Capítulo 1

Modelagem e Projeto de Redes sem Fio Heterogêneas Resilientes e Sobreviventes

Robson Melo, Aldri Santos, Michele Nogueira, Deep Mehdi

Abstract

Our society depends today on a diversity of network services to support requirements from entertainment to critical applications related to commerce, telecommunication and healthcare. Given the popularity of portable devices, industry and academia have intensified efforts to develop heterogeneous wireless networks, offering ubiquitous computing environment. These networks involve the interoperability among networks based on different dimensions and technologies, operating complementarily to support final user requests. However, heterogeneous wireless networks characteristics reinforce resilience and survivability issues due to their operational complexity. In face of this context, this short course aims to provide a general overview about the state-of-the-art of researches related to resilience and survivability in these networks based mainly on practical examples. Models and metrics to resilience and survivability analyses are emphasized, as well as research open questions.

Resumo

Nossa sociedade depende hoje de uma grande variedade de serviços de rede para apoiar novas exigências que vão desde o entretenimento até serviços críticos voltados ao comércio, telecomunicação e cuidados com a saúde. Impulsionados pelo crescente uso de dispositivos portáteis, a indústria e a academia têm feito esforços para desenvolver as redes sem fio heterogêneas, oferecendo um ambiente computacional ubíquo. Estas redes contemplam a interoperabilidade entre outras redes de dimensões e tecnologias diferentes, operando de forma complementar para suportar as necessidades dos usuários finais. Entretanto, suas características reforçam os desafios em garantir a resiliência e a sobrevivência dos serviços nessas redes devido à sua complexidade de operação. Diante deste contexto, este minicurso visa prover uma visão geral do estado da arte de pesquisas relacionadas à resiliência e à sobrevivência de redes sem fio heterogêneas com base em exemplos práticos. De forma complementar, os modelos e as métricas para análise da resiliência e da sobrevivência de redes são enfatizados, assim como os desafios de pesquisa.

1.1. Introdução

Nossa sociedade depende de uma grande variedade de serviços de rede para apoiar exigências que vão desde o entretenimento até serviços críticos voltados ao comércio, bancos, telecomunicação e cuidados com a vida. Os serviços críticos, como os sistemas de controle de tráfego de transportes, os sistemas de emergência e os serviços financeiros, têm exigências restritas de Qualidade de Serviço (QoS) relacionadas ao desempenho e à resiliência [Sterbenza et al. 2010]. Idealmente, as faltas em componentes ou em um subsistema devem ser imperceptíveis para os usuários finais, e não devem fazer qualquer diferença se o comprometimento do serviço é causado por ataques, acidentes ou falhas. No entanto, é muito dispendioso construir uma rede que mantém o nível de QoS inalterado para todos os serviços diante de eventos comprometedores. Desta forma, a preferência é dada aos serviços mais críticos, pois em geral usuários tendem a aceitar a indisponibilidade temporária de serviços menos críticos, enquanto os serviços críticos são gradualmente degradados ou, na melhor das hipóteses, afetados sem comprometimento [Sterbenza et al. 2010].

As dificuldades relacionadas à resiliência são ainda maiores quando se trata de uma rede sem fio heterogênea devido à complexidade em integrar diferentes tecnologias de comunicação e tipos de redes. Impulsionados pelo uso crescente de dispositivos portáteis, os quais aumentam as requisições pelos serviços de rede, a indústria e a academia têm investido no desenvolvimento de redes sem fio heterogêneas por serem de fácil instalação e flexíveis [Buljore et al. 2009, Ghosh et al. 2012, Lashgari and Avestimehr 2012]. Essas redes contemplam a interoperabilidade entre as redes de dimensões e tecnologias diferentes, operando de forma complementar para suportar as necessidades dos usuários finais, diminuir os problemas causados pelo crescimento da infraestrutura de rede e oferecer um ambiente de computação ubíquo. Essas redes possuem uma ampla variedade de aplicações, como as redes domésticas, militares e de emergência. Diferentes empresas têm investido em macrocélulas e microcélulas de redes de acesso por rádio e equipamentos de núcleo da rede [Ghosh et al. 2012], estimando que os investimentos cheguem a quase 57 milhões de dólares em 2017 nos Estados Unidos. A Figura 1.1 ilustra as redes sem fio heterogêneas e sua relação com a infraestrutura de rede existente e os serviços.

Entretanto, uma variedade de ameaças, como ataques, acidentes e falhas, pode causar degradações menores ou maiores nos serviços de rede de telecomunicações. As ameaças de ataque têm recebido bastante atenção depois dos acontecimentos terroristas em Nova York (2001), Madrid (2004) e Londres (2005). Muitos países consideram a sua infraestrutura de redes e telecomunicações como uma parte da infraestrutura crítica nacional, e consequentemente ela precisa ser protegida [Nat 2003]. A proteção também deve cobrir desastres naturais, como inundações, terremotos, tempestades e outros, que podem comprometer a infraestrutura de telecomunicação [Zandt]. Finalmente, através de regulamentos e acordos de nível de serviço (SLA, do inglês *Service-Level Agreement*), o governo e os clientes devem dar incentivos suficientes às operadoras e aos fornecedores de equipamentos e serviços para que estes tomem todas as precauções necessárias e evitem interrupções da rede, tais como a famosa queda do *frame relay* da rede da operadora AT&T [Neumann 1998], que durou até 26 horas e incluiu o fracasso de operações de missão crítica de bancos; ou a interrupção mais recente na rede da NTT no Japão [Duffy 2007], onde cerca de 4 mil roteadores pararam por aproximadamente sete horas,

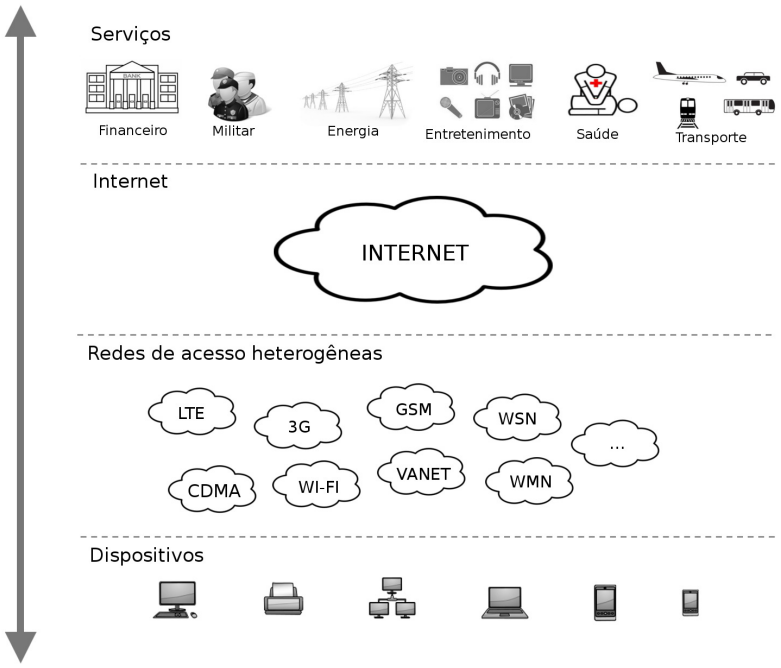

Figura 1.1. Ilustração de redes sem fio heterogêneas e sua relação entre a infraestrutura de rede legada e os serviços

desconectando milhões de usuários de Internet banda larga no Japão.

Em uma rede heterogênea é essencial prover comunicação entre pares de nós, mesmo utilizando tecnologias de comunicação diversas, com certas garantias de desempenho, como limites de atrasos na transferência de pacotes ou de perda, e ao mesmo tempo garantir um bom aproveitamento global dos recursos da rede. A gestão da comunicação deve ser projetada para lidar com *1)* as mudanças lentas e rápidas na carga e nos padrões de tráfego, *2)* as interrupções curtas e longas de menor importância para o funcionamento do nó, do enlace e/ou dos servidores, e *3)* as interrupções curtas e longas, porém severas, em um grande número de nós ou enlaces, ou em um servidor crítico na plataforma de serviços. A concepção, a construção e o gerenciamento da infraestrutura heterogênea de rede e de serviços são tarefas muito importantes e extremamente desafiadoras, principalmente quando o objetivo é garantir a resiliência e a sobrevivência dessas redes [Sterbenza et al. 2010]. Para isto, uma combinação de diferentes abordagens são tomadas, como: *1)* evitar as causas das falhas, *2)* projetar a rede para garantir que haja diversidade suficiente e capacidade de recuperação do serviço a fim de suportar a perda de parte da capacidade, e *3)* o desenvolvimento e a configuração pró-ativa, e técnicas e protocolos reativos de gestão de tráfego a fim de possibilitar a recuperação de serviços e sua continuidade em um nível de serviço solicitado. Técnicas de gerenciamento estão em desenvolvimento para atender a esses requisitos considerando as características dessas novas redes. Essas técnicas se aplicam a diferentes camadas da pilha de protocolos, usam abordagens reativas e pré-planejadas, e utilizam vários métodos de configuração.

Tanto os aspectos comportamentais, quanto os aspectos estruturais da rede precisam ser levados em conta para a modelagem (transitória) da resiliência no gerenciamento dos serviços e das comunicações. Isto significa que o modelo deve representar como o

desempenho dos serviços e das comunicações é afetado pelo encaminhamento e reencaminhamento de dados, por variações de carga de tráfego, e mudanças na capacidade da rede, e por exigências de serviços diferentes. Os modelos de confiabilidade estrutural geralmente se concentram em probabilidades de conectividade terminais, enquanto modelos comportamentais, tal como o modelo proposto em [Gan and Helvik 2006], levam em conta a dinâmica da rede na disponibilidade do serviço e na representação de um estado estável. A combinação de aspectos estruturais e de comportamento é tipicamente feito usando modelos de Markov ou modelos de filas para análise de desempenho. Além disso, o estudo combinado de desempenho e confiabilidade é realizado por modelos markovianos com recompensa.

Este minicurso apresenta uma visão geral sobre o estado da arte de pesquisas relacionadas à sobrevivência de redes sem fio heterogêneas, com base em exemplos práticos. Neste trabalho, considera-se a capacidade da rede de sobreviver a falhas maiores e menores na infraestrutura de rede sem fio e nas plataformas de serviços que são causadas por eventos indesejados externos (ambiente, desastres naturais, de acidentes, ataques maliciosos humanos e ataques eletrônicos), ou internos (congestionamento de tráfego, falha no enlace/nó, reparação e diferentes modos de falha). Sobreviver significa que os serviços críticos são fornecidos em conformidade com os requisitos de desempenho e confiabilidade, mesmo na presença de múltiplas falhas.

O minicurso inicia com uma descrição dos conceitos de resiliência e sobrevivência de redes, redes sem fio heterogêneas e outros correlacionados. Posteriormente, exemplos reais de desafios para resiliência e sobrevivência em redes e sistemas de telecomunicações são apresentados como forma de enfatizar a importância desses dois aspectos no projeto de redes sem fio heterogêneas. Os principais modelos de sobrevivência propostos na literatura para o escopo de redes em fio são descritos, ressaltando suas vantagens e desvantagens, além das oportunidades de pesquisa para estendê-los. As formas e métricas de avaliação de resiliência e sobrevivência de redes existentes são ressaltadas, assim como os avanços científicos necessários para consolidar a área. Desta forma, os desafios e as oportunidades de pesquisa serão explorados neste minicurso, enfatizando as necessidades para avançar nessa direção. Por fim, exemplos práticos e direções futuras concluirão este minicurso.

1.2. Fundamentos

Esta seção introduz as conceitos relacionados à resiliência e à sobrevivência de redes. Como esses conceitos foram desenvolvidos de forma independente ao longo de várias décadas, não há uma terminologia auto consistente estabilizada. Os conceitos são organizados dentro do domínio de resiliência, após apresentar o importante conceito de falta, erro e cadeia de falhas. Posteriormente, o conceito de redes sem fio heterogêneas é apresentado.

1.2.1. Falta → erro → falhas (tolerância a faltas)[1]

Uma falta é uma vulnerabilidade no sistema que pode causar um erro [Laprie 1994, Steinder and Sethi 2004]. Ela pode ser uma vulnerabilidade acidental no projeto (como um *bug*

[1] Alguns pesquisadores utilizam a terminologia falha, erro e defeito [Weber 2001].

de software), ou uma vulnerabilidade intencional devido a limitações que permitam uma ação externa ao sistema causar um erro, como, por exemplo, não projetar um sistema suficientemente robusto devido a restrições de custo. Uma falta latente pode ser desencadeada, levando a uma falta ativa, que pode ser perceptível como um erro. Um erro é um desvio entre um valor ou estado observados e seu valor ou estado especificados como corretos [Steinder and Sethi 2004, USA 1996] que pode conduzir a uma subsequente falha de serviço [Laprie 1994]. Uma falha de serviço (frequentemente abreviada por falha) é um desvio do serviço a partir do funcionamento desejado do sistema por não cumprir sua especificação ou expectativa [Laprie 1994, Group 2004, Group 2001, USA 1996, Laprie et al. 2004]. Consequentemente uma falta pode ser desencadeada a causar um erro perceptível, o que pode resultar em uma falha caso o erro se manifeste de uma forma que impeça ao sistema de atingir sua especificação de serviço. Essa relação é mostrada na Figura 1.2. As caixas denominadas *defender e detectar* são partes da estratégia de resiliência que será explicada ao longo desta seção. Observe que segundo a ilustração as defesas da rede podem prevenir que ações externas desencadeiem uma falta e que muitos erros perceptíveis não resultem em uma falha. A tolerância a faltas é um exemplo de redução dos impactos de faltas e erros na entrega de serviço. Além disso, os desafios e erros podem ser detectados, o que também proporciona uma base para ações tomadas como parte de uma estratégia de resiliência.

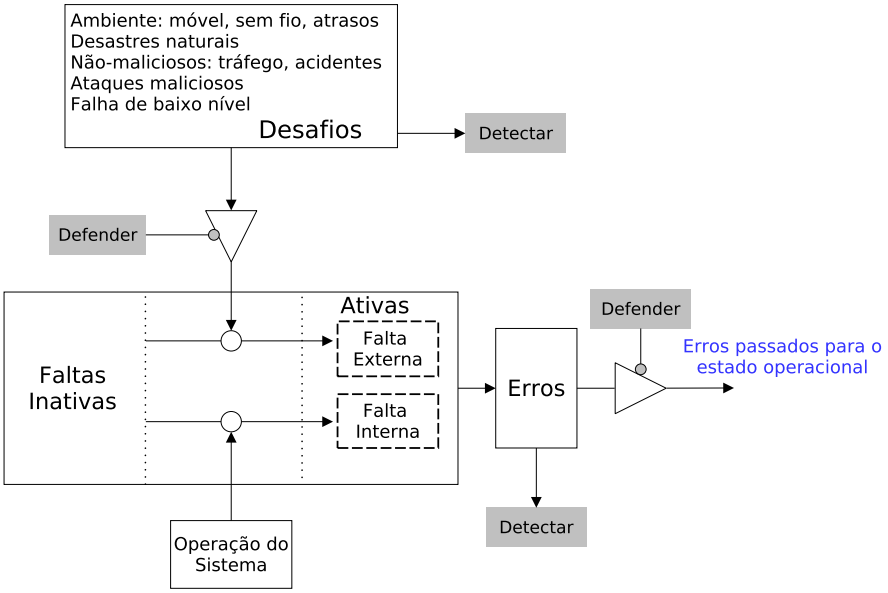

Figura 1.2. Falta → erro → falhas [Sterbenza et al. 2010]

1.2.2. Conceitos relacionados à tolerância a desafios

De forma geral, os conceitos relacionados à resiliência são divididos em duas categorias, como mostrado na Figura 1.3. No lado esquerdo, estão os conceitos de tolerância a desafios que lidam com o projeto e a engenharia de sistemas que possam continuar a prover serviço diante de desafios. No lado direto, estão os conceitos de confiabilidade que descrevem propriedades mensuráveis de sistemas resilientes. A relação entre essas duas categorias é a robustez, que formalmente é conceituada como o desempenho de um

sistema quando comprometido, ou em nosso contexto, a confiabilidade de um sistema quando desafiado. Neste trabalho utilizamos o termo desafios. Estes são quaisquer características ou condições que afetem a operação normal da rede, consistindo de uma má configuração ou erros operacionais não intencionais, desastres em larga escala naturais ou causados pelo homem, ataques maliciosos de adversários inteligentes, desafios ambientais (mobilidade, canais fracos, atraso prolongado não previsível, energia restrita), carga de tráfego anormal porém legítima, e falhas de serviço em baixo nível [Sterbenza et al. 2010, Heegaard and Trivedi 2009].

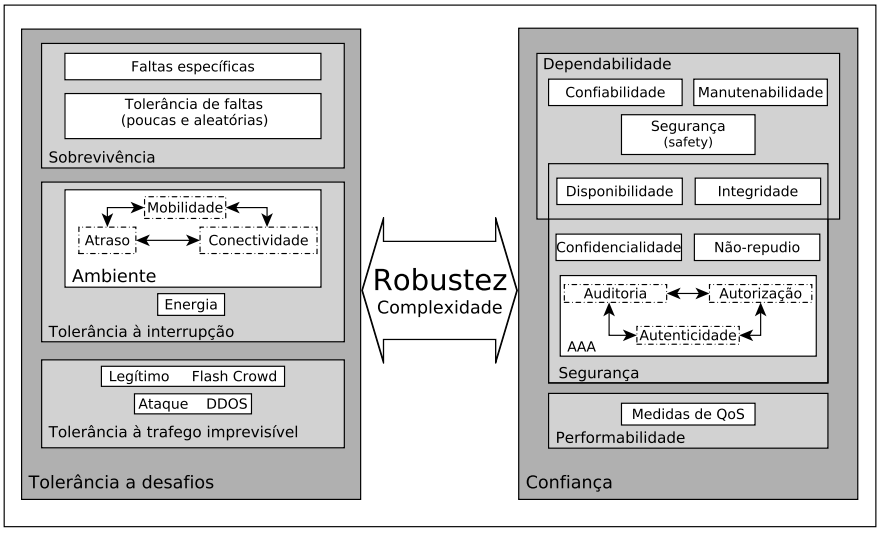

Figura 1.3. Disciplinas de resiliência (adaptado de [Sterbenza et al. 2010])

O primeiro subconjunto principal de conceitos relacionados à resiliência lida com o problema de como projetar sistemas que consigam dar continuidade do serviço mesmo na presença de desafios. Esses desafios podem ser subdivididos em *1)* faltas de componentes de sistemas, relacionados à tolerância a falhas e à capacidade de sobrevivência, *2)* rompimento de caminhos de comunicação relacionados à tolerância à interrupção, e *3)* desafios devido à introdução de tráfego na rede, relacionado à tolerância ao tráfego.

Tolerância a faltas

O conceito de tolerância a faltas é um dos mais antigos relacionados à resiliência, e é definido como a capacidade de um sistema tolerar faltas tal que não ocorram falhas de serviço [Group 2004, USA 1996]. Enquanto o uso de redundância para cobrir falhas em sistemas físicos remonta séculos, talvez a primeira referência em um contexto de computação foi a introdução da programação versão-N no contexto da máquina calculadora de Babbage [Lardner 1834] em meados de 1800. A tolerância a faltas emergiu como uma disciplina moderna nos anos 1950 quando Von Neumann e Shannon conceberam técnicas para projetar sistemas de comutação telefônica confiáveis em relés mecânicos relativamente não confiáveis [Moore and Shannon 1956]. A tolerância a faltas também foi aplicada ao projeto de sistemas de computadores na década de 60, particularmente

para sistemas de missão crítica, usados na defesa e espaço aéreo [Pierce 1965, Avizienis 1967].

A tolerância a faltas depende de redundância como uma técnica para compensar a falha aleatória e não correlacionada de componentes. As técnicas de tolerância a faltas podem ser aplicadas para ambos hardware e softwares, como por exemplo a redundância triplo-modular [Lyons and Vanderkulk 1962], a programação versão-N [Chen and Avizienis 1978] e os blocos de recuperação [Randell 1975], respectivamente. Estas técnicas são geralmente suficientes quando aplicadas a sistemas de escopo geográfico limitado. A tolerância a faltas não é suficiente para fornecer cobertura diante de múltiplas falhas correlacionadas, portanto é necessária, mas não suficiente para prover resiliência. Assim, a tolerância a faltas pode ser considerada um subconjunto da capacidade de sobrevivência, a qual trata de múltiplas falhas correlacionadas.

É importante enfatizar que a comunidade de redes óticas utiliza o termo capacidade de sobrevivência significando tolerância a faltas em nível físico e enlace. As técnicas como proteção automática de comutação SONET/SDH [Ellinas and Stern 1996] e ciclos-p [Grover and Stamatelakis 1998] são técnicas de tolerância a faltas aplicadas a um grafo da rede. Observe que grupos de risco de enlaces compartilhados (SLRGs, do inglês *Shared Link Risk Groups*) [Strand et al. 2001] provêem diversidade topológica, mas não necessariamente diversidade geográfica.

Sobrevivência

O surgimento da Internet e sua dependência levaram à percepção de que novas técnicas eram necessárias para redes afetadas por múltiplas falhas correlacionadas, para as quais as técnicas de projeto de tolerância a faltas não são suficientes. Em geral, a sobrevivência consiste na capacidade de um sistema cumprir sua missão, em tempo hábil, mesmo na presença de ameaças como ataques ou desastres naturais de grande escala. Essa definição captura o aspecto de falhas correlacionadas, resultantes de um ataque gerado por um adversário inteligente [Ellison et al. 1997, Sterbenz et al. 2002], bem como falhas em grande parte da infraestrutura de uma rede [Mahmood 2009, Bassiri and Heydari 2009]. Além da redundância requerida pela tolerância a faltas, a capacidade de sobrevivência requer diversidade tal que a mesma fatalidade seja improvável de ser compartilhada por partes do sistema sob falhas correlacionadas.

Além dessa definição geral de sobrevivência, são encontradas na literatura diversas outras especificadas por diferentes organizações, institutos de padrão industrial, e comunidades acadêmicas em muitos contextos. Essas definições são categorizadas como qualitativas e quantitativas, sendo a maioria delas consideradas qualitativas. Por exemplo, o Instituto Americano de Aeronáutica e Astronáutica (AIAA) define a capacidade de sobrevivência da aeronave como a "capacidade de uma aeronave evitar ou resistir a ambientes hostis, incluindo ambos os ambientes artificiais e naturais, como raios, em pleno ar, colisões e acidentes" [Heegaard and Trivedi 2009]. A Divisão Nacional de Tecnologia Sistema de Comunicação e Padrões criou uma definição de sobrevivência para sistemas de telecomunicações [Heegaard and Trivedi 2008b] como "a propriedade de um sistema, equipamento, subsistema, processo ou procedimento que fornece um grau definido de ga-

rantia de que a entidade nomeada continuará a função durante e após um distúrbio natural ou feito pelo homem, por exemplo, uma explosão nuclear". [Deutsch and Willis 1988] propuseram uma definição de sobrevivência para o sistema de software como "o grau em que as funções essenciais ainda estão disponíveis mesmo que alguma parte do sistema esteja comprometido". Neumann [Neumann and Barnes 1999] utilizou a sobrevivência para sistemas de computadores e redes como "a capacidade de um sistema satisfazer e continuar suas necessidades críticas diante de condições adversas". Ellison et al. do *Software Engineering Institute* (SEI) [Knight and Sullivan 2000], definiu sobrevivência de sistemas de informação críticos como "a capacidade de um sistema para cumprir sua missão, em tempo hábil, com a presença de ataques, falhas ou acidentes".

Embora cada uma das definições acima dê uma boa descrição conceitual de sobrevivência em um contexto específico, nenhuma delas é suficientemente clara e precisa para estabelecer os requisitos de sobrevivência mensuráveis e proceder a uma avaliação quantitativa. Os termos nessas definições como "a capacidade de", "o grau de", "a missão", "essencial" e outros dificilmente têm o mesmo significado em contextos diferentes ou leva a resultados consistentes em diferentes sistemas. Desta forma, tenta-se reduzir este problema estabelecendo definições quantitativas de sobrevivência. Uma definição que chega perto do que é essencialmente transmitido por todas as definições qualitativas apresentadas na pesquisa anterior é a seguinte.

A quantificação de sobrevivência pode ser realizada através de uma medida de interesse M contendo um valor m_0 pouco antes de ocorrer uma falha. O comportamento de sobrevivência pode ser representado pelos seguintes atributos: m_a é o valor de M logo após a falha ocorrer, m_u é a diferença máxima entre o valor corrente de M e m_a após a falha; m_r é o valor de M restaurado após um período de tempo t_r e t_R é o período de tempo para o sistema se restaurar o valor de M.

Esta definição foi proposta pelo grupo de trabalho de desempenho de rede T1A1.2 [Liu et al. 2004]. Por esta definição, a sobrevivência descreve o comportamento variável do sistema no tempo, depois de ocorrer uma falha.

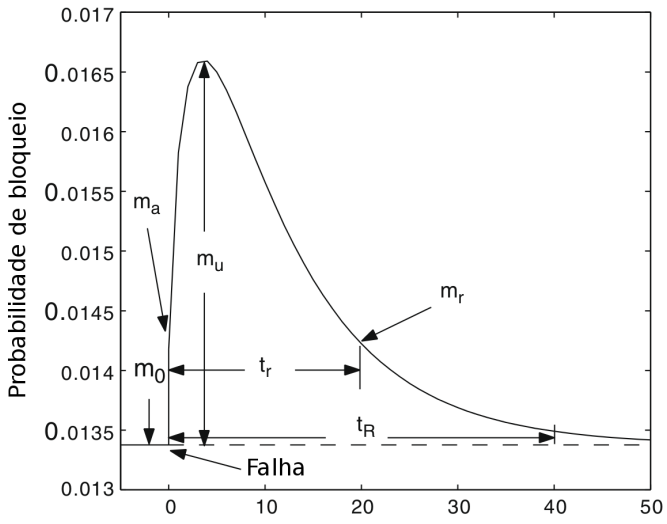

Figura 1.4. Sobrevivência após a primeira falha [Heegaard and Trivedi 2008a]

Tolerância à interrupção

Um tipo de desafio único em redes de comunicação sem fio vem do ambiente de comunicação. Este torna difícil manter conexões estáveis fim a fim entre os dispositivos, requisitando tolerância à interrupção, ou seja, a capacidade de um sistema tolerar interrupções na conectividade entre seus componentes. Exemplos de desafios no ambiente específicos dessas redes são: canais de comunicação fracos e intermitentes, mobilidade, atrasos imprevisivelmente longos, bem como limitações de energia.

Há três principais fatores que impulsionaram o campo de tolerância à interrupção. O primeiro é motivado pelo comportamento de redes dinâmicas e começou com redes sem fio por comunicação via rádio. Isso levou a mais pesquisas em redes móveis *ad hoc (MANETs*, do inglês *Mobile Ad Hoc NETworks)* que propuseram mecanismos de encaminhamento e roteamento para redes dinâmicas nas quais a conectividade entre os membros muda continuamente [Perkins and Bhagwat 1994, Johnson 1994]. O segundo foi motivado por atrasos muito grandes que os protocolos tradicionais de redes não podiam tolerar, especificamente em ambiente espacial e de satélites [Durst et al. 1996] e a Internet Interplanetária (IPN) [Burleigh et al. 2003]. Essa pesquisa levou a noções mais gerais de redes tolerantes a atrasos [Fall 2003] e redes tolerantes à interrupção em que os caminhos estáveis fim a fim podem nunca existir. As técnicas que suportam essas redes incluem comunicação tão longa quanto possível, mas revertem à técnica de armazenar e encaminhar, quando necessário, ou à técnica com base em nós móveis transportando informações, chamada de armazenar, transportar e encaminhar [Sterbenz et al. 2002, Li and Rus 2000, Zhao et al. 2004]. O terceiro fator está relacionado às restrições as quais estas redes estão sujeitas, como a restrição de energia, exemplificada por meio das redes de sensores sem fio [Akyildiz et al. 2002], em que os nós com suas baterias drenadas não podem mais contribuir com a conectividade da rede [Singh et al. 1998, Shah and Rabaey 2002].

Mais recentemente, as técnicas de redes tolerantes à interrupção têm encontrado aplicações em várias situações, incluindo as redes veiculares ad hoc *(VANETs*, do inglês *Vehicular Ad Hoc NETworks)* [Li and Wang 2007, Ott and Kutscher 2004], as redes meteorológicas tolerantes à interrupção [Jabbar et al. 2009], e as redes aéreas altamente dinâmicas [Rohrer et al. 2008].

Tolerância a tráfego imprevisível

A última categoria de desafios é aquela causada pela introdução de um tráfego imprevisível na rede. A tolerância a tráfego é a capacidade do sistema tolerar carga imprevisível oferecida sem uma queda significativa na carga transportada (incluindo colapso por congestionamento), bem como isolar efeitos de tráfego cruzado, outros fluxos, e outros nós. Ao definir tráfego como um desafio, considera-se tráfego além dos parâmetros de projeto da rede na sua operação normal. Os desafios de tráfego podem ser ou não esperados; ser legítimos, como de um momento de aglomeração [Jung et al. 2002]; ou maliciosos como de um ataque distribuído de negação de serviço (DDoS, do inglês *Distributed Denial of Service*) [Mirkovic and Reiher 2004]. É importante observar que embora a detecção

de DDoS seja um esforço importante, os recursos da rede são impactados independentemente do tráfego ser malicioso ou não. Além disso, um ataque DDoS suficientemente sofisticado é indistinguível do tráfego normal, e consequentemente mecanismos de tolerância a tráfego são importantes independentemente se os mecanismos de detecção de ataque forem bem sucedidos ou não.

1.2.3. Conceitos relacionados à confiança

A confiança é definida como a garantia de que um sistema funcionará conforme esperado [Laprie et al. 2004], o que deve ser de acordo com propriedades mensuráveis. Os conceitos relacionados à confiança medem consequentemente a entrega de serviço da rede, e consistem da (1) dependabilidade, (2) segurança, e (3) peformabilidade.

Dependabilidade

A dependabilidade é a disciplina que quantifica a confiança que pode ser disposta no serviço entregue por um sistema [Lee and Anderson 1990, Laprie 1994], e consiste de dois principais aspectos: disponibilidade e confiabilidade. Os valores esperados para as funções de densidade de falhas e reparos são importantes para ambos aspectos. As métricas básicas de confiança são MTTF (do inglês, *Mean Time to Failure*), que é o valor esperado da função de densidade de falhas, e o MTTR (do inglês, *Mean Time to Recovery*), que é o valor esperado da função de densidade de reparos. O tempo médio entre falhas MTBF (do inglês, *Mean Time Between Failure*) é a soma desses dois [Billinton and Allan 1992]:

$$MTBF = MTTF + MTTR \tag{1}$$

A disponibilidade é a prontidão para uso, que é a probabilidade de um sistema ou serviço estar operante quando necessário, e é calculado como:

$$A = \frac{MTTF}{MTBF} \tag{2}$$

A confiabilidade representa a continuidade do serviço, que é a probabilidade de um sistema ou serviço manter-se operante por um período específico de tempo (t):

$$R(t) = Pr[\text{não haver falhas em}[0,t]] = 1 - Q(t) \tag{3}$$

Onde $Q(t)$ é a função de distribuição acumulativa de falhas.

Essas noções de sistemas confiáveis foram registradas pela IFIP WG 10.4 [Laprie et al. 2004] e ANSI T1A1 [Group 2001] e são comumente aplicadas à dependabilidade da rede. Essas noções de sistemas confiáveis são também aplicadas a enlaces de fibra ótica como uma medida de tolerância a faltas [Clouqueur and Grover 2002, Grover 2003].

A importância da disponibilidade e da confiabilidade do serviço depende da aplicação. A disponibilidade é essencial para serviços de transações como navegação Web

com base no protocolo HTTP. Contanto que o servidor geralmente esteja disponível, importa menos se ele falha frequentemente, ou seja, se o MTTR for bastante curto. Por outro lado, a confiabilidade é de suma importância para serviços orientados à sessão e à conexão como teleconferências: para ser útil, a sessão deve permanecer disponível por um período específico de tempo requerendo um longo MTTF.

Além disso, há vários outros aspectos de confiabilidade [Laprie 1994]. A manutenabilidade é a aptidão do sistema submeter-se a reparos e evoluções. A segurança*(safety)* é a dependabilidade relacionada a falhas catastróficas [Nicol et al. 2004]. A segurança tem como preocupação particular as infraestruturas críticas como as redes de energia e plantas de energia nuclear, e sua interdependência na Internet e redes do tipo SCADA (do inglês, *Supervisory Control and Data Acquisition*). A integridade é um aspecto da confiabilidade que é mais comumente associada com segurança, descrita a seguir.

Segurança

A segurança é a propriedade do sistema, e de suas medidas tomadas, de tal forma que se proteja de acesso ou mudanças não autorizados, sujeitos a políticas [Landwehr 2001]. As propriedades de segurança incluem a Autenticação, Autorização e Auditoria (AAA), confidencialidade e a capacidade de não-repúdio. A segurança compartilha com a dependabilidade as propriedades de disponibilidade e integridade [Laprie et al. 2004]. No contexto de confiança, estamos preocupados com as propriedades mensuráveis dos aspectos de segurança [Savola 2007, Vaughn et al. 2003]. Também podemos considerar segurança relacionada à autoproteção [Shirey 2007], que é um princípio necessário à resiliência.

Performabilidade

A performabilidade [Meyer 1992] é a propriedade de um sistema de reproduzir o desempenho projetado pela especificação do serviço, como descrito por medidas de Qualidade de Serviço (QoS, do inglês *Quality of Service*), tais como atraso, vazão e taxa de entrega de pacotes [Campbell et al. 1994, Lazar and Pacifici 1991]. A performabilidade estende os conceitos da dependabilidade para uma gama de sistemas degradáveis.

1.2.4. Robustez e complexidade

Duas disciplinas ficam de fora da tolerância a desafios e da confiabilidade, mas descrevem suas relações umas com outras (robustez) e com características gerais (complexidade).

Robustez

A robustez é uma propriedade de controle teórico que relaciona a operação da rede com perturbações na mesma [Jen 2005, Willinger and Doyle 2005]. No contexto de resiliência, a robustez descreve a confiança (comportamento quantificável) em uma rede diante de desafios que mudem seu comportamento. Observe que o termo robustez é frequentemente usado de forma muito menos precisa como sinônimo de resiliência, capacidade de sobrevivência ou segurança.

Complexidade

A complexidade se refere às formas em que um grande número de redes interagem, resultando em um comportamento emergente [Prokopenko et al. 2009]. A ciência da complexidade tem uma relação importante com a resiliência e a robustez, pois mecanismos de resiliência como auto-organização e comportamento autônomo aumentam a complexidade, e a complexidade incrementada pode resultar em uma grande vulnerabilidade na rede.

1.2.5. Conceitos relacionados às redes sem fio heterogêneas

As redes heterogêneas sem fio, ou apenas redes heterogêneas, compreendem vários tipos de redes utilizando diferentes e independentes tecnologias de acesso sem fio. A rede de telefonia celular, a rede de transmissão de TV e a Internet, por exemplo, representam redes heterogêneas que oferecem serviços específicos sob tecnologias distintas. Mesmo as redes de comunicação de dados, como redes locais sem fio (WLAN, do inglês *Wireless Local Area Networks*), as redes de área metropolitana sem fio (WMAN, do inglês *Wireless Metropolitan Area Networks*) e as redes de telefonia celular são consideradas redes heterogêneas por aplicarem diversas tecnologias de redes de acesso sem fio, como Wi-Fi, Wi-Max, 3G, 4G e outras, e diferentes padrões de comunicação, tais como IEEE 802.11b, 802.11a e 802.11g[Lin et al. 2011]

Na literatura, também é encontrado o termo HetNet (*Heterogeneous Networks*) para se referenciar ao uso de múltiplos tipos de redes de acesso em redes sem fio. Uma rede de amplo alcance WAN (do inglês, *Wide Area Networks*) pode ser composta por macrocélulas, pico-células e/ou femto-células a fim de oferecer uma cobertura sem fio em uma ampla região com uma diversidade de áreas de cobertura sem fio, variando desde um ambiente externo aberto a um escritório em um prédio comercial, residências e áreas subterrâneas[L. et al. 2011, David et al. 2011].

Especialistas em mobilidade definem as HetNets como redes com uma interoperação complexa entre macrocélula, pequena célula e, em alguns casos, entre elementos de rede usando a tecnologia Wi-Fi para prover cobertura de acesso e capacidade de transferências (do inglês, *handoffs*) entre os elementos de rede. Um estudo recente da ARCchart estima que HetNets assistirá a direcionar o mercado de infraestrutura móvel e receberá investimentos globais de aproximadamente 57 bilhões de dólares até 2017.

Uma rede heterogênea prevê interoperabilidade entre as redes e tecnologias, sendo capaz de prover serviços através de qualquer tipo de rede sem fio e manter a disponibilidade dos mesmos, de forma transparente para o usuário final, mesmo quando o tipo de rede ou a tecnologia sejam alterados. Em geral, mudanças no tipo de rede ou de tecnologia são provenientes da mobilidade dos usuários finais e seus dispositivos, a qual altera as condições da rede e dos serviços. Entretanto, mudanças de tecnologia de rede de acesso também podem ocorrer em redes fixas, tais como na infraestrutura fixa das redes em malha sem fio.

O processo de transferências consiste na mudança de estação-base/antena a qual um dispositivo móvel está associado para uma outra durante uma chamada ou comunicação fim a fim. Existem várias razões possíveis para ocorrer transferência, incluindo

a deterioração do sinal entre a estação-base corrente e o usuário móvel e a sobrecarga de uma célula, dificultando o gerenciamento de múltiplas comunicações fim a fim. Este processo introduz perdas de pacotes e latência, podendo comprometer severamente a comunicação dos dados e diferentes tipos de aplicações, como as aplicações multimídia, que possuem limitações temporais para a transmissão dos pacotes.

Acessar de uma forma eficiente os recursos de redes sem fio exige novas abordagens para um uso eficaz das tecnologias de rede heterogêneas disponíveis. Para ilustrar este aspecto, vamos seguir o exemplo. Alguns usuários móveis (UMs) em um ônibus estão usando aplicações de alto consumo de banda, como *streaming* de vídeo e aplicações de vigilância. Para implementar tais aplicações, elas podem usar a rede celular, a qual prover serviços muito caros, cobrados pelo tempo de antena ou de volume de tráfego. Por outro lado, podem existir redes Wi-Fi WLAN nas proximidades, com alta largura de banda disponível (por exemplo, 54 Mbps) e acesso gratuito e ilimitado. No entanto, mesmo com a banda larga sem fios disponível a partir de um ponto de acesso público (PAP) nas proximidades, os UMs não podem acessar os PAP, pois não têm acesso às redes heterogêneas, ou eles são incapazes de acessar o serviço direito, porque os UMs não têm o suporte necessário ao acesso, ou os serviços podem ser desconectados com frequência, resultando em uma qualidade inaceitável dos serviços.

Os desafios acima podem ser abordados por agregar/agrupar as ligações heterogêneas sem fios disponíveis, ou seja, com a convergência em uma rede única onipresente. As redes 3G (aplicando, por exemplo, as tecnologias W-CDMA, TD-SCDMA, CDMA2000, HSPA), as WLANs e a Internet. Para alcançar a agregação de enlaces de redes sem fio heterogêneas, três grandes desafios devem ser abordados como seguem: heterogeneidade na interface do enlace - os usuários finais precisam acessar diferentes tipos de ligações móveis; interrupção do enlace de comunicação, devido à mobilidade dos usuário finais, sinais de rádios instáveis e cobertura limitada, e a vulnerabilidade do enlace de acesso - as ligações móveis são altamente vulneráveis a ataques.

O processo de convergência requer suporte a características como conectividade, mobilidade, localização e acesso ubíquo dos elementos em transição de uma rede para outra, exigindo uma estrutura robusta das redes de acesso. Os trabalhos recentes na área de redes heterogêneas têm focado principalmente em técnicas de sobreposição de rede para o descarregamento de tráfego de dados para as células menores. Enquanto os ganhos com esta abordagem são promissores, eles representam apenas um ponto de partida. A previsão é de que as redes heterogêneas irão desempenhar um papel central na evolução da banda larga móvel sem fio, e servirá como uma plataforma facilitadora para inovações tecnológicas disruptivas. Desta forma, garantir a conexão, a identificação e a segurança nestas redes e em seus dispositivos ainda representam desafios para grupos de pesquisa [Kafle et al. 2010].

1.3. Por que resiliência e sobrevivência são importantes?

Esta seção ilustra os principais tipos de desafios associados às redes sem fio heterogêneas, enfatizando a necessidade por resiliência e sobrevivência. Além disso, é apresentado um conjunto de eventos e falhas que afetaram seriamente diferentes redes e seus serviços no passado e, caso essas vulnerabilidades permaneçam, elas poderão também comprometer

as redes sem fio heterogêneas.

1.3.1. Carga de tráfego anormal, porém legítima

Uma requisição, não maliciosa, por serviço e geradora de uma carga maior (ou diferente) da operação normal esperada é um desafio para a rede. Isto é causado normalmente por um evento de *flash crowd* [Jung et al. 2002], que consiste de um grande volume de requisições de serviço além da carga normal. Além de afetar o alvo do evento, a rede como um todo pode ser afetada, particularmente por tráfego cruzado perto do alvo [Xie et al. 2008]. Um segundo efeito de muitos dos desafios listados a seguir é a produção de um evento de *flash crowd* devido à necessidade de respostas emergenciais e à população tentando obter informações sobre um acontecimento importante.

1.3.2. Acidentes e erros humanos

Os acidentes e erros não-maliciosos são gerados por pessoas que interagem com a rede, tais como a má configuração não intencional de dispositivos ou o não cumprimento de uma política correta. Estes eventos podem ocorrer durante o projeto ou a operação da rede e podem se tornar prejudiciais se as equipes envolvidas tentarem acobertar seus erros. Às vezes, equívocos e acidentes podem ter consequências bastante significativas, como descrito a seguir.

Um apagão em larga escala afetou grande parte do nordeste dos Estados Unidos e a província de Ontário em 14 de agosto de 2003, afetando 50 milhões de pessoas. Muitos fatores interligados foram responsáveis em tornar alguns poucos problemas operacionais em uma falha maior. A queda de três usinas de energia em um dia quente de verão resultou em reservas reativas insuficientes e indisponíveis. Por volta da mesma hora, os processos do software de alarme automático foram deixados desligados devido a erro humano. O resultado foi que quando as três linhas de alta voltagem caíram, a carga não foi apropriadamente reencaminhada e, em vez disso, causou o colapso de grande parte das redes de energia (mais 15 linhas de alta tensão) do norte de Ohio. Isto causou surtos e falhas em cascata nas redes vizinhas conectadas, espalhando-se mais rápido que os repasses de proteção automáticas eram capazes de propagar. Eventualmente, ocorreram *failsafes* e quedas de linhas suficientes, isolando assim as porções menos robustas da rede de energia e parando as falhas em cascatas, ao mesmo tempo particionando a rede em várias ilhas. O serviço total ficou indisponível por mais de uma semana em algumas áreas. Os custos, ambos em termos de reparos e perda de produtividade, foram na ordem de 10 bilhões de dólares [Liscouski and Elliot 2004]. Esta falha de energia em larga escala teve um impacto significativo na infraestrutura interligada da Internet. Mais de 2000 prefixos anunciados globalmente tiveram quedas severas de 2 horas ou mais, afetando por volta de 50% de todos os sistemas autônomos da Internet [Cowie et al. 2003]. Este exemplo de infraestrutura interdependente, com a Internet dependendo da rede de energia para os equipamentos continuarem operando, enquanto ao mesmo tempo muitos sistemas de controle de energia SCADA estão se comunicando através de serviços da Internet. Uma falta de conhecimento da complexidade da rede como um todo, assim como alguns erros humanos ambos em planejamento e decisões de remediação contribuíram para a extensão e severidade deste apagão.

Em 8 de maio de 1988, um incêndio na *Illinois Bell switching office* em Hinsdale causou danos graves aos serviços de telefone no norte de Illinois nos Estados Unidos. A queda afetou comunicações locais, a longa distância, de telefones móveis, de serviços 800, e também de comunicação de controle de tráfego aéreo entre os aeroportos de Midway e O'Hare em Chicago e o centro da Administração Federal da Aviação (FAA, do inglês *Federal Aviation Administration*) em Aurora, Illinois. Levou até o final de maio para restaurar o serviço. Apesar da rede de telefonia conter redundância de hardware e enlace, os serviços falharam porque ambos os sistemas primário e de backup estavam localizados no mesmo prédio, e foram ambos destruídos no incêndio resultante de um raio. O incêndio de Hinsdale é um exemplo canônico de como a tolerância a falta sozinha não é suficiente para resiliência. Uma fração significativa das falhas das redes de telefonia ocorrem devido a acidentes e erros humanos [Kuhn 1997].

Em 18 de julho de 2001, um trem descarrilhou no túnel da Howard Street em Baltimore, Maryland no nordeste estadunidense. O fogo subsequente causou disrupções nas fibras do *backbone* utilizadas por sete grandes provedores de serviço da Internet [Styron 2001, Carter et al. 2002]. Este túnel era um lugar conveniente para passar linhas de fibra por baixo da cidade de Baltimore. A maioria do tráfego foi redirecionado, mas resultou em congestionamento em conexões alternativas, causando uma queda de velocidade notável numa porção significante da Internet. Novos fios de fibra foram colocados, dentro de 36 horas, para restaurar a capacidade física. No caso de tanto o túnel de Baltimore como o incêndio de Hinsdale, as escolhas de projeto que foram feitas resultaram em implementações de sistemas redundantes; entretanto, sem diversidade geográfica, a infraestrutura redundante compartilhava o mesmo destino.

Enquanto a Pakistan Telecom tentava obedecer uma ordem do Governo de bloquear um certo vídeo do *YouTube* em 24 de fevereiro de 2008, eles anunciaram o seu próprio sistema anônimo como o caminho mais curto para uma porção do espaço de endereço IP do YouTube. Este anúncio saiu não só para os provedores dentro do Paquistão, mas também a PCCW (do inglês, *Pacific Century CyberWorks*). Naquele tempo, a PCCW não estava filtrando falsos anúncios de prefixo como este, e propagou o anúncio para o resto do mundo, fazendo com que maioria das requisições HTTP para o YouTube provenientes do mundo inteiro fossem redirecionadas para a Pakistan Telecom. Dentro das horas seguintes, muitas tentativas foram feitas pelo YouTube para competir com os falsos anúncios de rota usando prefixos mais específicos, mas a situação não retornou ao normal até a PCCW desconectar a Pakistan Telecom completamente. Enquanto o escopo global deste problema tenha sido, mas provavelmente, acidental, mostra claramente a vulnerabilidade do Protocolo de Roteamento de Borda (do inglês, *Border Gateway Protocol* ou BGP) para *route-spoofing* e ainda apresenta um grande desafio para operação de Internet resistente [Brown 2008, Meinel , Brown and Zmijewski 2008].

1.3.3. Desastres em larga escala

Os desastres em larga escala podem resultar de causas naturais ou de erros humanos. Em qualquer caso, eles são uma categoria única de desafios, pois resultam em falhas correlacionadas em uma grande área. Por exemplo, destruir o hardware, simultaneamente impedindo os operadores de efetuar funções normais e restringindo o acesso à informação por administradores, resultando em más escolhas de remediação. Os exemplos de

desastres em larga escala incluem furacões, terremotos, nevascas, tsunamis, enchentes e quedas de energia propagadas.

Em 29 de agosto de 2005, o furacão Katrina causou destruição em massa na Louisiana e Mississippi no sudeste dos Estados Unidos, e interrompeu comunicação com 134 redes [Davis et al. 2006, Cowie et al. 2005]. Muitas dessas interrupções foram causadas por quedas de energia, e a maioria delas foram restauradas em um período de dez dias. Uma conexão da rede de pesquisa da Abilene *Internet2* também ficou interrompida, mas havia capacidade disponível nos caminhos alternativos para redirecionar o tráfego. Teve também um rompimento significativo na rede de telefonia devido ao aumento de tráfego e à destruição das torres de celular. Os esforços de recuperação de desastres realizados após o furacão, também demonstraram os desafios causados por equipamentos de comunicação incompatíveis usados por socorristas (polícia local, bombeiros, polícia estadual, guarda costeira, guarda nacional, etc.).

Em 26 de dezembro de 2006, com continuidade pelos próximos dois dias, um grande terremoto acompanhado de fortes e pequenos tremores ocorreu perto de Hengchun, Taiwan que danificou os cabos submarinos que provêm conectividade de Internet entre Ásia e América do Norte [Kitamura et al. 2007]. 1200 faixas de prefixo tornaram-se temporariamente inalcançáveis, a capacidade de acesso a Internet da China foi reduzida em 74% e o acesso a Internet em Hong Kong foi completamente desativado. Apesar do BGP ter sido capaz de redirecionar automaticamente parte do tráfego, ele o fez sem ter conhecimento da topologia física por baixo ou da utilização de conexão, o que resultou no tráfego entre China e Taiwan cruzando o Oceano Pacífico duas vezes. A engenharia de tráfego teve que redirecionar manualmente o tráfego pela Coréia e Japão, em vez de pelos Estados Unidos.

Outros eventos que poderiam causar falhas catastróficas de rede elétrica, comunicação por rádio e comunicação de rede em uma grande área incluem tempestades geomagnéticas induzidas pelo Sol [NRC 2008] (com um pico previsto para 2012) e ataques de armas de pulso eletromagnético [CNI 2004]. Finalmente, uma epidemia viral séria poderia ter impactos severos se as pessoas forem incapazes ou receosas de operar e manter a infraestrutura crítica incluindo a Internet [DHS 2006, DHS 2007]. Esta foi uma preocupação que felizmente não ocorreu com a epidemia da gripe H1N1 de 2009-2010, mas poderá ser inevitável numa futura epidemia de gripe suína ou aviária, por exemplo.

1.3.4. Ataques maliciosos

Os ataques maliciosos são geralmente originados por criminosos cibernéticos, por razões que incluem terrorismo, guerra de informações entre nações, grupos políticos ou empresas concorrentes, assim como crackers, incluindo *script kiddies*. Estes desafios podem destruir ou danificar componentes críticos na infraestrutura da rede com intenção de indisponibilizar os serviços da rede, ou desativar a infraestrutura da rede como dano colateral.

Os ataques terroristas de 11 de setembro de 2001 em New York foram relativamente localizados e não almejavam as infraestruturas das redes por si só, mas muitas linhas de assinantes de acesso à Internet foram afetadas, pois 1 a 2% dos blocos de prefixo da Internet estavam inalcançáveis no pico do rompimento [Partridge et al. 2003]. A

maioria destes blocos foram restaurados depois de um dia ou dois, e o impacto em serviços essenciais da Internet, como o DNS e o BGP, foi mínimo. Os efeitos mais notáveis foram devidos a *flash crowds* em sites de notícias, que foram rapidamente sobrecarregados, chegando em alguns casos a 350% da carga de tráfego normal. No entanto, o DNS teve apenas com a carga normal devido a cache nos sistemas finais, e a própria Internet teve quantias agregadas de tráfego abaixo do normal, pois provavelmente muitas pessoas estavam preocupadas assistindo à cobertura televisiva do evento e usando menos a Web. Os ataques terroristas de 7 de julho de 2005 contra o sistema de transporte de Londres não afetou diretamente a estrutura da rede, exceto aquela utilizada pelos metrôs, mas também induziu a um aumento no uso da rede de celulares e da Internet pelo fato das pessoas estarem tentando obter notícias, e a capacidade prejudicada dos socorristas de comunicarem uns com os outros [GLA 2006].

Os ataques maliciosos que exploram vulnerabilidades de softwares ou protocolos incluem ataques de negação de serviço distribuídos que normalmente têm a intenção de prejudicar um indivíduo, organização, corporação ou nação. Quando o governo da Estônia decidiu mover um memorial de guerra soviético, um ataque DDoS foi lançado com endereços IP de dentro da Rússia. Esses ataques, motivados politicamente, tinham como alvo sites comerciais e governamentais estonianos. Em resposta, a Estônia cortou sua conexão de Internet com o resto do mundo a fim de parar os efeitos dos ataques [Lesk 2007].

1.3.5. Desafios ambientais

Os desafios resultantes do ambiente de comunicação incluem a conectividade fraca, assimétrica e esporádica dos canais sem fio; a alta mobilidade dos nós e das sub-redes; os atrasos prolongados não previsíveis devido à distância ou à conectividade intermitente. Estes desafios são associados às redes com tolerância a interrupções (do inglês, *Disruption-Tolerant Networks* ou DTNs), mas também a redes VANETs e a redes que utilizam tecnologias de acesso oportunista ao espectro, como as redes de rádio cognitivo.

1.3.6. Falhas em baixo nível

Se algum desses desafios causa uma falha de serviço em uma camada particular, essa falha se transforma em um desafio para qualquer serviço de alto nível que dependa do funcionamento correto da camada que falhou. Este tipo de falhas pode induzir em defeitos recursivos de serviços até que o erro possa ser contido por um serviço de alto nível. Desta forma, os sintomas de um desafio podem ser percebidos por múltiplos serviços. Por exemplo, um corte em uma fibra faz com que o serviço de cominho físico falhe, resultando assim na falha de todas as conexões na camada de enlace que dependem desse caminho. Contudo, a remediação pode ocorrer em uma camada mais alta apenas redirecionando o tráfego por outra fibra.

Em janeiro de 2008, cortes em cabos no Mar Mediterrâneo causaram interrupções substanciais na conexão ao Egito e Índia, e interrupções menores no Afeganistão, Bahrein, Bangladesh, Kuwait, Maldivas, Paquistão, Qatar, Arábia Saudita e Emirados Árabes. Mais de 20 milhões de usuários da Internet foram afetados. A causa dos cortes foi assumida por ser não intencional, podendo ser por naufrágio de barco, desgaste natural

e abrasão dos cabos submarinos contra rochas no fundo do mar [A. Popescu 2008, Wik 2008, RIP 2008].

Os cortes em cabos submarinos ocorrem frequentemente, com uma média de um corte ocorrendo a cada três dias no mundo inteiro. A maioria destes cortes ocorrem como eventos aleatórios não correlacionados e passam despercebidos pelos usuários finais devido à redundância na infraestrutura. Entretanto, em casos como no terremoto de Taiwan e os cortes do cabos Mediterrâneo (assim como o incêndio do túnel de Baltimore), as múltiplas falhas de conexões correlacionadas causaram grandes apagões, enfatizando que redundância para tolerância a faltas não é o suficiente para resiliência, mostrando que a diversidade geográfica para a sobrevivência também é necessária.

1.4. Modelos de resiliência e sobrevivência

Um modelo é uma representação da realidade de forma simplificada, com o objetivo de permitir a realização de análises de eventos passados ou futuros. Portanto, a modelagem implica na criação de modelos que expliquem as características, o funcionamento e o comportamento de um fenômeno (computacional ou não), auxiliando em seu entendimento e permitindo a realização de prognósticos futuros. A modelagem permite uma avaliação qualitativa e quantitativa de fenômenos físicos, que ocupam um lugar no espaço, e de elementos mais abstratos. As técnicas de modelagem, mais precisamente a modelagem analítica é uma poderosa ferramenta para a construção de soluções numéricas para elementos abstratos como é o caso da resiliência e sobrevivência. Assim, além qualificar as redes como resilientes e sobreviventes, é possível quantificar os índices de resiliência e sobrevivência das mesmas.

Esta seção apresenta um conjunto de modelos propostos na literatura para qualificar ou quantificar resiliência e sobrevivência de redes. Apresentam-se esses modelos com o foco em redes sem fio heterogêneas, porém muitos deles foram desenvolvidos com um escopo mais genérico. Esta seção não pretende ser exaustiva, mas sim descrever os principais modelos existentes na literatura que possam ser aplicados para avaliar a resiliência e a sobrevivência de redes sem fio heterogêneas. Discutimos brevemente as vantagens e desvantagens em aplicar os modelos e as métricas apresentados considerando o contexto dessas redes.

1.4.1. Modelos qualitativos

Os modelos qualitativos de resiliência e sobrevivência podem ser classificados como *modelos de estados estacionários* ou *modelos de estados transitórios*. Os modelos de resiliência no estado estacionário se caracterizam pela permanência em um estado de execução sem sofrer alterações bruscas. Neste estado, o *desempenho* do sistema pode ser avaliado considerando que o mesmo esteja continuamente em execução. Uma outra avaliação possível no estado estacionário é a *disponibilidade* do sistema que indica sua permanência em um estado de execução ou em um estado de erro. A *performabilidade* consiste em outra característica inerente ao estado estacionário. A performabilidade avalia a capacidade do sistema em reproduzir o desempenho projetado pela especificação do serviço. Assim, os sistemas resilientes em estado estacionário devem permanecer estáveis em sua execução e na prestação de seus serviços [Trivedi 2002, Amiri and Ghassemi-Tari 2007].

Os modelos qualitativos de resiliência de sistemas na classe transitória são caracterizados pela constante mudança de seu estado de execução até atingir um estado de estabilidade. Nesta classe os modelos são avaliados por um *modelo de desempenho transitório* que qualifica o sistema dependendo do estado de execução ao qual o mesmo esteja; pelo *modelo de disponibilidade transitória* que determina se em um dado instante o sistema pode estar disponível ou não; e pelo *modelo de desempenho transitório condicionado a cenários de falhas* em que a avaliação é influenciada pelos cenários de falhas que podem comprometer o desempenho do sistema ou não [Trivedi 2002, Amiri and Ghassemi-Tari 2007]. Uma taxonomia para os modelos qualitativos de sobrevivência anteriormente discutidos é ilustrada na Figura 1.5.

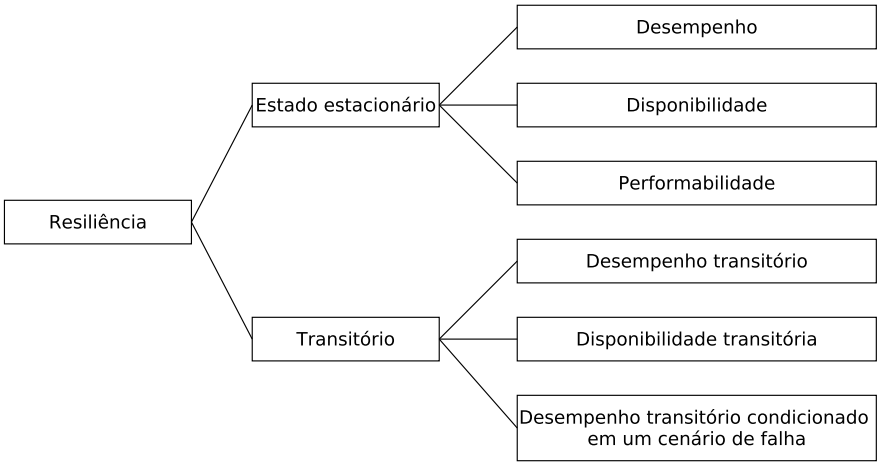

Figura 1.5. Taxonomia dos modelos qualitativos de resiliência

1.4.2. Modelos quantitativos

Os modelos quantitativos de resiliência têm como base os métodos analíticos que permitem mensurar os valores que representam determinadas grandezas de uma propriedade como desempenho e disponibilidade, por exemplo. Existe na literatura um conjunto de modelos utilizados para a quantificação de resiliência e sobrevivência. A Figura 1.6 ilustra uma síntese da classificação desses modelos.

Como visto na Figura 1.6, a avaliação quantitativa tem como base a *medição* ou *modelos*. A avaliação com base em *medição* utiliza *logs, traces e outros* do sistema como valores de entradas para a construção de uma avaliação representativa. A avaliação com base em *modelos* pode ser realizada por meio de *simulação de eventos discretos* utilizando simuladores que tomam como entrada variáveis aleatórias independentes, com valores randômicos ou pseudorandômicos para a construção de suas representações de cenários reais, como é o caso do conhecido simulador *Network Simulator (NS2)*. Outra forma de avaliação quantitativa com base em modelos são os *modelos analíticos*. Esses modelos utilizam métodos matemáticos abstratos para construção de representações variáveis de acordo com parâmetros específicos do sistema. Uma terceira forma de avaliação com base em modelo é a *híbrida*. Os modelos híbridos combinam a simulação de eventos discretos com os modelos analíticos. Neste trabalho o foco principal são os modelos analíticos

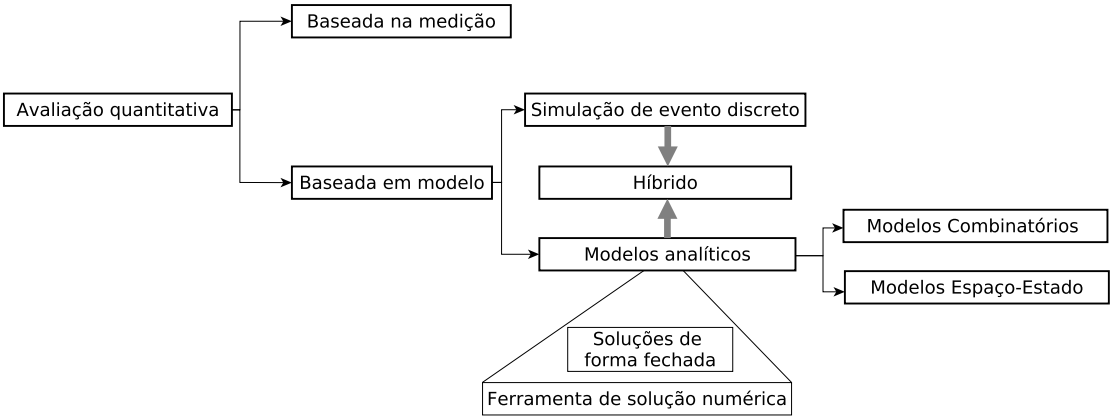

Figura 1.6. Taxonomia de modelos quantitativos

que compõem soluções fechadas utilizando ferramentas numéricas ou algébricas. Esses modelos, permitem compor e mensurar índices que representam grandezas quantificáveis de determinadas propriedades com é o caso da resiliência e da sobrevivência.

De acordo com [Trivedi 2002], as principais classes de modelagem analítica para resiliência são os *modelos do tipo combinatórios* e *modelos do tipo espaço-estados*. Deste modo, iniciaremos um resumo destas classes utilizadas para construção de modelos analíticos a fim de quantificar a resiliência e a sobrevivência, iniciando pelos modelos do tipo combinatório que têm como base a verificação do número de possibilidade de ocorrência de um determinado evento, sem a necessidade de descrição de todas as possibilidades.

Modelos do tipo combinatórios

Os modelos combinatórios também conhecidos como modelos de não espaço-estado são os tipos mais simples de técnicas analíticas/numérica e podem ser usadas para a modelagem de confiabilidade e disponibilidade sob determinados pressupostos. Esses modelos assumem suposições de que as falhas de componentes são independentes, bem como seus reparos.

Os modelos do tipo combinatório podem ser representados por diagramas de blocos de confiabilidade, grafo de confiabilidade, árvores de falhas e árvores de falhas com eventos repetidos. Uma ilustração das categorias de representação dos modelos combinatórios é apresentada na Figura 1.7. O modelo de representação de *diagrama de blocos de confiabilidade ou RBD*, define um relacionamento lógico entre componentes de um sistema. Os relacionamentos típicos são em *Série* ou *Paralelo* (para mais detalhes recomenda-se a leitura de [Kuo and Zuo 2003]). O modelo de representação de *grafo de confiabilidade* é um dos modelos mais intuitivos para a análise de confiabilidade, entretanto, possui um poder de expressão limitada pois em sua composição utiliza apenas representações de vértices e arestas [Kim and Seong 2001]. A *árvore de falha* relaciona diversas possíveis causas de falhas de um sistema usando lógica booleana para inferir parâmetros de dependabilidade, por meio de um processo dedutivo [Kuo and Zuo 2003].

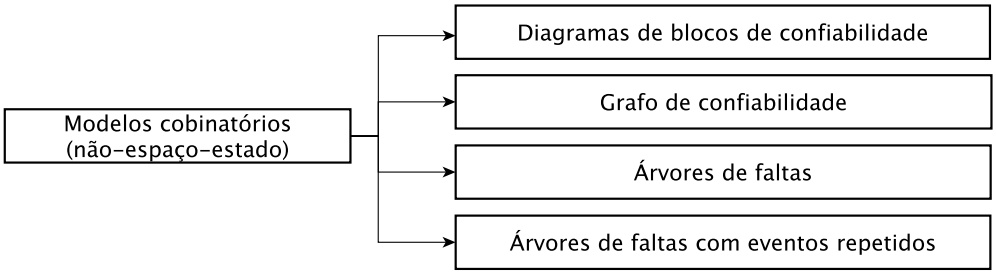

Figura 1.7. Modelos combinatórios

Os modelos combinatórios, apesar de serem modelos quantitativos, são amplamente utilizados apenas para a representação de um sistema de forma analítica. Contudo, modelos utilizados para mensurar índices que representam os valores de determinadas grandezas de uma propriedade são os modelos de espaço-estado.

Modelos do tipo espaço-estado

A representação em modelos espaço-estado também conhecidos como modelos discretos, fornecem uma maneira prática e compacta para modelar e analisar sistemas com múltiplas entradas e saídas. Esses modelos são classificados como *markovianos, não-markovianos e markovianos não homogêneos*.

Os modelos markovianos se subdividem em *cadeia de Markov de tempo discreto (DTMC), cadeia de Markov de tempo contínuo (CTMC)* e *modelos de recompensa markovianos (MRM)*. Já os modelos não-markovianos se subdividem em *processos semi-markovianos (SMP)* e *processo regenerativo de Markov* [Trivedi 2002]. A Figura 1.8 ilustra uma classificação dos modelos de espaço-estado. No entanto, este trabalho tem como foco principal as cadeias de Markov de tempo contínuo, devido sua ampla utilização na literatura para modelagem analítica.

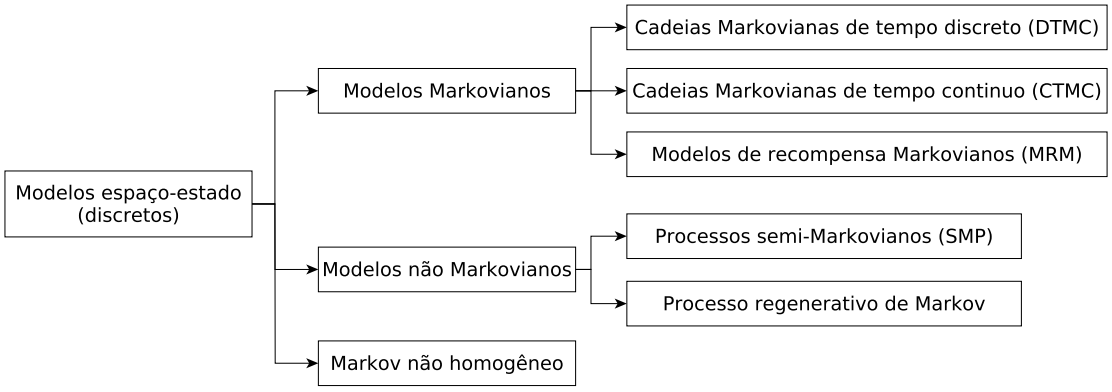

Figura 1.8. Modelos de espaço-estado

As CTMC são caraterizadas por possuir n estados e n transições entre esses estados. Uma propriedade importante é a falta de memória *(memoryless)*, assim, os estados

anteriores são irrelevantes para a predição dos estados seguintes, desde que o estado atual seja conhecido. Deste modo, o tempo em que o processo passa em cada estado de uma CTMC é exponencialmente distribuído, dado a propriedade da falta de memória dos processos markovianos. Com essas características a CTMC permite definir a probabilidade de o sistema estar em determinado estado em um instante de tempo. Em seguida, serão apresentados exemplos de modelo de desempenho e modelo de disponibilidade utilizando cadeia de Markov de tempo contínuo [Trivedi 2002].

- **Modelos de desempenho**: Considere uma rede sem fio heterogênea que implica em n enlaces com uma quantidade de nós infinita. Uma conexão será perdida (referida como bloqueio) quando encontrar todos os n enlaces desconectados após a sua chegada. O processo de chegada de conexão segue uma distribuição de Poisson com taxa λ. Assumimos que as conexões em espera são exponencialmente distribuídas com taxa μ. Sem considerar as falhas da ligação, o modelo *puro* de desempenho será uma cadeia de Markov de tempo contínuo homogênea (CTMC), como mostrado na Figura1.9, em que as conexões em curso j estão presentes no estado j [Trivedi 2002].

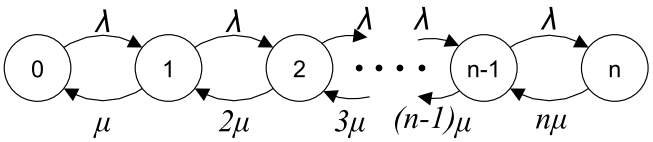

Figura 1.9. Modelo puro de desempenho

- **Modelos de disponibilidade**: Quando falhas nos enlaces são consideradas, assume-se que a falha e os tempos de reparo de cada enlace são exponencialmente distribuídos com taxas γ e τ, respectivamente. Assume-se também que um único reparo pode ser compartilhado por todos os enlaces da rede. O modelo *puro* de disponibilidade do sistema é também uma CTMC homogênea, como mostra a Figura 1.10, em que o estado i indica que existem i nós não falhos nos enlaces da rede [Trivedi 2002].

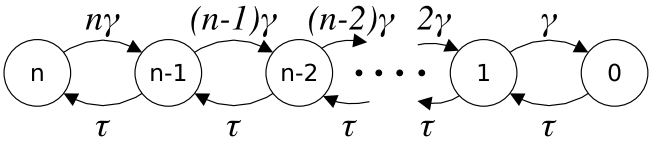

Figura 1.10. Modelo puro de disponibilidade

De acordo com [Heegaard and Trivedi 2008a] a partir do modelo puro de disponibilidade e do modelo puro de desempenho é possível definir um modelo para a quantificação da resiliência de uma rede seguindo os 5 passos a seguir. Salientamos apenas que

apesar dos autores em [Heegaard and Trivedi 2008a] utilizarem o termo sobrevivência, o substituímos neste texto pelo termo resiliência, para estar de acordo com as definições apresentadas na seção 1.2.

Passo 1: Desenvolver o modelo de disponibilidade puro, no qual os recursos (hardware e/ou software) falham e obtém reparação (ou são reiniciado).

Passo 2: Desenvolver um modelo de desempenho puro e obter os resultados dos estados estacionários do modelo, o que reflete o uso de recursos e outras informações de estado do sistema antes de uma falha acontecer.

Passo 3: Combinar os modelos de disponibilidade e desempenho obtidos nos dois primeiros passos para um modelo composto.

Passo 4: Escolher uma medida de sobrevivência de interesse. Forçar uma falha específica no sistema e construir um modelo truncado. A fim de refletir o uso de recursos do sistema antes da falha acontecer.

Passo 5: Realizar a análise transiente do modelo truncado composto

1.4.3. Abordagem para geração de modelos de resiliência qualitativos

Uma abordagem sistemática para elaboração de modelos qualitativos de resiliência para redes é apresentada em [Trivedi et al. 2009]. Esta abordagem tem como base os modelos de confiabilidade, confidencialidade, disponibilidade, desempenho, manutenabilidade, segurança e sobrevivência, e a combinação dos mesmos. Na abordagem, disponibilidade é representada por A, confidencialidade por C, integridade por I, desempenho por P, confiabilidade por R, sobrevivência por S, segurança por SF e manutenabilidade por M.

A Figura 1.11 mostra a classificação dos modelos e suas possíveis combinações. Na raiz da árvore é feita uma distinção entre modelos do $Tipo-R$ e do $Tipo-A$. Com o modelo $Tipo-R$, uma vez que o sistema falhar (devido a falhas de hardware, erros de software, ataques e acidentes) ocorre um retorno permitido, enquanto no $Tipo-A$ a composição dos modelos de falha do sistema é incluída na análise. A confiabilidade foca no evento quando um erro torna-se visível (como uma falha) na interface do serviço. Não existem níveis de qualidade de serviço contemplados pelos modelos de confiabilidade e disponibilidade típicos. Assim, os modelos de desempenho podem ser vistos como estados *ativos* de um modelo $Tipo-R$ ou $Tipo-A$ em vários níveis de serviço.

Combinando desempenho e falha/recuperação, obtém-se o modelo de performabilidade. Ao considerar a análise transiente de desempenho do sistema, imediatamente após a ocorrência de uma falha, um ataque ou um acidente, obtém-se o modelo $Tipo-S$ (sobrevivência). Os modelos $Tipo-I$ representam o grau em que a informação é correta (integra). Os modelos $Tipo-C$ capturam a probabilidade de que a informação é acessível e apenas pessoas autorizadas tenham acesso (confidencial). Os modelos $Tipo-C$ e $Tipo-I$ podem, ser vistos como outra classificação de estados *inativos* de um $Tipo-R$ ou $Tipo-A$.

Os modelos do $Tipo-SF$ podem ser vistos como um refinamento dos estados *inativos* de um modelo $Tipo-R$ em estados seguros e não seguros. Os modelos $Tipo-R$

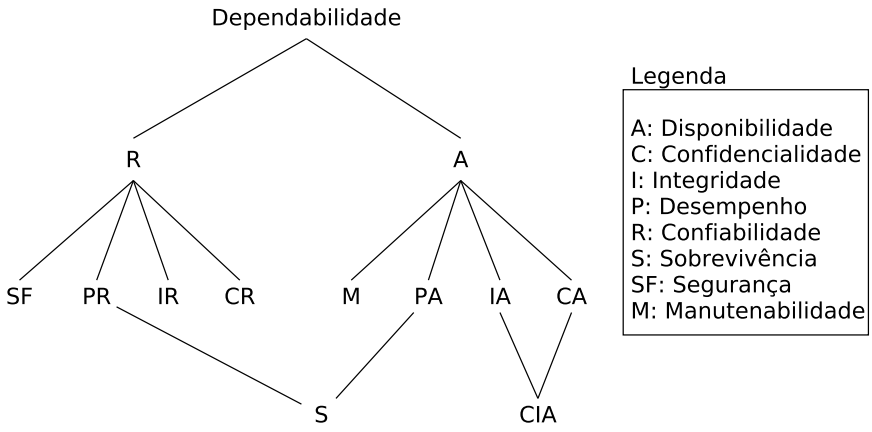

Figura 1.11. Abordagem para geração de modelos qualitativos

podem ser divididos em modelos do $Tipo - SF$ (segurança), *PR* (Desempenho e Confiabilidade), *IR* (integridade e confiabilidade) e *CR* (Confidencialidade e Confiabilidade). Se considerarmos o comportamento transiente do $Tipo - PR$ imediatamente após a ocorrência de uma das ameaças, o modelo $Tipo - PR$ torna um modelo $Tipo - S$.

O modelo $Tipo - M$ pode ser visto como uma elaboração de um modelo $Tipo - A$, onde o foco está em diferentes estratégias de manutenção: corretiva *vs* preventiva; com base no tempo ou nas condições preventivas. O modelo $Tipo - A$ pode ser dividido em modelos do tipo *M* (Manutenabilidade), *PA* (Desempenho e Disponibilidade), *IA* (Integridade e Disponibilidade), e *CA* (confidencialidade e disponibilidade). Se considerarmos o desempenho transitório do modelo $Tipo - PA$ imediatamente após a ocorrência de uma das ameaças, o $Tipo - PA$ torna-se $Tipo - S$. Tanto modelos do $Tipo - IA$ ou *CA* podem ser ainda mais generalizados a modelos $Tipo - CIA$ (confidencialidade, integridade e disponibilidade), incorporando modelos do $Tipo - C$.

Usando as categorias de tipos de modelos apresentados, pode-se abordar uma série de métricas e técnicas qualitativas relacionadas à segurança, a sobrevivência e a confiabilidade juntas. A vantagem é que esses modelos ajudam a compreender e avaliar o impacto sobre os sistemas globais e redes de forma sistemática.

1.4.4. Exemplo de modelagem quantitativa de resiliência e sobrevivência

Os modelos quantitativos de resiliência e sobrevivência de rede neste trabalho consideraram redes expostas a eventos indesejáveis que podem causar falhas nas ligações e nos nós e normalmente são seguidos por uma súbita mudança na disponibilidade de recursos de rede, como largura de banda, enlaces de transmissão, posições de enfileiramento (memória), e capacidade do processador. Gradualmente os recursos são restaurados através de reencaminhamento e pela restauração dos enlaces e nós falhos, o que resulta em restauração do desempenho.

A primeira parte desta seção descreve o princípio da modelagem de sobrevivência, seguido por detalhes de modelos de desempenho, que são construídos para avaliar taxa de perda esperada, vazão, e atraso (média e distribuição). Na sequência são descritas

duas abordagens para modelar a propagação de falhas e recuperação gradual com base ou conhecimento dos processos de propagação e mecanismos de recuperação [Heegaard and Trivedi 2008a] ou com base em seus traços reais [Heegaard and Trivedi 2008b].

Abordagem do modelo de sobrevivência

O modelo de sobrevivência não considera a frequência de eventos indesejáveis, pois o foco é dado ao que causou um evento indesejado, ou seja, à natureza de degradação do desempenho logo após o evento até que o sistema se estabilize novamente. Os modelos de sobrevivência são construídos através da combinação da cadeia de Markov de tempo contínuo (CTMC), modelos de desempenho e modelos de diferentes falhas de propagação e fases de recuperação do sistema. A Figura 1.12 ilustra o princípio de modelagem onde as falhas de propagação e recuperação são modeladas como uma sequência de fases, com cada uma das fases sendo um estado em um modelo CTMC, e as transições são causadas por eventos como a detecção de falhas, reencaminhamento completo e outros. Em cada estado do sistema, assume-se estar em um estado de equilíbrio de desempenho com condições operacionais inalteradas. Também em cada estado, as métricas de desempenho,como perda esperada, vazão e atraso (média e distribuição), são obtidos a partir de medidas de recompensa de uma cadeia de Markov de tempo contínuo (CTMC).

No tempo t, um conjunto de eventos indesejados são assumidos para tomar lugar de onde o período transitório de interesse começa. A mudança no estado do sistema é acionada e a evolução do sistema é seguida pelas fases de falha de propagação, detecção, recuperação e restauração/reparação. Observe que não precisa-se saber a frequência de eventos indesejáveis, como, por exemplo, o tempo até a falha, pois o fracasso é forçado ou desencadeado (linha tracejada na Figura 1.12).

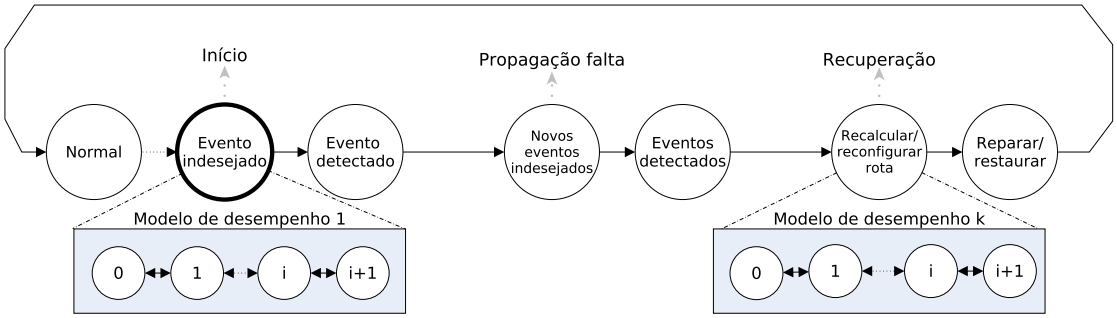

Figura 1.12. Sequência de falha de propagação e recuperação

Modelo de desempenho da rede

A rede é representada por um grafo $\overrightarrow{G} = (V,E)$, em que V é o conjunto de nós e E é o conjunto de enlaces. O único ou multi-caminho de roteamento de uma conexão virtual entre nó origem s e destino d reduz G para um grafo direcionado $\overrightarrow{G}_{[s,d]}$. O modelo de rede é markoviano com as chegadas externas de *Poisson* e distribuição de serviço de

tempo exponencial com uma disciplina de serviço *FCFS* em cada nó e/ou enlace. O encaminhamento entre o nó i e j é estocástico com probabilidade de tempo independente r_{ij}. Dependendo se o processamento do nó ou do enlace de transmissão é o gargalo no encaminhamento de pacotes, uma abordagem centrada no nó ou enlace é tomada:

- Centrada no nó: se o processamento de cada pacote no nó é o gargalo, cada nó é modelado como uma fila independente $M/M/1$;
- Centrada no enlace: se a ligação de transmissão é o gargalo, cada enlace (ou a interface de rede correspondente) é uma fila $M/M/n$, onde n é o número de canais de transmissão (tempos de propagação são incluídos ou adicionados ao tempo de serviço do enlace);
- Centrada nó e enlace: se o gargalo é flutuante entre o processamento e ligação de transmissão do nó, cada nó e enlace são modelados separadamente.

Na Figura 1.13, são ilustradas três possíveis modelos de desempenho da rede. Cada pacote ocupa um servidor ou a posição no *buffer* no nó ou de uma interface de transmissão. O estado x_i no CTMC é o número de pacotes em estado i. A variável de estado i refere-se a uma única enumeração dos nós e as ligações que são incluídas no modelo. Na seção seguinte, supõe-se uma visão de modelo de nó central, ou seja, a variável de estado i corresponde ao nó i. Em seguida, a capacidade do nó n_i é o número de posições do *buffer*. Com n_i finito, a taxa de chegada Γ_i é obtido através da resolução do sistema linear de equações de tráfego [Jackson 1957].

$$\Gamma_i^{(v)} = \gamma_i^{(v)} + \sum_{j \neq i} r_{ji}^{(v)} \Gamma_j^{(v)} (1 - \pi_j(n_j)); \text{ Para } j = 1, ..., n, \tag{4}$$

Onde $\gamma_i^{(v)}$ é a taxa de chegada de tráfego externo para o nó i por um caminho v e $\pi_j(n_j)$ é a probabilidade de estado estacionário que o nó j irá rejeitar um pacote de entrada. No caso de infinito n_j, o $\pi_j(n_j) = 0; \forall j$, e que o modelo é do tipo aberto *BCMP*[2] de redes de filas [Baskett et al. 1975]. Um único caminho de conexão é modelado como um caso especial em que cada salto tem um único enlace j com $r_{ij} = 1$ e 0 para todos os outros enlaces. O tráfego total externo γ_i para o nó i é $\gamma_i = 0$ para todo $i \neq s$ e $\gamma_s = \gamma$. A taxa de chegada total para o nó i é:

$$\Gamma i = \sum_{v=1}^{\vartheta \varphi} \Gamma_i^{(v)}. \tag{5}$$

[2]O teorema BCMP, desenvolvido por Baskett, Chandy, Muntz e Palacios, especifica a combinação de distribuições de tempo de serviço e disciplinas de escalonamento que produz redes de fila multi classe na forma de produto. São permitidas redes abertas, fechadas e mistas.

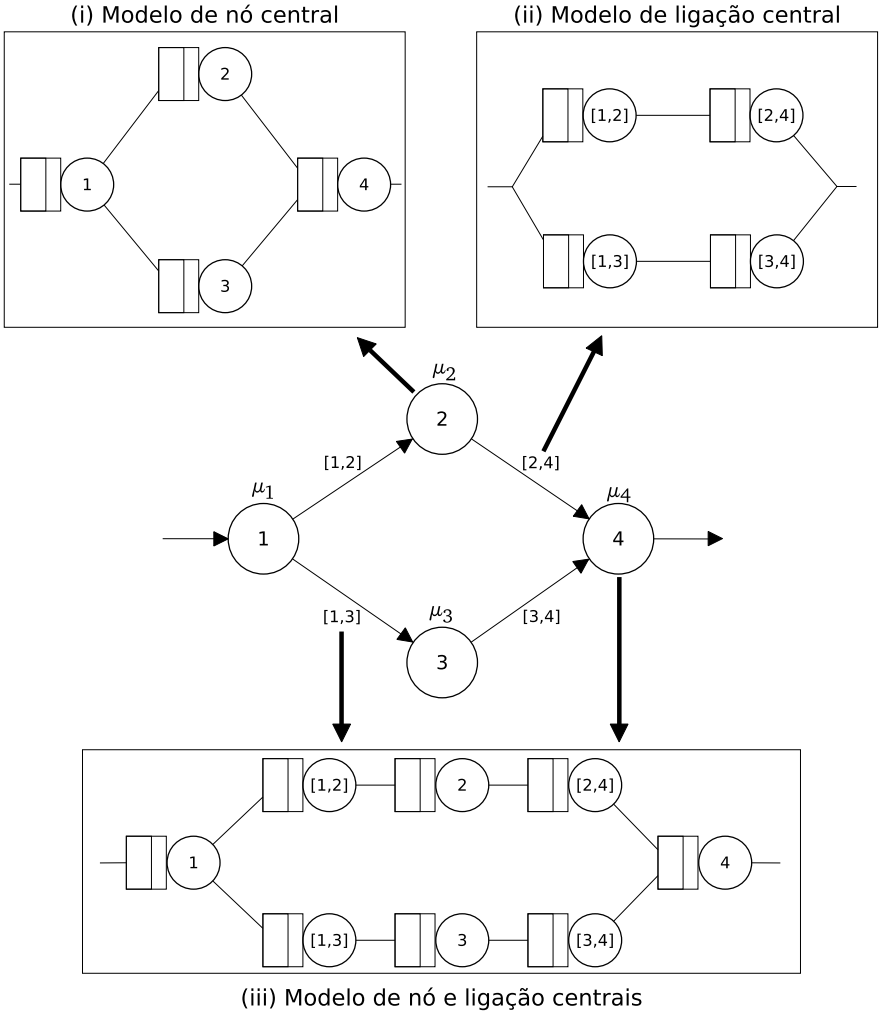

Figura 1.13. Modelo centrado no nó e/ou enlace, exemplo com 4 nós

As medidas de desempenho: os valores esperados

A medida de interesse, M, da seção 1.2, é obtida como medidas de recompensa de utilização de recurso de um modelo de cadeia de Markov de tempo contínuo (CTMC). Suponha que cada nó é uma fila $M/M/1/n_i$. Com Γ_i e μ_i independente do estado x_i e seja $\rho_i = \Gamma_i/\mu_i$, então a probabilidade de estado estacionário $\pi_i(x_i)$ do estado x_i no nó i tem a seguinte fórmula de solução fechada [Trivedi 2002] (se $n_i \to \infty$, então $\rho^{n_i} \to 0$).

$$\mu_i(x_i) = \frac{1-\rho_i}{1-\rho^{n_i+1}}\rho_i^{x_i}. \tag{6}$$

As métricas de desempenho são: perdas esperadas, vazão e atraso (média e distribuição). A taxa de perda transitória é denotada por $L(t)$ no instante t, a taxa de perda (probabilidade aka) $l(t)$ no instante t, o número de pacotes do sistema de $N(t)$ no instante t, e a média de atraso fim a fim de pacotes que não são perdidos no caminho, $D(t)$. Nos modelos de resiliência nas seções seguintes o CTMC de estado composto $(y, \overrightarrow{x})$ constitui

a fase y e $\overrightarrow{x} = (x_1, ..., x_n)$, onde x_i é o número de pacotes no nó i. As taxas de recompensa são atribuídas ao nó i da seguinte maneira:

1. Para cálculo da taxa de perda: $f_{L_i}(t, y, x_i) = \begin{cases} \Gamma_i(y) & \text{se } (x_i = n_i) \text{ no tempo } t, \\ 0 & \text{caso contrário} \end{cases}$
2. Para o cálculo da média do número de pacotes: $f_{N_i}(t, y, x_i) = x_i$,

Onde n_i é o número máximo de pacotes e $\Gamma_i(y)$ é a taxa de chegada, para o nó i em fase y. Uma vez que as probabilidades transientes, $p(t, y, \overrightarrow{x})$, são obtidas a partir dos modelos compostos CTMC, então as métricas de desempenho se tornam:

1. Taxa de perda total esperada no tempo t:

$$E[L(t)] = \sum_{y=1}^{IV} \sum_{i=1}^{n} \sum_{x_i=0}^{n_i} f_{L_i}(t, y, x_i) p(t, y, \overrightarrow{x}) \tag{7}$$

2. Probabilidade de perda total esperada no tempo t:

$$E[L(t)] = E[L(t)]/\gamma \tag{8}$$

3. Número total de pacotes esperados na rede no tempo t:

$$E[N(t)] = \sum_{y=1}^{IV} \sum_{i=1}^{n} \sum_{x_i=0}^{n_i} f_{L_i}(t, y, x_i) p(t, y, \overrightarrow{x}) \tag{9}$$

4. Atraso total esperado por pacotes não perdidos no tempo t:

$$E[D(t)] = E[N(t)]/\gamma(1 - E[l(t)]) \tag{10}$$

As medidas de desempenho: distribuição de atraso

Para avaliar a distribuição de atraso unidirecional de pacotes em uma conexão virtual usa-se o bloco de tempo de resposta, abordagem descrita em [Trivedi 2002]. A ideia básica consiste em utilizar o conhecimento do *atraso de sentido único* (também conhecido como o tempo de resposta, em [Trivedi 2002]) a distribuição de um único nó na rede para construir um CTMC onde a distribuição de atraso de sentido único para um caminho entre dois pontos é igual ao tempo para absorção desta CTMC. O encaminhamento de cada caminho v é dado em uma matriz de encaminhamento, $\mathbf{R}_{(v)} = \{r_{ij}^{(v)}\}$ em que $r_{ij}^{(v)}$ é a probabilidade de roteamento de pacotes de um nó i para j por um caminho $v, v = 1, ..., \vartheta\varphi$. Em um esquema de roteamento de estado do enlace, tipicamente de acordo com o número mínimo de saltos de custo mínimo, ou com base em métricas estáticas. Em esquemas de roteamento estocásticos, os caminhos são direcionados ao longo de múltiplos caminhos pelos quais a carga é compartilhada. Roteamento estocástico não é mais usado nas redes

atuais, mas estão sob considerações em engenharia de tráfego MPLS [Awduche et al. 1999] e em exames baseados em esquemas de roteamento [Liu and Trivedi 2006, Liu 2008].

A abordagem de bloco que é descrito nesta seção assume uma rede de fila markoviana aberta $M/M/1$. As aproximações para redes markovianas , e não-markovianas também existem, ver [Trivedi 2002] para mais detalhes. Em [Heegaard and Trivedi 2009] estende-se o método de bloco para aplicar a rede com múltiplos caminhos. O método é apresentado em quatro etapas:

1. Cálculo das taxas de chegada de cada nó na rede de filas para v caminhos, resolvendo o sistema de equações lineares de tráfego da equação 4.

2. Criação de uma CTMC com estados $S = \left\{S_f, S_l, \cup_{i=1}^{n}, S_i\right\}$ onde S_f é a absorção de estado para qualquer caminho o S_l é o estado de perda em que os pacotes são roteados, se $r_{ij}^{(v)} > 0$, enquanto o enlace $[i, j]$ ou o nó j é inativo, e S_i é o estado do nó i. A CTMC tem $n+2$ estados onde n é o número de nós na rede. Na Figura 1.14 um exemplo de CTMC de rede com 11 nós e dois caminhos são dados.

3. A distribuição de atraso no nó i pode ser representado por uma distribuição exponencial com taxa $\mu_i - \Gamma_i$, dado que $\Gamma_i < \mu_i$, onde Γ_i é a taxa de chegada de todos os caminhos e μ_i é a taxa de serviço para i. O gerador da Matriz $\mathbf{Q}_{(v)}$ da CTMC acima é construída por:

   (a) O roteamento do nó i para o nó j por um caminho v é:
   $\mathbf{Q}_{S_i,S_j}^{(v)} = (\mu_i - \Gamma_i) r_{ij}^{v}$

   (b) A saída da rede no destino $i = d_v$ de um caminho v é:
   $\mathbf{Q}_{S_i,S_j}^{(v)} = (\mu_i - \Gamma_i) r_{ij}^{v}$

   (c) Se $r_{ij}^{(v)} > 0$ e o enlace $[i, j]$ ou o nó j estão inativos, então a taxa de transmissão do estado de perda é: $\mathbf{Q}_{S_i,S_j}^{(v)} = (\mu_i - \Gamma_i) r_{ij}^{v}$ e $\mathbf{Q}_{S_i,S_j}^{(v)} = 0$

   (d) Todas as outras entradas são 0, exceto as diagonais :
   $\mathbf{Q}_{S_i,S_j}^{(v)} = -\Sigma_{S_j \in S_i, S_j \neq S_i} \mathbf{Q}_{S_i,S_j}^{(v)}$

4. A função de distribuição de atraso cumulativa de um caminho v é igual à probabilidade $\mathbf{P}_{Q^{(v)}}(S_f; u)$ de estar no estado de absorção em S_f no tempo u sob a matriz geradora $\mathbf{Q}^{(v)}$. A probabilidade $\mathbf{P}_{Q^{(v)}}(s; u), s \in S$ são encontradas resolvendo $(\mathbf{P}_{Q^{(v)}}(u) = \left\{\mathbf{P}_{Q^{(v)}}(s; u)\right\}) \frac{d}{du}\mathbf{P}_{Q^{(v)}}(u) = \mathbf{P}_{Q^{(v)}}(u) \cdot \mathbf{Q}^{(v)}$

   As condições iniciais são $\mathbf{P}_{Q^{(v)}}(s; 0) = 1$ se $s = s_v$ é o nó origem S_v de um caminho v, e $\mathbf{P}_{Q^{(v)}}(s; 0) = 0$ caso contrário. Observe que se os pacotes são perdidos e a distribuição de atraso dos pacotes é defeituosa, isto é $\mathbf{P}_{Q^{(v)}}(S_f, u) < 1$ porque $\mathbf{P}_{Q^{(v)}}(S_f; u) > 0$ quando $u \to \infty$ [Trivedi 2002].

Os cálculos acima são exatos no caso de uma rede aberta de forma que o produto é alimentado para a frente, e todos os caminhos estão livres de ultrapassagem. Como

exemplo,o caminho *C1* na Figura 1.14 é de ultrapassagem livre, enquanto *C2* tem um caminho reconvergente, portanto, não é de ultrapassagem livre.

Com intervalo de chegadas não exponencial ou tempo de serviço, as suposições a abordagem é aproximada. Um exemplo está incluído neste trabalho para ilustrar a aproximação. Mais detalhes sobre este método e a generalização para filas $M/M/c/b$ e redes não-markovianas são encontrados em [Trivedi 2002].

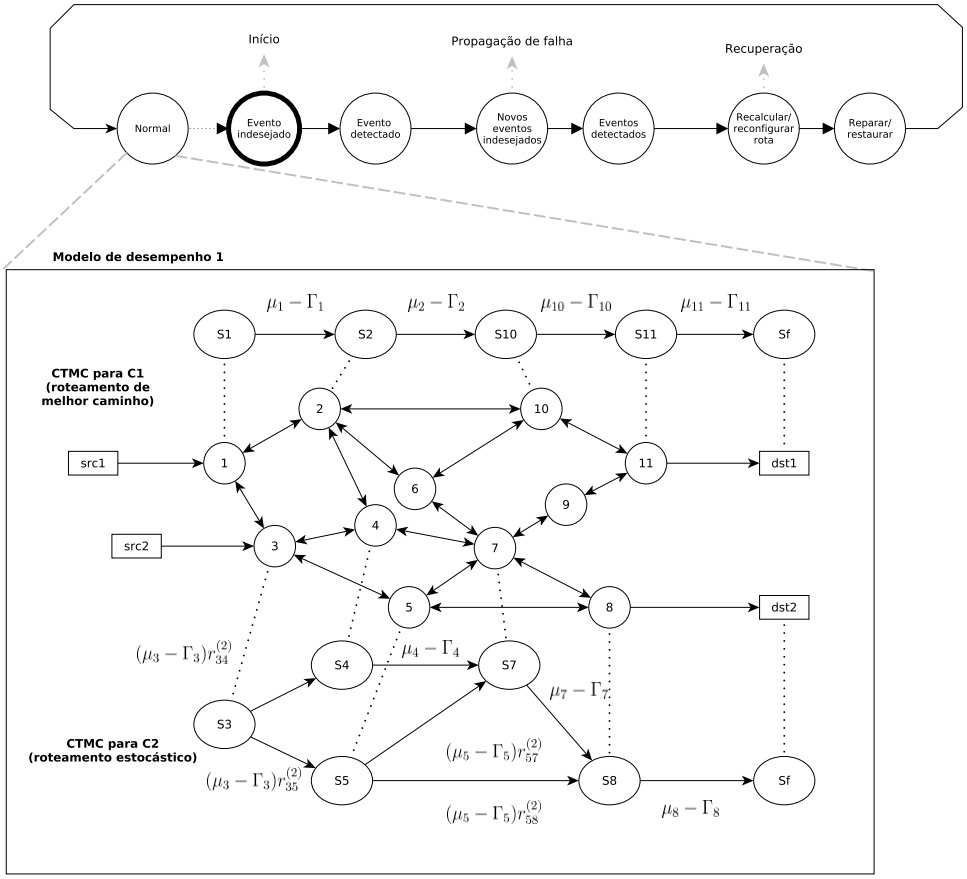

Figura 1.14. Especificação de um CTMC para uma abordagem bloco de atraso em uma rede com duas conexões virtuais

Modelo de propagação de falhas e de recuperação

A seguir, são exemplificadas as abordagens para modelar a propagação de falhas e a recuperação gradual com base no conhecimento dos processos de propagação e mecanismos de recuperação [Heegaard and Trivedi 2008a], ou em traços reais [Heegaard and Trivedi 2008b]. Em [Heegaard and Trivedi 2008a], o modelo de fase de recuperação e retransmissão foi introduzido. A ideia é descrever a sequência de eventos com o modelos CTMC em 4 fases em que cada uma delas representa diferentes estágios de roteamento na matriz de atualização. Isto requer o conhecimento do mecanismo de detecção e processo de (re) encaminhamento.

O modelo de fase de descoberta descreve o "ciclo" iniciando de um evento instável que causa falha em um ou múltiplos enlaces e nós, até que o sistema volte a um estado um pouco antes deste evento. Isso pode ser modelado por fases onde cada fase pode ter um conjunto diferente de recursos disponíveis para caminhos, representados por probabilidades de encaminhamento estacionário fase-dependente $\left\{r_{ij}(y)\right\}$ com correspondente taxa de chegada fase-dependente $\Gamma_i(y)$. Uma abordagem similar é encontrada em [Wang et al. 1996]. Na Figura 1.15, o ciclo de vida de uma falha e o reencaminhamento é descrito em quatro fases, I, II, III, IV.

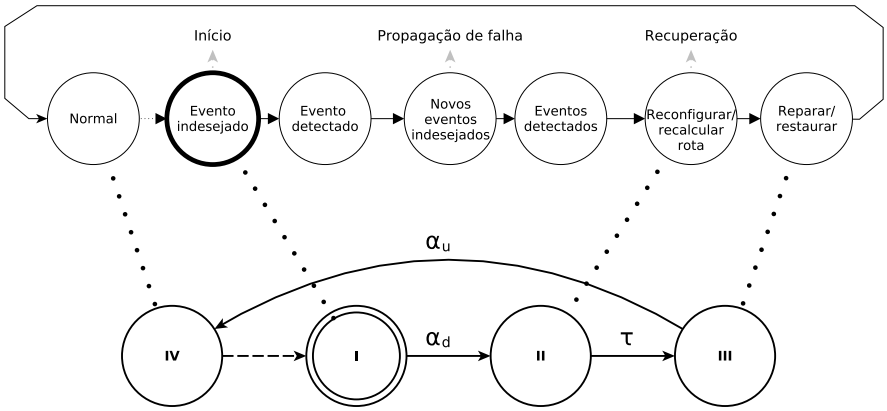

Figura 1.15. Modelo de fases de recuperação de reencaminhamento e restauração

Fase I É ativada imediatamente depois da falha de reencaminhamento mas isso leva algum tempo antes do reencaminhamento ser eficaz. Ao mesmo tempo, os pacotes são encaminhados de acordo com o esquema de rota original, $r_{ij}(I) = r_{ij}(IV)$, exceto pela falha do nó i e enlace $[i, j]$ onde $r_{ij}(I) = 0$, isto é, nenhum pacote é encaminhado para frente a partir deste nó e enlace. O tempo de reencaminhamento é assumido como sendo exponencialmente distribuído com taxa α_d.

Fase II Quando o reencaminhamento é eficaz e o nó ou enlace ainda é falho. Os pacotes são encaminhados de acordo com o novo esquema de encaminhamento e irá evitar esses enlaces e nós falhos (se possível). Depois de reparar o sistema com o tempo exponencialmente distribuído com taxa τ entra na Fase III.

Fase III Após a conclusão da reparação o sistema retorna para um estado livre de falhas, mas o encaminhamento ainda pode mudar. Após o tempo de reencaminhamento com distribuição exponencial e taxa μ_u, o sistema volta para a o encaminhamento normal na fase IV. As fases II e III podem ter probabilidades idênticas de encaminhamento com $r_{ij}(II) = r_{ij}(III)$. Neste caso os dois estados podem ser agrupados para reduzir o espaço de estados e o modelo será semi-markoviano.

Fase IV Depois do encaminhamento ser restaurado a rede funciona em modo livre de falha, que é um estado de absorção para fins de análise de sobrevivência [Liu and Trivedi 2006].

Esse modelo não deve implicar que somente um evento de falha em um tempo possa ocorrer, um evento pode ser uma combinação de múltiplas e simultâneas falha de nós e enlaces. Entretanto, o modelo de fase de recuperação pode ser facilmente modificado para refinar as fases de reencaminhamento para também modelar mudanças graduais nas probabilidades de roteamento ou nas etapas do gerenciamento de caminhos ou para modelar outros modos de falha, como falhas de enlaces intermitente.

No modelo de fase de recuperação na Figura 1.15, o sistema inicia na fase I e caminha sobre todas as fases antes de retornar para a fase IV. A probabilidade transiente $p(t,y)$ no tempo t das quatro fases $y = I,...,IV$ pode ser obtida de uma forma fechada pela abordagem de integração de convolução [Trivedi 2002] em suposição $\alpha_d \neq \tau \neq \alpha_u$. Para o caso de $\alpha_d = \alpha_u$ a solução é simples.

$$\begin{aligned} p(t,I) &= e^{-t\alpha_d}, \\ p(t,II) &= \frac{\alpha_d}{\alpha_d - \tau}(e^{-t\tau} - e^{-t\alpha_d}), \\ p(t,III) &= \frac{\alpha_d \tau}{\alpha_d - \tau}\left(\frac{e^{-t\alpha_d} - e^{-t\alpha_u}}{\alpha_d - \alpha_u} - \frac{e^{-t\alpha_u - e^{-t\tau}}}{\alpha_u - \tau}\right), \\ p(t,IV) &= 1 - p(t,I) - p(t,II) - p(t,III) \end{aligned} \tag{11}$$

Com um exemplo de como o modelo de desempenho relacionado com o modelo de fase de recuperação gradual, considere o caminho *1(C1)* na Figura 1.14. Na Figura 1.16 o enlace $[2,10]$ falha. A falha no enlace é parte do caminho *C1* e causará 100% de perda de pacotes neste estágio. Como mostra a Figura 1.17, depois de algum tempo a existência de caminhos alternativos é provida, o protocolo de roteamento obtém um novo caminho para *C1* que não contem falha no enlace. A perda de pacotes é 0% novamente pois os *buffers* são infinitos.

Modelo de recuperação de rastreamento

Em [Heegaard and Trivedi 2008b], a fase de recuperação é modelada por uma simples monitoração e gravação da matriz de roteamento, $\mathbf{R}^{v,p}$, para o caminho $v(v = 1,...,\vartheta\varphi)$, na fase $p(p = 1,...,\wp\Im)$ em diferentes instantes de tempo iniciando no tempo $t = 0$ de eventos indesejados. A matriz pode ser registrada a partir de uma rede operacional ou a partir de simulações.

O modelo de recuperação consiste de $\wp\Im$ fases, onde cada fase representa uma condição da rede. Seja $S_p = \{1,...,\wp\Im\}$ sendo o estado em um modelo de fase de recuperação. A solução transiente $P(S_p;t), p = 1,...\wp\Im$, é obtida pela resolução da equação 12.

$$\frac{d}{dt}P(S_p;t) = -\alpha_p P(S_p;t) + \alpha_{p-1}P(S_{p-1};t), \tag{12}$$

Em que cada fase de tempo é exponencialmente distribuída com taxa $\alpha_p(\alpha_0 = 0)$ e condição inicial $P(S_1;0) = 1$ e $P(S_p;0) = 0 \forall p > 1$.

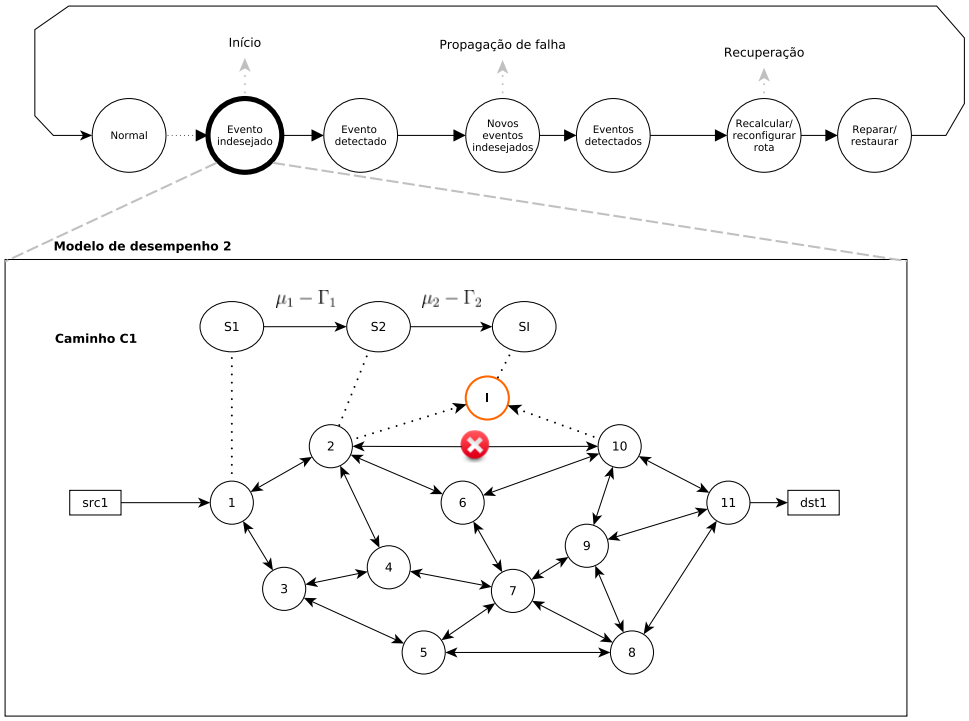

Figura 1.16. CTMC de uma rede com duas conexões virtuais onde o enlace [2,10] falham e todos os pacotes são perdidos

Modelo de decomposição de espaço

É desafiante obter a probabilidade transiente $p(t,y,\overrightarrow{x})$ pois o espaço de estado torna-se enorme com o aumento do tamanho da rede. Desta forma aplica-se a decomposição de aproximação do espaço e a abordagem de modelo transiente para cada nó separadamente. As probabilidades globais são aproximadas pelo produto $p(t,y,\overrightarrow{x}) = p_i(t,y,x_1)\cdots p_n(t,y,x_n)$ onde $p_i(t,y,x_i)$ é a probabilidade transiente de x_i pacotes no nó i no tempo t na fase y. Isso é similar para a solução da forma produto de Jackson [Jackson 1957] ou rede BCMP [Baskett et al. 1975]. A decomposição de espaço separa o modelo de sobrevivência em modelos de nó (e enlace) independentes e obtém a taxa de chagada Γ_i para o nó i pela resolução de um conjunto de equações (1) de tráfego. A dinâmica do nó depende de um enlace ligado a este nó, ou o próprio nó, falhou ou não. Para dar um exemplo considere o nó 4 na Figura 1.18 onde o nó $j = 2$ falha. Para isso os dois seguintes modelos CTMC descrever o nó falho $(j = 2)$ e os nós não falhou $(i = 1,3,4)$.

1. O modelo CTMC para *nós falhos* (ver Figura 1.20). Imediatamente depois de um evento indesejado todos os pacotes enviados para o nó j são perdidos, portanto todas as transmissões levadas ao estado do nó j serão perdidas, deste modo, todas as transmissões levadas ao estado (I,n_j) onde todos os recursos estão indisponíveis e nenhum pacote será entregue.

2. O Modelo CMTC para nós não falhos (ver Figura 1.21). Imediatamente após um evento indesejado o estado da rede é trocado de (IV,x_i) para (I,x_i). Deste modo,

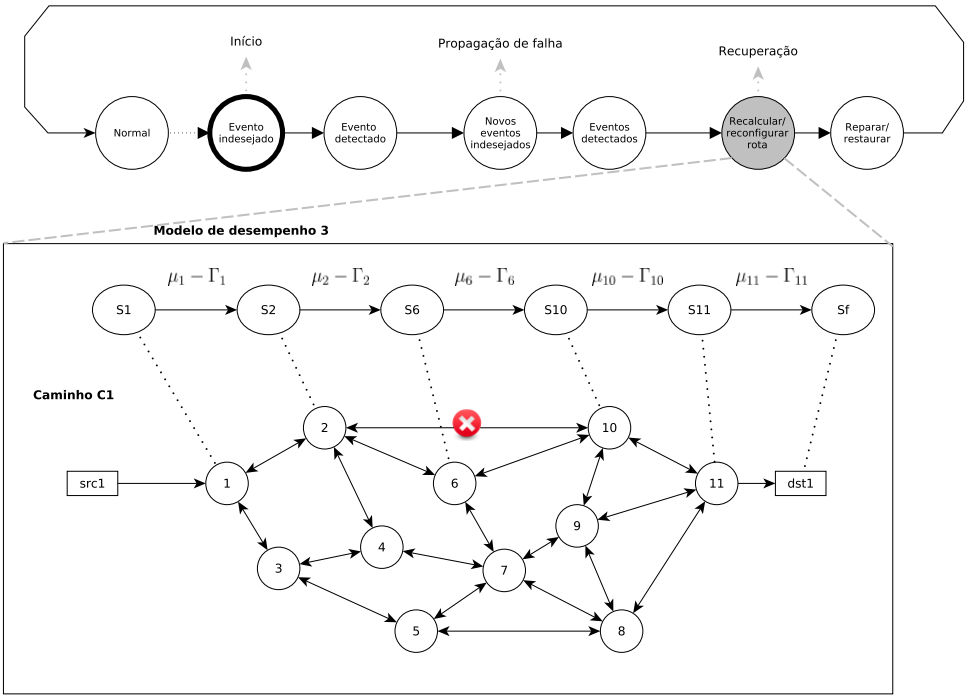

Figura 1.17. CTMC de uma rede com duas conexões virtuais onde o enlace [2,10] falham e os pacotes são reencaminhados

nenhum pacote será perdido mas a taxa de chegada $\Gamma i(IV)$ será alterada para $\Gamma i(I)$. Para alguns nós, as taxas de chegadas serão inalteradas, mas para os nós usados para receber pacotes do nó j a taxa de chegada será reduzida. Na fase II o reencaminhamento é completado e a taxa de chegada $\Gamma_i(II)$ dependendo da posição do i relativo a j e a probabilidade de encaminhamento são dadas pela equação (4).

No modelo CTMC de falha do nó j a $p_i(t,y,x_j)$ é obtida com a condição inicial $p_j(0,I,n_j) = 1$. Enquanto para os nós não falhos $(i \neq j)$, a $p_i(t,Y,x_i)$ é obtida com a condição inicial $p_i(0,I,x_i) = \pi_i(x_i)$. O $\pi_i(x_i)$ é a probabilidade de estados estáveis da equação 6. Finalmente a probabilidade global de estado são obtidas pela fórmula de aproximação do produto da equação 13.

$$p(t,y,\overrightarrow{x}) \approx \prod_{i=1}^{n} p_i(t,y,x_i) \tag{13}$$

Não há nenhuma maneira fácil de obter soluções de forma fechada de $p_i(t,y,x_i)$ dos modelos na Figuras 1.20 e 1.21. Mas, soluções podem ser obtidas por meio da ferramenta SHARP [Sahner et al. 1996, Hirel et al. 2000] e SPNP [Ciardo et al. 1993] para sistemas muito grande. Entretanto, com o tamanho da rede ou incrementando o número de *buffers* n_i o modelo do nó aumenta causando um modelo de recuperação mais sofisticado, a solução torna-se muito difícil e com recursos lento.

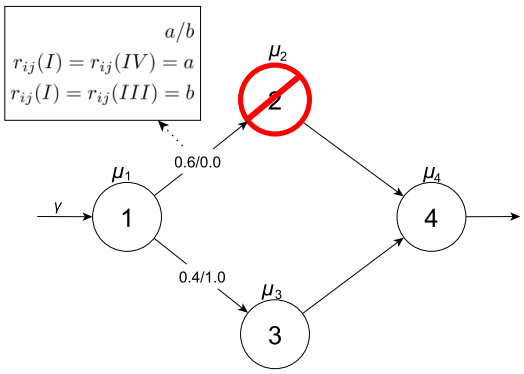

Figura 1.18. Exemplo de rede com 4 nós

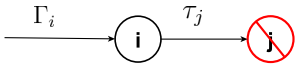

Figura 1.19. Exemplo com falha no nó j

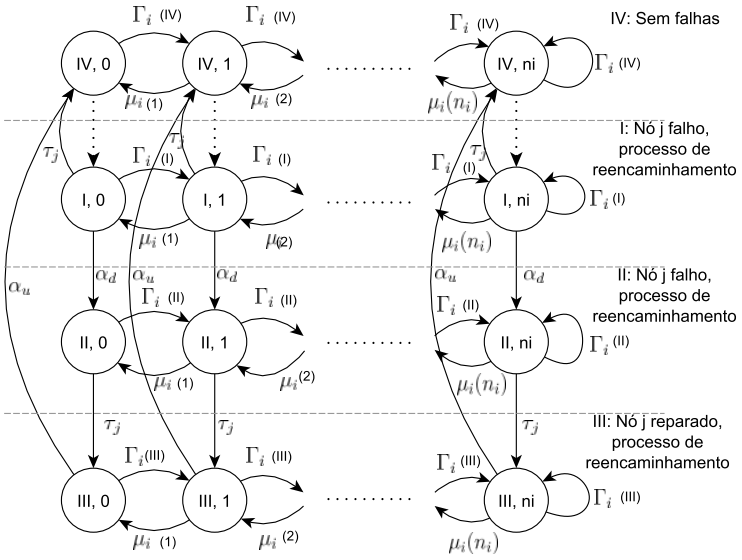

Figura 1.20. Modelo CTMC de nós não falhos

Modelo de decomposição de tempo-espaço

A decomposição do espaço melhora a escalabilidade do modelo significativamente. Entretanto, mesmo o modelo CTMC de um único nó (enlace) pode ser muito complexo para uma solução fechada simbólica, e também muito grande para uma solução numérica. A seguir descreve-se a decomposição do tempo [Meyer 1980], onde assumimos o desempenho do estado estável em cada estado pelo qual requer somente solução transiente de um modelo de recuperação de fase.

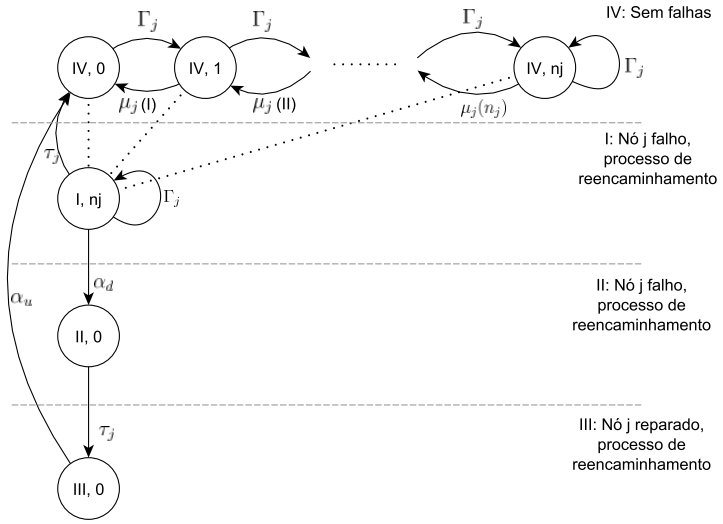

Figura 1.21. Modelo CTMC de nós falhos

A decomposição do tempo é uma dissociação do desempenho e modelo de recuperação. Isto significa que as probabilidades de estado estáveis nos modelos de desempenho e da solução transiente do modelo de recuperação gradual são obtidos separadamente e independentemente uns dos outros, e $p_i(t,y,x_i) \approx p(t,y) \cdot \pi_i(x_i,y)$. As probabilidades transiente $p(t,y)$ são da equação 11, enquanto as probabilidades do estado estável $\pi_i(x_i,y)$ são $\pi_i(x_i)$ da equação6 para diferentes taxas de chegada fase-dependente $\Gamma_i(y)$ e matriz de roteamento fase-dependente. Observe que assumimos desempenho de estado estável em casa fase. A aproximação é boa quando ocorre uma mudança de fase, o desempenho do estado estável em uma nova fase é alcançado rapidamente em comparação com a duração da fase. Neste caso, em nosso modelo de rede quando $(\Gamma_i, \mu_i) \gg (\alpha_d, \tau, \alpha_u)$. A qualidade da decomposição do tempo da aproximação do modelo de Markov depende do grau de acoplamento entre o "bloco de desempenho" e o "bloco de fase " na nossa matriz de Markov [Meyer 1980, Courtois and Courtois 1977, Bobbio and Trivedi 1986]. Então as probabilidades globais são obtidas pela aproximação usando tanto a decomposição do espaço, equação 13, e a decomposição do tempo e espaço.

$$p(t,y,\overrightarrow{x}) \approx \prod_{i=1}^{n} p_i(t,y,x_i) \approx p(t,y) \cdot \prod_{i=1}^{n} \pi_i(x_i,y) \tag{14}$$

O modelo de recuperação de fases é exemplificado na Figura 1.15 ou tem como base em *traces* reais. Isso permite a avaliação eficiente de redes muito grandes, e até mesmo infinitos servidores e capacidades de *buffer*, n_i. Com um modelo de recuperação de fases muito mais complexo, uma solução de forma fechada transiente é difícil ou impossível, as abordagens de decomposição de tempo ainda são vantajosos uma vez que a solução numéricas são significativamente mais rápida em comparação com o modelo de nó, porque o número de estados é reduzido.

A decomposição do tempo permite que as métricas de desempenho sejam obtidas

por determinação independente do desempenho $P_{Q^{v,p}}(S;u)$ de um caminho em cada fase, com a matriz geradora $\mathbf{Q}^{(v,p)}$, e a probabilidade transiente $P(S_p,t)$ da equação 12. Isso reduz significativamente a complexidade computacional e permite mais fases serem incluídas no modelo de recuperação, e ao mesmo tempo, pode aumentar o número de nós e dos enlaces da rede.

Modelos escalabilidade e hipóteses

O principal objetivo das aproximações propostas neste trabalho é reduzir o esforço computacional de obtenção de soluções transitórias em modelos de grandes redes sem perda indevida na precisão. O modelo subjacente CTMC da rede estocástica de recompensa tem um espaço de estado, que é proporcional para $\prod_{i=1}^{n} n_i \times n_p$, onde n_p é o número de fases e n é o número de componentes da rede, isto é, os nós ou enlaces. O modelo CTMC de espaço decomposto reduzirá a solução de espaço de estado transiente para $\sum_{i=1}^{n} n_i \times n_p$. Enquanto, para o modelo de decomposição de tempo-espaço um modelo com somente n_p estados necessita ser resolvido.

1.5. Frameworks e estratégias existentes

Nesta seção descreveremos uma seleção de *frameworks* e estratégias existentes na literatura para resiliência de redes. Primeiramente, muitos *frameworks* analisaram a confiabilidade, a capacidade de sobrevivência e a performabilidade. Assim, serão apresentadas algumas dessas estratégias anteriores e as novas abordagens utilizadas para as redes sem fio heterogêneas, como é o caso da estratégia ResiliNets.

1.5.1. Estratégias anteriores

Existem várias estratégias sistemáticas de resiliência. Contudo, devido a época de concepção, essas estratégias possuíam desafios diferentes dos desafios atuais, principalmente se tratando de redes heterogêneas. Deste modo, serão apresentadas as estratégias anteriores ANSA, T1, CMU-CERT e SUMOWIN, que apesar de serem referência ainda não se preocupavam com a o fator da heterogeneidade da rede.

ANSA

O projeto ANSA (do inglês, *Advanced Networked Systems Architecture*) [Edwards 1994] cobriu vários aspectos de projeto de grandes sistemas, incluindo confiabilidade. A estratégia de gerenciamento de confiabilidade consiste de oito estágios [D.P. Siewiorek 1992]: reclusão de faltas, detecção de faltas (propriamente detecção de erro/falha), diagnóstico de faltas, reconfiguração, recuperação, reinício, reparo, e reintegração. O *framework* ANSA define regiões de expectativa em um espaço bidimensional valor x tempo para descrever serviço aceitável, e consequentemente considera a performabilidade. As falhas de serviço são as incompatibilidades entre uma ocorrência nesse espaço e nas regiões de expectativas.

T1

O grupo de pesquisa T1A1.2 em desempenho da capacidade de sobrevivência de redes, da ATIS (do inglês, *Alliance for Telecommunications Industry Solutions*), desenvolveu um *framework* multi nível para a sobrevivência da rede [Group 1993] com quatro camadas: *1)* física, que consiste em infraestrutura com diversidade geográfica para capacidade de sobrevivência; *2)* sistema, que consiste em nós e enlaces com comutação de proteção para capacidade de sobrevivência; *3)* lógica, que consiste em capacidade na camada de sistema; e *4)* serviço, que consiste em circuitos de voz e dados com reconfiguração e roteamento dinâmicos para prover sobrevivência. Esse *framework* quantifica falhas de serviço como uma tripla (I, D, E), em que I é a incapacidade de confiança (uma métrica inversa à confiabilidade, como indisponibilidade ou falha), D é a duração em tempo, E é a extensão sobre vários parâmetros incluindo área geográfica, população, e serviços. As categorias de severidade dessa tripla mapeada em um espaço tridimensional são categorizadas como, pequena, grande ou catastrófica.

CMU-CERT

O centro de coordenação CERT na CMU propôs uma estratégia de quatro passos [Ellison et al. 1997] consistindo de três $R's$: *1)* resistência segurança tradicional, diversidade, redundância, especialização, validação de confiança e propriedades estocásticas observadas; *2)* reconhecimento redundância e teste analíticos, monitoramento de intrusão, comportamento do sistema, monitoramento de integridade; *3)* recuperação redundância, localização variada dos recursos de informação, planejamento de contingente e equipes de resposta; seguidos por *4)* adaptação e evolução.

SUMOWIN

O projeto SUMOWIN (do inglês *Survivable Mobile Wireless Networking* explorou mecanismos e estratégias para tolerância a interrupção (antes do termo ser generalizado) em ambientes onde conectividade fim a fim estável não era possível [Sterbenz et al. 2002]. A estratégia consiste de três componentes: *1)* manter a conectividade quando possível usando técnicas como controle de poder de transmissão adaptativa e técnicas de MANETs altamente dinâmicas (estabilidade eventual); *2)* usar novas técnicas de encaminhamento e roteamento que não requeiram a convergência de roteamento para mover dados em direção ao destino, usando técnicas de armazenar e encaminhar, como armazenar e encaminhar via *bufferização* e armazenar e transmitir (armazenar, transportar e encaminhar ou transportar) em direção ao destino quando caminhos fim a fim estáveis não puderem ser mantidos (conectividade eventual); *3)* usar tecnologias inovadoras como redes de satélites e redes ativas (programáveis adaptavelmente) para estabilizar conectividade e manter reservas.

1.5.2. Estratégia recente - ResiliNets

A iniciativa ResiliNets [Sterbenz and Hutchison 2008] desenvolveu um framework para redes resilientes [Scholler et al. 2006], inicialmente como parte dos projetos ANA (do inglês, *Autonomic Network Architecture*) [ANA 2006, Bouabene et al. 2010] e PoMo (do inglês, *Post-modern Internet Architecture*) [Bhattacharjee et al. 2006, Sterbenz et al. 2008], servindo como a base para o projeto ResumeNet (do inglês, *Resilience and Survivability for Future Networking: Framework, Mechanisms, and Experimental Evaluation*) [Res 2009, Scholler et al. 2010]. Essa iniciativa foi fortemente influenciada pelos *frameworks* descritos acima, e pode ser vista como um sucessor e síntese de todos eles. O *framework* ResiliNets é descrito por um conjunto de axiomas e uma estratégia.

Axiomas do ResiliNets

Os axiomas do ResiliNet provêm a base para qualquer *framework* sistemático, em [Sterbenza et al. 2010] são apresentados quatro pontos que formam a base para a estratégia do ResiliNets.

A1. Faltas são inevitáveis; não é possível construir sistemas perfeitos, nem é possível evitar desafios e ameaças.

A2. É necessário entender a operação normal, incluindo o ambiente e as demandas da aplicação. É somente pelo entendimento da operação normal que temos qualquer esperança de determinar quando a rede é desafiada ou ameaçada. Define-se operação normal do estado da rede quando não há condições adversas presentes.

A3. Expectativa e preparação são necessários para eventos e condições adversas, de modo que defesas e detecção de desafios que interrompam operações normais possam ocorrer.

A4. Resposta a eventos e a condições adversas é requerida para resiliência, por remediação garantindo operação correta e degradação, restauração à operação normal, diagnóstico da raiz da causa de faltas, e refinação de futuras respostas.

Embora seja necessário esperar eventos e condições adversas, é importante tomar medidas quando desafios realmente ocorrem. Isso motiva o aspecto de remediação da estratégia de resiliência.

Estratégia do ResiliNets

Os axiomas de resiliência motivam a estratégia para resiliência, desenvolvidas como parte dos projetos ResiliNets [Sterbenz and Hutchison 2008], ANA [ANA 2006] e ResumeNet [Res 2009]. A estratégia consiste de duas fases denominadas de $D^2R^2 + DR$, como mostrado na Figura 1.22. No núcleo estão defesas estruturais passivas. A primeira fase ativa, D^2R^2, defender, detectar, remediar, recuperar, compõe o laço de controle interno e descreve um conjunto de atividades que são aplicadas a fim de que um sistema rapidamente

adapte-se a desafios e a ataques e mantenha um nível de serviço aceitável. A segunda fase ativa DR: diagnosticar, refinar, é o laço externo que habilita evolução do sistema a longo prazo a fim de aprimorar as abordagens para as atividades da fase um.

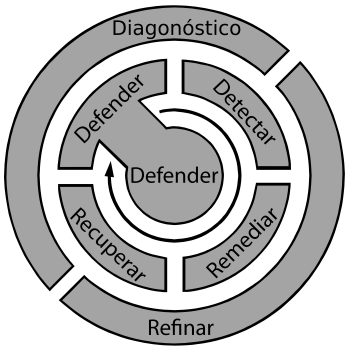

Figura 1.22. Estratégia de duas fases [Sterbenza et al. 2010].

1.6. Implementações existentes na literatura

Com o objetivo de ilustrar implementações de resiliência existentes na literatura, esta seção descreve a avaliação experimental usando um ambiente de testes programável de larga escala chamado de GpENI (do inglês, *Great Plains Environment for Network Innovation*).

1.6.1. Visão geral

O GpENI é um ambiente de testes internacional de redes programáveis. Seu objetivo é construir uma infraestrutura de pesquisa colaborativa permitindo a comunidade de pesquisa a conduzir experimentos na arquitetura da Internet do Futuro. Duas características chave do GpENI necessárias para avaliação experimental são a programabilidade em todos os níveis e a topologia flexível em larga escala.

- A programabilidade é permitida em todos os níveis. Na camada mais alta, o Gush [Gus 2009] provê controle do experimento e o Raven [Rav 2009] distribui o código; ambos são softwares desenvolvidos como parte do programa GENI. A programabilidade das camadas 4 e 7 são providos pela versão do PlanetLab do GENIwrapper [Pla 2009]. Na camada 3, roteadores programáveis são implementados em Quagga [Qua 2009], XORP [XOR 2009], e Click [Cli 2009], complementada por qualquer outra tecnologia que as instituições GpENI decidam implantar. Na camada 2, as configurações dinâmicas VLAN são providas pelo *DCN-enabled managed Gigabit-Ethernet switches* no centro de cada nó *cluster* GpENI. Na camada 1, a arquitetura permite ainda programabilidade em camadas fotônicas em *switches* que provêm tal suporte. Além disso, cada instituição GpENI pode conectar sites de ambientes de testes de redes específicas; planos incluem ambientes de testes de redes sem fio, de sensores e rádio cognitivo. Os usuários externos em comunidades de pesquisa maiores podem pedir contas GpENI com os quais podem executar experimentos de redes.

- A topologia do GpENI é construído em torno do núcleo do *backbone* ótico interligado por universidades e instituições de pesquisa. O GpENI estende-se pela Europa através da Internet2 para GÉANT2 e NORDUnet e então para redes regionais ou nacionais. Atualmente, a conectividade é alcançada usando L2TPv3 e túneis IP. Uma ligação de fibra direta sobre JANET está implantada entre as Universidades de Lancaster e de Cambridge. De modo semelhante, o GpENI estende-se pela Ásia através de Internet2 para APAN, e então para infraestrutura de rede de pesquisa nacional incluindo ERNET, na Índia. Além disso, o GpENI é interligado ao *cluster* ProtoGENI baseado no Emulab [Pro 2010] em Utah, e está implantando diversos pequenos *clusters* ProtoGENI próprios. Assim, o GpENI provê uma topologia rica e de larga escala no qual são feitos experimentos relacionados à resiliência.

1.6.2. Avaliação experimental da resiliência

Com a utilização deste ambiente de testes, pode-se então avaliar o desempenho quando partes do GpENI são desafiadas por falhas correlacionadas de nós e enlaces, medindo a conectividade, a razão de entrega de pacotes, a vazão e o atraso, quando sujeito a taxa de fluxo de dados constante (do inglês, *constant bit rate* ou CBR), transferência de dados em massa e tráfego transacional (HTTP). Pode-se também caracterizar a probabilidade de perda de pacotes de enlaces sem fios na Utah Emulab [Emu 2009], e as capacidades de emular nós problemáticos dentro do *Emulab-federated CMU wireless emulator*.

O objetivo é verificar as configurações idênticas de topologias simuladas e protocolos discutidos para os experimentos GpENI. Estes experimentos terão a vantagem de incorporar redes reais que não são fáceis de emular, embora ainda em escalas menores que grandes topologias de simulação.

1.7. Considerações finais e direções futuras

Esse trabalho apresentou os conceitos e fundamentos relacionados a sobrevivência e resiliência, bem como, exemplos práticos de como essas propriedades são importantes para manutenção da estabilidade dos serviços oferecidos pelas redes e sistemas reais. Outro ponto abordado, foi o desenvolvimento de modelos e métricas para a uma avaliação qualitativa e quantitativa. As taxonomias e exemplos de modelagem foram apresentados, assim como exemplos de estratégias e de as implementações existentes na literatura.

Contudo, o foco principal de discussão dos modelos de resiliência e sobrevivência abordados neste trabalho, foram as redes sem fio heterogêneas. Essas redes, em razão de sua complexidade necessitam de atenção especial por reunirem todos os desafios individuais de cada rede. Com as técnicas de modelagem descritas é possível desenvolver modelos representativos de redes reais, que dificilmente poderiam ser analisadas, em função da heterogeneidade. Os modelos analíticos discutidos, possibilitam mensurar tanto o desempenho dessas redes, como quantificar os índices que representam propriedades mais abstratas, como é o caso da sobrevivência, resiliência e até mesmo a segurança.

Direções futuras

Apesar da existência de muitos trabalhos aplicando técnicas de modelagem direcionadas às redes de computados, poucos são os que se dedicam as redes sem fio heterogêneas. A

modelagem e o projeto de redes sem fio heterogêneas resilientes e sobreviventes, ainda possuem muito espaço para pesquisas e evolução do tema. A criação de modelos analíticos para qualificar e quantificar as propriedades mais abstratas dessas redes também são campos em aberto para discussões futuras, principalmente no que se refere à sobrevivência.

Agradecimentos

O trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), edital MCT/CNPq 09/2010-PDI, número do processo 560184/2010-7.

Este trabalho não poderia ter sido realizado sem a assistência dos integrantes do Núcleo de Redes sem Fio e Redes Móveis - NR2 (http://www.nr2.ufpr.br/) e da equipe do projeto Redes em Malhas Robusta - REMAR (http://www.nr2.ufpr.br/∼remar/), em especial sem o auxílio do Felipe A. N. de Oliveira, Flávio Arieta, Gregory S. Santos e Rodrigo de T. M. Dumont.

Referências

[Nat 2003] (2003). National strategy for the physical protection of critical infrastructures and key assets. `http://www.dhs.gov/xlibrary/assets/Physical_Strategy.pdf`. Último acesso: Novembro 2012.

[CNI 2004] (2004). Report of the commission to assess the threat to the united states from electromagnetic pulse (emp) attack. Technical report, Critical National Infrastructures.

[ANA 2006] (2006). Autonomic network architecture. `http://www.anaproject.org`. Último acesso: Dezembro 2012.

[DHS 2006] (2006). Pandemic influenza preparedness, response, and recovery guide for critical infrastructure and key resources. CI/KR Guide.

[GLA 2006] (2006). Report of the 7 july review committee. Technical report, Greater London Authority.

[DHS 2007] (2007). Pandemic influenza impact on communications networks study. Technical report, Department of Homeland Security (DHS). Unclassified.

[RIP 2008] (2008). Mediterranean fibre cable cut - a ripe ncc analysis. `http://www.ripe.net/projects/reports/2008cable-cut/`. Último acesso: Dezembro 2012.

[NRC 2008] (2008). Severe space weather events: Understanding societal and economic impacts. Technical report, National Research Council. Workshop Report.

[Wik 2008] (2008). Wikipedia: 2008 submarine cable disruption. `http://en.wikipedia.org/wiki/2008_submarine_cable_disruption`. Último acesso: Novembro 2012.

[Cli 2009] (2009). The click modular router project. `http://read.cs.ucla.edu/click/`. Último acesso: Janeiro 2013.

[Emu 2009] (2009). EMULAB: Network emulation testbed. `http://www.emulab.net/`. Último acesso: Janeiro 2013.

[Gus 2009] (2009). Gush: Geni user shell. http://gush.es.williams.edu/trac/gush. Último acesso: Dezembro 2012.

[Pla 2009] (2009). PLANETLAB. http://www.planet-lab.org/. Último acesso : Janeiro 2013.

[Qua 2009] (2009). Quagga routing route. http://www.quagga.net/. Último acesso : Janeiro 2013.

[Rav 2009] (2009). Raven provisioning service. http://raven.cs.arizona.edu/. Último acesso: Dezembro 2012.

[Res 2009] (2009). Resumenet. http://www.resumenet.eu/project/index. Último acesso: Novembro 2012.

[XOR 2009] (2009). XORP: Extensible open-source routing platform. http://www.xorp.org/. Último acesso: Dezembro 2012.

[Pro 2010] (2010). Protogeni wiki. http://www.protogeni.net/trac/protogeni. Último acesso : Janeiro 2013.

[A. Popescu 2008] A. Popescu, B. Premore, E. Z. (2008). Impact of the middle east cable breaks: A global BGP perspective. Presentation.

[Akyildiz et al. 2002] Akyildiz, I., Su, W., Sankarasubramaniam, Y., and Cayirci, E. (2002). Wireless sensor networks: a survey. *In Proceedings of the Computer Networks*, 38(4):393–422.

[Amiri and Ghassemi-Tari 2007] Amiri, M. and Ghassemi-Tari, F. (2007). A methodology for analyzing the transient availability and survivability of a system with repairable components. *Applied Mathematics and Computation*, pages 300–307.

[Avizienis 1967] Avizienis, A. (1967). Design of fault-tolerant computers. In *Proceedings of the Fall Joint Computer Conference*, pages 733–743. American Federation of Information Processing Societies (AFIPS) Spring Joint Computing Conference, Thompson Books.

[Awduche et al. 1999] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and McManus, J. (1999). Requirements for traffic engineering over MPLS - (rfc 2702).

[Baskett et al. 1975] Baskett, F., Chandy, K. M., Muntz, R. R., and Palacios, F. G. (1975). Open, closed, and mixed networks of queues with different classes of customers. *Journal of the Assoclatlon for Computing Machinery (ACM)*, 22(2):248–260.

[Bassiri and Heydari 2009] Bassiri, B. and Heydari, S. (2009). Network survivability in large-scale regional failure scenarios. In *Proceedings of the ACM Second Canadian Conference on Computer Science and Software Engineering (C3S2E)*, pages 83–87.

[Bhattacharjee et al. 2006] Bhattacharjee, B., Calvert, K., Griffioen, J., Spring, N., and Sterbenz, J. (2006). Postmodern internetwork architecture. Technical report, Information and Telecommunication Center.

[Billinton and Allan 1992] Billinton, R. and Allan, R. (1992). *Reliability Evaluation of Engineering Systems*. Plenum Press.

[Bobbio and Trivedi 1986] Bobbio, A. and Trivedi, K. S. (1986). An aggregation technique for the transient analysis of stiff markov chains. *IEEE Transaction on Computer*, 35(9):803–814.

[Bouabene et al. 2010] Bouabene, G., Jelger, C., Tschudin, C., Schmid, S., and May, A. K. M. (2010). The autonomic network architecture (ANA). *IEEE Journal on Selected Areas in Communications (JSAC)*, 28(1):4–14.

[Brown 2008] Brown, M. (2008). Pakistan hijacks youtube. `http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml`. Último acesso: Dezembro 2012.

[Brown and Zmijewski 2008] Brown, M. and Zmijewski, E. (2008). Pakistan telecom hijacks youtube: Or how to syn-flood DoS yourself while annoying everyone on the planet. In *Asia Pacific Regional Internet Conference on Operational Technologies (APRICOT)*. Presentation.

[Buljore et al. 2009] Buljore, S., Harada, H., Filin, S., Houze, P., Tsagkaris, K., Holland, O., Nolte, K., Farnham, T., and Ivanov, V. (2009). Architecture and enablers for optimized radio resource usage in heterogeneous wireless access networks. *IEEE Communications Magazine Magazine*, 47(1):122 –129.

[Burleigh et al. 2003] Burleigh, S., Hooke, A., Torgerson, L., Fall, K., Cerf, V., Durst, B., Scott, K., and Weiss, H. (2003). Delay-tolerant networking: an approach to interplanetary internet. *IEEE Communications Magazine*, 41(6):128–136.

[Campbell et al. 1994] Campbell, A., Coulson, G., and Hutchison, D. (1994). A quality of service architecture. *SIGCOMM Computer Communication*, 24(2):6–27.

[Carter et al. 2002] Carter, M., Howard, M., Owens, N., Register, D., Kennedy, J., Pecheux, K., and Newton, A. (2002). Effects of catastrophic events on transportation system management and operations. Technical report, US Department of Transportation.

[Chen and Avizienis 1978] Chen, L. and Avizienis, A. (1978). N-version programming: a fault tolerance approach to the reliable software. In *Proceedings of the Eighth International Symposium on Foult-Tolerant Computing*.

[Ciardo et al. 1993] Ciardo, G., Blakemore, A., Chimento, P. F., Muppala, J. K., and Trivedi, K. S. (1993). Automated generation and analysis of markov reward models using stochastic reward nets. *IMA Volumes in Mathematics and its Applications: Linear Algebra, Markov Chains, and Queueing Models / Meyer, C.; Plemmons, R.J*, 48:145–191.

[Clouqueur and Grover 2002] Clouqueur, M. and Grover, W. (2002). Availability analysis of span-restorable mesh networks. *IEEE Journal on Selected Areas in Communications*, 20(4):810–821.

[Courtois and Courtois 1977] Courtois, P. and Courtois, P. (1977). *Decomposability: queueing and computer system applications*. ACM monograph series. Academic Press.

[Cowie et al. 2003] Cowie, J., Ogielski, A., Premore, B., Smith, E., and Underwood, T. (2003). Impact of the 2003 blackouts on Internet communications. Technical report, Renesys Corporation.

[Cowie et al. 2005] Cowie, J., Popescu, A., and Underwood, T. (2005). Impact of hurricane katrina on internet infrastructure. Technical report, Renesys.

[David et al. 2011] David, L., Ismail, G., and k. Marios (2011). Enhanced intercell interference coordination challenges in heterogeneous networks. *IEEE Wireless Communications*, pages 22–30.

[Davis et al. 2006] Davis, T., Rogers, H., Shays, C., and Other (2006). A failure of initiative: The final report of the select bipartisan committee to investigate the preparation for and response to hurricane katrina. Technical report, Congressional Report H. Rpt. US House of Representatives.

[Deutsch and Willis 1988] Deutsch, M. and Willis, R. (1988). *Software quality engineering: a total technical and management approach*. Prentice-Hall series in software engineering. Prentice Hall.

[D.P. Siewiorek 1992] D.P. Siewiorek, R. S. (1992). *Reliable Computer Systems: Design and Evaluation*. Digital Press.

[Duffy 2007] Duffy, J. (2007). Cisco routers caused major outage in Japan: Report. `http://www.networkworld.com/news/2007/051607-cisco-routers-major-outage-japan.html`. Último acesso: Dezembro 2012.

[Durst et al. 1996] Durst, R., Miller, G., and Travis, E. (1996). TCP extensions for space communications. In *Proceedings of the ACM Second Annual International Conference on Mobile Computing and Networking*, pages 15–26. ACM Press.

[Edwards 1994] Edwards, N. (1994). Building dependable distributed systems. Technical report, ANSA.

[Ellinas and Stern 1996] Ellinas, G. and Stern, T. (1996). Automatic protection switching for link failures in optical networks with bi-directional links. In *Proceedings of the IEEE Global Telecommunications Conference (IEEE GLOBECOM)*, pages 152–156.

[Ellison et al. 1997] Ellison, R., Fisher, D., Linger, R., Lipson, H., Longstaff, T., and Mead, N. (1997). Survivable network systems: an emerging discipline. Technical report, Software Engineering Institute, Carnegie Mellon University.

[Fall 2003] Fall, K. (2003). A delay-tolerant network architecture for challenged internets. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pages 27–34. ACM.

[Gan and Helvik 2006] Gan, Q. and Helvik, B. (2006). Dependability modelling and analysis of networks as taking routing and traffic into account. In *Proceedings of the IEEE Second EuroNGI Conference on Next Generation Internet Design and Engineering, Valencia, Spain*.

[Ghosh et al. 2012] Ghosh, A., Mangalvedhe, N., Ratasuk, R., Mondal, B., Cudak, M., Visotsky, E., Thomas, T., Andrews, J., Xia, P., Jo, H., Dhillon, H., and Novlan, T. (2012). Heterogeneous cellular networks: From theory to practice. *IEEE Communications Magazine*, 50(6):54 –64.

[Group 1993] Group, T. W. (1993). Network survivability performance. Technical report, Alliance for Telecommunications Industry Solutions (ATIS).

[Group 2001] Group, T. W. (2001). American national standard for telecommunications. *Alliance for Telecommunications Industry Solutions (ATIS)*, page T1.523.

[Group 2004] Group, T. W. (2004). Reliability-related metrics and terminology for network elements in evolving communications networks. *Alliance for Telecommunications Industry Solutions (ATIS)*.

[Grover 2003] Grover, W. (2003). *Mesh-based Survivable Networks*. Prentice-Hall PTR Pearson.

[Grover and Stamatelakis 1998] Grover, W. and Stamatelakis, D. (1998). Cycle-oriented distributed preconfiguration: ring-like speed with mesh-like capacity for selfplanning network restoration. In *Proceedings of the IEEE International Conference on Communications*, pages 537–543. IEEE Computer Society.

[Heegaard and Trivedi 2008a] Heegaard, P. and Trivedi, K. (2008a). Survivability quantification of communication services. In *Proceedings of the IEEE International Conference on Dependable Systems and Networks With FTCS and DCC (DSN)*, pages 462–471.

[Heegaard and Trivedi 2008b] Heegaard, P. E. and Trivedi, K. S. (2008b). Survivability quantification of real-sized networks including end-to-end delay distributions. In *Proceedings of the 2008 Third International Conference on Systems and Networks Communications*, ICSNC '08, pages 50–55. IEEE Computer Society.

[Heegaard and Trivedi 2009] Heegaard, P. E. and Trivedi, K. S. (2009). Network survivability modeling. *Computer Networks.*, 53:1215–1234.

[Hirel et al. 2000] Hirel, C., Sahner, R. A., Zang, X., and Trivedi, K. S. (2000). Reliability and performability modeling using sharpe 2000. In *Proceedings of the 11th International Conference on Computer Performance Evaluation: Modelling Techniques and Tools*, (TOOLS), pages 345–349. Springer-Verlag.

[Jabbar et al. 2009] Jabbar, A., Rohrer, J., Oberthaler, A., Çetinkaya, E., Frost, V., and Sterbenz, J. (2009). Performance comparison of weather disruption-tolerant cross-layer routing algorithms. In *Proceedings of the IEEE INFOCOM 2009, Conference on Computer Communications*, pages 1143–1151. IEEE Computer Society.

[Jackson 1957] Jackson, J. R. (1957). Networks of Waiting Lines. *Operations Research*, 5(4):518–521.

[Jen 2005] Jen, E. (2005). *Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*. Oxford University Press.

[Johnson 1994] Johnson, D. (1994). Routing in ad hoc networks of mobile hosts. In *Proceedings of the First Workshop on Mobile Computing Systems and Applications (WMCSA)*, pages 158–163. IEEE Computer Society.

[Jung et al. 2002] Jung, J., Krishnamurthy, B., and Rabinovich, M. (2002). Flash crowds and denialof-service attacks: characterization and implications for cdns and web sites. In *Proceedings of the 11th International Conference on World Wide Web (WWW)*, pages 293–304. ACM.

[Kafle et al. 2010] Kafle, V. P., Otsuki, H., and Inoue, M. (2010). An id/locator split architecture for future networks. *IEEE Communications Magazine*, pages 138–144.

[Kim and Seong 2001] Kim, M. C. and Seong, P. H. (2001). Reliability graph with general gates: an intuitive and practical method for system reliability analysis. *Reliability Engineering and System Safety*, pages 239–246.

[Kitamura et al. 2007] Kitamura, Y., Lee, Y., Sakiyama, R., and Okamura, K. (2007). Experience with restoration of asia pacific network failures from taiwan earthquake. *IEICE Transactions on Communications*.

[Knight and Sullivan 2000] Knight, J. C. and Sullivan, K. J. (2000). On the definition of survivability. Technical report.

[Kuhn 1997] Kuhn, D. (1997). Sources of failure in the public switched telephone network. pages 31–36.

[Kuo and Zuo 2003] Kuo, W. and Zuo, M. J. (2003). *Optimal Reliability Modeling: Principles and Applications*. John Wiley & Sons.

[L. et al. 2011] L., H., Hajipour, J., Attar, A., and Leung, V. C. M. (2011). Efficient HetNet implementation using broadband wireless access with fiber-connected massively distributed antennas architecture. *IEEE Wireless Communications*, pages 72–78.

[Landwehr 2001] Landwehr, C. (2001). Computer security. *International Journal of Information Security*, 1(1):3–13.

[Laprie 1994] Laprie, J.-C. (1994). Dependability: basic concepts and terminology. *Dependable Computing and Fault Tolerance*.

[Laprie et al. 2004] Laprie, J.-C., Randell, B., and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transaction Dependable Security Computer*, 1(1):11–33.

[Lardner 1834] Lardner, D. (1834). *Babbage's calculating engine*, pages 263–327. Edinburgh, 120 edition.

[Lashgari and Avestimehr 2012] Lashgari, S. and Avestimehr, A. S. (2012). Timely throughput of heterogeneous wireless networks: Fundamental limits and algorithms. *CoRR*, abs/1201.5173.

[Lazar and Pacifici 1991] Lazar, A. and Pacifici, G. (1991). Control of resources in broadband networks with quality of service guarantees. *IEEE Communications Magazine*, 29(10):66–73.

[Lee and Anderson 1990] Lee, P. and Anderson, T. (1990). *Fault Tolerance: Principles and Practice*. Springer-Verlag New York.

[Lesk 2007] Lesk, M. (2007). The new front line: Estonia under cyberassault. *IEEE Security and Privacy 5*, pages 76–79.

[Li and Wang 2007] Li, F. and Wang, Y. (2007). Routing in vehicular ad hoc networks: a survey. *IEEE Vehicular Technology Magazine*, 2(2):12–22.

[Li and Rus 2000] Li, Q. and Rus, D. (2000). Sending messages to mobile users in disconnected ad hoc wireless networks. In *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 44–55. ACM.

[Lin et al. 2011] Lin, P., Zhang, J., Chen, Y., and Zhang, Q. (2011). Macro-femto heterogeneous network deployment and management: from business models to technical solutions. *IEEE Wireless Communications*, pages 64–70.

[Liscouski and Elliot 2004] Liscouski, B. and Elliot, W. (2004). Final report on the august 14, 2003 blackout in the united states and canada: causes and recommendations. Technical report, US - Canada Power System Outage Task Force.

[Liu 2008] Liu, Y. (2008). *Survivability of Networked Systems*. Duke University.

[Liu et al. 2004] Liu, Y., Mendiratta, V. B., and Trivedi, K. S. (2004). Survivability analysis of telephone access network. In *Proceedings of the International Symposium on Software Reliability Engineering*, ISSRE, pages 367–378, Washington, DC, USA. IEEE Computer Society.

[Liu and Trivedi 2006] Liu, Y. and Trivedi, K. S. (2006). Survivability quantification: The analytical modeling approach. In *Internation Journal of Performability Engineering*, pages 29–44. RAMS Consultants.

[Lyons and Vanderkulk 1962] Lyons, R. and Vanderkulk, W. (1962). The use of triple-modular redundancy to improve computer reliability. Journal of Research and Development 6, IBM.

[Mahmood 2009] Mahmood, R. (2009). Simulating challenges to communication networks for evaluation of resilience. Master's thesis, School of Engineering, University of Kansas.

[Meinel] Meinel, C. Attacking and defending the internet with border gateway protocol (BGP).

[Meyer 1992] Meyer, J. (1992). Performability: a retrospective and some pointers to the future. *Performance Evaluation*, 14(3):139–156.

[Meyer 1980] Meyer, J. F. (1980). On evaluating the performability of degradable computing systems. *IEEE Transactions on Computers*, 29(8):720–731.

[Mirkovic and Reiher 2004] Mirkovic, J. and Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. *SIGCOMM Computer Communication*, 34(2):39–53.

[Moore and Shannon 1956] Moore, E. and Shannon, C. (1956). Reliable circuits using less reliable relays. *Journal of the Franklin Institute*, 262(3):191–208.

[Neumann 1998] Neumann, P. G. (1998). The risks digest. *Forum on Risks to the Public in Computers and Related Systems-Committee on Computers and Public Policy*, 19.

[Neumann and Barnes 1999] Neumann, P. G. and Barnes, A. (1999). Practical architectures for survivable systems and networks: Phase-one final report. Technical report.

[Nicol et al. 2004] Nicol, D., Sanders, W., and Trivedi, K. (2004). Model-based evaluation: from dependability to security. *IEEE Transactions on Dependable and Secure Computing*, 1(1):48–65.

[Ott and Kutscher 2004] Ott, J. and Kutscher, D. (2004). Drive-thru internet: Ieee 802.11b for "automobile" users. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, page 373. IEEE Computer Society.

[Partridge et al. 2003] Partridge, C., Barford, P., Clark, D., Donelan, S., Paxon, V., Rexford, J., and Vernon, M. (2003). The Internet under crisis conditions: Learning from september 11. Technical report, National Research Council.

[Perkins and Bhagwat 1994] Perkins, C. and Bhagwat, P. (1994). Highly dynamic destination-sequenced distance-vector routing (dsdv) for mobile computers. In *Proceedings of the Conference on Communications Architectures, Protocols and Applications*, pages 234–244. ACM.

[Pierce 1965] Pierce, W. (1965). *Failure-tolerant Computer Design*. Academic Press.

[Prokopenko et al. 2009] Prokopenko, M., Boschetti, F., and Ryan, A. (2009). An information-theoretic primer on complexity, self-organization and emergence. *Complexity*, 15(1):11–28.

[Randell 1975] Randell, B. (1975). System structure for software fault tolerance. In *Proceedings of the International Conference on Reliable Software*, pages 437–449. ACM.

[Rohrer et al. 2008] Rohrer, J., Jabbar, A., Perrins, E., and Sterbenz, J. (2008). Cross-layer architectural framework for highly-mobile multihop airborne telemetry networks. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*.

[Sahner et al. 1996] Sahner, R. A., Trivedi, K. S., and Puliafito, A. (1996). *Performance and reliability analysis of computer systems: an example-based approach using the SHARPE software package*. Kluwer Academic Publishers.

[Savola 2007] Savola, R. (2007). Towards a taxonomy for information security metrics. In *Proceedings of the ACM workshop on Quality of Protection (QoP)*, pages 28–30. ACM.

[Scholler et al. 2010] Scholler, M., Smith, P., Rohner, C., Karaliopoulos, M., and Jabbar, A. (2010). On realising a strategy for resilience in opportunistic networks. *Future Network and Mobile Summit*, pages 1–8. Proceedings of the EU Future Network and Mobile Summit.

[Scholler et al. 2006] Scholler, M., Sterbenz, J., and A. Jabbar, D. H. (2006). First draft of the resilience and security framework. `http://www.anaproject.org/deliverables/2006/D.3.2.-Resilience.pdf`. Último acesso: Janeiro 2013.

[Shah and Rabaey 2002] Shah, R. and Rabaey, J. (2002). Energy aware routing for low energy ad hoc sensor networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 350–355. IEEE Computer Society.

[Shirey 2007] Shirey, R. (2007). *Internet Security Glossary, Version 2, RFC 4949*. IETF.

[Singh et al. 1998] Singh, S., Woo, M., and Raghavendra, C. (1998). Power-aware routing in mobile ad hoc networks. In *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 181–190. ACM.

[Steinder and Sethi 2004] Steinder, M. and Sethi, A. (2004). A survey of fault localization techniques in computer networks. *Science of Computer Programming*, 2(53):165–194.

[Sterbenz et al. 2008] Sterbenz, J., Bhattacharjee, B., Calvert, K., Griffioen, J., and Spring, N. (2008). oMo: Postmodern internetwork architecture. `http://wiki.ittc.ku.edu/pomo`. Último acesso: Janeiro 2013.

[Sterbenz and Hutchison 2008] Sterbenz, J. and Hutchison, D. (2008). Resilinets: Multilevel resilient and survivable networking initiative. `http://wiki.ittc.ku.edu/resilinets`. Último acesso: Janeiro 2013.

[Sterbenz et al. 2002] Sterbenz, J., Krishnan, R., Hain, R., Jackson, A., Levin, D., Ramanathan, R., and Zao, J. (2002). Survivable mobile wireless networks: issues, challenges, and research directions. In *Proceedings of the ACM Third ACM Workshop on Wireless Security*, pages 31–40. ACM Press.

[Sterbenza et al. 2010] Sterbenza, J. P. G., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Scholler, M., and Smith, P. (2010). Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*, 58(1):1245–1265.

[Strand et al. 2001] Strand, J., Chiu, A., and Tkach, R. (2001). Issues for routing in the optical layer. *IEEE Communications Magazine*, 39(2):81–87.

[Styron 2001] Styron, H. (2001). Csx tunnel fire: Baltimore, md, us fire administration technical report. Technical report, Federal Emergency Management Administration.

[Trivedi et al. 2009] Trivedi, K., Kim, D. S., Roy, A., and Medhi, D. (2009). Dependability and security models. In *Proceedings of the 7th International Workshop on Design of Reliable Communication Networks (DRCN)*.

[Trivedi 2002] Trivedi, K. S. (2002). *Probability and statistics with reliability, queuing and computer science applications*. John Wiley and Sons Ltd., 2nd edition edition.

[USA 1996] USA, N. (1996). *Federal Standard 1037C, Telecommunications: Glossary of Telecommunication Terms1037C, Telecommunications: Glossary of Telecommunication Terms*. National Communicatons system-NCS USA.

[Vaughn et al. 2003] Vaughn, R., Henning, R., and Siraj, A. (2003). Information assurance measures and metrics: state of practice and proposed taxonomy. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS)*. IEEE Computer Society.

[Wang et al. 1996] Wang, C.-Y., Logothetis, D., Trivedi, K. S., and Viniotis, I. (1996). Transient behavior of atm networks under overloads. In *Proceedings of the Fifteenth annual joint conference of the IEEE computer and communications societies conference on The conference on computer communications - Volume 3*, (INFOCOM), pages 978–985. IEEE Computer Society.

[Weber 2001] Weber, T. (2001). `http://www.inf.ufrgs.br/~taisy/disciplinas/textos/ConceitosDependabilidade.PDF`. Último acesso: Janeiro 2012.

[Willinger and Doyle 2005] Willinger, W. and Doyle, J. (2005). *Robustness and the Internet: Design and Evolution*. Oxford University Press.

[Xie et al. 2008] Xie, L., Smith, P., Hutchison, D., Banfield, M., Leopold, H., Jabbar, A., and Sterbenz, J. (2008). From detection to remediation: a self-organized system for addressing flash crowd problems. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pages 5809–5814.

[Zandt] Zandt, D. What sandy has taught us about technology, relief and resilience. `http://www.forbes.com/sites/deannazandt/2012/11/10/what-sandy-has-taught-us-about-technology-relief-and-resilience/`. Último acesso: Novembro, 2012.

[Zhao et al. 2004] Zhao, W., Ammar, M., and Zegura, E. (2004). A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of the ACM Fifth ACM International Symposium on Mobile Ad hoc Networking and Computing (MobiHoc)*, pages 187–198.