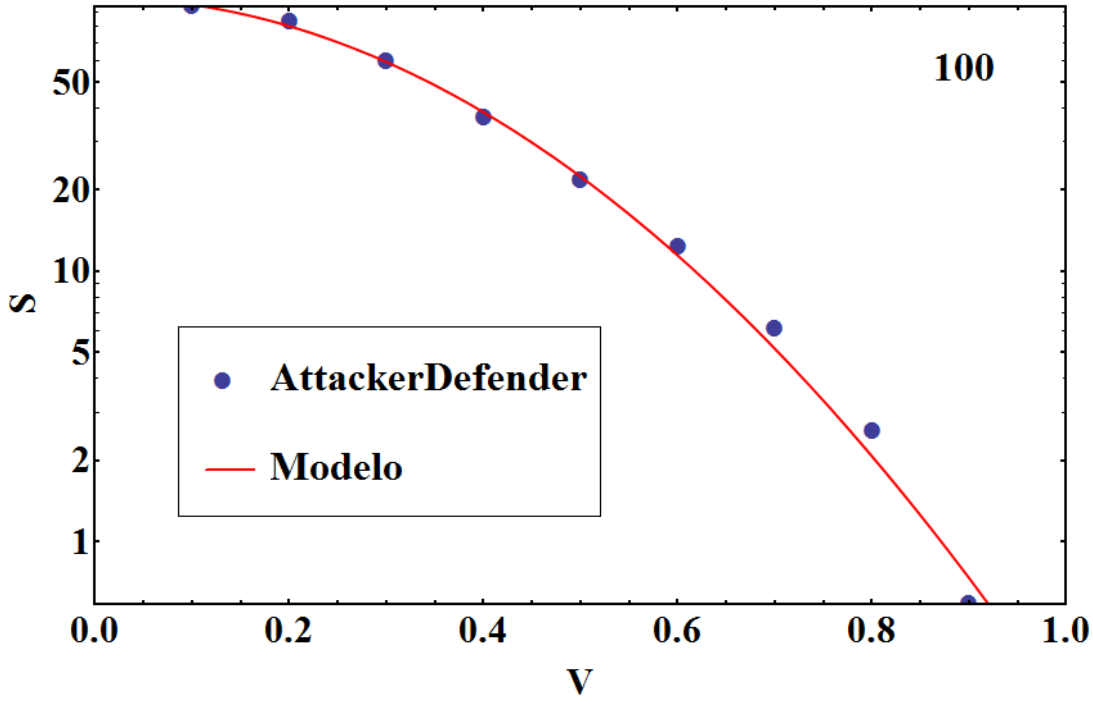

Figura 18: Sobreviventes à falha aos *hubs* na rede de cem nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

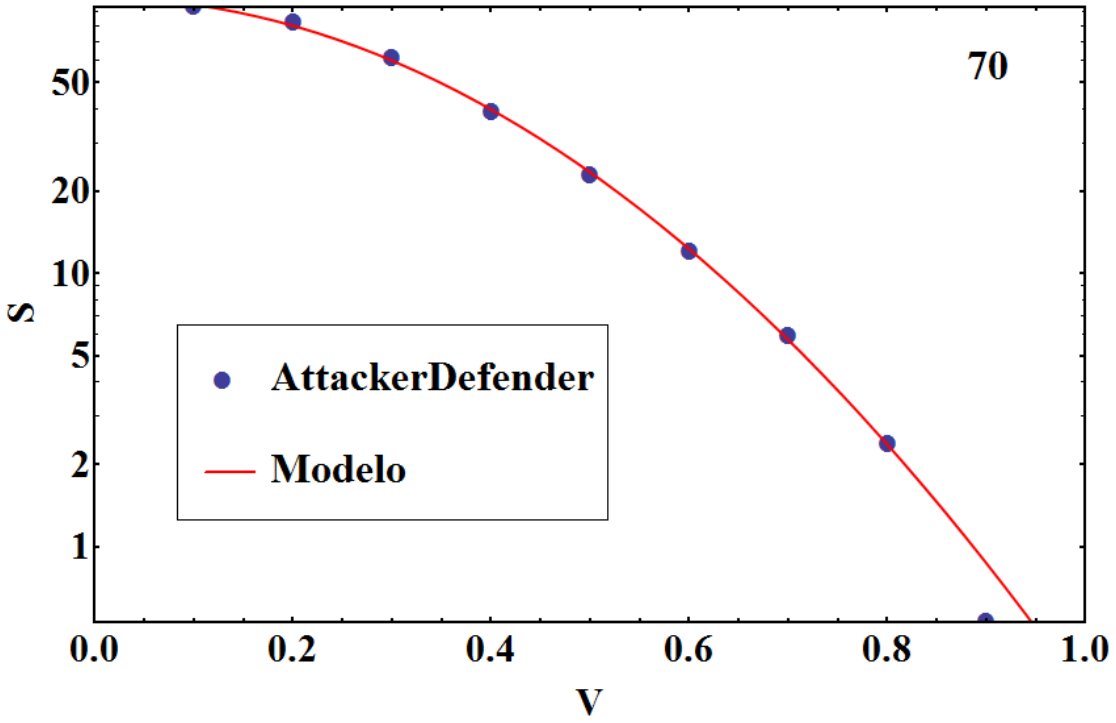

Figura 19: Sobreviventes à falha aos *hubs* na rede de setenta nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

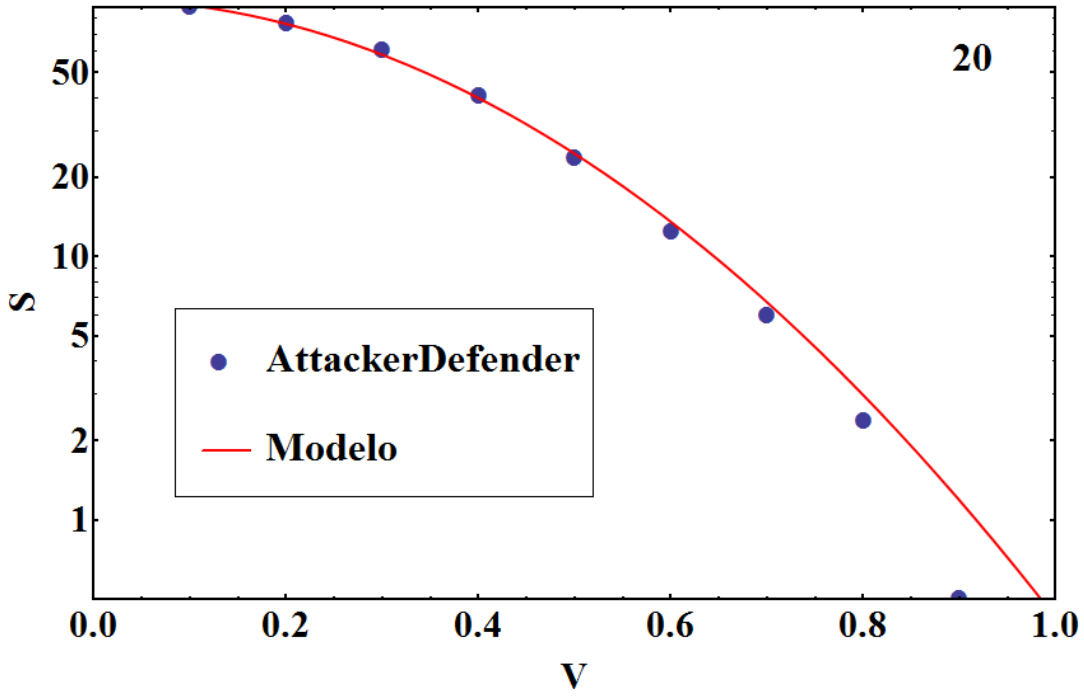

Figura 20: Sobreviventes à falha aos *hubs* na rede de vinte nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

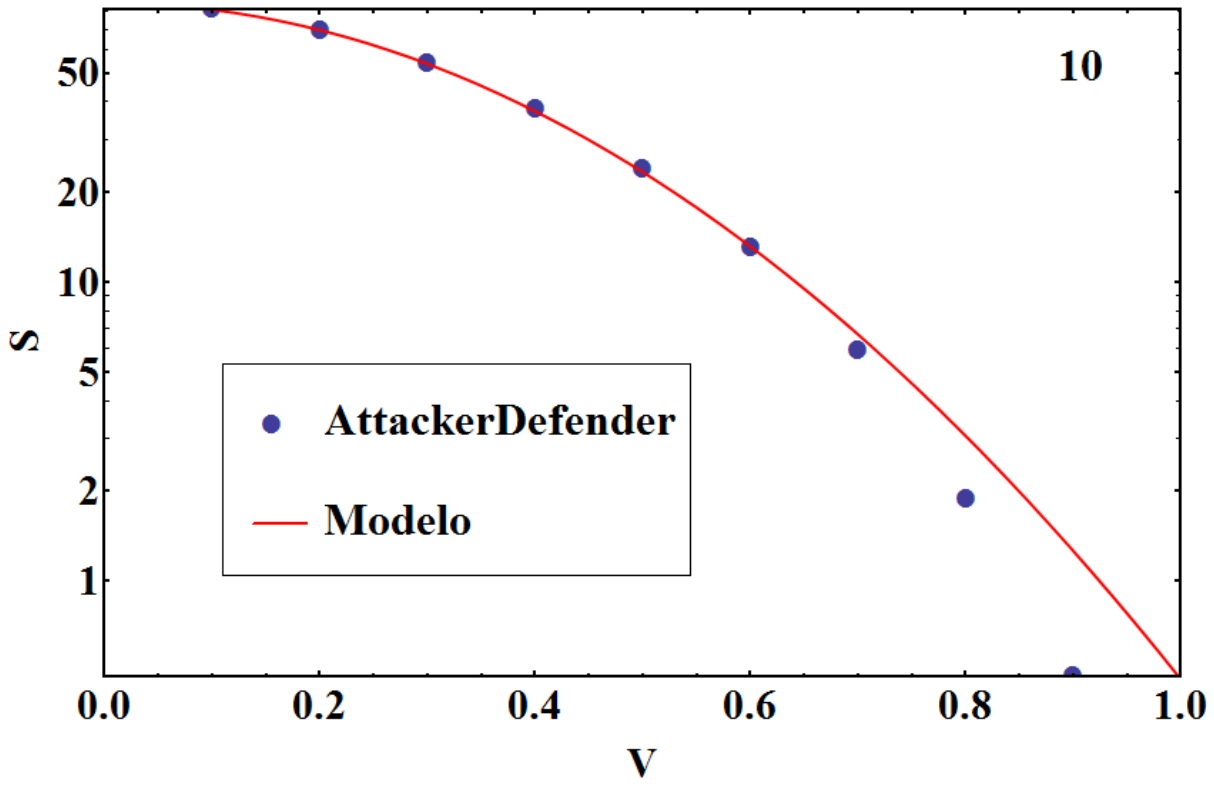

Figura 21: Sobreviventes à falha aos *hubs* na rede de dez nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

Após a adequação do novo modelo para as redes nas falhas aos *hubs* tentam-se adequar esses mesmos modelos encontrados para os casos da falha aleatória. Os gráficos são apresentados nas Figuras de 22 a 28.

Observa-se que os modelos não se aproximam adequadamente, então se refazem os cálculos para os parâmetros *a, b* e é adicionado o parâmetro *c*. A partir de valores específicos para as curvas das falhas aleatórias, aí então se encontram os modelos ideais, nas Figuras 29 a 35.

No mesmo gráfico de cada rede, desenham-se a curva da falha original (em cor azul) e a curva da função exponencial encontrada para o respectivo tamanho de rede (em cor vermelha).

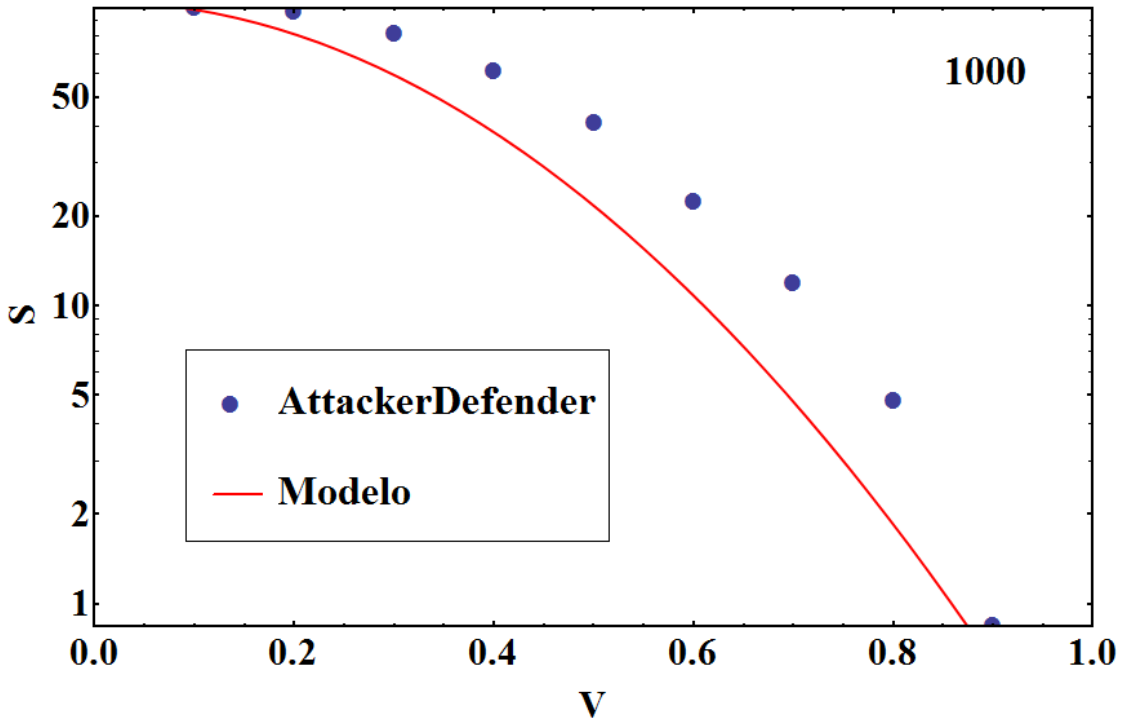

Figura 22: Sobreviventes à falha aleatória na rede de mil nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

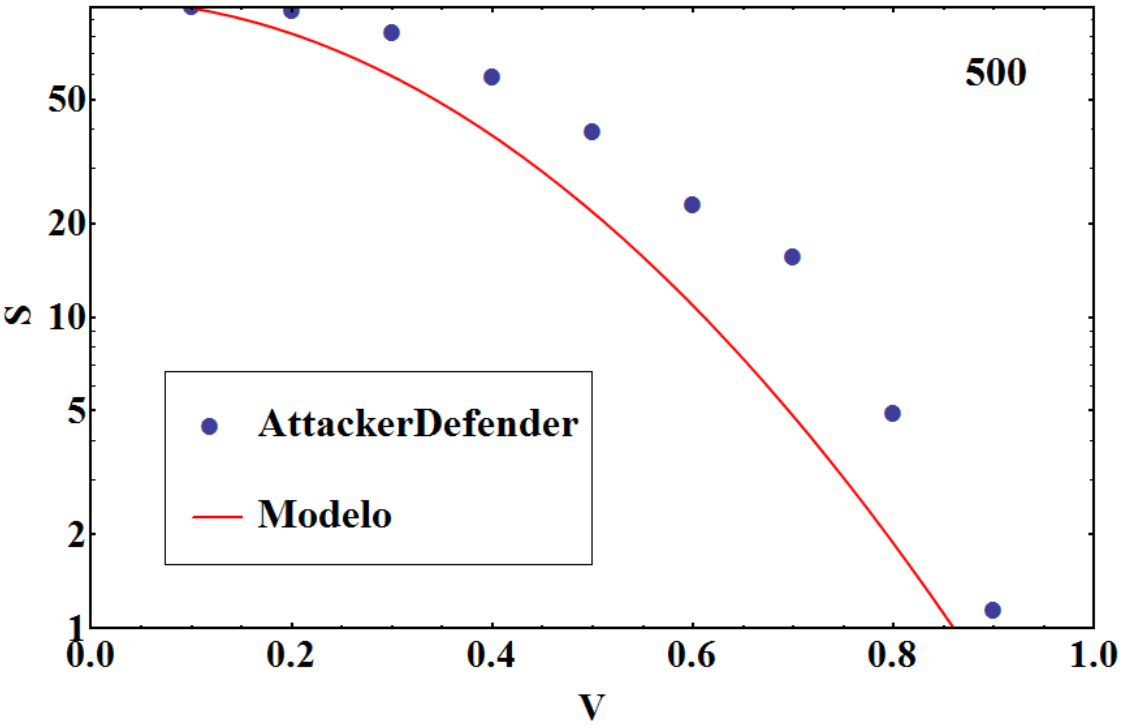

Figura 23: Sobreviventes à falha aleatória na rede de quinhentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

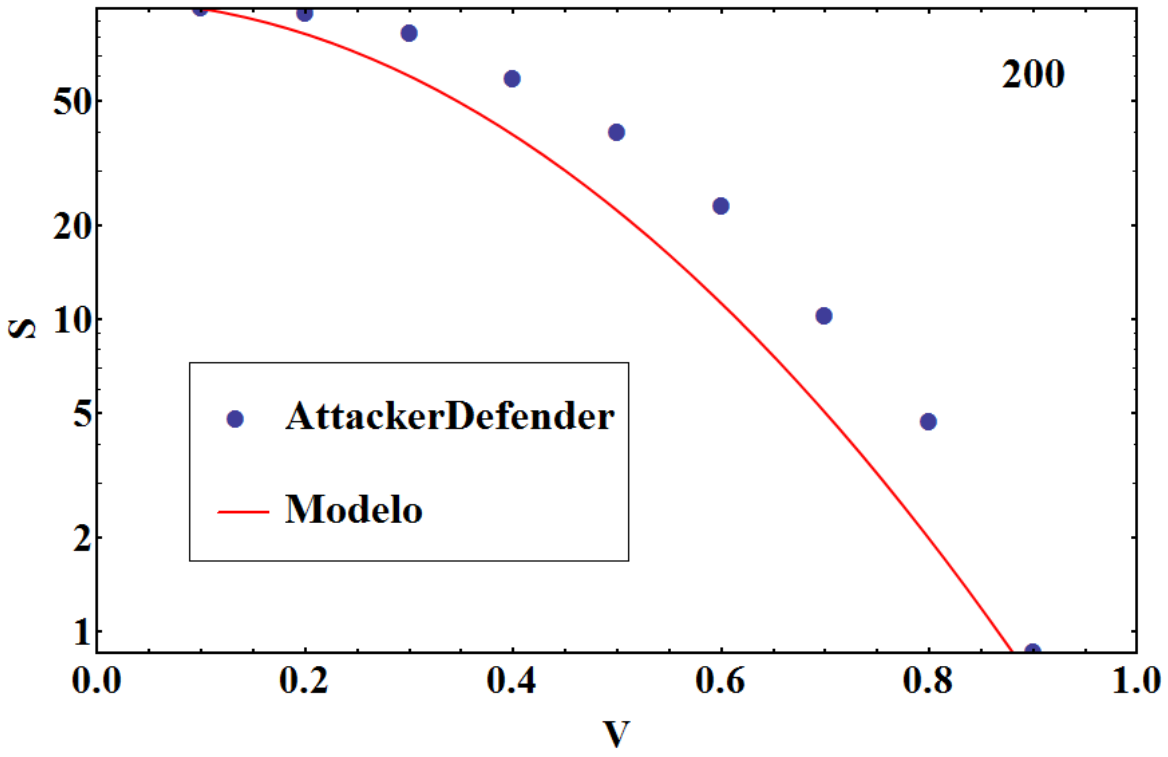

Figura 24: Sobreviventes à falha aleatória na rede de duzentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

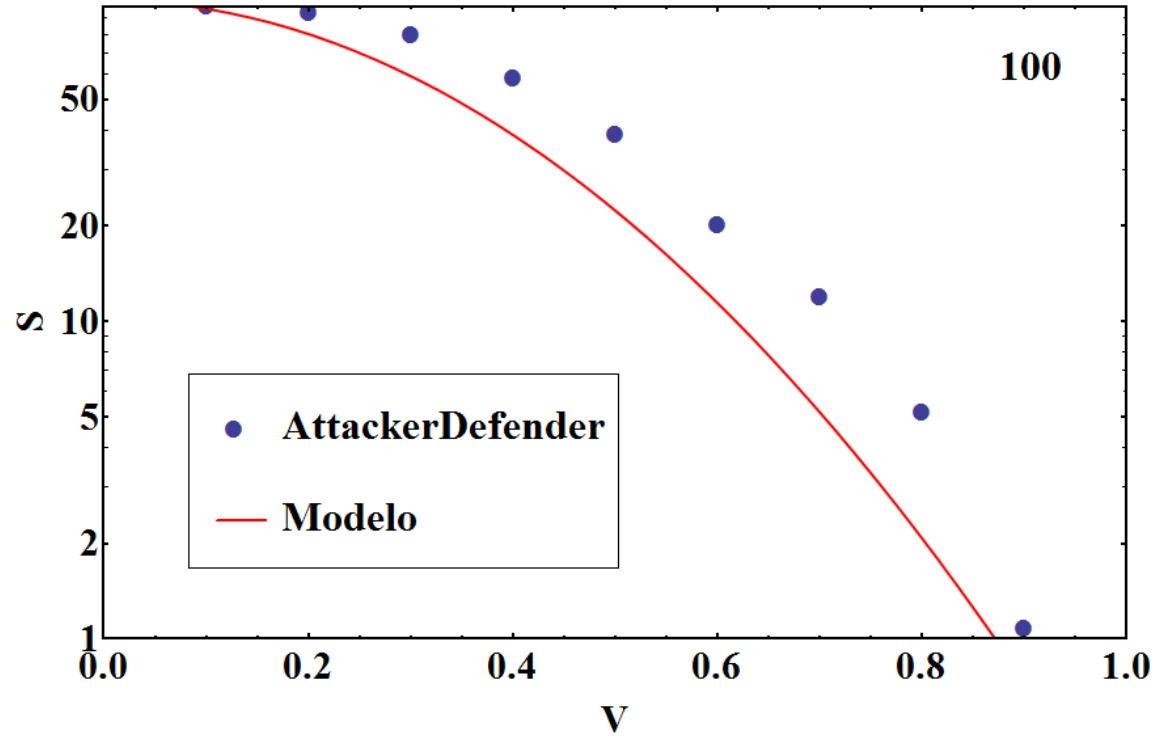

Figura 25: Sobreviventes à falha aleatória na rede de cem nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

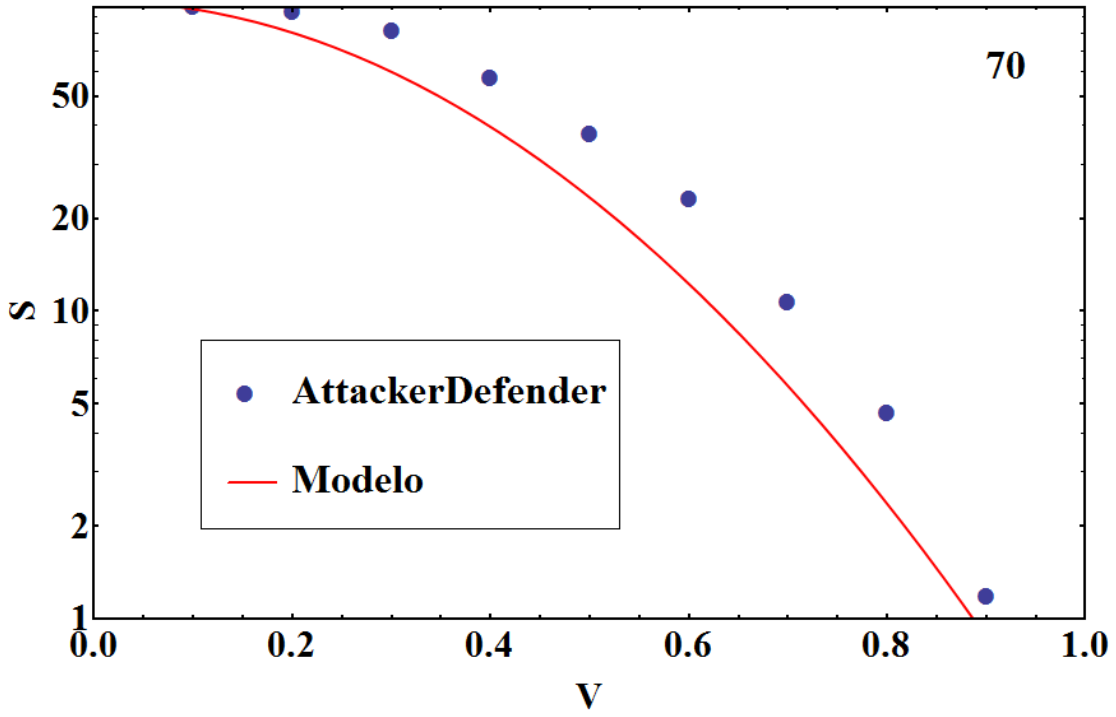

Figura 26: Sobreviventes à falha aleatória na rede de setenta nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

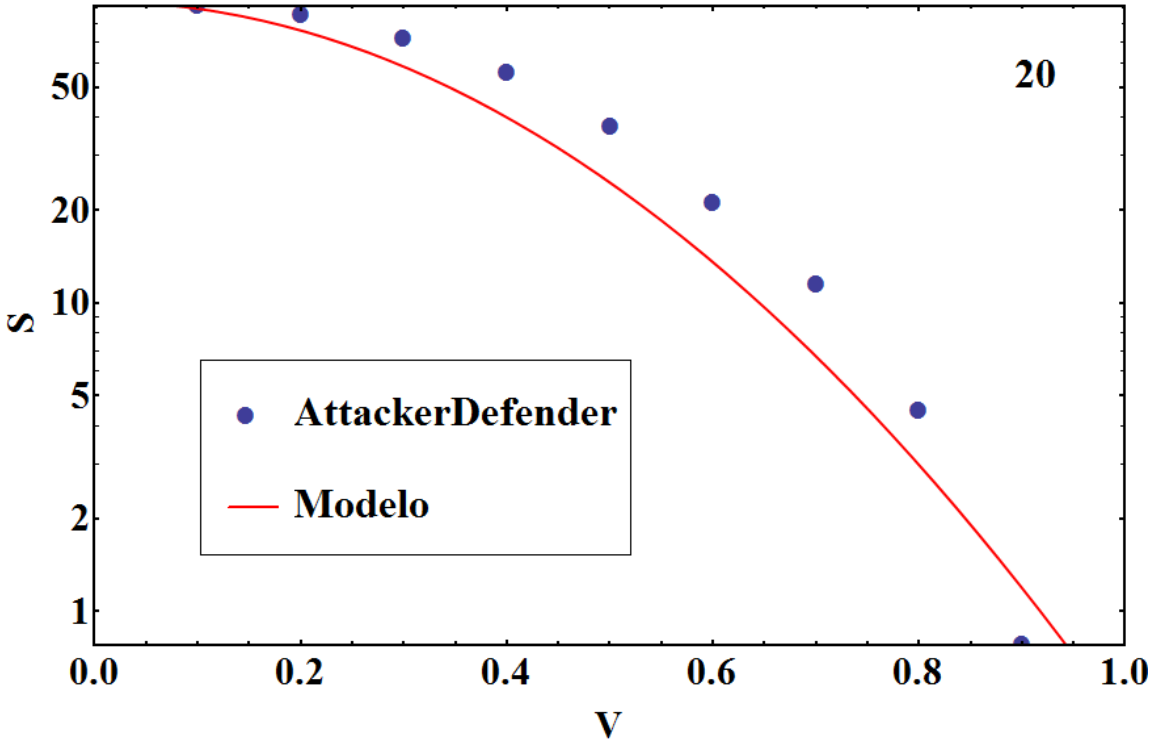

Figura 27: Sobreviventes à falha aleatória na rede de vinte nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

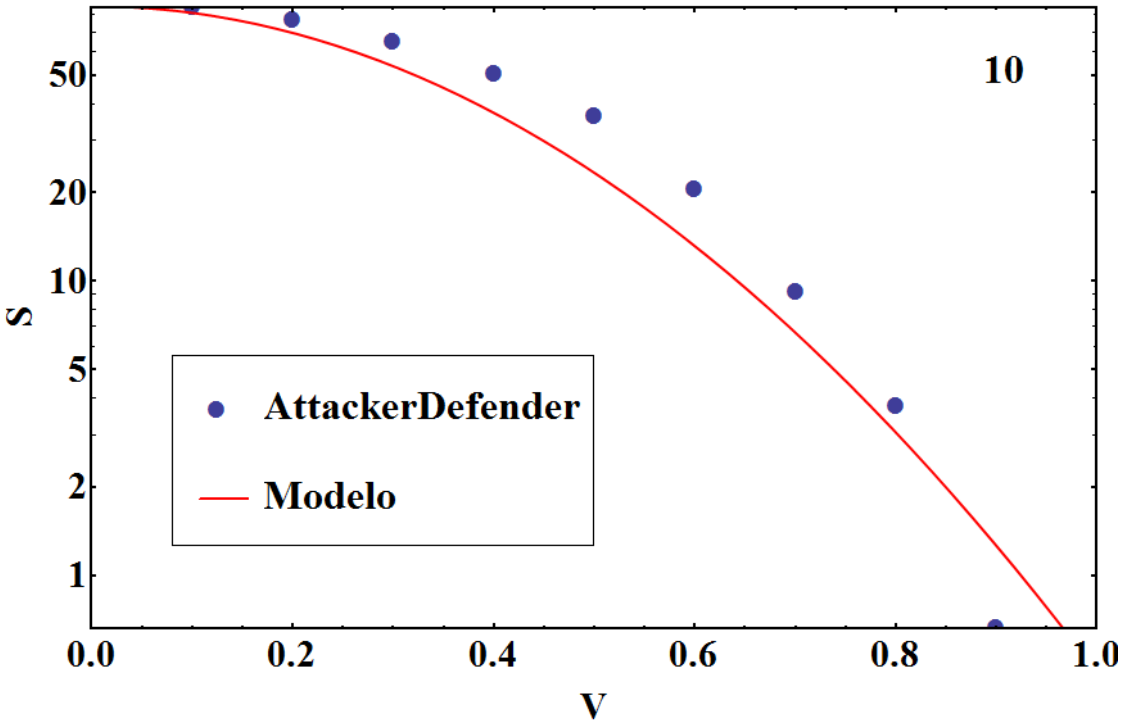

Figura 28: Sobreviventes à falha aleatória na rede de dez nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

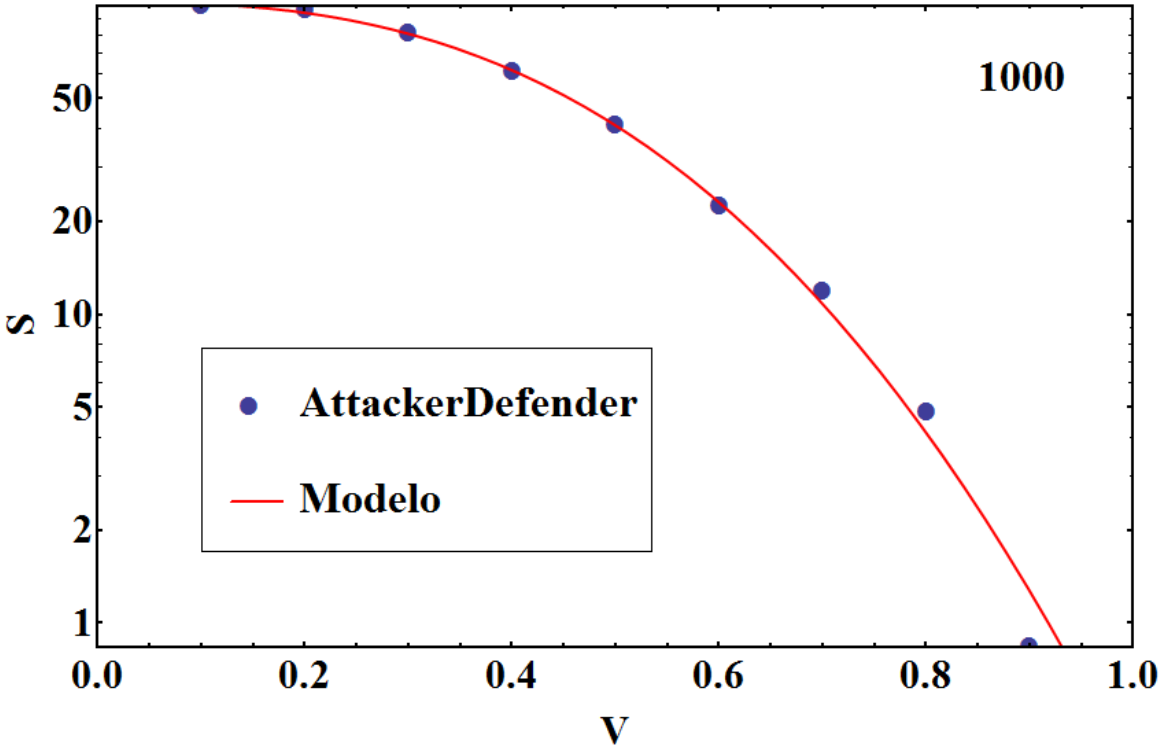

Figura 29: Sobreviventes à falha aleatória na rede de mil nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

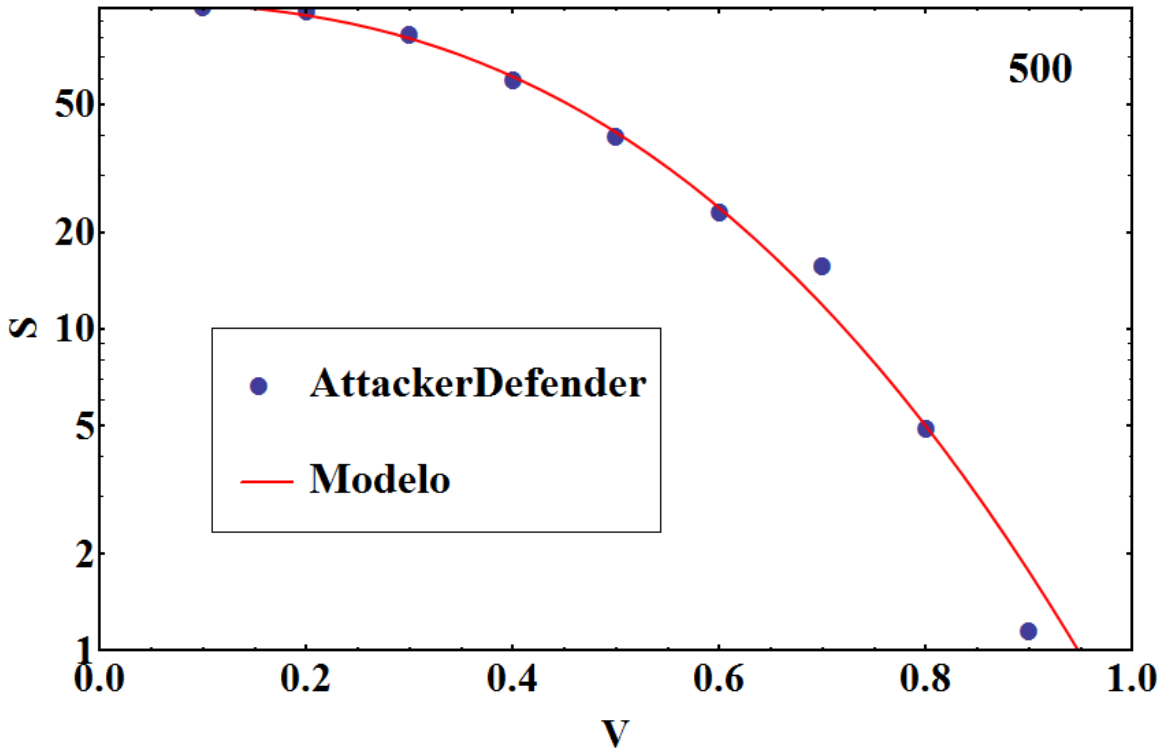

Figura 30: Sobreviventes à falha aleatória na rede de quinhentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

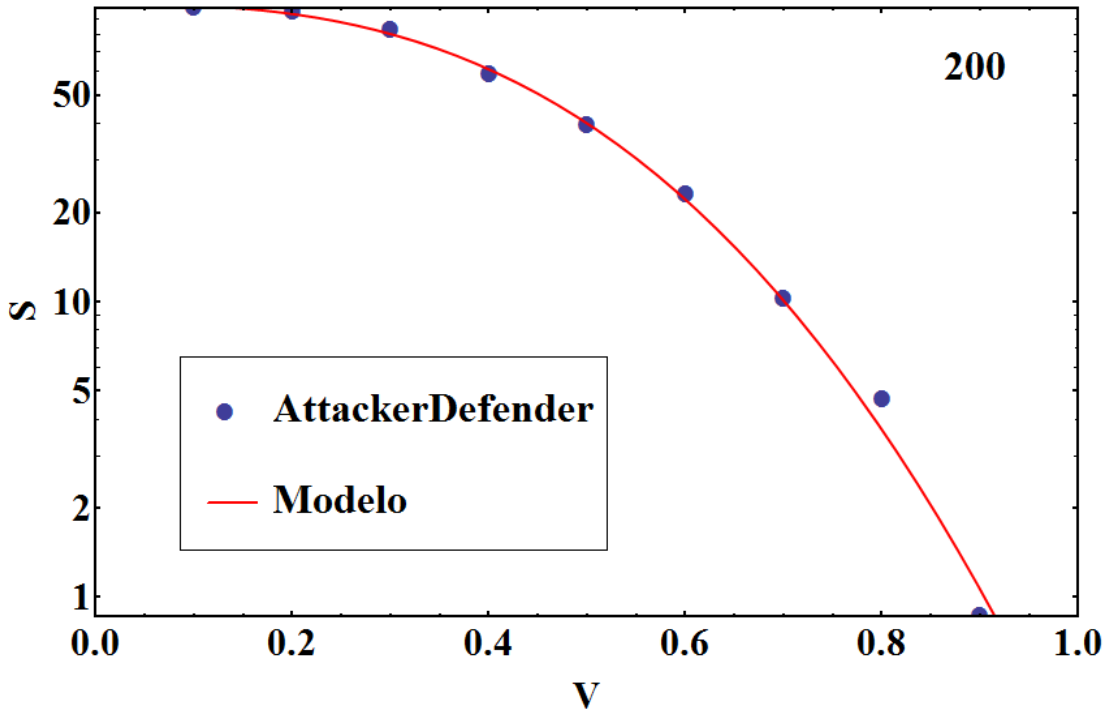

Figura 31: Sobreviventes à falha aleatória na rede de duzentos nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

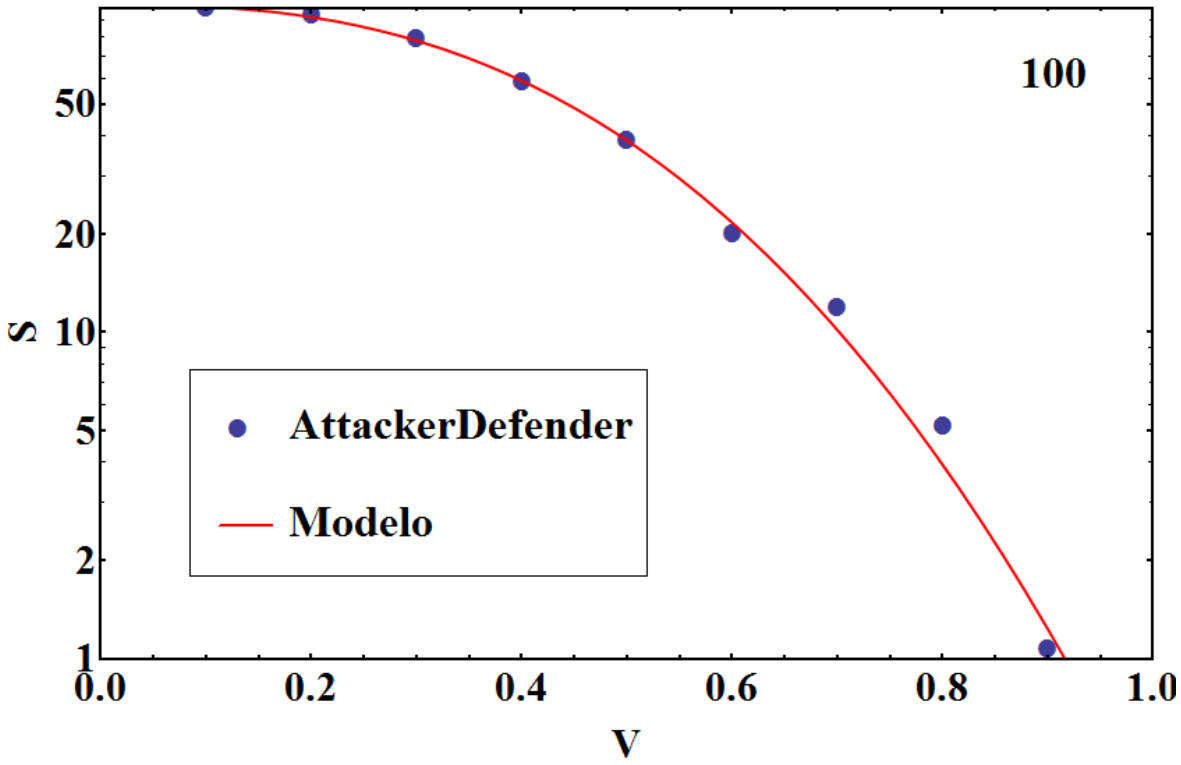

Figura 32: Sobreviventes à falha aleatória na rede de cem nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

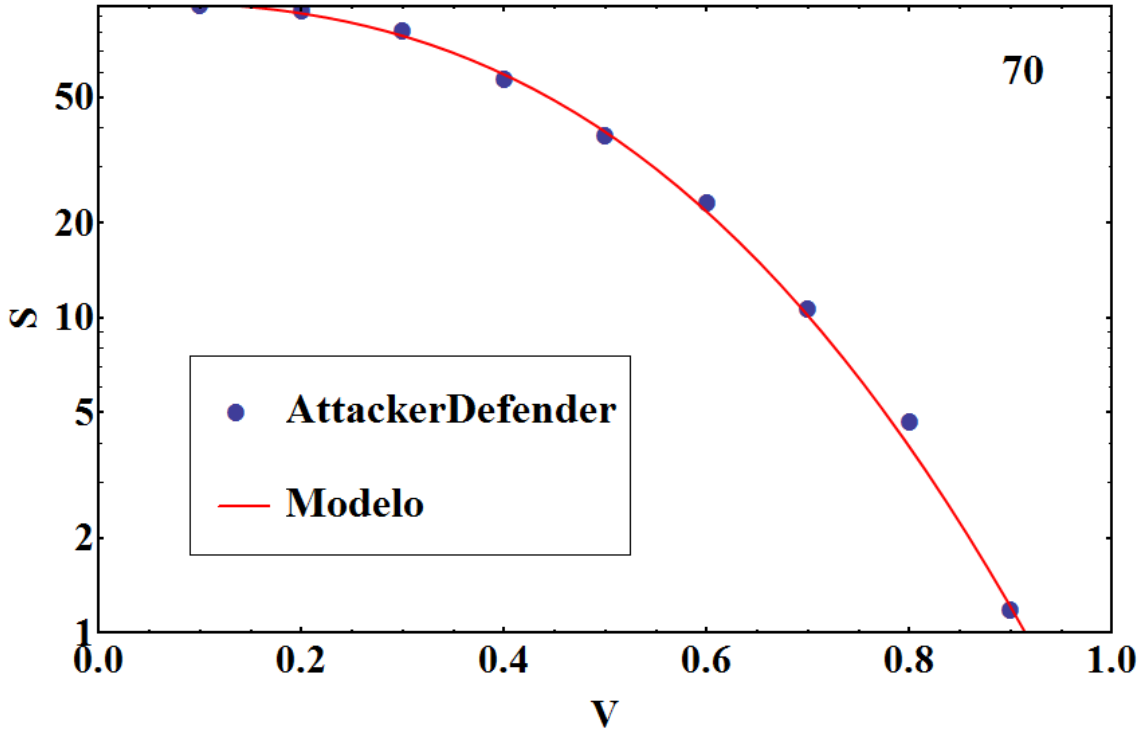

Figura 33: Sobreviventes à falha aleatória na rede de setenta nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

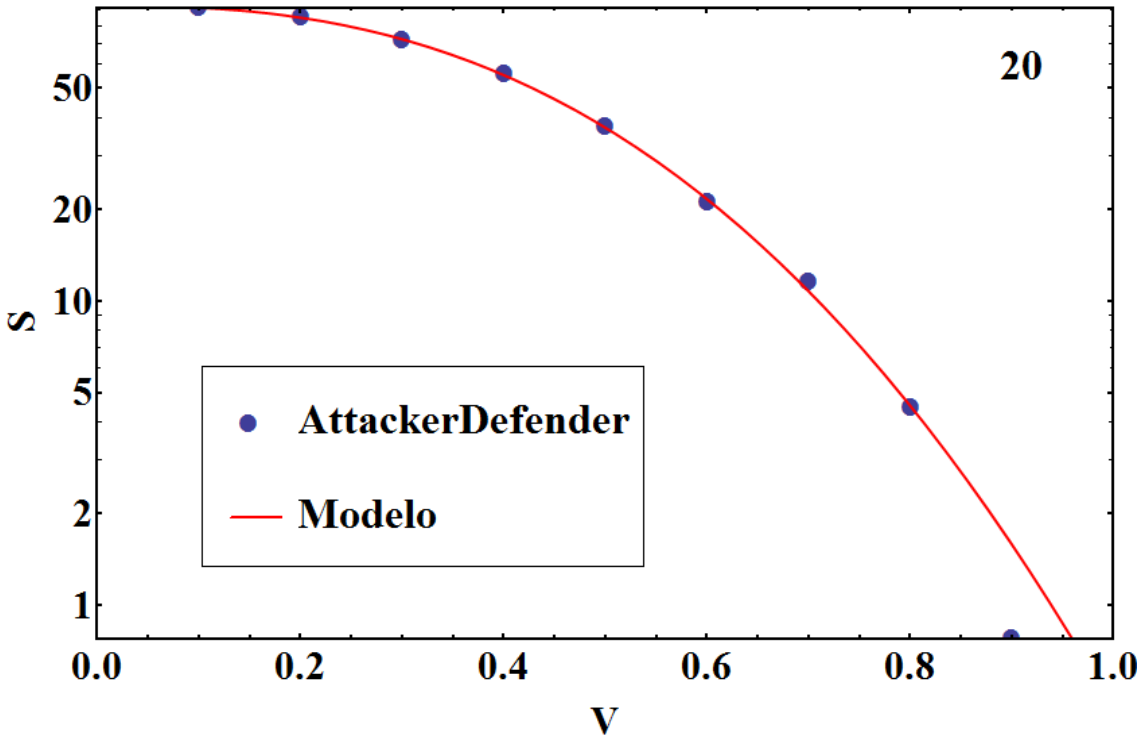

Figura 34: Sobreviventes à falha aleatória na rede de vinte nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

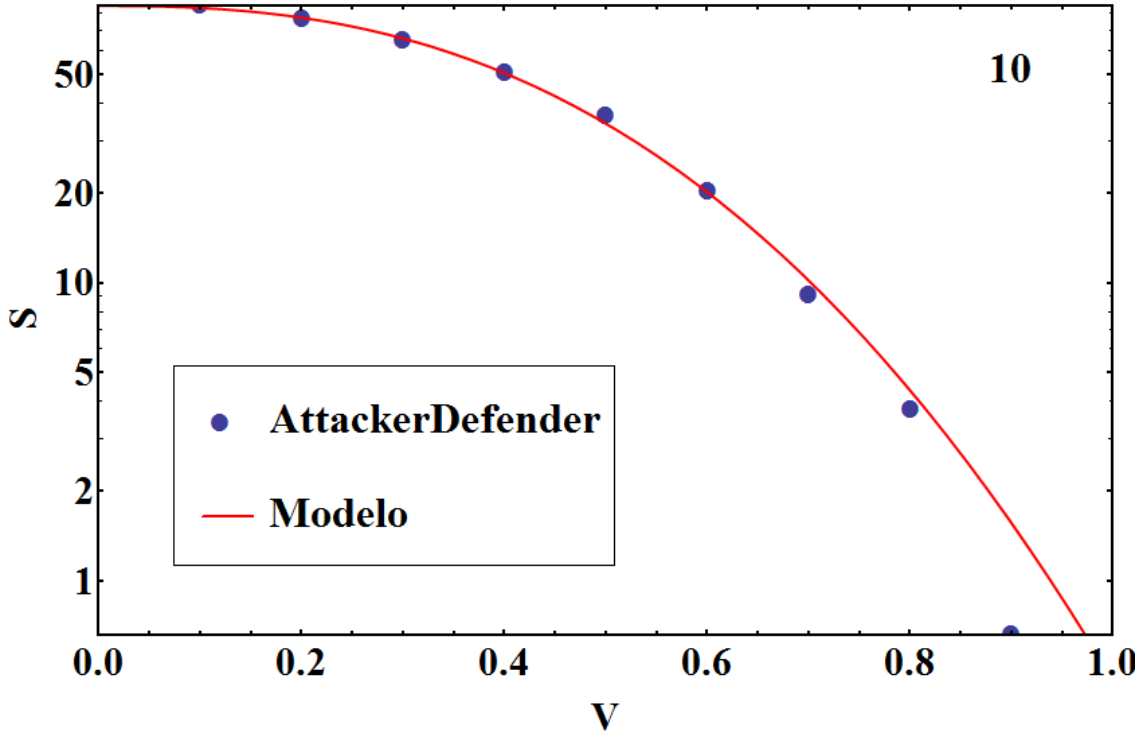

Figura 35: Sobreviventes à falha aleatória, na rede de dez nós, resultado do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

Além dos dados relativos às porcentagens de sobreviventes nas redes *Scale-Free*, também se coletam as porcentagens de atingidos pela cascata em cada uma das redes, nos dois cenários de falhas. Tais valores são graficamente representados e também se aplicam os mesmos passos para adequação das curvas às possíveis funções de definição do comportamento destas. Os gráficos também são elaborados no *software Wolfram Mathematica*.

A Tabela 11, a partir dos dados dos atingidos pela cascata em cada rede, relaciona os valores das variáveis substituíveis na função exponencial citada que descreve a curva apresentada nos gráficos anteriores em escala logarítmica.

Tabela 11: Variáveis e funções que descrevem o comportamento das simulações dos nós atingidos.

Atingidos

Simulações utilizadas	Nós	Ataque	a	b	c	F(v), v = [0,1;1,0]
81ª a 90ª Simulação	10	Hub	4,45141	1,17071	-	$100 - e^{a(1-bv^2)}$
101ª a 110ª Simulação	20		4,55363	1,18756		
121ª a 130ª Simulação	70		4,62180	1,27201		
141ª a 150ª Simulação	100		4,63104	1,31698		
161ª a 170ª Simulação	200		4,65923	1,33271		
181ª a 190ª Simulação	500		4,65369	1,35289		
201ª a 210ª Simulação	1000		4,65175	1,35710		
91ª a 100ª Simulação	10	Aleatório	4,44183	1,17263	2,51885	$100 - e^{a(1-bv^c)}$
111ª a 120ª Simulação	20		4,53832	1,17244	2,52544	
131ª a 140ª Simulação	70		4,61326	1,26296	2,60895	
161ª a 160ª Simulação	100		4,61499	1,25503	2,59426	
171ª a 180ª Simulação	200		4,62226	1,31230	2,70485	
191ª a 200ª Simulação	500		4,64025	1,14506	2,51106	
211ª a 220ª Simulação	1000		4,63083	1,25527	2,65949	

As Figuras 36 a 42 mostram a distribuição de sobreviventes quando da falha aos *hubs* e o modelo encontrado para cada tamanho de rede. As Figuras 43 a 49 mostram os modelos adequados às redes para a falha aleatória.

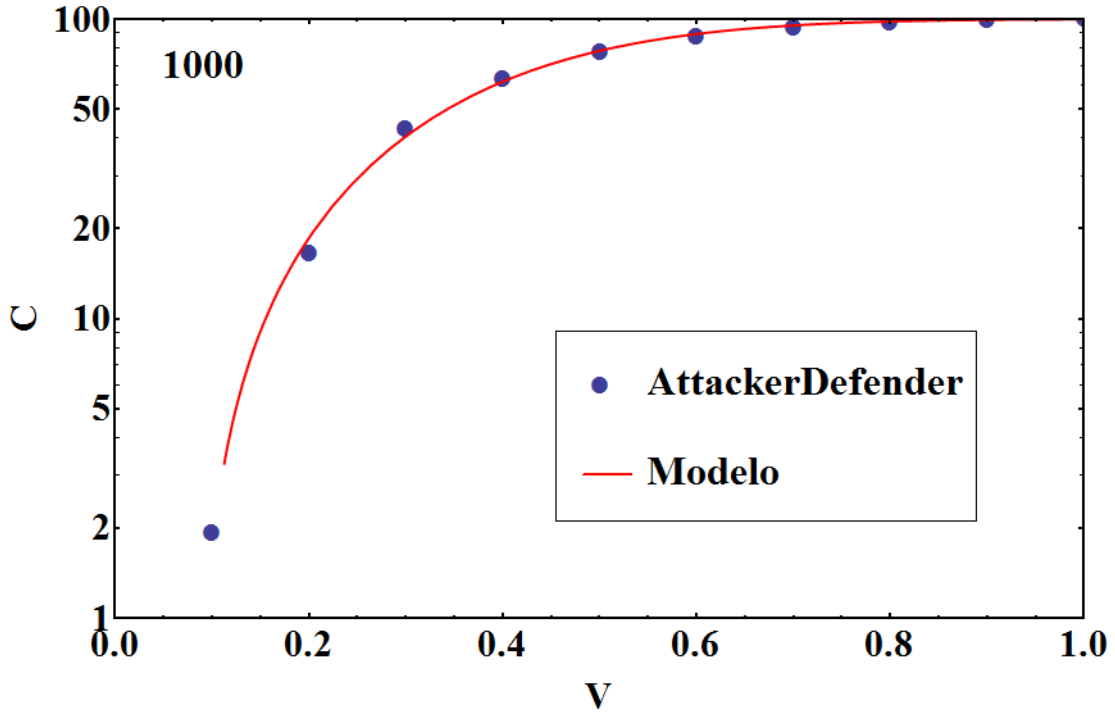

Figura 36: Atingidos na falha aos *hubs* na rede de mil nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

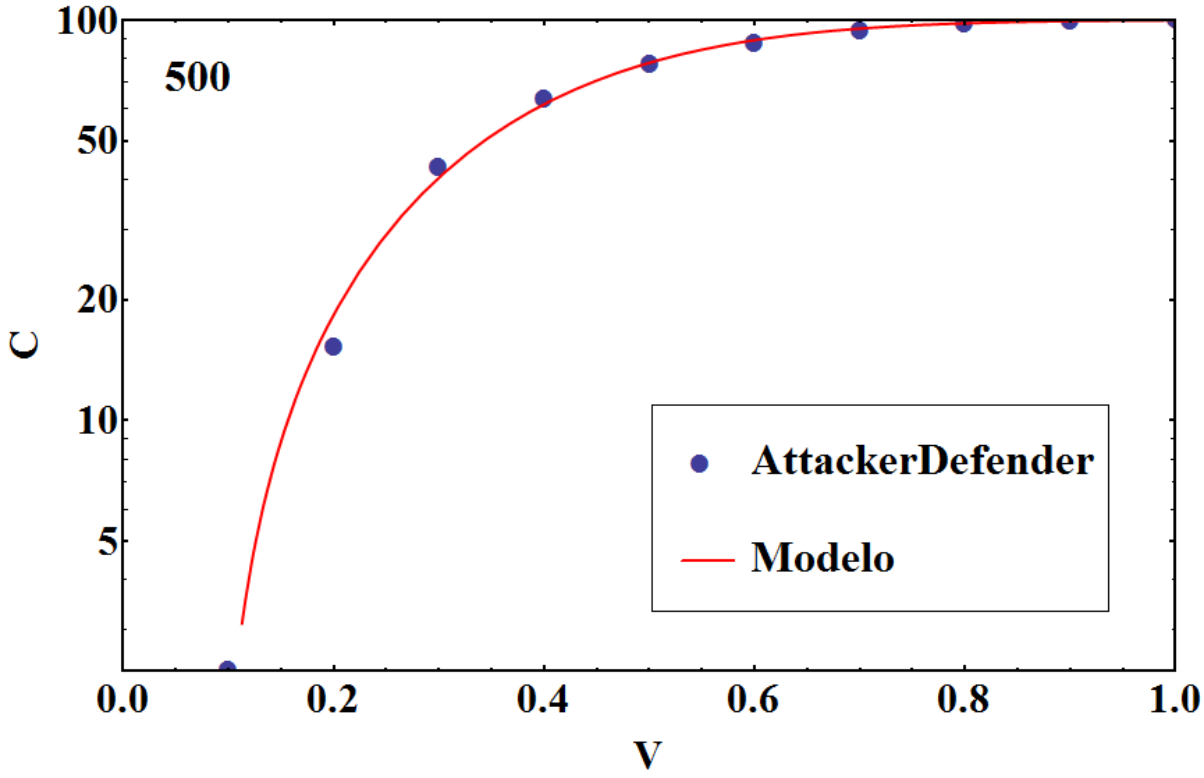

Figura 37: Atingidos na falha aos *hubs* na rede de quinhentos nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

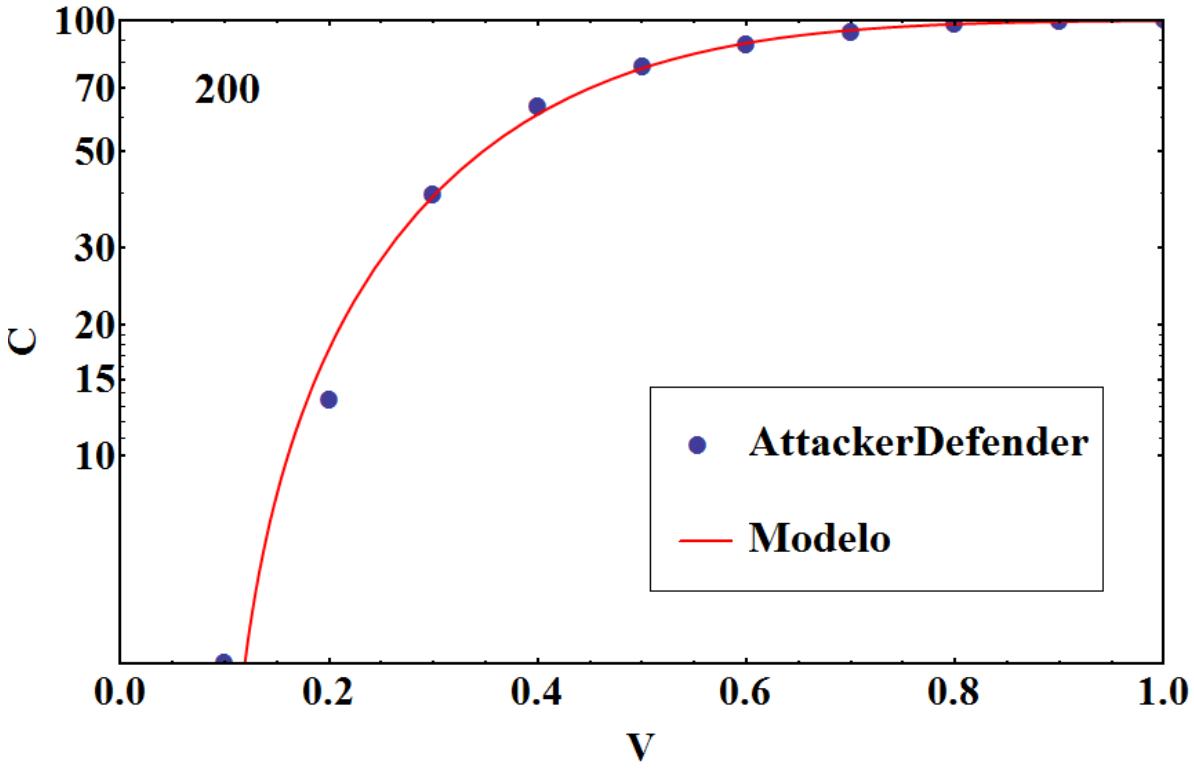

Figura 38: Atingidos na falha aos *hubs* na rede de duzentos nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

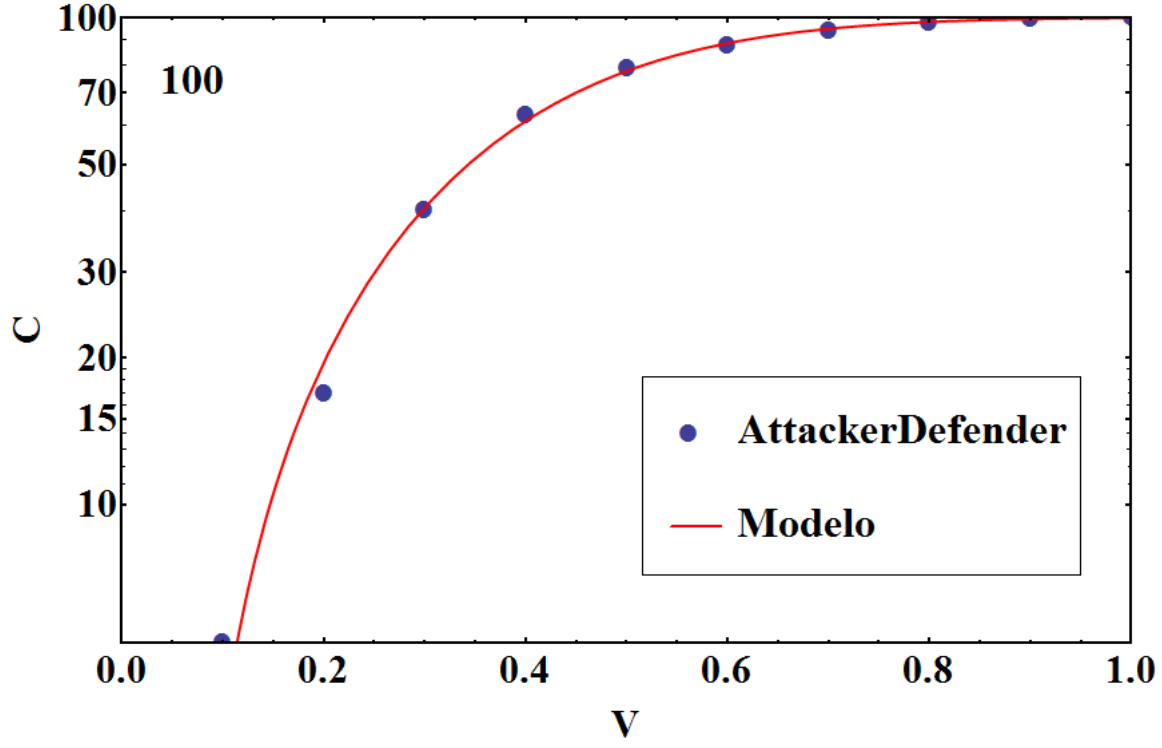

Figura 39: Atingidos na falha aos *hubs* na rede de cem nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

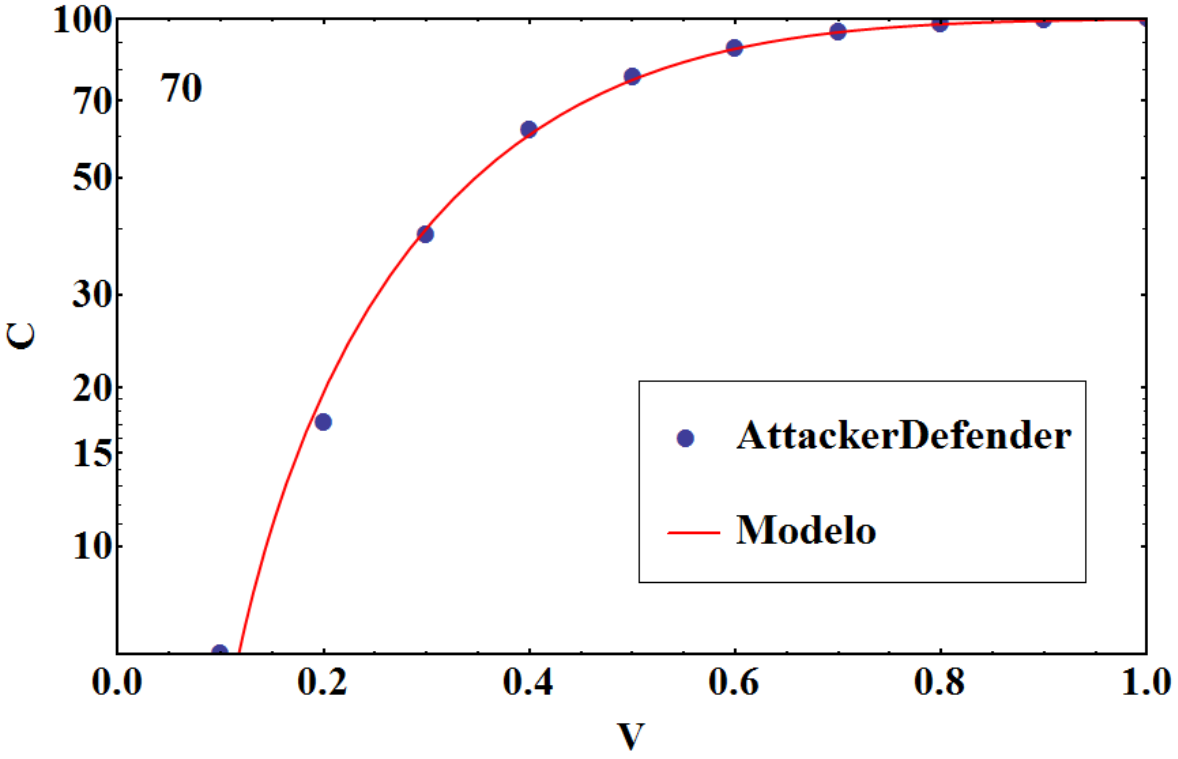

Figura 40: Atingidos na falha aos *hubs* na rede de setenta nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

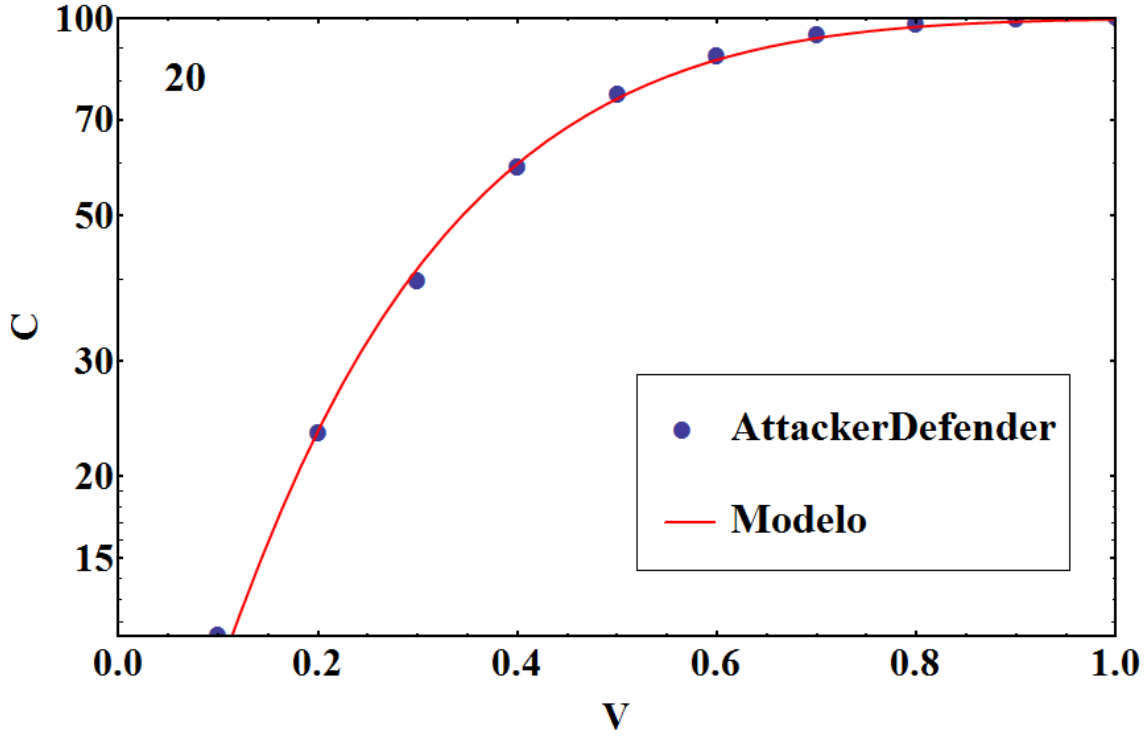

Figura 41: Atingidos na falha aos *hubs* na rede de vinte nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

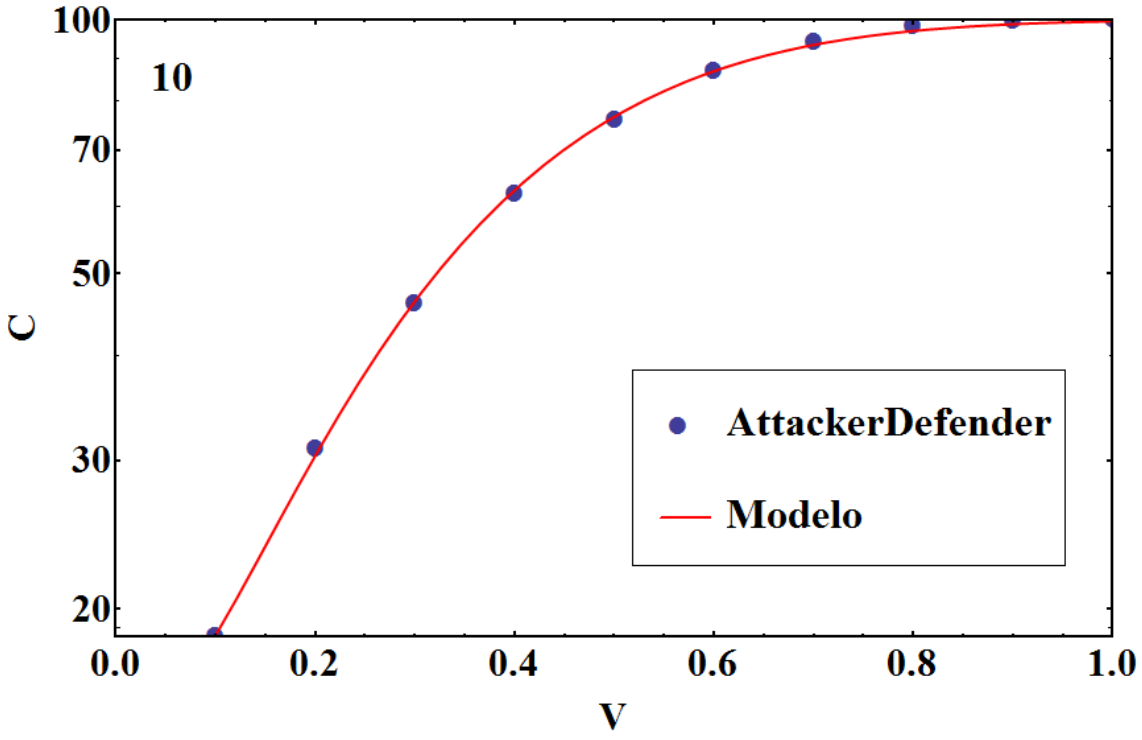

Figura 42: Atingidos na falha aos *hubs* na rede de dez nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

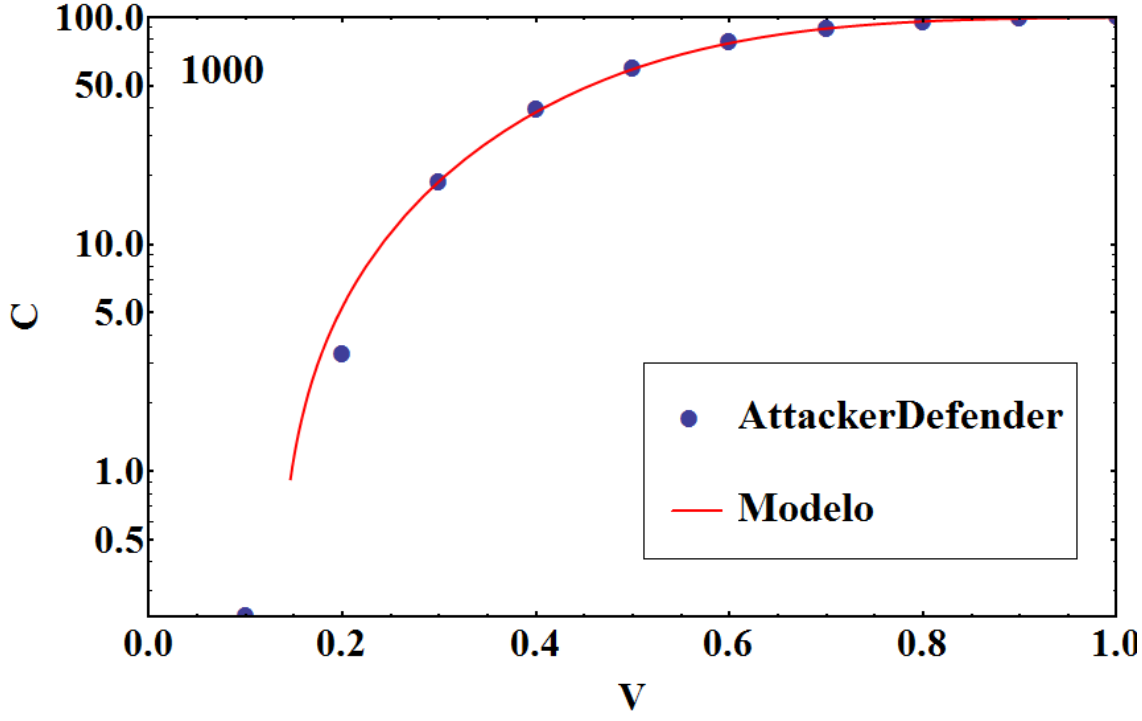

Figura 43: Atingidos na falha aleatória na rede de mil nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

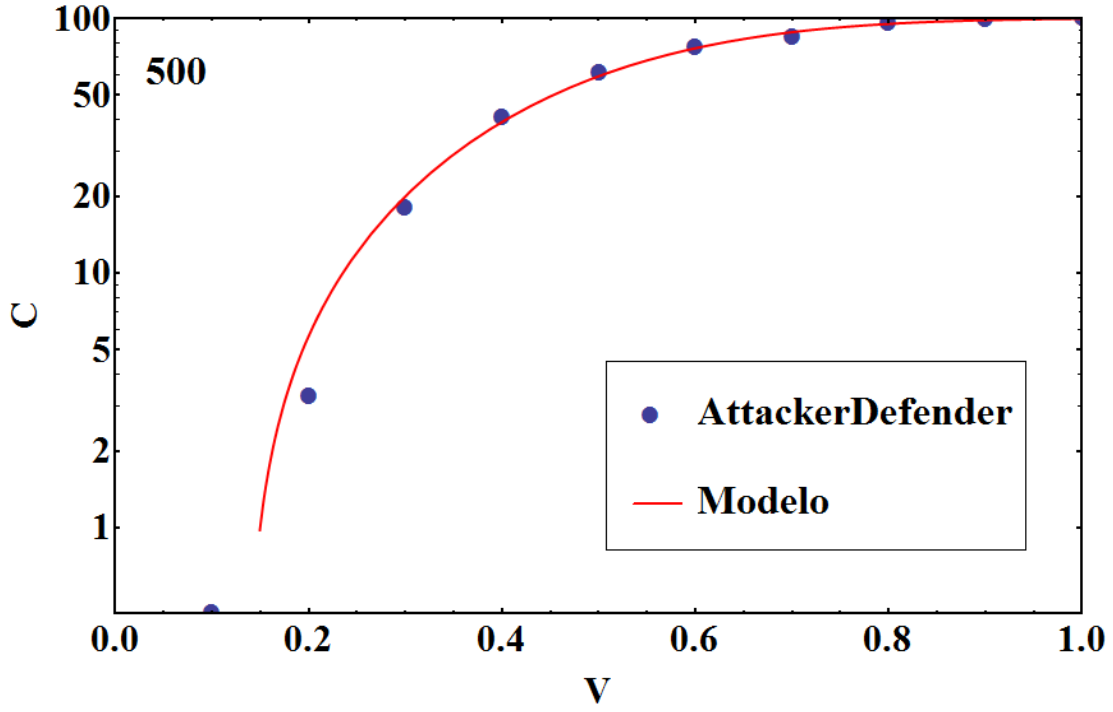

Figura 44: Atingidos na falha aleatória na rede de quinhentos nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

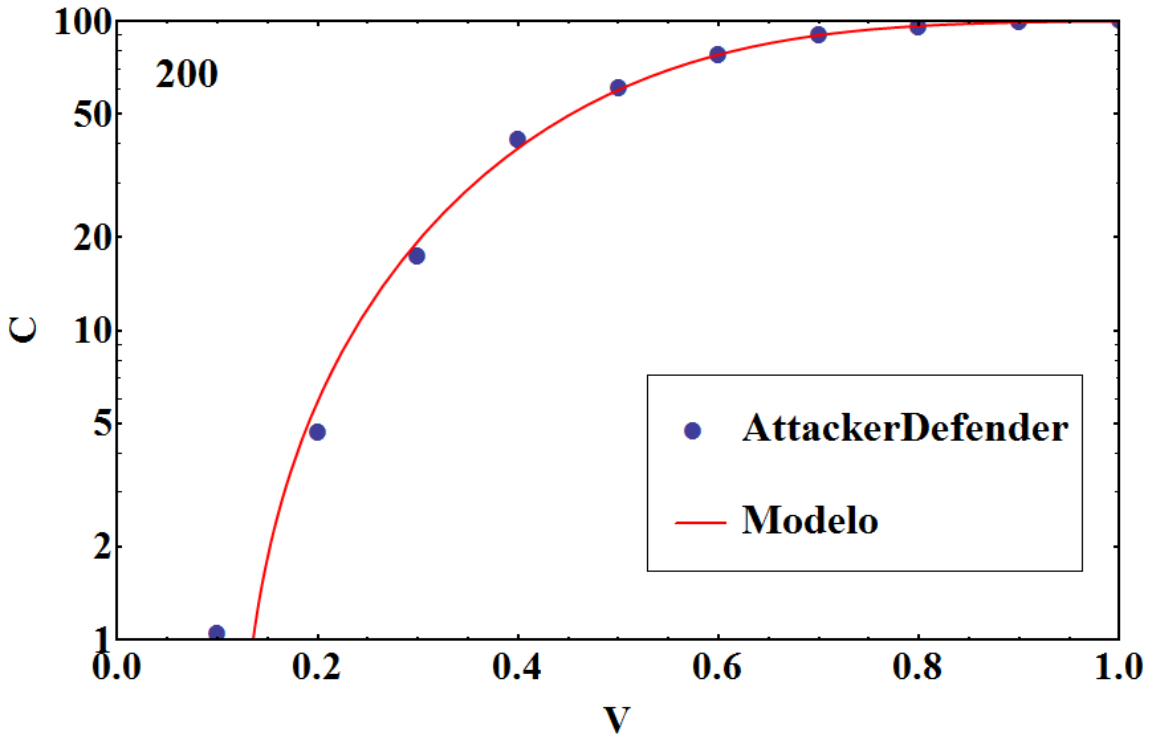

Figura 45: Atingidos na falha aleatória na rede de duzentos nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

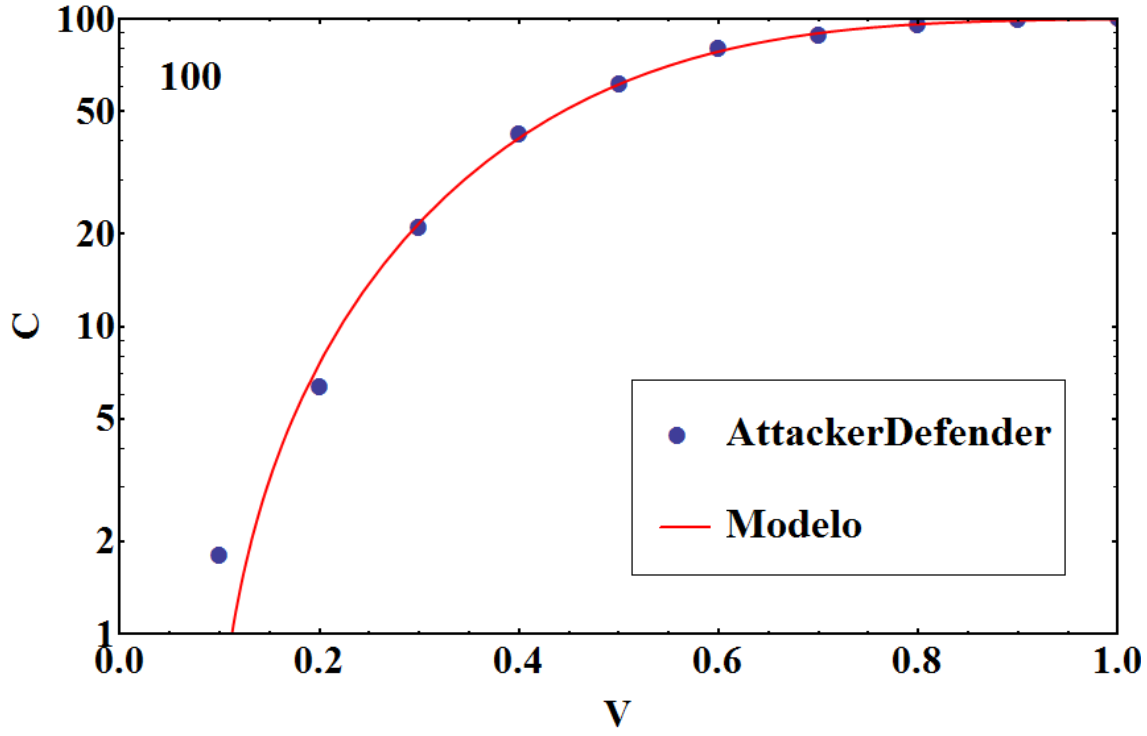

Figura 46: Atingidos na falha aleatória na rede de cem nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

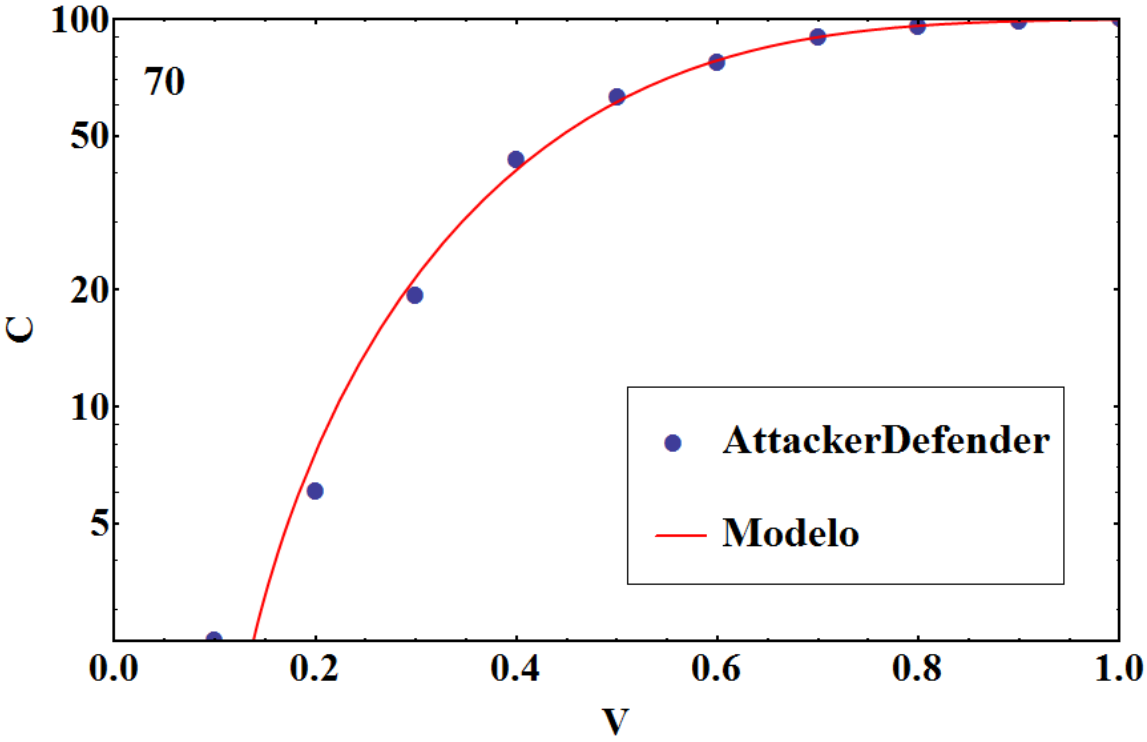

Figura 47: Atingidos na falha aleatória na rede de setenta nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

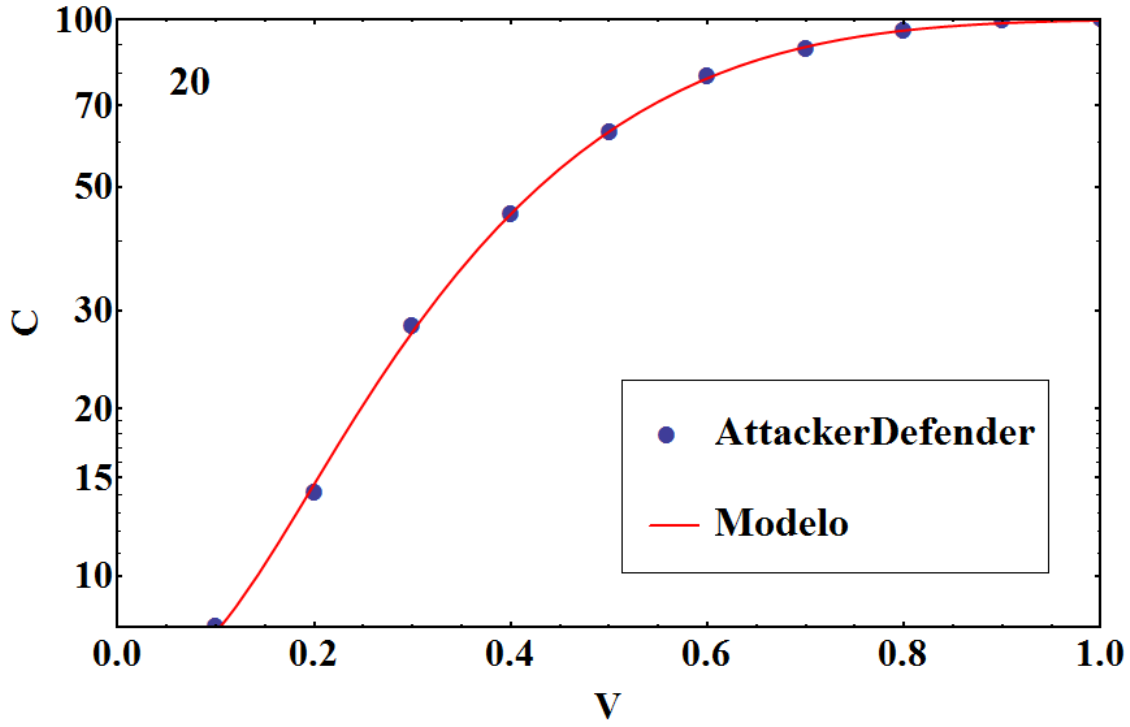

Figura 48: Atingidos na falha aleatória na rede de vinte nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

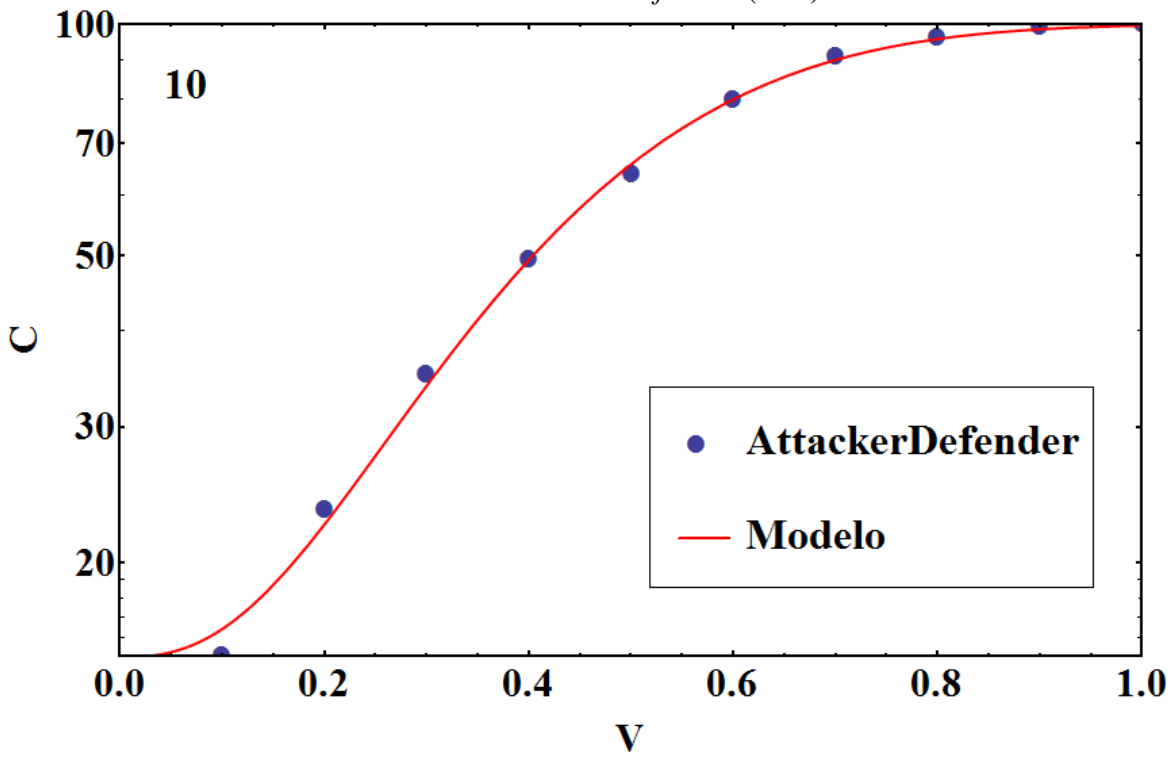

Figura 49: Atingidos na falha aleatória na rede de dez nós, curva do novo modelo (vermelho) e do software *Attacker-Defender* (azul).

6 CONCLUSÕES

Dentre outras contribuições, este trabalho provê o estudo e a análise da dinâmica das falhas, dos efeitos da alteração de dimensões das redes *Scale-Free* nas falhas em cascata disparadas por ataques aleatórios e ataques direcionados aos *hubs*.

A partir da primeira etapa de simulações, através das simulações do comportamento da falha em diferentes tamanhos de redes, da simulação na rede normal, da rede com Estratégia 1, da rede com Estratégia 2 e com ambas observa-se um desempenho inferior da Estratégia 2 (adição de *links*), que tanto no cenário aleatório quanto no ataque aos *hubs* reduziu a quantidade de sobreviventes.

Conforme as redes aumentam em tamanho, aumentam também o número de nós de baixo grau. Logo, o número de *links* adicionados na Estratégia 2 também aumenta em cada rede. A centralidade máxima aumenta em todos os casos, devido ao aumento no número de caminhos possíveis e da ocorrência do vértice nestes caminhos, exceto na rede de dez nós, em que essa ocorrência foi reduzida devido a adição de um único *link* não fazer efeito no aumento.

Observou-se também que a medida do coeficiente de agrupamento médio é reduzida em todas as redes submetidas à adição de *links*. Essa medida mostra o quão agrupados os vizinhos de cada um dos nós da rede se apresentam. Nos casos analisados, esta medida é sempre inferior a rede original, sem a adição de *links*, já que aumentou o número de nós vizinhos que não são vizinhos entre si.

Através dos experimentos, também se observa que a robustez das redes é reduzida conforme se adicionam *links,* uma vez que os sobreviventes diminuem. Pode-se inferir, em ordem de benefícios para a rede, que a adição de capacidade é a melhor opção na contenção ou redução da falha em cascata, a associação das estratégias fica em segundo lugar, a rede sem estratégia em terceiro e a situação menos vantajosa é a adição de *links*, por reduzir o número de sobreviventes até mesmo da rede sem estratégia alguma.

Outra comprovação relacionada a redução da robustez da rede pela adição de *links* é o cálculo do Raio Espectral das redes antes e após a adição de *links*. O Raio Espectral aumenta após a adição de *links* em todos os casos, e o aumento do raio espectral em vários trabalhos

significa redução da robustez da rede, e, portanto, o aumento de sua vulnerabilidade a ataques. (LEWIS, 2009; JAMAKOVIC et al., 2007; KOOJI e VAN DAM, 2007; GANESH et al., 2005)

Isto pode ser consequência do fato de mais conexões significarem persistência no espalhamento da falha, que se espalha de forma semelhante ao de epidemias a uma dada taxa de infecção.

Outra hipótese se confirma, a de que as redes desta topologia são mais robustas a ataques aleatórios e mais vulneráveis a ataques intencionais, pelo número de sobreviventes ser reduzido no ataque aleatório.

Na segunda etapa de simulações, observa-se que o ataque aos *hubs*, onde se coloca nos gráficos os nós sobreviventes versus vulnerabilidades, conforme a vulnerabilidade ou probabilidade de espalhamento aumenta, a porcentagem de sobreviventes tende a se igualar em todos os tamanhos de redes. Já quando a vulnerabilidade é menor, a tendência é a quantidade de nós não atingidos ser coerente com o tamanho da rede: quanto maior o tamanho, maior a porcentagem, além de próximas entre si.

Já na análise do ataque aleatório, é evidente o aumento na porcentagem de nós não atingidos para todos os tamanhos, se comparada com as mesmas probabilidades de espalhamento nos casos de ataque aos *hubs*. Isso devido ao fato de a falha ser aleatória, o que acaba por atingir parcela menor da rede quando ocorrida. Observa-se também uma queda um pouco mais suave no número de nós sobreviventes, se comparada ao ataque aos *hubs* conforme se aumenta a probabilidade de espalhamento da falha. Isso pode ser consequência do tipo da falha, que neste caso é aleatória e ocasiona menor impacto na rede.

Observando os resultados das simulações em escala logarítmica, é encontrado um modelo matemático capaz de descrever a trajetória das curvas. Utilizando uma função exponencial para cada tipo de falha, encontra-se o modelo de função exponencial que se adéqua a cada curva.

Comparando as curvas da falha original e as curvas do modelo encontrado, já com os valores de variáveis encontrados pelo *software* em cada um dos tamanhos de rede, é possível visualizar o comportamento da cascata. Comportamento este, muito próximo ao da função exponencial parametrizada com os valores de variáveis cabíveis para cada cenário de falha, em cada tamanho de rede.

A descoberta da ineficaz adição de *links* e das funções que definem as curvas do comportamento da cascata é de fundamental importância para o entendimento, previsão e análise

de seu comportamento. Os achados também iluminam a direção de trabalhos futuros em estudos de redes complexas em presença de falhas e como as falhas no cenário aleatório e no cenário de ataque aos *hubs* pode se comportar na topologia S*cale-Free*.

Para outros trabalhos, podem-se simular outras estratégias de contenção de falhas, em separado, verificar o desempenho das mesmas, bem como associá-las para ver os resultados e impactos na falha em cascata. Além disso, podem-se inferir outras medidas para comportamento de outras propriedades das redes, diferentes do *betweenness,* do coeficiente de agrupamento e do raio espectral para verificação das variações em cada estratégia. Os impactos no mecanismo da cascata, a análise de outros aspectos como custos e velocidade de propagação também podem ser estudados.

Referências

ADAMIC, L.A. e HUBERMNA, B.A. Power-law Distribution of the Word Wide Web. **Science**. vol. 287, p. 2115, 2000.

ALBERT, R.; JEONG, H.; BARABÁSI, A.L. Diameter of the Word-Wide Web. **Nature**, vol. 401, p. 130-131, 1999.

ALMAAS, E.; KULKARNI, R.V.; STROUD, D. Characterizing the Structure of Small-Word Networks. **Physics Review Letters**., vol. 88., p. 98101, 2002.

ARIANOS, S., BOMPARD, E., CARBONE, A., XUE, E., Power grid vulnerability: A complex network approach. **Chaos** 19, 2009.

ASH, J., NEWTH, D. Optimizing complex networks for resilience against cascading failure. **Physica A** 380, 673–683, 2007.

BANKS, D.L. e CARLEY, K.M. Models for Network Evolution. **Journal of Mathematical Sociology**, vol. 21, p. 173-196, 1996.

BARABÁSI, A.L. **Linked: How Everything Is Conected to Everything and What It Means for Business, Science, and Everyday Life**. Plume. 2003. 294p.

BARABÁSI, A.L., ALBERT, R. Emergence of scaling in random networks. **Science 286**: 509–512, 1999.

BARABÁSI, A.L. e BONABEAU, E. Scale-Free Networks, **Scientific American 288**, 50-59 2003.

BARRAT, A. e WEIGT, M. On the Properties of Small-Word Network models. **Eur. Phys. J.B**., vol. 13, p. 547-560. 2000.

BRANDES, U: A Faster Algorithm for Betweenness Centrality. **Journal of Mathematical Soctology** 2001 , 25**:**163-177.

CÂMARA, G. **Grandes Desafios da Computação: A construção de uma terceira cultura.** Apresentação no Seminário Grandes Desafios, Maio de 2006.

CARRERAS, B. A., LYNCH, V. E., DOBSON, I. and NEWMAN, D. E., Critical points and transitions in an electric power transmission model for cascading failure blackouts, **Chaos**, vol. 12, p. 985-994, 2002.

COSTA, L.F. e DIAMBRA, L. Topographical Maps as Complex Networks. **Physical Review E**, vol. 71, p. 21901, 2004.

CRUCITTI, P., LATORA, V., MARCHIORI, M., Model for cascading failures in complex networks, **Physical Review E**, vol. 69, p. 45104, 2004.

DeMARCO, C. L., A phase transition model for cascading network failure**, IEEE Control Systems Magazine**, vol. 21, p. 40-51, 2001.

DIESTEL, R. **Graph Theory**, 2nd Ed. Springer, GTM 173, NewYork, 2000.

DOBSON, I., CARRERAS, B. A., NEWMAN, D. E., A loading-dependent model of probabilistic cascading failure, **Probability in the Engineering and Informational Sciences**, vol. 19, p. 25-32, 2005.

ERDÖS, P., RÉNYI, A., "On Random Graphs. I." **Publicationes Mathematicae 6**: 290–297, 1959.

GANESH, A., MASSOULIÉ, L., TOWSLEY, D., The effect of network topology on the spread of epidemics. **INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE**, vol. 2, p. 1455 – 1466, 2005.

GOH, K.I.; KAHNG, B. e KIM, D. Universal Behavior of Load Distribution in Scale-Free Networks. **Physics Review Letters**, vol. 87, p. 278701, 2001.

GUIMARÃES, J. O. Teoria dos Grafos, Departamento de Computação, UFSCar, 75p. http://www2.dc.ufscar.br/~jose/courses/tg/btg.pdf.

HINES, P., COTILLA-SANCHEZ, E., BLUMSACK, S. Do topological models provide good information about vulnerability in electric power networks? **Physics and Society**, 5p. 2010.

JAMAKOVIC, A., KOOIJ, R.E., VAN MIEGHEM, P. VAN DAM, E.R., Robustness of networks against viruses: the role of the spectral radius. **Symposium on Communications and Vehicular Technology,** p. 35 – 38, 2007.

KOOIJ, R.E., VAN DAM, E.R., The minimal spectral radius of graphs with a given diameter. **Linear Algebra and its Applications,** Volume 423, Issues 2-3, Pages 408-419, 2007.

KURANT, M. THIRAN, P., Layered complex networks, **Physics Review Letters**,, vol. 96, p. 138701, 2006.

LEWIS, T. G., Cause-and-Effect or Fooled by Randomness? **Homeland Security Affairs VI**, no. 1, January 2010, http://www.hsaj.org/?article=6.1.6

LEWIS, T. G., Telecommunications in the United States. Wiley **Handbook of Science and Technology for Homeland Security**, John Wiley & Sons, Inc., 2009.

LUCE, R.D., PERRY, A.D. A method of matrix analysis of group structure. **Psychometrika 14** (1): 95–116. 1949.

MANNAI, W.I.A., LEWIS, T.G. A general defender-attacker risk model for networks. Journal of Risk Finance, **Emerald Group Publishing**, vol. 9, issue 3, pages 244-261, 2008.

MASLOV, S., SNEPPEN, K. and ZALIZNYAK, A. Detection of Topological Patterns in Complex Networks: Correlation Profile of the Internet, **Physica A 333**, 529-540, 2004.

MEDEIROS, C. B. Modelagem Computacional de sistemas complexos artificiais, naturais e socioculturais e da interação homem-natureza, **Computação Brasil, Ano VII – nº 23** Setembro/Outubro e Novembro de 2006, Sociedade Brasileira de Computação, p. 5.

MILGRAM, S.Behavioral Study of Obedience. **Journal of Abnormal and Social Psychology**, Vol. 67, No. 4, 371-378, 1963.

NEWMAN, M.E.J. The Structure and Function of Complex Networks. **SIAM Review**, vol. 45, p. 167-256, 2003.

OPSAHL, T., AGNEESSENS, F., SKVORETZ, J., Node centrality in weighted networks: Generalizing degree and shortest paths. **Social Networks**, 2010.

RIOS, M. A., KIRSCHEN, D.S., JAYAWEERA, D., NEDIC, D. P. and ALLAN, R. N., Value of Security: Modeling Time-Dependant Phenomena and Weather Conditions, **IEEE Power Engineering Review**, vol. 22, p. 53, 2002.

SANSAVINI, G., HAJJ, M.R., PURI, I.K., ZIO, E., A deterministic representation of cascade spreading in complex networks. **EPL journal** 87, 2009.

SUN, H.J., ZHAO, H., WU, J.J. A robust matching model of capacity to defense cascading failure on complex networks, **Physica A**, vol. 387, p. 6431-6435, 2008.

WANG, J. RONG, L., A model for cascading failures in scale-free networks with a breakdown probability, **Physica A**, vol. 388, p. 1289-1298, 2009.

WANG, J., LIU, Y., JIAO, Y., A new cascading failure model with delay time in congested complex networks. **Systems Engineering Society of China, co-published with Springer-Verlag** GmbH, Volume 18, Number 3, 369-381, 2009.

WANG, J.,RONG, L., ZHANG, L., A model for cascading failure in complex networks with tunable parameter. **Modern Physics Letters B**, Vol. 23, No. 10, 2009.

WANG, X. F., CHEN, G., Complex Networks: Small-World, Scale-Free and Beyond, **IEEE Circuits and Systems Magazine**, vol. 3, p. 6-20, 2003.

WATTS, D.J., STROGATZ, S.H., Collective dynamics of 'small-world' networks. **Nature** 393 (6684): 409–10, 1998.

WU, J., GAO, Z. SUN, H. Cascade and breakdown in scale-free networks with community structure. **Physical Review E**, 74, 2006.

ZHAO J., XU K., Enhancing the robustness of scale-free networks, **Journal of Physics A: Mathematical and Theoretical**, vol. 42, p. 195003, 2009.

ZHAO, L.; PARK, K.; YING-CHENG, L. e CUPERTINO, T. H. Attack Induced Cascading Breakdown in Complex Networks, **Journal of the Brazilian Computer Society JBCS**, Ed. SBC, p. 68-76, 2005.