

Otimização multi-objetivo aplicada à alocação dinâmica de rotas LSP em redes MPLS

Fernando A. Santos¹, Geraldo R. Mateus¹

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Av. Antônio Carlos - 6629 – Belo Horizonte – MG

{fsantos,mateus}@dcc.ufmg.br

Abstract. *The MPLS architecture allows to define explicit routes in the network, called LSP, where the data flow. To define these routes can be a difficult task, especially considering multiple QoS metrics. This paper addresses the LSP routes dynamic assignment problem using a multi-objective optimization approach. The results show a good performance of this approach, especially taking into account environments where is necessary to consider different QoS metrics to assign the routes in the network.*

Resumo. *A arquitetura MPLS permite a definição de rotas explícitas na rede, denominadas LSP, por onde os dados devem trafegar. Determinar estas rotas pode ser uma tarefa complexa, principalmente quando múltiplos indicadores de QoS devem ser considerados. Este trabalho trata do problema da alocação dinâmica de rotas LSP através de uma abordagem de otimização multi-objetivo. Os resultados mostram um bom desempenho desta abordagem, principalmente em ambientes onde é necessário considerar diferentes indicadores de QoS para alocar as rotas na rede.*

1. Introdução

A grande utilização da Internet e o desenvolvimento de tecnologias aplicadas às redes de computadores trouxeram como consequência o aumento no número de aplicações que disponibilizam seus recursos em rede. Em particular, aplicações de tempo real, como a transmissão de som e vídeo, tem aumentado expressivamente o tráfego de dados nas redes.

Para o perfeito funcionamento destas aplicações, são definidos parâmetros de qualidade de serviço - QoS (*Quality of Service*), que determinam o comportamento da rede com respeito a alguns indicadores, como a latência da transmissão, o percentual de rejeição de pacotes, a taxa de utilização dos links da rede, entre outros. No entanto, o protocolo IP é baseado no paradigma de *best-effort*, inviabilizando a satisfação dos parâmetros de QoS necessários a algumas aplicações. Neste contexto, foi proposta a arquitetura MPLS (*MultiProtocol Label Switching*) [Rosen et al. 2001].

A arquitetura MPLS introduz nas redes IP uma capacidade avançada de controle do roteamento, como a característica de roteamento explícito, e auxilia na elaboração de redes com suporte a QoS [Awduche 1999]. São adicionados rótulos aos pacotes que trafegam no domínio MPLS, permitindo o controle das rotas, que são denominadas LSP (*Label Switching Path*).

Uma das vantagens proporcionadas pela arquitetura MPLS é a possibilidade de se considerar vários indicadores de QoS no roteamento dos pacotes, o que não é considerado por protocolos de roteamento comuns na internet, como o OSPF (*Open Shortest Path First*). No entanto, a otimização das rotas pode ser uma tarefa complexa. O problema de se determinar rotas LSP sujeitas a mais de duas restrições de QoS é provado ser NP-Completo [Shao et al. 2006].

A definição de rotas LSP considerando mais de um indicador de QoS sugere um problema multi-objetivo, onde cada indicador de QoS será representado por um objetivo, existindo conflitos em satisfazê-los simultaneamente. Ao solucionar este problema, obtém-se como resultado um conjunto de soluções ótimas denominado conjunto de Pareto.

As soluções do conjunto de Pareto podem privilegiar um objetivo específico ou mesmo buscar um equilíbrio entre eles, mas todas são consideradas ótimas. A escolha da “melhor” solução do conjunto de Pareto está em função do contexto da otimização, ou seja, quais indicadores de QoS são mais relevantes para a comunicação. Portanto, esta escolha deve ser feita de maneira dinâmica, levando em consideração informações da rede e das aplicações envolvidas na comunicação.

Este trabalho apresenta uma abordagem multi-objetivo para o problema de roteamento dinâmico em redes MPLS. Serão definidas rotas LSP explícitas com garantia de reserva de banda, ou seja, ao alocar uma requisição na rede é definido um caminho e reservada uma banda nos links por onde os dados devem trafegar. Através da otimização multi-objetivo é obtido o conjunto de Pareto, e a escolha da melhor solução deste conjunto será feita dinamicamente.

O restante deste artigo está estruturado da seguinte maneira: na seção 2 estão descritas as contribuições anteriores à deste trabalho nesta área, destacando-se alguns pontos fundamentais em cada trabalho; na seção 3 está a representação matemática do problema multi-objetivo proposto que foi realizada através de uma modelagem de programação linear inteira mista; na seção 4 é apresentada uma solução heurística para obter aproximações do conjunto de Pareto com qualidade e eficiência, enquanto na seção 5 são apresentados os resultados de simulações utilizando a abordagem proposta. Finalmente as considerações finais são apresentadas na seção 6.

2. Trabalhos relacionados

Desde o surgimento da arquitetura MPLS, diversos algoritmos vêm sendo propostos afim de otimizar a alocação das rotas LSP na rede. O objetivo principal é a utilização dos recursos da rede com máxima eficiência. Dentre os trabalhos existem aqueles que realizam a otimização estática, que consideram que tanto os recursos da rede quanto as requisições são previamente conhecidos, e outros que otimizam as rotas dinamicamente, onde podem existir variações no comportamento da rede ao longo do tempo e novas requisições podem surgir, bem como outras deixar de existir.

Uma característica comum aos trabalhos é a busca pela otimização de múltiplos indicadores de QoS simultaneamente. Diferentes alternativas são utilizadas para a modelagem dos problemas, e neste contexto são aplicadas técnicas de otimização mono-objetivo e multi-objetivo, normalmente modeladas utilizando programação linear inteira.

Devido a este problema ser NP-Completo [Shao et al. 2006], em todos os trabalhos são propostas heurísticas para obter soluções aproximadas com maior eficiência.

Um dos primeiros a tratar deste problema foi [Kar et al. 2000] que formulou o problema mono-objetivo utilizando programação linear inteira. Como heurística para aproximação dos resultados, propôs o algoritmo MIRA (*Minimum Interference Routing Algorithm*), que procura por rotas que minimizem a interferência nos demais caminhos da rede através da estratégia de fluxo máximo.

Também utilizando otimização mono-objetivo através de programação linear inteira, [Dias et al. 2003] propôs uma relaxação lagrangeana com heurísticas - RLH para solucionar o problema. Os resultados apresentados pela RLH não foram suficientes para sua aplicação em um contexto dinâmico, mas em [Dias et al. 2005] é apresentada uma nova heurística que possibilita sua utilização no roteamento dinâmico de rotas LSP.

No trabalho apresentado por [Oliveira and Mateus 2006] foram propostos dois modelos conjugados de otimização mono-objetivo e a otimização das rotas é realizada em dois passos. O primeiro passo busca balancear a distribuição da carga na rede, isto é feito ao minimizar o percentual de utilização do link mais sobrecarregado da rede. Já o segundo passo utiliza como parâmetro de entrada o resultado do primeiro passo e tem como função objetivo o número de links utilizados na comunicação somados ao número de rejeições multiplicado por uma penalidade. Este objetivo foi proposto com o intuito de minimizar os dois indicadores de QoS utilizando uma única função objetivo.

Como o tempo computacional necessário para realizar a otimização de dois problemas de programação linear inteira é elevado, foi proposto um algoritmo genético que obteve aproximações dos resultados eficientemente, no entanto, os resultados não foram utilizados na otimização dinâmica das rotas LSP.

Ao analisar a abordagem de [Oliveira and Mateus 2006] para o problema, é possível observar que a otimização multi-objetivo seria mais adequada para a alocação das rotas. Neste caso, os indicadores de QoS poderiam ser modelados e resolvidos utilizando apenas um problema multi-objetivo, ao invés de dois problemas mono-objetivo.

Um dos primeiros a utilizar técnicas de otimização multi-objetivo aplicadas a problemas de roteamento, [Knowles et al. 2000] utilizou os algoritmos PAES (*Pareto Archived Evolution Strategy*) [Knowles and Corne 2000b] e M-PAES (*Memetic Pareto Archived Evolution Strategy*) [Knowles and Corne 2000a] para solucionar o problema de roteamento estático. Os resultados apresentaram boas aproximações do conjunto de Pareto, em particular a busca local incluída no algoritmo memético trouxe melhorias significativas às soluções, mas a eficiência dos algoritmos não foi discutida.

Já a utilização de técnicas de otimização multi-objetivo aplicadas à arquitetura MPLS foi proposta por [Koyama et al. 2004] que considerou dois indicadores de QoS como função objetivo para serem minimizados: (i) o custo e (ii) o atraso da comunicação. Um algoritmo genético multi-objetivo foi proposto para obter aproximações do conjunto de Pareto, mas o roteamento dinâmico não foi abordado nas simulações.

Embora a utilização de técnicas de otimização para a definição de rotas LSP seja um assunto amplamente estudado, observa-se que a abordagem multi-objetivo em um contexto dinâmico de otimização ainda não foi explorada, sendo o grande diferencial

proposto neste trabalho. Além das dificuldades encontradas na otimização multi-objetivo, a escolha de uma única solução do conjunto de Pareto aumenta o desafio desta abordagem. Neste trabalho serão consideradas 3 funções objetivo de minimização: (i) o número de links utilizados, (ii) o número de requisições rejeitadas e (iii) o balanceamento da carga nos links.

3. Formulação matemática

A rede a ser otimizada será representada por um grafo $G = (V, A)$, onde $V = \{1, 2, \dots, ||V||\}$ são os roteadores do domínio MPLS, também conhecidos como LSR (*Label Switch Router*), e $A = \{(i, j)\}$ são os links existentes entre os roteadores LSR. A capacidade de banda de cada link (i, j) será representada no modelo como B_{ij} .

As requisições de comunicação são representadas por (o^k, d^k, b^k) , sendo que o^k e d^k representam respectivamente os roteadores LSR de origem e de destino do tráfego para a k -ésima requisição. Enquanto b^k indica a quantidade de banda que deve ser reservada para esta requisição. O conjunto de todas as requisições a serem otimizadas é denotado por R .

Entre as variáveis de decisão do modelo, tem-se x_{ij}^k que assume valor 1 caso a k -ésima requisição passe pelo link (i, j) ou 0 caso contrário. Há também a variável de decisão r^k que indica se a k -ésima requisição foi atendida ou não, assumindo valor 1 em caso positivo, senão 0. Finalmente, a variável de decisão α assumirá valores contínuos no intervalo entre $[0, 1]$ e representa o percentual de utilização do link mais utilizado da rede.

O parâmetro L_r representa o limite de rejeições permitido pelo modelo. Mesmo que o número de rejeições seja um dos objetivos a ser minimizado, a imposição de um limite de rejeição pode ser útil em alguns contextos, sendo um parâmetro a ser definido antes da otimização. O modelo multi-objetivo, utilizando programação linear inteira, pode ser exibido como:

$$\text{Min} \begin{cases} F_1 = \sum_{k \in R} \sum_{(i,j) \in A} x_{ij}^k \\ F_2 = \sum_{k \in R} (1 - r^k) \\ F_3 = \alpha \end{cases}$$

sujeito a:

$$\sum_{k \in R} b^k x_{ij}^k \leq \alpha B_{ij} \quad \forall (i, j) \in A \quad (1)$$

$$\sum_{(i,j) \in \Gamma_i^+} x_{ij}^k - \sum_{(l,i) \in \Gamma_i^-} x_{li}^k = \begin{cases} r^k & (o_k) \\ -r^k & (d_k) \\ 0 & \end{cases} \quad \forall i \in V, k \in R \quad (2)$$

$$\sum_{k \in R} (1 - r^k) \leq L_r \quad (3)$$

$$x_{ij}^k \in 0, 1 \quad \forall (i, j) \in A, \forall k \in R \quad (4)$$

$$r^k \in 0, 1 \quad \forall k \in R \quad (5)$$

$$\alpha \in [0, 1] \quad (6)$$

Quanto às funções objetivo, tem-se que F_1 representa o número de links utilizados para atender as requisições do conjunto R . Quanto menos links forem utilizados mais rapidamente os dados trafegarão da origem ao destino, logo este objetivo será minimizado. F_2 representa a quantidade de requisições de R que serão rejeitadas durante a otimização. Espera-se que este número também seja minimizado, de modo a atender o maior número possível de requisições. E F_3 diz respeito ao percentual de utilização dos links da rede, e como pretende-se balancear a carga, espera-se que este valor também seja minimizado para que a quantidade de dados trafegando nos links esteja equilibrada.

É possível observar os conflitos em atender simultaneamente a todas as funções objetivo, principalmente ao considerar F_1 e F_3 , uma vez que ao se otimizar uma delas a outra tende a ter seu resultado comprometido. Ao verificar a influência de F_2 em F_1 e F_3 observa-se também uma relação de compromisso. Ao minimizar F_2 , tanto F_1 quanto F_3 tendem a ter seus valores aumentados uma vez que a alocação de uma requisição na rede consumirá seus recursos. A não alocação de uma requisição privilegiará F_1 e F_3 mas não F_2 . Esta impossibilidade de se obter valores mínimos para as três funções objetivo que justificam a diversidade de soluções no conjunto Pareto, existindo mais de uma solução ótima, que deverá ser escolhida em função do contexto de otimização.

Quanto às restrições, em (1) verifica-se que a quantidade de dados a trafegar em um link (i, j) não pode ultrapassar sua capacidade de banda B_{ij} . As restrições clássicas de conservação de fluxo estão presentes em (2), que determinam que o tráfego a entrar em um vértice deve ser o mesmo a sair, exceto na origem e no destino da requisição. A restrição (3) permite limitar o número de rejeições, o que pode ser de grande relevância em alguns ambientes. Finalmente as restrições (4), (5) e (6) limitam o espaço das variáveis de decisão envolvidas.

4. Solução heurística - Algoritmo Genético Multi-Objetivo

A complexidade computacional para solucionar o problema multi-objetivo de roteamento em redes MPLS é elevada. O modelo matemático proposto permite a obtenção de soluções ótimas, no entanto, para obter todas aquelas soluções contidas no conjunto de Pareto seria necessário um grande tempo computacional, o que inviabilizaria sua aplicação em um contexto dinâmico de otimização.

Portanto, um algoritmo genético multi-objetivo foi proposto, para realizar aproximações do conjunto de Pareto com um tempo computacional viável para sua aplicação em um ambiente dinâmico de otimização. Este algoritmo se baseia no arcabouço NSGA-II [Deb et al. 2000], que tem como característica principal a eficiência, além de incorporar elitismo na seleção dos indivíduos da próxima geração.

As subseções a seguir, a saber 4.1, 4.2 e 4.3, descrevem os passos da implementação do algoritmo genético multi-objetivo, e na subseção 4.4 o desempenho computacional do algoritmo bem como a qualidade das aproximações são avaliados.

4.1. Codificação e população inicial

Na codificação, cada indivíduo armazenará as rotas para ligar a origem ao destino de cada requisição atendida. Portanto, os indivíduos serão compostos por soluções viáveis para o problema. A ausência de soluções inviáveis é justificada pela necessidade de eficiência devido ao ambiente dinâmico de otimização, uma vez que ao processar soluções inviáveis

o tempo de execução do algoritmo aumenta consideravelmente, embora isto pudesse melhorar a qualidade das soluções.

Para gerar a população inicial, é utilizado o algoritmo de Dijkstra [Dijkstra 1959] para a geração das rotas viáveis entre origem e destino das requisições. Este algoritmo se destaca pela eficiência para calcular o caminho mínimo em um grafo. A geração de cada indivíduo da população inicial é feita atribuindo-se custos aleatórios aos links e executando-se o algoritmo de Dijkstra para calcular as rotas de cada requisição a ser atendida. Observe que os custos aleatórios atribuídos aos links são utilizados apenas pelo algoritmo de Dijkstra, pois ao computar os valores da função objetivo F_1 , considera-se que cada link tem custo unitário, ou seja, deseja-se minimizar o número de links utilizados. Esta estratégia contribui para aumentar a diversidade das soluções na população inicial.

As requisições são selecionadas aleatoriamente para que sua rota seja definida pelo algoritmo de Dijkstra. Após definir a rota, a respectiva banda exigida pela requisição é retirada dos links utilizados, simulando o processo de alocação de rotas explícitas com reserva de recurso das redes MPLS. Sendo assim, caso um link esteja com sua capacidade de banda no limite, nenhuma requisição passará por ele, evitando a perda de pacotes na rede por sobrecarga dos links.

Assim que as rotas de um indivíduo são geradas, os valores correspondentes às funções objetivo F_1 , F_2 e F_3 são calculados. Calcula-se o número de links utilizados pelas rotas, a quantidade de requisições rejeitadas e o percentual de utilização do link mais utilizado, respectivamente.

4.2. Seleção e cruzamento

Para entender o procedimento de seleção, um importante conceito é o de dominância entre os indivíduos. Considere dois indivíduos X e Y com seus respectivos valores das funções objetivo F_1 , F_2 e F_3 . Diz-se que X domina Y caso o valor de pelo menos uma função objetivo de X seja melhor (para este problema, entenda-se menor) que de Y , e o valor das demais não seja pior (maior).

Assim, os indivíduos são selecionados aleatoriamente aos pares, e aquele que domina o seu respectivo par é selecionado para o cruzamento. Caso não exista relação de dominância entre eles, ambos são selecionados. Este método de seleção é denominado de Torneio.

Seja P_1 e P_2 indivíduos selecionados para o cruzamento. Como ambos armazenam soluções viáveis para o problema, o indivíduo C_1 , resultante do cruzamento entre P_1 e P_2 , herdará as características de ambos, mantendo a viabilidade das rotas. A figura 1 ilustra o cruzamento.

4.3. Mutação e próxima geração

Na mutação uma pequena alteração é incluída em um indivíduo afim de mudar a sua posição no espaço de busca, evitando mínimos locais. No caso específico deste algoritmo genético, a mutação em um indivíduo retira um link que estava sendo utilizado no roteamento das requisições, para que novas rotas sejam encontradas, fazendo com que o algoritmo explore melhor o espaço de busca. Um exemplo deste procedimento pode ser visto na figura 2, onde apenas duas requisições são consideradas.

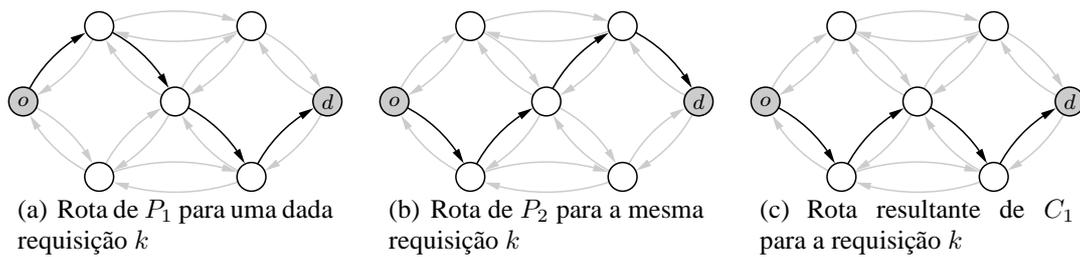


Figura 1. Processo de cruzamento no algoritmo genético multi-objetivo

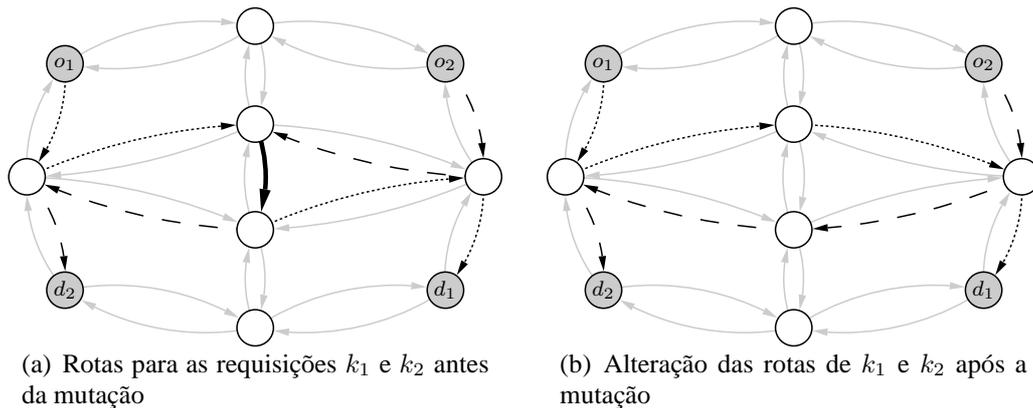


Figura 2. Exemplo de mutação no algoritmo

Conforme sugere o NSGA-II, para definir os indivíduos da próxima geração é necessário calcular o valor de dominância associado a cada indivíduo. Este valor é obtido ao se confrontar um indivíduo com todos os outros da população quanto à dominância. Aqueles indivíduos com menor valor de dominância, estão mais próximos do conjunto Pareto e devem ter preferência para continuarem na próxima geração.

4.4. Eficiência e qualidade das aproximações

Esta solução heurística foi proposta com o intuito de obter aproximações do conjunto de Pareto para o problema de roteamento em redes MPLS. Quanto mais próximas estiverem as soluções do conjunto de Pareto, considera-se maior sua qualidade. No entanto, as aproximações devem ser obtidas eficientemente, pois no contexto dinâmico de otimização a eficiência das respostas é um fator de grande importância.

Para realizar os experimentos, foram utilizadas as mesmas topologias de rede adotadas em [Oliveira and Mateus 2006], conforme pode ser visto na figura 3. Estas topologias permitem avaliar o comportamento do algoritmo em diferentes cenários, permitindo sua melhor validação. Quanto às requisições a serem otimizadas, seleciona-se aleatoriamente o par (origem, destino) e a respectiva banda necessária para sua alocação varia também aleatoriamente no intervalo de 200K a 400K.

Medir a qualidade das soluções de um problema multi-objetivo é uma tarefa complexa. Isto se justifica pela necessidade de comparar conjuntos de soluções, levando em consideração a proximidade dos pontos ao conjunto de Pareto e também sua distribuição no conjunto. Para mensurar a qualidade das aproximações, foram adotadas neste trabalho duas estratégias.

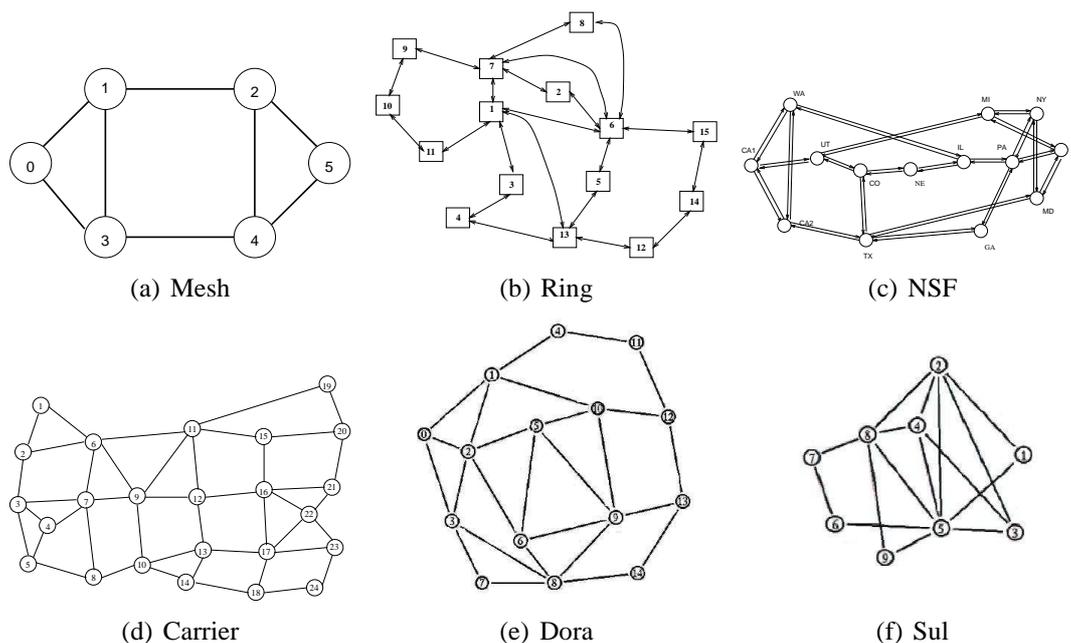


Figura 3. Topologias de rede utilizadas nos experimentos

A primeira delas é a solução utópica. Esta solução corresponde ao ponto cujas coordenadas são representadas pelo melhor valor de cada objetivo, ou seja, seria a solução ideal com os valores ótimos para cada uma das funções objetivo, mas não é obtida devido à existência de conflitos para se obter os valores ótimos de todas as funções objetivo simultaneamente. Para obter a solução utópica, foi utilizado o software CPLEX [Ilog Incorporation 2006] que solucionou a formulação matemática apresentada na seção 3 considerando-se uma função objetivo por vez. O resultado da otimização para cada função objetivo representará uma coordenada da solução utópica.

Para comparar a qualidade das aproximações obtidas pelo algoritmo genético, será utilizada uma aproximação da solução utópica, denominada *p-utópica*. Esta solução será representada por um ponto onde cada coordenada terá o melhor valor de cada função objetivo obtido ao se executar o algoritmo genético multi-objetivo. Os resultados exibidos na tabela 1 foram calculados para 10 requisições definidas aleatoriamente na rede.

instância	utópica			<i>p-utópica</i>		
	objetivo F_1	objetivo F_2	objetivo F_3	objetivo F_1	objetivo F_2	objetivo F_3
Mesh	8	2	0.84	8	2	0.84
Ring	15	2	0.81	17	2	0.81
NSF	15	1	0.86	16	1	0.86
Carrier	19	1	0.88	20	1	0.88
Dora	17	0	0.72	17	0	0.72
Sul	6	5	0.78	7	5	0.78

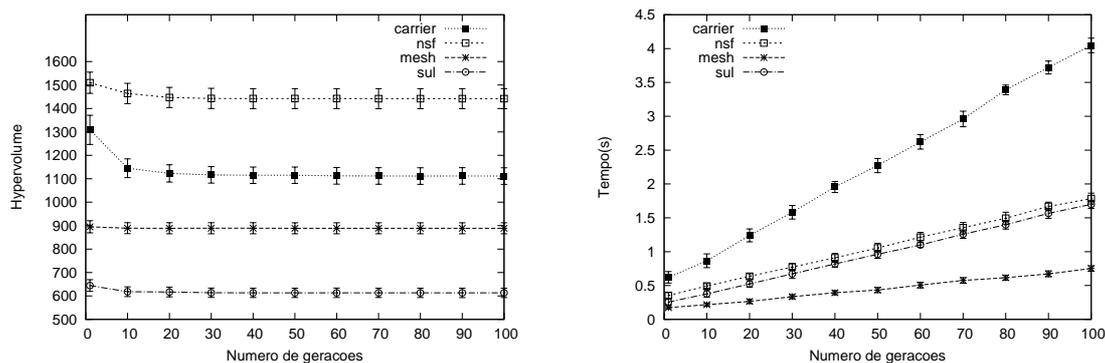
Tabela 1. Comparação entre a solução utópica e sua aproximação (*p-utópica*)

Os resultados da tabela 1 mostram que os pontos das soluções *p-utópica* e utópica estão próximos, inclusive atingindo-se valores ótimos para as topologias Mesh e Dora. Isto indica uma boa proximidade entre as soluções ótimas e as aproximadas.

A segunda estratégia para mensurar a qualidade das aproximações foi a utilização do *hypervolume* [While et al. 2006]. Esta medida calculará no espaço tri-dimensional das funções objetivo o volume existente entre os pontos aproximados do conjunto de Pareto até a origem. Por se tratar de um problema de minimização, quanto menor o volume, maior a qualidade das soluções.

O desempenho do algoritmo genético multi-objetivo está em função de dois parâmetros: o tamanho da população e o número de gerações. Quanto maior uma população de indivíduos, maior a possibilidade de se obter soluções de qualidade. E também à medida em que os indivíduos da população evoluem de geração, a qualidade das soluções tende a aumentar. Entretanto, ambos têm como consequência o aumento no tempo de execução do algoritmo.

Portanto, os próximos experimentos utilizam a medida *hypervolume* para avaliar a qualidade das aproximações relacionada ao número de gerações e também ao tamanho da população. É realizada também uma avaliação da eficiência do algoritmo. Os gráficos da figura 4 ilustram o comportamento do algoritmo com relação ao número de gerações. O tamanho da população é fixado em 50 indivíduos e a taxa de mutação é de 0.02.



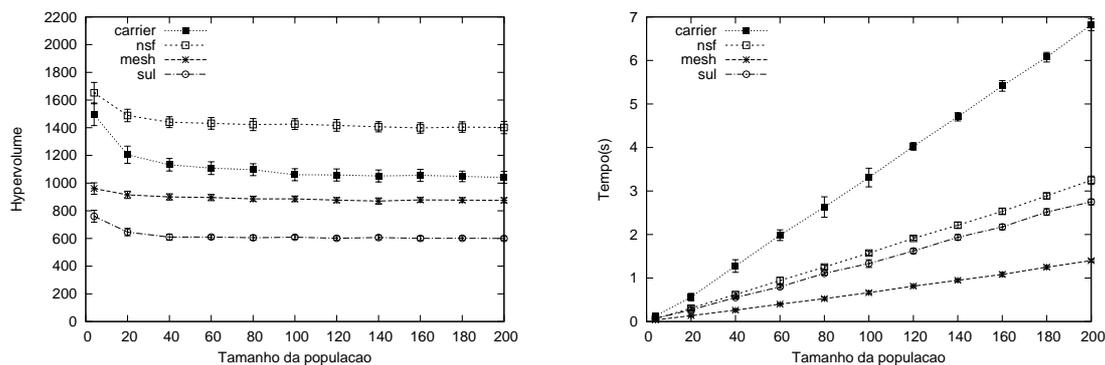
(a) Melhoria da qualidade das soluções em função do número de gerações

(b) Relação entre o número de gerações e o tempo de execução

Figura 4. Influência do número de gerações no desempenho do algoritmo

Observa-se uma melhora na qualidade das soluções à medida em que o número de gerações do algoritmo aumenta, no entanto, esta melhora deixa de ser significativa quando o número de gerações ultrapassa 30, como pode ser observado no gráfico 4(a). As topologias menores apresentam uma melhora pouco significativa, devido ao limite de soluções existentes, como é o caso das topologias Mesh e Sul. Isto sugere que alterações podem ser realizadas nos operadores genéticos para melhorar o desempenho do algoritmo. Quanto à eficiência, analisando-se o gráfico 4(b) observa-se que o tempo de execução do algoritmo é suficiente para sua aplicação em um contexto dinâmico de otimização. Considerando-se 30 gerações, o algoritmo apresenta respostas em até 1,5 segundos.

Considerando agora o desempenho do algoritmo em função do tamanho da população, ao utilizar um número fixo de 50 gerações e uma taxa de mutação novamente de 0.02, os resultados são exibidos nos gráficos da figura 5. A influência do tamanho da população no valor do *hypervolume* é mais acentuada, quando comparada com o parâmetro número de gerações, conforme pode ser visto no gráfico 5(a), porém o tempo computacional também apresenta um maior crescimento, como visto em 5(b).



(a) Melhoria da qualidade das soluções em função do tamanho da população

(b) Relação entre o tamanho da população e o tempo de execução

Figura 5. Medidas de desempenho e qualidade das soluções do algoritmo genético

Considera-se que o algoritmo genético proposto obteve um bom desempenho tanto na qualidade das aproximações quanto em eficiência. Portanto, na próxima seção será avaliado o desempenho da abordagem multi-objetivo utilizando um ambiente de simulação, que permitirá conclusões a respeito do comportamento da rede ao considerar diferentes indicadores de QoS na comunicação, em um contexto dinâmico de otimização.

5. Resultados de simulações

Para avaliar o desempenho da otimização multi-objetivo na alocação dinâmica de rotas LSP em redes MPLS, foi utilizado o ambiente de simulação *Network Simulator NS-2* [Fall and Varadhan]. Assim, será possível observar o comportamento do algoritmo considerando diferentes indicadores de QoS na rede. Particularmente, serão considerados como indicadores de QoS a quantidade de rejeições de requisições, a latência média e o percentual de carga do link mais utilizado da rede.

Um grande desafio encontrado na alocação dinâmica dos LSPs na rede foi o processo de tomada de decisão. A execução do algoritmo genético multi-objetivo resulta em uma aproximação para o conjunto de Pareto, portanto, um conjunto de soluções para o problema. A escolha de qual solução é a mais adequada é uma decisão difícil de ser tomada.

Para o processo de decisão do problema multi-objetivo foi utilizado o método AHP - *Analytic Hierarchy Process* [Saaty 1980]. Este método recebe como parâmetro uma matriz de preferências preenchida pelo decisor que aponta o grau de importância existente entre as funções objetivo do problema. Então utiliza operações matriciais e vetoriais para definir qual a solução que mais se adequa àquelas preferências previamente definidas.

Para elaborar os cenários de simulação foram utilizadas as topologias de rede ilustradas na figura 3. As requisições são definidas selecionando-se aleatoriamente um par origem-destino e sua exigência de banda é escolhida aleatoriamente dentro de um intervalo definido. Cada simulação tem duração de 300 segundos e as requisições surgem aleatoriamente em intervalos de tempo, podendo estar agrupadas ou não. Como parâmetros do algoritmo genético tem-se que o número de gerações é 30, o tamanho da

população é 50 e a taxa de mutação 0,02.

Para avaliar o desempenho da abordagem multi-objetivo esta foi comparada com as tradicionais técnicas mono-objetivo. A primeira simulação utiliza um cenário de rede com tráfego de baixa demanda, onde a exigência de todas as requisições não excede a capacidade dos links da rede. Cada requisição tem sua exigência de banda definida aleatoriamente no intervalo de 50K a 100K, e o número de requisições varia entre 10, 30 e 50.

Os resultados mostrados nas tabelas 2 e 3 foram obtidos ao se alocar as requisições utilizando uma otimização mono-objetivo, considerando como função objetivo o número de links (tabela 2) e o percentual de carga no link mais utilizado da rede (tabela 3). Já a tabela 4 contém os resultados da otimização multi-objetivo, sendo que a importância das funções objetivo no método AHP são iguais. Os indicadores de QoS medidos são a latência média da rede e o percentual de carga do link mais utilizado.

instância	10 requisições		30 requisições		50 requisições	
	% carga	latência (ms)	% carga	latência (ms)	% carga	latência (ms)
Mesh	49	38.5	52	66.1	54	85.2
Ring	62	30.1	67	59.1	71	80.2
NSF	60	27.2	62	57.5	69	74.0
Carrier	59	24.8	65	50.0	76	62.1
Dora	71	25.7	78	52.2	83	71.1
Sul	50	27.1	59	61.3	68	80.1

Tabela 2. Alocação mono-objetivo das requisições, otimizando o número de links

Analisando os resultados da tabela 2, observa-se que os valores obtidos para a latência são baixos, no entanto, o percentual de carga no link mais utilizado apresenta valores mais elevados. Este comportamento condiz com o procedimento de otimização para alocar as requisições na rede. Uma vez que deseja-se minimizar apenas o número de links, o valor do balanceamento de carga pode ser prejudicado.

instância	10 requisições		30 requisições		50 requisições	
	% carga	latência (ms)	% carga	latência (ms)	% carga	latência (ms)
Mesh	43	41.6	48	70.4	52	94.0
Ring	54	33.6	64	59.1	69	82.9
NSF	48	31.7	59	61.5	66	75.2
Carrier	55	28.3	62	53.8	67	67.0
Dora	59	26.4	75	54.5	81	78.1
Sul	47	32.4	54	67.7	59	84.1

Tabela 3. Alocação mono-objetivo das requisições, otimizando o balanceamento de carga

Os resultados da tabela 3 apresentam um comportamento oposto daqueles exibidos na tabela 2. Neste caso a otimização das rotas é feito minimizando-se apenas a carga no link mais utilizado da rede, não privilegiando o número de links utilizados na comunicação. Por exemplo, o atendimento de uma requisição pode ser feito por um caminho mais longo, para desviar o fluxo de um link sobrecarregado.

instância	10 requisições		30 requisições		50 requisições	
	% carga	latência (ms)	% carga	latência (ms)	% carga	latência (ms)
Mesh	45	39.1	50	67.4	54	92.1
Ring	59	29.6	66	60.7	72	79.6
NSF	53	29.0	61	56.4	67	71.1
Carrier	60	25.3	64	50.5	70	61.3
Dora	69	22.5	80	50.2	85	74.8
Sul	48	36.2	57	62.1	65	77.9

Tabela 4. Alocação das requisições utilizando a abordagem multi-objetivo

Finalmente, a tabela 4 mostra os resultados obtidos ao se realizar a alocação das rotas através da abordagem multi-objetivo. O procedimento de otimização procurará por rotas que utilizem um número mínimo de links, mas sem desprezar o balanceamento de carga da rede. Portanto, se trata de uma estratégia que busca o equilíbrio dos indicadores de QoS da rede. Mas caso seja necessário priorizar um determinado indicador de QoS, é possível fazê-lo ao atribuir uma prioridade maior à função objetivo correspondente.

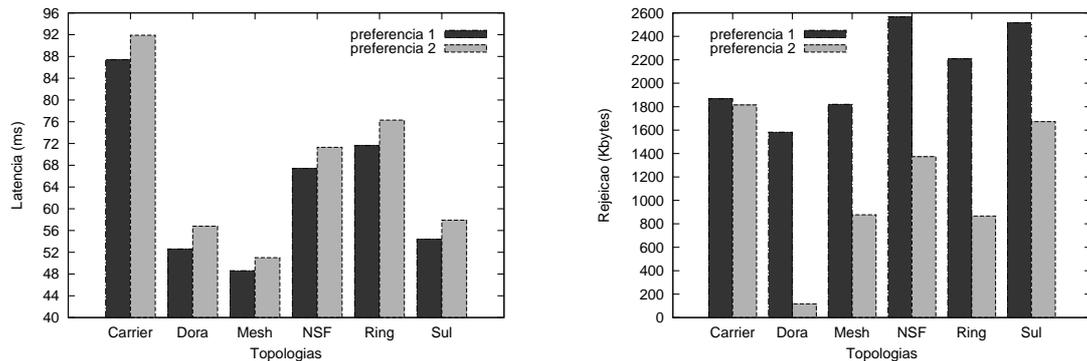
Além da possibilidade de considerar diferentes indicadores de QoS simultaneamente no processo de otimização, a estratégia de roteamento multi-objetivo apresenta também como vantagem a flexibilidade. A solução do problema de roteamento multi-objetivo fornece como resultado um conjunto de soluções viáveis e a escolha da solução mais adequada será definida no processo de tomada de decisão.

É possível privilegiar as soluções quanto ao atraso, ao balanceamento de carga ou mesmo à quantidade de rejeições na rede apenas escolhendo-se soluções adequadas do conjunto. Ao utilizar o método AHP, é possível alterar os valores de prioridade das funções objetivo direcionando o roteamento sem alterar o problema de otimização.

Neste contexto, o próximo cenário de simulação foi elaborado para demonstrar a flexibilidade da otimização multi-objetivo no roteamento dinâmico em redes MPLS. O mesmo procedimento de otimização será utilizado neste cenário considerando duas matrizes de preferências distintas para o processo de tomada de decisão. O rótulo 'preferencia 1' identifica a utilização da matriz de preferências que prioriza baixos valores de atraso na rede, enquanto o rótulo 'preferencia 2' diz respeito à matriz de preferências que evita rejeições. O surgimento das requisições na rede segue uma distribuição de Poisson com o parâmetro $\lambda = 4$. A exigência das requisições varia de 300K a 600K, como consequência a rede apresentará um tráfego intenso e requisições poderão ser rejeitadas devido à indisponibilidade de banda. Os resultados podem ser vistos na figura 6.

A melhoria obtida nos valores de atraso em contraposição à quantidade de rejeições observada nos gráficos 6(a) e 6(b) pode ser útil em alguns contextos de rede, por exemplo em aplicações de video-conferência ou VoIP. Diversas outras particularidades podem ser observadas em outros contextos de rede, e a otimização multi-objetivo proporciona flexibilidade para lidar com elas, sendo necessário apenas alguns ajustes no modelo de decisão para que novos resultados sejam obtidos.

Através destes resultados é possível observar que a utilização da abordagem multi-objetivo para a alocação de rotas LSP em redes MPLS apresentou benefícios quando comparada com a tradicional abordagem de otimização mono-objetivo. Com a sua utilização



(a) Variação da latência em função do número de requisições alocadas na rede

(b) Quantidade de bytes rejeitados das requisições na simulação

Figura 6. Índices de latência e rejeição ao utilizar diferentes abordagens de otimização

é possível considerar diferentes indicadores de QoS na alocação das rotas LSP, tornando-a mais equilibrada e flexível, proporcionando melhor desempenho em cenários distintos de comunicação.

6. Considerações finais

Este trabalho apresentou uma abordagem multi-objetivo para a alocação de rotas LSP em redes MPLS em um cenário dinâmico de otimização. Foi proposta uma formulação matemática para a solução do problema, mas devido à sua complexidade computacional elevada foi proposta uma heurística para obter soluções aproximadas eficientemente.

A heurística implementada foi um algoritmo genético multi-objetivo, que obteve aproximações com qualidade para o problema. A eficiência apresentada pelo algoritmo tornou viável sua utilização em um contexto dinâmico de otimização. Por se tratar de uma otimização multi-objetivo, seu resultado consiste em um conjunto de soluções para o problema. A escolha de qual a solução mais adequada para realizar a alocação das rotas LSP é uma tarefa complexa e foi solucionada utilizando o método AHP. Os resultados confirmaram a viabilidade de se utilizar a abordagem multi-objetivo, principalmente em ambientes onde é necessário considerar diferentes indicadores de QoS simultaneamente e flexibilidade na escolha das soluções em diferentes cenários.

Como continuidade deste trabalho, o processo de simulação será aprimorado afim de ampliar a validade dos resultados apresentados para outros cenários de simulação e diferentes contextos. Além disto, outras alternativas para a tomada de decisão podem ser utilizadas, como considerar informações sobre o histórico do comportamento da rede para a tomada das decisões, o que poderia melhorar o desempenho desta abordagem.

Referências

- Awduche, D. (1999). MPLS and traffic engineering in IP networks. *IEEE Communications Magazine*, 37(12):42–47.
- Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In *Proceedings*

- of the Parallel Problem Solving from Nature VI Conference*, pages 849–858, Paris, France.
- Dias, R., Camponogara, E., Farines, J.-M., Willrich, R., and Campestrini, A. (2003). Otimização Lagrangeana em engenharia de tráfego para redes IP sobre MPLS. *XXI Simpósio Brasileiro de Redes de Computadores*, pages 475–490.
- Dias, R., Farines, J.-M., Camponogara, E., Willrich, R., and Campestrini, A. (2005). Engenharia de tráfego dinâmica em redes IP sobre tecnologia MPLS: Otimização baseada em heurísticas. *XXIII Simpósio Brasileiro de Redes de Computadores*, pages 191–204.
- Dijkstra, E. W. (1959). A note on two problems in connection with graphs. *Numer. Math.*, 1:269–271.
- Fall, K. and Varadhan, K. The NS manual. <http://www.isi.edu/nsnam/ns/ns-documentation.html>, acessado em 04/12/2008.
- Ilog Incorporation (2006). Solver CPLEX v10. www.ilog.com/products/cplex/ (acessado em 06/06/2008).
- Kar, K., Kodialam, M., and Lakshman, T. (2000). Minimum interference routing of bandwidth guaranteed tunnels with MPLS traffic engineering applications. *IEEE Journal on Selected Areas in Communications*, 18:2566–2579.
- Knowles, J. and Corne, D. (2000a). M-PAES: A memetic algorithm for multiobjective optimization. In *Proceedings of the 2000 Congress on Evolutionary Computation CEC00*, pages 325–332, California, USA. IEEE Press.
- Knowles, J., Oates, M., and Corne, D. (2000). Advanced multi-objective evolutionary algorithms applied to two problems in telecommunications. *BT Technology Journal*, 18(4):51–65.
- Knowles, J. D. and Corne, D. W. (2000b). Approximating the nondominated front using the Pareto archived evolution strategy. *Evolutionary Computation*, 8(2):149–172.
- Koyama, A., Barolli, L., Matsumoto, K., and Apduhan, B. O. (2004). A GA-based multi-purpose optimization algorithm for QoS routing. In *AINA '04: Proceedings of the 18th International Conference on Advanced Information Networking and Applications*, page 23, Washington, DC, USA. IEEE Computer Society.
- Oliveira, A. and Mateus, G. R. (2006). Using genetic algorithms to LSP setup in MPLS networks. *XXIV Simpósio Brasileiro de Redes de Computadores*, pages 705–721.
- Rosen, E., Viswanathan, A., and Callon, R. (2001). Multiprotocol Label Switching Architecture. *RFC Editor*, RFC 3031.
- Saaty, T. (1980). *The Analytic Hierarchy Process, Planning, Priority Setting, Resource Allocation*. McGraw-Hill, New York.
- Shao, H., Chen, X., and Wang, W. (2006). A multiobjective optimization algorithm for LSP setup in diffserv and MPLS networks. *First International Conference on Communications and Networking*, pages 1–5.
- While, R. L., Hingston, P., Barone, L., and Huband, S. (2006). A faster algorithm for calculating hypervolume. *IEEE Transactions on Evolutionary Computation*, 10(1):29–38.