

ENCAMINHAMENTO DE CHAMADAS TELEFÔNICAS:
ANÁLISE DE DESEMPENHO.

Amaury Kruei Budri 859

DISSERTAÇÃO DE DOUTORADO
submetida à Faculdade de Engenharia Elétrica
da Universidade Estadual de Campinas
como requisito parcial ao título de
doutor em engenharia elétrica.

Banca Examinadora:

Presidente: Prof. Dr. Ivanil Sebastião Bonatti 859
Membros: Prof. Dr. Anilton Salles Garcia
Dra. Eunice Luvizotto Medina Pissolato
Prof. Dr. Jurandir Fernando Ribeiro Fernandes
Prof. Dr. Raul Vinhas Ribeiro
Suplentes: Prof. Dr. João Bosco Ribeiro do Val
Prof. Dr. Akebo Yamakami

Este exemplar corresponde à edição final da tese
defendida por Amaury Kruei Budri
da Comissão
Julgadora em 15.07.94
Bonatti Orientador

Outubro 1994

Pro Brasilia fiant eximia.

Pelo Brasil façam-se grandes coisas.

Divisa do brasão do estado de São Paulo.

Banca Examinadora

Prof. Dr. Ivanil Sebastião Bonatti - Graduou-se em 1973 em engenharia elétrica e obteve o título de mestre em engenharia elétrica em 1975 pela UNICAMP. Em 1981 concluiu seu doutorado no Laboratoire d'Automatique et d'Analyse des Systèmes, Université Paul Sabatier, Toulouse, França. Atualmente é professor da Faculdade de Engenharia Elétrica da UNICAMP.

Prof. Dr. Anilton Salles Garcia - Graduou-se em engenharia mecânica pela Universidade Federal do Espírito Santo (UFES) em 1976. Em 1978 recebeu o título de mestre em matemática aplicada e em 1987 o título de doutor em engenharia elétrica pela UNICAMP. Atualmente é professor da Universidade Federal do Espírito Santo.

Dra. Eunice Luvizotto Medina Pissolato - Graduou-se em 1977 em engenharia elétrica e obteve o título de mestre em engenharia elétrica em 1982 pela UNICAMP. Em 1986 concluiu seu doutorado no Laboratoire d'Automatique et d'Analyse des Systèmes, Université Paul Sabatier, Toulouse, França. Atualmente é pesquisadora no Centro de Pesquisa e Desenvolvimento da TELEBRÁS.

Prof. Dr. Jurandir Fernando Ribeiro Fernandes - É engenheiro mecânico, formado pelo Instituto Tecnológico da Aeronáutica em 1972. Recebeu os títulos de mestre em 1974 e doutor em 1979 pela UNICAMP. Atualmente é professor da Faculdade de Engenharia Elétrica da UNICAMP e secretário de transportes do município de Campinas.

Prof. Dr. Raul Vinhas Ribeiro - Graduou-se em engenharia elétrica em 1973 pelo Instituto Tecnológico da Aeronáutica. Recebeu os títulos de mestre em 1974 e doutor em 1980 pela UNICAMP. Atualmente é professor da Faculdade de Engenharia Elétrica da UNICAMP.

Prof. Dr. João Bosco Ribeiro do Val - Graduou-se em 1977 em engenharia elétrica e obteve o título de mestre em engenharia elétrica em 1981 pela UNICAMP. Recebeu o título de doutor em 1986 pelo Imperial College of Science and Technology, Inglaterra. Atualmente é professor da Faculdade de Engenharia Elétrica da UNICAMP.

Prof. Dr. Akebo Yamakami - Graduado em engenharia elétrica pela Pontifícia Universidade Católica do Rio de Janeiro em 1973. Recebeu os títulos de mestre em 1977 e doutor em 1985 pela UNICAMP. Atualmente é professor da Faculdade de Engenharia Elétrica da UNICAMP.

Resumo.

Este trabalho apresenta propostas de solução para alguns problemas decorrentes da determinação do plano de encaminhamento de redes telefônicas. Basicamente são abordados dois problemas: determinação do encaminhamento de custo mínimo e determinação do encaminhamento de desempenho ótimo.

Numa primeira abordagem, o problema de dimensionamento da rede de troncos é formulado como um problema de fluxo multi-produto de custo mínimo. A solução deste problema resulta num plano de encaminhamento que otimiza o custo de implantação de novos equipamentos na rede. É estudada uma abordagem outra para dimensionamento da rede de troncos utilizando-se o princípio de Moe.

É apresentado um novo método para determinação de um plano de encaminhamento ótimo dependente do estado para uma rede telefônica. O método é fundamentado na teoria de decisão markoviana. A otimização é realizada por meio de simulação.

Este trabalho também apresenta algumas melhorias para modelos de Bloqueio para tráfego telefônico.

Palavras-chave:

- Planejamento Telefônico.
- Encaminhamento Telefônico.
- Encaminhamento Dinâmico.
- Encaminhamento Dependente do Estado.
- Processo de Decisão Markoviano.
- Tráfego Telefônico.

Abstract.

This work presents a study about telephone networks routing plan assignment problems. Basically, two problems are studied: minimal cost routing plan assignment and optimal performance routing plan assignment.

The junction network dimensioning problem is formulated as a multi-commodity minimal cost flux problem. The solution of this problem leads to a routing plan that optimizes the equipment cost of the network. The same problem is formulated, in a second aproach, using the Moe's principle.

A new method to determine the optimal state dependent routing for a telephone network is presented. The method is based on the Markovian decision process theory. The optimization is performed by simulating the network.

The work also presents improvements for bloking models for telephone network traffics.

Keywords:

- Network Planning.
- Routing Plan.
- Dynamic Routing.
- State Dependent Routing.
- Markov Decision Process.
- Telephonic Traffic.

Agradecimentos.

Ao professor Ivanil Sebastião Bonatti, pelo seu apoio incondicional e dedicação à tarefa de orientação, e à sua esposa Nícia, que teve de aturar em dose dupla, orientador e orientado, nos últimos cinco anos.

Aos membros da banca, pela sua atenção e paciência.

Aos planejadores da TELESP, que muito contribuíram com sua experiência e conhecimento para o enriquecimento deste trabalho.

Aos pesquisadores Luiz Del Fiorentino, Jorge Moreira e Marcos Trindade do CPqD da TELEBRÁS, que ofereceram um suporte para minha participação em cursos e seminários, que muito contribuíram para o desenvolvimento deste trabalho.

Aos professores José Claudio Geromel e Edson Moschim que participaram da banca de exame de qualificação.

À FAPESP que apoiou financeiramente este trabalho.

À UNICAMP pelo apoio de infra-estrutura.

Prefácio.

Esta dissertação de doutorado trata do plano de encaminhamento de redes telefônicas sob diferentes aspectos. Basicamente são tratados dois problemas:

- Determinação do encaminhamento de mínimo custo de implantação de novos equipamentos na rede. Este problema consiste no dimensionamento das rotas entre centrais e determinação do plano de encaminhamento de forma integrada, tendo como objetivo minimizar o custo de implantação dos equipamentos.
- Determinação do plano de encaminhamento de ótimo desempenho. É o problema de escolher o plano de encaminhamento que minimiza o bloqueio dos tráfegos da rede.

O capítulo 1 apresenta os objetivos inicialmente estabelecidos para o trabalho de doutorado. Também apresenta a motivação de escolha deste tema.

O capítulo 2 é uma breve revisão bibliográfica a respeito do plano de encaminhamento e da estrutura das redes telefônicas.

No capítulo 3 é analisado o atual estágio tecnológico das redes brasileiras e sua relação com o processo de planejamento de redes.

O capítulo 4 aborda aspectos relativos ao desenvolvimento de ferramentas computacionais de planejamento e operação. São definidas as características necessárias para que estas ferramentas sejam eficientes. É definido um conjunto ideal de ferramentas relacionadas com o plano de encaminhamento da rede.

O capítulo 5 apresenta uma contribuição aos modelos de bloqueio de tráfego. São propostas modificações do modelo para tratar as técnicas de reserva de circuitos e reencaminhamento.

No capítulo 6 é abordada uma formulação tipo fluxo multi-produto de custo mínimo que pode ser aplicada tanto ao problema de determinação do encaminhamento de mínimo custo de implantação de novos circuitos como de determinação do plano de encaminhamento de ótimo desempenho.

O capítulo 7 trata da aplicação da teoria de processos de decisão de Markov ao problema de encaminhamento de redes telefônicas. É apresentado um mecanismo para determinação do encaminhamento dependente de estado ótimo.

O capítulo 8 trata da aplicação do princípio de Moe para o dimensionamento de entroncamento.

Conteúdo

Banca Examinadora	iii
Resumo.	iv
Abstract.	v
Agradecimentos.	vi
Prefácio.	vii
1 Introdução.	1
1.1 Objetivos.	1
1.2 Motivação.	2
1.3 Planejamento Telefônico	3
2 Plano de Encaminhamento.	6
2.1 Estrutura Lógica.	6
2.2 Técnicas de Encaminhamento.	7
2.3 Encaminhamento para a Rede Nacional de Telefonia.	10
2.4 Encaminhamento não Hierárquico	12
2.5 Conclusão	12
3 Contexto.	13
3.1 Novo Contexto de Planejamento.	13
3.2 Dimensionamento do Entroncamento.	15
3.2.1 Caminhos Diretos e Indiretos.	16
3.2.2 Caminhos de Transbordo.	19
3.2.3 Fraca Dependência Entre Roteamento e Entroncamento.	21
3.2.4 As Modularidades dos Meios de Comutação e de Transmissão.	23
3.2.5 Distribuição de Tráfego e Dimensionamento Multi-horário.	25
3.2.6 História da Rede.	25
3.3 Gerência de Rede.	26
3.4 Conclusões.	26

4	Desenvolvimento de Ferramentas.	28
4.1	Quem Toma a Decisão?	28
4.2	Conclusões.	30
5	Modelos de Redes.	33
5.1	Reserva de Circuitos.	33
5.2	Encaminhamento com Retorno.	40
5.3	Conclusão.	41
6	Formulação Tipo Fluxo.	42
6.1	Encaminhamento de Mínimo Custo.	42
6.1.1	Formulação do Problema de Otimização.	44
6.1.2	Algoritmo de Resolução.	46
6.1.3	Experimento Computacional.	51
6.1.4	Aplicação em Redes de Grande Porte.	56
6.2	Encaminhamento de Desempenho Ótimo.	57
6.2.1	Formulação do Problema de Otimização.	57
6.2.2	Algoritmo de Resolução.	60
6.3	Conclusões	63
7	Formulação Tipo Cadeia de Markov.	65
7.1	Cadeia Semi-Markoviana	66
7.2	Políticas de Encaminhamento	68
7.3	Política Ótima	69
7.4	Modelamento de Redes Telefônicas por meio de Cadeias Semi-Markovianas.	70
7.5	Algoritmo de Otimização de Encaminhamento	72
7.6	Infactibilidade de Cadeias Representando Redes Telefônicas.	75
7.7	Modelos Separados por Rota.	76
7.8	Algoritmo de Encaminhamento.	78
7.9	Modelamento de Uma Fila GI/M/C/C.	79
7.10	Modelamento Individual das Rotas.	80
7.10.1	Rede Elementar.	81
7.10.2	Modelo semi-markoviano.	82
7.10.3	Modelo agregado.	83
7.10.4	Modelo GI/M/C/C.	85
7.11	Comparação dos Resultados.	87
7.12	Proposta de Encaminhamento Dependente do Estado Separado por Rotas.	88
7.12.1	Estimativa dos Custos Relativos	89
7.12.2	Discussão do Método de Encaminhamento	92
7.12.3	Resultados Experimentais	93

7.13 Conclusão	99
8 Formulação Tipo Rentabilidade.	100
8.1 Dimensionamento clássico.	100
8.2 Rentabilidade.	101
8.3 Dimensionamento de uma Rota.	102
8.4 Dimensionamento de Rede.	104
8.5 Garantia do Grau de Serviço.	105
8.6 Método Heurístico.	106
8.7 Experimento Computacional.	107
8.8 Conclusão	108
9 Conclusão.	109
Bibliografia	114

Lista de Tabelas

1	Tabela de fatores de concentração de tráfego. Apresenta uma comparação entre duas rotas com probabilidades de bloqueio de 1%. Às rotas são oferecidos os tráfegos A e f.A, onde f é o fator de concentração de tráfego. Os campos preenchidos por ">" indicam um fator superior a 100.	18
2	Fatores de concentração de tráfego na rota alternativa e seus correspondentes ganhos de escala.	20
3	Resultados numéricos para o sistema. São apresentados os parâmetros número de circuitos nas rotas (N_0 e N_1), o limiar de reserva da rota 1 (R_1), Os tráfegos oferecidos (A_0 e A_1). São apresentados os tráfego cursados nas rotas 0 e 1 obtidos a partir do modelo analítico e de simulação.	39
4	Resultados numéricos para o sistema. São apresentadas a diferença normalizada entre os resultados do modelo de simulação e analítico para o tráfego cursado na rota 1, a probabilidade de bloqueio na rota 1 e a probabilidade conjunta de bloqueio para o tráfego A_0 . São apresentadas a probabilidade de bloqueio da rota 1 e a diferença normalizada entre os resultados dos modelos para o bloqueio conjunto para o tráfego A_0	39
5	Esta tabela apresenta, para cada segmento da função custo de tipo contínuo, o seu limite superior, o custo linear por Erlang e o seu custo fixo.	53
6	Esta tabela apresenta, para cada segmento da função custo tipo Modular, o seu limite superior, o custo linear por Erlang e o seu custo fixo.	54
7	Tempo de execução para diferentes portes de redes.	54
8	Parâmetros da rede artificial para cada tipo de arco.	55
9	Resultados obtidos utilizando custos do tipo linear e modular comparados aos resultados obtidos por uma ferramenta clássica de planejamento. Não se conhece precisamente o tempo de execução para o programa OTELO, mas sabe-se que é da ordem de poucos minutos.	56
10	Grandes Números.	76
11	Mapeamento da ocupação das rotas para cada estado da rede.	82

12	Esta tabela apresenta para cada estado da rede quais os estados correspondentes nas rotas 1 e 2, o custo calculado para a rede como um todo, a soma dos custos calculados para as rotas separadamente pelo modelo GI/M/C/C e o erro em relação ao modelo da rede como um todo.	87
13	A tabela apresenta para cada rota o número de circuitos e o tráfego cursado para cada simulação.	94
14	A tabela apresenta o plano de encaminhamento hierárquico, para o qual a rede foi dimensionada (FAR). Para cada tráfego são listados, em ordem de preferência, os caminhos para cada tráfego.	94
15	A tabela apresenta os tráfegos oferecidos e o bloqueio ponto-a-ponto (%).	94
16	A tabela apresenta o número de circuitos para cada rota e o tráfego oferecido entre cada par de centrais da rede. Existem três cenários de simulação. Um cenário de sobrecarga (S1), um cenário onde simula-se uma pane na rota 4-5, que passa de 192 para 72 circuitos (S2), e um cenário onde os tráfegos 1-4 e 4-6 são menores que na condição de sobrecarga (S3).	95
17	A tabela apresenta a sequência de escolha dos caminhos para os encaminhamentos fixos FAR2 (FAR4). As simulações foram realizadas com dois tipos de sequências. Uma que permite apenas dois caminhos alternativos no máximo e outra que permite sempre quatro caminhos. Por exemplo, uma chamada entre as centrais 1 e 6, no encaminhamento FAR2 tenta inicialmente a rota direta 1-6. Se ela estiver congestionada, então é tentado o caminho 1-5-6. Se congestionado, a chamada é bloqueada. Já no encaminhamento FAR4 tenta-se ainda os demais caminhos (pela ordem 1-2-6, 1-3-6, 1-4-6). As mesmas seqüências são utilizadas inicialmente para o encaminhamento dependente do estado, quando ainda não há custos estimados.	96
18	Resultados das simulações para a rede não hierárquica. A tabela apresenta a probabilidade de bloqueio total da rede (%).	96
19	Resultados das simulações para a rede não hierárquica, com os custos calculados para S1. A tabela apresenta a probabilidade de bloqueio total da rede (%).	96
20	Comparação da rentabilidade de duas rotas, com tráfegos oferecidos de 500 e 22 Erlangs respectivamente, com pequenas variações do número de circuitos nas rotas. Pequenas variações no número de circuitos provocam grandes alterações na rentabilidade das rotas.	106
21	Número de circuitos originalmente planejado e redimensionado a partir da abordagem tipo rentabilidade e probabilidade de bloqueio nas rotas no segundo caso.	108
22	Probabilidade de bloqueio ponto-a-ponto.	108

Lista de Figuras

1	Este gráfico apresenta a probabilidade de bloqueio para uma rota com dois tráfegos oferecidos poissonianos, um tráfego nominal e um tráfego intruso. A rota é dimensionada para uma probabilidade de bloqueio de 1% para o tráfego nominal, sem levar em conta o tráfego intruso. O tráfego nominal ocupa a rota sem restrições, enquanto que o tráfego intruso sofre um mecanismo de reserva de circuitos. A figura apresenta gráficos da probabilidade de bloqueio do tráfego nominal, quando o tráfego intruso tem uma intensidade muito grande (infinito), para diferentes valores de reserva de circuitos. São apresentados gráficos para tráfegos nominais de 10, 100 e 1000 Erlangs.	9
2	Caso elementar para estudo do mecanismo de reserva de circuitos.	10
3	Nesta topologia, as centrais estão interligadas por rotas diretas. O tráfego oferecido a cada rota é de dois Erlangs. Para uma probabilidade de bloqueio de 1%, cada rota deve ser dimensionada com 7 circuitos. Como existem 10 rotas na rede, a rede possui um total de 70 circuitos.	17
4	Nesta topologia, as centrais são interligadas por uma central de trânsito. O tráfego oferecido à cada rota é de 8 Erlangs (Fator de concentração de 4 vezes). Para uma probabilidade de bloqueio de 1%, cada rota deve ser dimensionada com 15 circuitos. Como existem 5 rotas na rede, a rede possui um total de 75 circuitos.	17
5	Ganho de Escala para um fator de concentração $f=100$.	18
6	Rede lógica com rotas diretas interligando as centrais. Cada rota possui 53 circuitos, totalizando 159 circuitos para a rede.	21
7	Rede lógica com centrais interligadas via central trânsito. Cada rota possui 96 circuitos, totalizando 192 circuitos para a rede.	22
8	Rede física que interliga as três centrais.	22
9	Gráfico do custo por Erlang cursado na rota, supondo-se uma modularidade de 30 canais.	24
10	Caso elementar de encadeamento com transbordo mútuo e reserva de circuitos.	34
11	Arranjo básico de uma sequência de transbordo.	36
12	Arranjo de uma rota de primeira escolha e uma rota alternativa.	38
13	Caso elementar com "crankback".	40

14	Algoritmo para Resolução do Problema de Encaminhamento de Custo Mínimo de Implantação.	52
15	Algoritmo para Resolução do Problema de Encaminhamento de Ótimo Desempenho.	62
16	Número médio acumulado de clientes que encontram uma fila M/M/1/1 ocupada ao longo do tempo. O limite para tempo infinito da diferença entre as duas curvas é o custo relativo do estado ocupado em relação ao estado livre.	68
17	As centrais A e B e C são interligadas por duas rotas. Existem tráfegos com origem em A e destino em B e C, e com origem em B com destino em C. O tráfego entre A e C utiliza a central B como trânsito e possui uma taxa de geração de chamadas α . Os demais tráfegos possuem uma taxa β	69
18	Rede com seis centrais. A definição dos estados deve ser preferivelmente do tipo tráfego caminho.	72
19	Rede com duas centrais de trânsito. O tráfego com origem em A e destino em D possui três caminhos possíveis. A escolha de qual caminho utilizar depende do algoritmo de encaminhamento dependente do estado da rede.	78
20	Rede elementar. O tráfego AC utiliza a rota 1 em primeira escolha e as rotas 2 e 3 como caminho de transbordo.	81
21	Diagrama dos estados e transições de uma cadeia semi-markoviana que representa a rede elementar.	82
22	Diagrama de transição de estados de uma cadeia de Markov para calcular o número médio de chamadas bloqueadas durante uma chamada, que é o número de chamadas bloqueadas até que o processo alcance o estado 0. O estado 0 representa o fim da chamada.	90
23	O gráfico apresenta, para os estados C-1 e C (onde C é o número de circuitos), as diferenças de custos relativos exatas ($\diamond$) e estimadas (+) para uma rota de 10 circuitos; exatas ($\square$) e estimadas ($\times$) para uma rota com 90 circuitos.	91
24	O gráfico apresenta, para os estados M e M+1 (onde M é dado por $M=C.(10-B)$, B é a probabilidade de bloqueio), as diferenças de custos relativos exatas ($\diamond$) e estimadas (+) para uma rota de 10 circuitos; exatas ($\square$) e estimadas ($\times$) para uma rota com 90 circuitos.	91
25	O tráfego originado em A (a) pode escolher entre dois caminhos alternativos. O caminho do tráfego b contém duas rotas que pertencem a estes caminhos alternativos.	92
26	Rede hierárquica de dois níveis.	93
27	A figura apresenta os valores dos custos relativos do estado 324 circuitos ocupados da rota 1-4, isto é, o número médio de chamadas bloqueadas na rota 1-4 durante a vida de uma chamada que, no instante de nascimento, levou a rota para o estado de 324 circuitos ocupados.	97
28	A figura apresenta os valores dos custos relativos do estado 314 circuitos ocupados da rota 1-4.	97

29	A figura apresenta os valores dos custos relativos do estado 96 circuitos ocupados da rota 2-3.	97
30	A figura apresenta os valores dos custos relativos do estado 71 circuitos ocupados da rota 2-3.	98
31	Rede exemplo para discussão da função rentabilidade.	102
32	Gráfico da rentabilidade da rede em função do número de circuitos na rota 1 da rede exemplo.	103
33	Rede hierarquica de dois níveis, dimensionada pelo método heurístico.	107

Capítulo 1

Introdução.

Legis Cautio Contra Ineptos Criticos.

Qui legent hosce versus nature censunto;

Prophanum vulgus et inscium ne attrectato;

Omnesque Astrologi, Blenni, Barbari procul sunt.

Qui aliter, is, rite, sacer esto.

Nostradamus (1503-66).

1.1 Objetivos.

O presente trabalho é um aprofundamento da temática tratada na dissertação de mestrado “Encaminhamento de Chamadas Telefônicas: Modelamento e Simulação”, onde foram desenvolvidos um modelo analítico e um modelo de simulação a eventos discretos para o cálculo de bloqueio de redes telefônicas [10].

Este trabalho foi desenvolvido, tendo como objetivos:

1. Melhorar os modelos existentes para redes telefônicas, buscando uma maior precisão dos resultados e a ampliação das técnicas de encaminhamento e das características do tráfego oferecido.
2. Desenvolver técnicas de análise de redes telefônicas considerando variações da intensidade de tráfego oferecido ou em situações de contingência.
3. Desenvolver uma metodologia de determinação do plano de encaminhamento de chamadas telefônicas buscando a minimização das probabilidades de bloqueio.
4. Contribuir para o melhoramento da metodologia de planejamento de redes telefônicas, permitindo uma escolha de um plano de encaminhamento com minimização do custo da rede.

1.2 Motivação.

O estudo do encaminhamento de chamadas telefônicas encontra aplicação no desenvolvimento de métodos e ferramentas de planejamento de redes telefônicas e no controle e gerenciamento de redes.

Historicamente as redes telefônicas utilizam uma estrutura hierárquica e a técnica de encaminhamento de transbordo, onde há uma rota de primeira escolha e rotas alternativas. Quando a rota de primeira escolha está inteiramente ocupada, o tráfego passa a ser direcionado para a primeira rota alternativa, e assim sucessivamente até o congestionamento da última rota alternativa. A partir de então as chamadas são bloqueadas [41].

Recentemente os equipamentos de comutação instalados na rede de telefonia passaram a incorporar a tecnologia CPA (controle a programa armazenado), que possui grande flexibilidade em relação às técnicas de encaminhamento, muito superior à encontrada em antigas tecnologias. Nestes equipamentos, é possível implementar, com maior facilidade, novas técnicas de encaminhamento. Novas técnicas de encaminhamento tem sido sugeridas e estudadas com o intuito de melhorar o desempenho das redes telefônicas, considerando inclusive situações de contingência, como falha em equipamentos ou sobrecarga de tráfego.

As redes telefônicas apresentam uma matriz de tráfego variante ao longo do dia, da semana e do ano. Por exemplo, no caso de redes locais, existem pelo menos duas situações distintas ao longo do dia. No horário comercial a população se desloca para as áreas comerciais e industriais da cidade. Há portanto um tráfego telefônico intenso nestas áreas. No horário não comercial as pessoas estão nas áreas residenciais. Se esta rede empregar um único plano de encaminhamento do tipo hierárquico, o seu dimensionamento, mesmo que leve em conta a existência dos dois horários (dimensionamento multi-horário), tornará a rede ociosa. No horário comercial os circuitos que interligam as centrais das regiões residenciais estarão superdimensionados. O mesmo ocorre para as áreas comerciais no outro horário. Este problema pode ser minimizado com a aplicação de novas técnicas de encaminhamento.

Entre as facilidades incorporadas às centrais modernas está a de permitir uma mudança do plano de encaminhamento ao longo do tempo, em horários pré-determinados ou mesmo de acordo com as condições de tráfego medidas em um dado período (encaminhamento adaptativo) [3], [4], [5], [8].

Por outro lado, as atividades de operação da rede telefônica tornam-se cada vez mais complexas. O gerenciamento de redes é um aspecto que tem ganho grande importância nos últimos anos. Desde 1963 a AT&T tem treinado profissionais nesta área. Na década de 80 foram propostos e implementados diversos sistemas de gerência de redes em tempo real [19].

Esta preocupação em gerenciamento de redes tem aumentado nas principais administrações. Algumas redes possuem um controle de tráfego em tempo real, que realizam mudanças automáticas no plano de encaminhamento.

A privatização do sistema de telecomunicação ou o fim de monopólios, como ocorreu na Inglaterra, Japão e EUA, trouxe uma competição entre empresas telefônicas, particularmente no que se

refere ao tráfego interurbano de longa distância. O usuário pode escolher a rota para sua chamada, utilizando diferentes códigos de acesso. Estes eventos trouxeram uma preocupação maior voltada para a opinião dos usuários. Estudos de mercado mostram que a opinião do usuário sobre o desempenho do sistema telefônico é particularmente influenciada pelos períodos de pior desempenho devido ao congestionamento ou falhas [2]. Como consequência as redes telefônicas passaram a ser avaliadas não apenas nas condições nominais, mas também em condições de sobrecarga, contingência e variação de tráfego [2], [11].

Neste contexto, o plano de encaminhamento de uma rede telefônica é um elemento chave, tanto do ponto de vista de planejamento como de operação.

O presente trabalho apresenta um estudo de problemas que envolvem o plano de encaminhamento de uma rede telefônica, tendo em vista a melhoria da qualidade de serviço e o planejamento de redes telefônicas.

1.3 Planejamento Telefônico

Devido a sua grande complexidade, o planejamento de redes telefônicas exige a decomposição do problema em partes interrelacionadas, mas que possam ser tratadas separadamente. É feita uma decomposição do problema no tempo e no espaço [15].

A decomposição no tempo é feita fixando-se um horizonte de planejamento, que deve ser escolhido de acordo com a previsão de alterações tecnológicas da rede. Este período (10 a 20 anos) é subdividido em estágios de curto e médio prazo a partir do ano base (ano inicial) até atingir o ano horizonte.

A decomposição no espaço consiste em planejar separadamente segmentos da rede. A decomposição no espaço inicialmente divide a rede telefônica em redes internacional, interurbana e local.

A rede local é a seção da rede que compreende as centrais locais (centrais que possuem terminais de assinantes), as centrais tandem (centrais de trânsito locais, isto é, que comutam tráfegos locais que são originados e terminados em outras centrais da rede local) e as centrais de trânsito interurbanas (centrais que comutam tráfegos que são originados e terminados em outras centrais de redes locais diferentes) interligadas às centrais da rede local.

A rede interurbana é a seção da rede que compreende as centrais de trânsito interurbanas de uma dada região geográfica.

A rede internacional é a rede formada pelas centrais de trânsito internacionais dos diversos países integrados ao sistema internacional de telecomunicações.

Para as redes locais, a decomposição no espaço é feita planejando-se, num primeiro passo, a localização das centrais locais, suas áreas de atendimento e as redes primária e secundária. As redes primária e secundária constituem as ligações dos terminais de assinante aos equipamentos de comutação. A localização das centrais, e a determinação de suas áreas de atendimento, consiste em determinar quantos equipamentos de comutação serão implantados, sua capacidade e localização e na identificação dos assinantes filiados a cada equipamento com o objetivo de minimizar o custo de

implantação da interface com os assinantes. Esta etapa do planejamento envolve a maior parcela dos custos de implantação. Isto permite que ela seja executada de forma independente das demais etapas.

Para redes interurbanas, o primeiro passo também é determinar a instalação das centrais de trânsito. No entanto, neste caso sua localização visa minimizar o custo de ligação entre a central trânsito e as centrais locais a ela subordinadas. Outro fator importante na localização das centrais são problemas relativos à facilidades existentes nas centrais. Nem todas as centrais são capazes de realizar funções como bilhetagem de chamadas e inserção de supressores de eco. As chamadas que necessitam destes recursos devem ser encaminhadas para centrais apropriadas, o que deve ser levado em conta na localização das centrais.

O segundo passo, tanto para redes locais como para redes interurbanas, é determinar como as várias centrais estarão interligadas e como as chamadas serão encaminhadas pela rede. Para este segundo passo existe uma abordagem clássica e uma abordagem moderna [32], [16], [26].

A abordagem clássica subdivide o problema em três etapas de planejamento que tratam do encaminhamento, entroncamento e roteamento.

- A etapa de encaminhamento consiste em determinar o trajeto para cada tipo de chamada na rede. O plano de encaminhamento é o conjunto de regras para escolha do caminho lógico (sequência de centrais) para estabelecimento das chamadas. Atualmente a técnica de encaminhamento utilizada na rede nacional é de encaminhamento fixo com transbordo [1]. Em geral esta etapa tem sido tratada de forma empírica e não otimizada.
- A etapa de entroncamento tem como objetivo dimensionar o número de canais lógicos interligando cada central, de forma a garantir a qualidade de serviço especificada. O principal item de qualidade de serviço a ser observado é a probabilidade de bloqueio.
- A etapa de roteamento determina os meios de transmissão e sua localização na rede física. O roteamento determina os meios de transmissão e os equipamentos de multiplex a serem utilizados e o caminho físico onde estes meios de transmissão serão instalados [16].

Esta metodologia foi criada para o dimensionamento de redes com tecnologia de comutação eletromecânica e tecnologia de transmissão analógica. Neste cenário as centrais eletromecânicas não possuem flexibilidade quanto ao encaminhamento. Usualmente estas centrais utilizam a técnica de transbordo com uma ou duas rotas alternativas. O custo de comutação para centrais tandem ou trânsito é elevado. Como consequência o encaminhamento é bastante conservador e do tipo hierárquico. A transmissão por sua vez consiste em cabos de pares para curtas distâncias ou rádio para longas distâncias. Assim, para uma dada rede é possível prever com razoável precisão os custos de transmissão a utilizar na fase de entroncamento. Conseqüentemente, entroncamento e roteamento podem ser tratados separadamente [34], [35].

Atualmente, os novos equipamentos de comutação e transmissão implantados utilizam tecnologia digital. Juntamente com a introdução desta nova tecnologia, surgiu a utilização de satélite e

fibras ópticas para chamadas nacionais, criando assim um novo cenário. As centrais tiveram seu tamanho físico reduzido, não exigindo grandes prédios para sua instalação, em comparação às centrais eletromecânicas. O custo para implementação da função tandem ou trânsito nestas centrais foi também diminuído. As possibilidades de encaminhamento foram ampliadas e sofisticadas com a introdução de novas técnicas de encaminhamento e facilidades como por exemplo o encaminhamento dinâmico.

No cenário digital os equipamentos de transmissão oferecem uma ampla escolha entre rádio, cabos coaxiais ou fibra óptica. Cada um destes meios suporta diversas hierarquias. Cada combinação, meio de transmissão e hierarquia, pode resultar em custos muito diferentes e conseqüentemente o custo por canal é não linear e é modular em relação ao número de canais utilizados. Torna-se difícil determinar o custo por canal para o cálculo do entroncamento sem levar em conta o roteamento, que por sua vez depende dos resultados do entroncamento.

Desta forma as etapas de encaminhamento, roteamento e entroncamento ganham uma maior interdependência, tornando menos eficaz a abordagem clássica de planejamento de redes.

Entretanto, os custos dos equipamentos de transmissão por unidade de tráfego cursado são muito menores que o custo dos equipamentos de comutação nas redes locais. O custo de interligação das centrais praticamente independe do caminho físico escolhido. Isto torna possível resolver o problema de roteamento após os problemas de encaminhamento e entroncamento.

O impacto causado pela tecnologia digital, tem maiores conseqüências nas redes locais. As redes interurbanas são menos afetadas porque neste caso a escolha entre vários meios de transmissão é mais limitada e o encaminhamento é mais rígido. A escolha dos meios de transmissão para redes interurbanas depende preponderantemente do volume de tráfego, da distância a ser percorrida e da localização da central. Isto permite que a escolha dos meios de transmissão e o cálculo dos custos de transmissão sejam feitos anteriormente às etapas de encaminhamento, entroncamento e roteamento. Neste caso, o sistema de planejamento clássico é eficiente. Todavia, as redes interurbanas podem se beneficiar do emprego de novas técnicas de encaminhamento, que exigem novos métodos de planejamento.

Atualmente estão sendo propostas novas metodologias, mais adequadas às novas características das redes telefônicas, que procuram resolver as etapas de planejamento clássicas de forma mais integrada, dando origem a uma abordagem moderna do problema [26]. Os trabalhos apresentados neste contexto se concentram no roteamento e entroncamento de redes.

Este trabalho apresenta propostas para solução do problema de determinação do plano de encaminhamento de forma otimizada e integrada com as demais etapas de planejamento.

Capítulo 2

Plano de Encaminhamento.

Sabedoria é uma coisa.

*É conhecer o pensamento pelo qual todas as coisas
são encaminhadas através de todas as coisas.*

Heráclito (500 A.C.).

Este capítulo apresenta uma breve introdução sobre o plano de encaminhamento para redes telefônicas. São apresentados os dois tipos de estrutura lógica de rede: rede hierárquica e não hierárquica. São apresentados também algumas novas técnicas de encaminhamento propostas. Finalmente é apresentado o plano de encaminhamento atualmente empregado no sistema de telefonia nacional.

Para estabelecer uma conexão entre dois terminais de uma rede telefônica, é preciso determinar um caminho, ou seja determinar os equipamentos de comutação que serão utilizados para estabelecimento da ligação. Em geral, as redes telefônicas admitem mais de um caminho possível para interligar dois pontos. A determinação deste caminho é feita a partir de um conjunto de regras.

O plano de encaminhamento é o conjunto de regras para o estabelecimento de uma chamada numa rede telefônica. O plano de encaminhamento pode ser classificado quanto a estrutura lógica da rede e quanto à técnica de encaminhamento empregada.

2.1 Estrutura Lógica.

As redes telefônicas possuem uma estrutura física e uma estrutura lógica. A estrutura física é determinada pela maneira que os equipamentos da rede são instalados e interligados fisicamente. A estrutura lógica determina como as centrais de comutação estão interligadas, sem levar em conta os caminhos físicos utilizados nesta ligação. Também determina quais os caminhos que podem ser escolhidos ou não para o estabelecimento de uma dada chamada. As redes podem ser classificadas, quanto a sua estrutura lógica, como redes hierárquicas ou não hierárquicas.

As redes hierárquicas apresentam, para as centrais de trânsito ou tandem, uma estrutura em níveis. Uma central de trânsito ou tandem é aquela cuja principal função é interligar outras centrais

entre si. A central tandem é uma central de trânsito utilizada no encaminhamento de chamadas locais. Nas redes hierárquicas há uma subordinação dos níveis inferiores aos níveis superiores. Esta subordinação se traduz por meio de regras que restringem o encaminhamento. Basicamente, uma central de trânsito não deve encaminhar chamadas originadas ou terminadas em centrais a ela não subordinadas.

A rede nacional de telefonia possui quatro níveis hierárquicos. As centrais são denominadas centrais de trânsito classe I, II, III ou IV de acordo com o nível a que pertencem. As centrais de classe I são as de maior hierarquia. A área geográfica atendida por uma central de trânsito e por todas as centrais de trânsito de classe inferior a ela subordinada direta ou indiretamente é denominada área de grupamento. De acordo com a classe da central a área de grupamento é chamada de primária, secundária, terciária ou quaternária. A área atendida por uma rede local ligada a uma central de qualquer classe também pode ser chamada de área quaternária.

A estrutura não hierárquica permite a livre interligação entre centrais, sem qualquer tipo de subordinação. A estrutura não hierárquica ainda mantém níveis de comutação. As redes não hierárquicas propostas, e algumas implementadas, em geral apresentam um ou dois níveis de comutação. A principal razão do emprego da estrutura hierárquica é que as centrais analógicas apresentavam limitações para o encaminhamento e tarifação de chamadas. Para as redes que empregam centrais de tecnologia CPA, o emprego da estrutura não hierárquica é preferível, pois permite reduzir o custo de implantação de circuitos e melhorar a qualidade de serviço.

2.2 Técnicas de Encaminhamento.

Define-se técnica de encaminhamento como o processo pelo qual um caminho é selecionado para o estabelecimento de uma chamada numa rede telefônica, dentre aqueles caminhos possíveis de acordo com a estrutura lógica da rede.

A técnica clássica de encaminhamento é a técnica de transbordo. No encaminhamento por transbordo uma chamada, ao encontrar uma determinada rota congestionada, pode tentar obter um circuito em uma outra rota chamada rota alternativa. A escolha das rotas é feita a partir de uma tabela fixa, que determina a ordem de transbordo.

Com a evolução da tecnologia de comutação estão sendo criadas novas técnicas de encaminhamento de chamadas pela rede, com o objetivo de maximizar o aproveitamento dos equipamentos instalados na rede e proporcionar um melhor serviço aos usuários.

Uma das primeiras técnicas apresentadas na literatura foi o encaminhamento por repartição de cargas. Este tipo de encaminhamento distribui o tráfego pelas diversas rotas (ou caminhos) possíveis de acordo com percentuais pré-estabelecidos. Se houver congestionamento em uma das rotas é permitido o transbordo, também distribuído, pela demais rotas [9] [13], [8]. Quando um tráfego poissoniano é repartido aleatoriamente, chamada a chamada, em diversos fluxos, de acordo com probabilidades fixas, cada fluxo também é poissoniano. Define-se tráfego poissoniano como o processo de estabelecimento de chamadas onde o tempo entre nascimento de chamadas e o tempo de

duração das chamadas são variáveis aleatórias independentes de distribuição exponencial negativa. Esta propriedade torna os modelos matemáticos mais simples.

A introdução na rede de centrais telefônicas de tecnologia tipo CPA-T permite a utilização de técnicas de encaminhamento dinâmico. No contexto deste trabalho, considera-se encaminhamento dinâmico o plano de encaminhamento que pode ser modificado automaticamente ao longo do tempo. Esta mudança pode ser feita de acordo com um calendário ou de acordo com as condições da rede. O encaminhamento dinâmico torna a rede mais robusta e flexível.

O emprego de técnicas de encaminhamento dinâmico permite reduzir ao máximo a ociosidade de equipamentos instalados na rede, reduzindo o custo e melhorando a qualidade de serviço telefônico. Estas técnicas ainda não tem sido largamente implementadas, mas algumas empresas já iniciam sua utilização com sucesso. As principais propostas são [5], [8], [12], [24], [29], [36], [42]:

- DNHR - Dynamic Nonhierarquical Routing - AT&T.
- DTM - Dynamic Traffic Management - Bell-Northern Research.
- STAR - System to Test Adaptive Routing - CNET.
- DAR - Dynamic Alternate Routing - British Telecom.
- STR - State and Time Dependent Routing - NTT.

O encaminhamento dinâmico está fundamentado nos seguintes mecanismos de encaminhamento:

- Agendamento. Esta técnica consiste em alterar automaticamente o plano de encaminhamento em horários pré-estabelecidos. Uma vez que o perfil de tráfego varia ao longo do dia, ou mesmo ao longo dos dias da semana e do ano, o agendamento permite que o encaminhamento acompanhe esta variação, otimizando a utilização dos recursos da rede.
- Reserva de circuitos. É mecanismo simples, porém muito eficiente, de encaminhamento dinâmico. Esta técnica consiste em permitir que determinados tráfegos ocupem circuitos de uma rota apenas se o número de circuitos vagos for superior a um certo limiar. Como regra geral, estes tráfegos são aqueles transbordados de outras rotas, enquanto que o tráfego oferecido em primeira escolha não sofre restrições. Neste caso, o tráfego de primeira escolha é protegido, permitindo uma distribuição mais homogênea da probabilidade de bloqueio entre as fontes de tráfego.

Por exemplo, seja uma rota com dois tráfegos oferecidos poissonianos, um tráfego nominal e um tráfego intruso. A rota é dimensionada para uma probabilidade de bloqueio de 1% para o tráfego nominal, sem levar em conta o tráfego intruso. O tráfego nominal ocupa a rota sem restrições, enquanto que o tráfego intruso sofre um mecanismo de reserva de circuitos. A figura 1 apresenta gráficos da probabilidade de bloqueio do tráfego nominal, quando o tráfego intruso tem uma intensidade muito grande (infinito), para diferentes valores de reserva de circuitos.

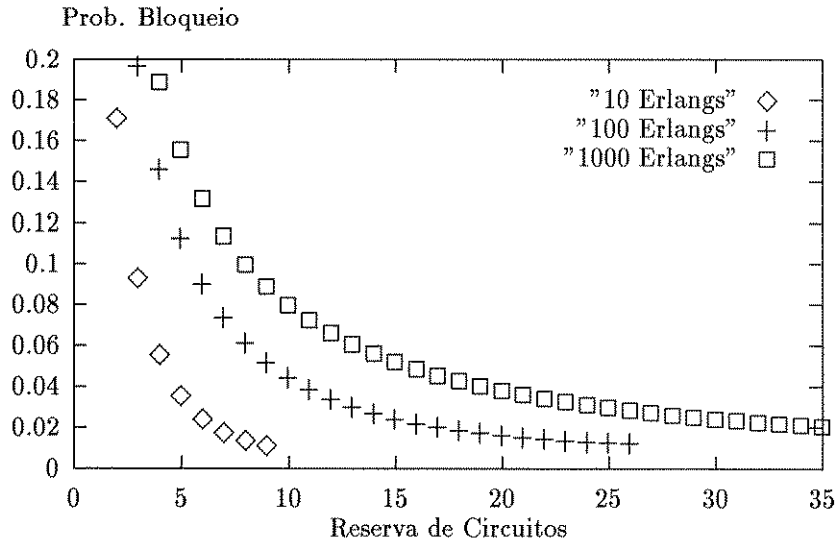


Figura 1: Este gráfico apresenta a probabilidade de bloqueio para uma rota com dois tráfegos oferecidos poissonianos, um tráfego nominal e um tráfego intruso. A rota é dimensionada para uma probabilidade de bloqueio de 1% para o tráfego nominal, sem levar em conta o tráfego intruso. O tráfego nominal ocupa a rota sem restrições, enquanto que o tráfego intruso sofre um mecanismo de reserva de circuitos. A figura apresenta gráficos da probabilidade de bloqueio do tráfego nominal, quando o tráfego intruso tem uma intensidade muito grande (infinito), para diferentes valores de reserva de circuitos. São apresentados gráficos para tráfegos nominais de 10, 100 e 1000 Erlangs.

São apresentados gráficos para tráfegos nominais de 10, 100 e 1000 Erlangs. O tráfego intruso apresenta uma probabilidade de bloqueio muito alta (cerca de 100%) para qualquer reserva. Neste exemplo, procura-se mostrar no pior caso, a efetividade do mecanismo de reserva de circuitos.

A aplicação desta técnica pode ser feita de forma descentralizada na rede, exigindo apenas que o programa armazenado das centrais seja capaz de reservar os circuitos. Graças a suas características, a técnica de reserva de circuitos tem sido incorporada em diversas propostas de novos esquemas de encaminhamento.

Para ilustrar esta técnica, seja por exemplo o arranjo da figura 2. A rota 1 é de primeira escolha para o tráfego A1. A rota 2 é a rota de primeira escolha para o tráfego A2 e rota alternativa para o tráfego transbordado da rota 1.

Este arranjo foi simulado para $A1 = A2 = 5$ Erl, e $N1 = N2 = 5$. Verifica-se que o tráfego A1 sofre um bloqueio de 14% das chamadas, enquanto que o tráfego A2 sofre um bloqueio de 38% das chamadas. Utilizando-se a técnica de proteção na rota 2, é possível aumentar a probabilidade de bloqueio do tráfego A1 e diminuir o bloqueio para A2, de forma que o bloqueio das duas fontes seja aproximadamente igual. Por exemplo, para um limiar de proteção da rota 2 de 3 circuitos, isto é o tráfego A1 somente transborda para a rota 2 se restarem pelo menos 3 circuitos vagos em $N2$, o bloqueio do tráfego A2 passa para 28% enquanto que o do tráfego A1 passa a 29%.

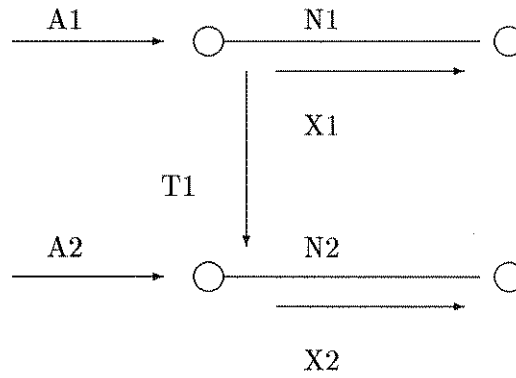


Figura 2: Caso elementar para estudo do mecanismo de reserva de circuitos.

- **Encaminhamento Dependente do Estado.** No encaminhamento dependente do estado, a escolha do caminho para estabelecimento das chamadas é executada levando-se em conta o estado de ocupação das rotas da rede. Este tipo de encaminhamento é abordado no capítulo 7.
- **Maximização da capacidade residual.** Neste caso as chamadas devem escolher os caminhos de forma a manter as rotas com o menor número de circuitos ocupados. Esta técnica não pode ser implementada de forma totalmente descentralizada. É preciso que haja uma coordenação entre as centrais, uma vez que é preciso conhecer a cada instante quantos circuitos estão ocupados (ou desocupados) em cada rota da rede. É uma forma particular de encaminhamento dependente do estado.
- **Autômatos aprendizes.** A técnica de autômatos aprendizes ("learning automatas") consiste num método de tentativa e erro. Para cada chamada existe um conjunto de planos de encaminhamento possíveis. Sempre que uma chamada é bloqueada, substitui-se o plano utilizado por um outro pertencente a este conjunto. Por exemplo, seja uma rede onde, para uma dada chamada, há um caminho de primeira escolha e uma dezena de caminhos alternativos. Sempre que o caminho de primeira escolha está congestionado é tentado um dos caminhos alternativos. Se a chamada for bem sucedida, este caminho é mantido como alternativo. Se a chamada for bloqueada então é escolhido outro caminho, entre os outros nove caminhos possíveis.

2.3 Encaminhamento para a Rede Nacional de Telefonia.

As redes brasileiras atuais se caracterizam por possuir uma estrutura hierárquica e utilizar a técnica de transbordo. Este tipo de rede surgiu com a introdução do encaminhamento semi-automático e automático nas redes telefônicas. Estas características estão relacionadas ao desenvolvimento tecnológico da época.

O estudo do congestionamento deste tipo de estrutura de rede historicamente utilizou uma abordagem por rotas. Isto significa que os principais elementos de planejamento e análise da rede

são voltados para as rotas. O parâmetro principal considerado é a probabilidade de bloqueio nas rotas. A partir de limites de bloqueio nas rotas e regras de encaminhamento procura-se garantir um desempenho mínimo ponto a ponto, que é o parâmetro percebido pelos usuários da rede. Neste contexto são estabelecidas as seguintes definições [1]:

- Rota de primeira escolha: É a rota na qual um dado tráfego realiza a primeira tentativa de alocação de circuitos.
- Rota alternativa: É a rota que recebe o transbordo de outras rotas que podem ser de primeira escolha ou alternativas.
- Rota Direta: Rota que interliga duas centrais sem subordinação hierárquica.
- Rota Final: Rota que interligue uma central a outra, de ordem hierárquica superior, à qual esteja subordinada ou que interligue duas centrais de trânsito classe I. Esta rota tem baixa probabilidade de congestionamento nominal (bloqueio $< 1\%$) e não admite rota alternativa.
- Rota de Alto Uso: Rota direta com alta probabilidade de congestionamento nominal (tipicamente, $1\% \leq \text{bloqueio} \leq 10\%$) que possui uma ou mais rotas alternativas.
- Rota de Proteção: Sub-grupo de circuitos de uma rota final que são ocupados somente por chamadas para as quais esta rota constitui a primeira escolha (e portanto única). Estas chamadas tentam ocupar inicialmente a rota de proteção e posteriormente os demais circuitos da rota final. A rota de proteção é dimensionada com alta probabilidade de congestionamento.

O plano de encaminhamento para uma rede hierárquica com técnica de transbordo possui os seguintes princípios de encaminhamento:

- As centrais de trânsito classe I devem ser interligadas duas a duas através de rotas finais. Em condições normais uma chamada não deve ser encaminhada por mais de duas centrais classe I. Em situações anormais de sobrecarga e em períodos de tempo curtos pode-se empregar uma terceira central classe I como medida de proteção de caráter temporário.
- As centrais de trânsito classe I devem ser ligadas às centrais internacionais através de rotas finais.
- Uma central tandem local pode exercer as funções de trânsito interurbana. As centrais locais podem ser subordinadas a centrais trânsito de qualquer classe.
- As rotas finais devem ter baixas perdas ($< 1\%$).
- Sempre que o tráfego e o custo justificarem poderão ser abertas rotas diretas entre quaisquer centrais. As chamadas encaminhadas por rotas diretas somente poderão ser comutadas por centrais pertencentes à cadeia de rotas finais que interligue origem e destino

- Uma chamada não deve ser encaminhada por central de classe superior se houver uma rota equivalente em número de comutações através de centrais de classe inferior.
- Sempre que uma rota final receber transbordo considerável de tráfego originado em rotas de alta perda, as rotas finais devem possuir rotas de proteção. Este procedimento evita que o tráfego de primeira escolha seja prejudicado pelo de transbordo.

2.4 Encaminhamento não Hierárquico

O encaminhamento hierárquico, empregado atualmente na rede nacional de telefonia, deve dar lugar ao encaminhamento não hierárquico. O encaminhamento hierárquico surgiu dentro de um contexto diferente das redes digitais. No contexto antigo, as centrais possuíam limitações técnicas quanto ao plano de encaminhamento. O custo da função trânsito era bastante elevado.

A tendência para as redes digitais é de interligar diretamente um grande número de centrais, e de permitir que qualquer central da rede realize a função trânsito. A escolha dos caminhos indiretos e alternativos é feita do ponto de vista de otimização do uso dos equipamentos da rede, de forma não hierárquica.

Este tipo de plano de encaminhamento, para que possa ser eficiente, exige o emprego da técnica de reserva de circuitos e do reencaminhamento (crankback). O reencaminhamento consiste em permitir que uma central tente estabelecer uma chamada num caminho alternativo, mesmo quando o congestionamento ocorre num rota de saída de uma outra central à frente. O capítulo 5 apresenta um exemplo de reencaminhamento.

2.5 Conclusão

Nas redes formadas por centrais telefônicas do tipo CPA, é preferível empregar uma estrutura de rede não hierárquico em lugar do encaminhamento hierárquico clássico, que está intimamente relacionado à técnica de transbordo. A estrutura hierárquica impõe restrições de encaminhamento, necessárias nas redes analógicas por motivos tecnológicos, que nas redes digitais não contribuem para a melhoria da qualidade de serviço e do custo de implantação da rede. Novas técnicas de encaminhamento tem sido propostas e implementadas para a melhoria destes aspectos.

Capítulo 3

Contexto.

Teorema fundamental da sistemática:

Novos sistemas geram novos problemas.

John Gall.

Este capítulo apresenta aspectos tecnológicos das redes telefônicas atuais que contribuem para uma evolução do método de planejamento clássico. Dentro deste contexto, é analisado o problema de dimensionamento da rede de troncos e determinação do plano de encaminhamento.

3.1 Novo Contexto de Planejamento.

A rede de telefonia tem evoluído de maneira sistemática para uma rede digital integrada, onde foram introduzidas novas tecnologias, como por exemplo, a telefonia móvel e a sinalização por canal comum, que alteraram significativamente as potencialidades de comunicação e os custos relativos dos componentes da rede.

Na prestação do serviço básico de telefonia a técnica de comutação empregada é a de comutação de circuitos e os novos equipamentos de comutação implantados são do tipo centrais de comutação temporal controladas a programas armazenados (CPA-T).

As centrais CPA-T possuem dois tipos diferentes de arquitetura interna que influenciam significativamente o planejamento telefônico: centrais de processamento centralizado e de processamento descentralizado.

As centrais de processamento centralizado possuem um único processador (processador central) que concentra certas funções, em geral as mais complexas (por exemplo, análise de encaminhamento). Nenhum outro processador pode realizar estas funções. Nesta estrutura, podem existir processadores periféricos, que realizam as demais funções (por exemplo, detectar se um assinante retira o fone do gancho). Para centrais de médio e grande porte, o processador central possui grande capacidade, velocidade e confiabilidade. Conseqüentemente seu custo é elevado, exigindo um grande investimento inicial. Como estratégia para minimizar este efeito, adotam-se estágios de linha remotos, que utilizam o processador da central mãe até que seu porte justifique a instalação

de um processador próprio, tornado-o uma central autônoma.

As centrais de processamento descentralizado não concentram nenhuma função num único processador. Portanto não há necessidade de um processador central caro. A capacidade de processamento é distribuída entre vários processadores. O número destes processadores pode variar de acordo com o porte instalado da central. Consequentemente, o custo é aproximadamente proporcional ao porte da central.

Nas redes locais (redes urbanas e rurais), a interface com os assinantes é analógica (cabos de pares) ou via rádio (no caso de assinantes móveis ou assinantes rurais). Nesta interface concentra-se a maior parcela do custo de implantação. A minimização deste custo é um dos principais objetivos do planejamento da expansão das redes. Esta minimização é obtida a partir da localização das centrais, dos estágios de linha remotos e dos concentradores e da filiação dos assinantes a estes elementos da rede.

As redes estão evoluindo para uma interface com o assinante digitalizada, com uma tendência de aumento de custo e de capacidade de comunicação para novos serviços. O custo de comutação e o custo de interface com os assinantes representa da ordem de 85% do custo total de implantação. O custo restante é relativo aos equipamentos de transmissão [21] [23].

Na rede de telefonia, a interface da central local com as demais centrais é feita por meio de enlaces digitais. Atualmente o entroncamento (ligação lógica entre centrais) é feito ao nível da hierarquia mais baixa de transmissão (30 canais).

Os novos equipamentos de transmissão implantados empregam fibras ópticas para as redes locais e fibras ópticas, rádio digital e satélite para as redes de longa distância. Atualmente, no caso de meio óptico, são utilizados equipamentos a 34 Mbits/s e 140 Mbits/s. Hierarquias mais elevadas deverão entrar em operação ao longo da década. O custo de transmissão é da ordem de 15% do custo total de implantação.

Devido a complexidade da tarefa de planejamento da expansão das redes de telefonia e dada a distribuição dos custos de implantação entre interface com assinantes, comutação e transmissão, o processo de planejamento é subdividido em diversas etapas [15].

- Localização de centrais, estágios de linha remotos e concentradores e determinação das áreas de filiação das centrais.
- Determinação do plano de encaminhamento (definição dos caminhos das chamadas na rede).
- Dimensionamento do entroncamento (dimensionamento da interligação lógica entre as centrais de comutação).
- Determinação do roteamento (definição dos meios físicos e equipamentos de transmissão)

3.2 Dimensionamento do Entroncamento.

A metodologia de dimensionamento do entroncamento clássica, ainda hoje utilizada, foi desenvolvida a partir dos trabalhos de Rapp e Walström [34], [35], [40]. Esta metodologia foi desenvolvida para um cenário analógico e para uma topologia hierárquica da rede, com a utilização da técnica de encaminhamento por transbordo. O método dimensiona as rotas a partir de um plano de encaminhamento pré-determinado. O modelamento de tráfego adotado foi o modelo equivalente de Wilkinson [41]. A formulação matemática de Rapp introduz aproximações que permitem desacoplar o problema em minimizações locais, permitindo a abordagem de problemas de grande porte com um esforço computacional moderado.

A base da metodologia de Rapp consiste em obter uma relação ótima entre o número de circuitos em uma rota direta e o número de circuitos em um caminho alternativo, em geral composto por duas rotas. Embora o caminho alternativo utilize duas vezes mais circuitos por conexão, seu custo pode ser atrativo, uma vez que suas rotas podem concentrar vários tráfegos. Isto proporciona um ganho de escala que pode ser compensador, particularmente no ambiente analógico onde o custo em função da distância (meios de transmissão) é significativo.

O cenário de telecomunicações para os próximos anos aponta para redes telefônicas onde os novos equipamentos de transmissão utilizarão em sua maioria fibras ópticas, fazendo com que os custos de transmissão sejam quase invariantes com a distância entre as centrais locais. Por outro lado espera-se que o custo de comutação seja muito superior ao de transmissão [21], [23], [31].

O custo de entroncamento ponto a ponto (comutação e transmissão) deve continuar sendo heterogêneo, uma vez que convivem na rede centrais de diferentes tecnologias, portes e fabricantes. Entretanto, se para análise do efeito da aplicação da metodologia clássica a este cenário, supor-se um custo de comutação por circuito igual para todas centrais, pode-se verificar que o resultado será uma rede interligada totalmente por rotas diretas, exceto no caso de interesses de tráfego muito pequenos.

Eventualmente estas rotas diretas poderão ser também utilizadas como rotas alternativas para outros tráfegos se forem utilizadas técnicas de encaminhamento com reserva de circuitos. Isto é possível graças ao baixo custo para utilizar centrais CPA-T como centrais locais/tandem.

Paradoxalmente esta abordagem do problema choca-se com a prática de utilização intensiva de estágios de linha remotos, que é uma tendência no planejamento das redes atuais. Os estágios de linha podem ser compreendidos como centrais de pequeno e médio porte que não possuem autonomia de encaminhamento e não possuem entroncamento a não ser com a central "mãe". São centrais que não permitem o estabelecimento de rotas diretas com as demais centrais. Este paradoxo é explicado por fatores externos ao processo de dimensionamento das rotas, tais como o custo de implantação das centrais e o custo de operação de um sistema com um grande número de rotas.

3.2.1 Caminhos Diretos e Indiretos.

A seguir são discutidos alguns pontos importantes para elaboração de uma nova metodologia de dimensionamento da rede de entroncamento. Na discussão que se segue, foi adotada uma hipótese em relação à rede quanto aos aspectos de custo do entroncamento que não pode ser aplicada às redes em geral. Supõe-se que o custo de interligar duas centrais na rede de entroncamento é invariante com sua distância e independe do tipo ou porte das centrais interligadas, e portanto o custo de interligação de duas centrais é proporcional ao número de circuitos do enlace de entroncamento. O objetivo desta hipótese é permitir observar os aspectos de ganho de escala envolvendo a concentração de tráfego.

Serão comparados os custos de implantação de um caminho por rota direta e de um caminho utilizando duas rotas, via central tandem. Se o custo por circuito for idêntico para todas as rotas, o caminho direto tem um custo por circuito igual à metade do custo do caminho indireto. Portanto, para ser economicamente viável, o caminho indireto deve apresentar um ganho de escala superior a dois ($G = 2$).

Por exemplo, seja uma rede telefônica com cinco centrais. Todas as centrais são indênticas e possuem um interesse de tráfego bidirecional de dois Erlangs entre cada par de centrais. Deseja-se dimensionar o entroncamento da rede de forma otimizada. Devido à simetria do problema, existem duas soluções para o entroncamento. Uma interliga todas as centrais por rotas diretas, como mostrado na figura 3 e exige 70 circuitos no total. A outra interliga todas as centrais por uma central de trânsito como na figura 4, exigindo 75 circuitos no total. Neste caso, cada chamada estabelecida ocupa dois troncos, enquanto que na solução anterior ocupa apenas um tronco na rota direta. Em contrapartida, as rotas no segundo caso são maiores. Isto leva a um ganho (ganho de escala) de eficiência. Para uma probabilidade de bloqueio de 1%, rotas com tráfego oferecido de dois Erlangs exigem 7 circuitos, enquanto rotas com tráfego oferecido de oito Erlangs exigem 15 circuitos (quatro vezes mais tráfego para cerca do dobro de circuitos). Neste caso, para um fator de concentração de quatro vezes, obteve-se um ganho de escala ligeiramente inferior a dois. Se o ganho fosse superior a dois, então a solução com a central de trânsito seria mais barata.

Pode-se expressar este problema de uma forma um pouco mais genérica. Seja o seguinte arranjo: dois tráfegos, de intensidade A cada um, são encaminhados por duas rotas separadas. Para uma dada probabilidade de bloqueio (por exemplo 1%), serão necessários $N1$ circuitos em cada rota. Por outro lado se os dois tráfegos fossem encaminhados por uma mesma rota, seriam necessários $N2$ circuitos, número inferior a soma de circuitos do caso anterior ($N2 < 2.N1$). A diferença de circuitos nos dois casos é devido ao ganho de escala.

A figura 5 apresenta o ganho de escala para tráfegos entre 1 e 10 Erlangs e um fator de concentração $f=100$. O ganho de escala decresce assintoticamente, atingindo o valor limite unitário. Para tráfegos acima de 7 Erlangs o ganho de escala é inferior a 2.

A tabela 1 apresenta uma comparação entre duas rotas, nas quais são oferecidos dois tráfegos

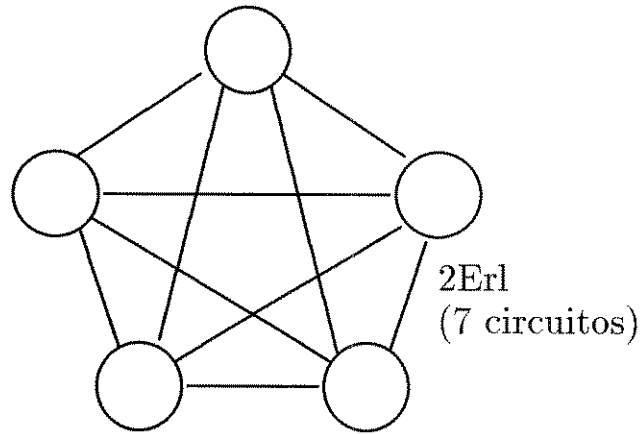


Figura 3: Nesta topologia, as centrais estão interligadas por rotas diretas. O tráfego oferecido a cada rota é de dois Erlangs. Para uma probabilidade de bloqueio de 1%, cada rota deve ser dimensionada com 7 circuitos. Como existem 10 rotas na rede, a rede possui um total de 70 circuitos.

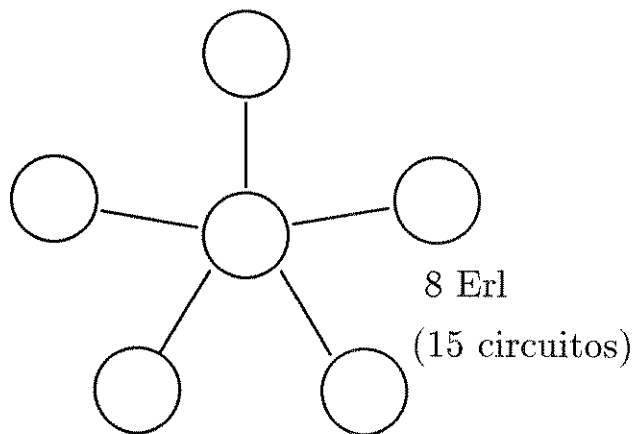


Figura 4: Nesta topologia, as centrais são interligadas por uma central de trânsito. O tráfego oferecido à cada rota é de 8 Erlangs (Fator de concentração de 4 vezes). Para uma probabilidade de bloqueio de 1%, cada rota deve ser dimensionada com 15 circuitos. Como existem 5 rotas na rede, a rede possui um total de 75 circuitos.

TRÁF.	GANHO DE ESCALA				
	1.6	1.7	1.8	1.9	2.0
2	2.5	2.9	3.6	4.3	5.1
4	4.0	5.3	7.0	9.7	14.0
6	5.4	8.1	12.6	21.9	45.7
8	12.3	25.6	82.8	>	>
10	15.6	43.4	>	>	>

Tabela 1: Tabela de fatores de concentração de tráfego. Apresenta uma comparação entre duas rotas com probabilidades de bloqueio de 1%. Às rotas são oferecidos os tráfegos A e f.A, onde f é o fator de concentração de tráfego. Os campos preenchidos por ">" indicam um fator superior a 100.

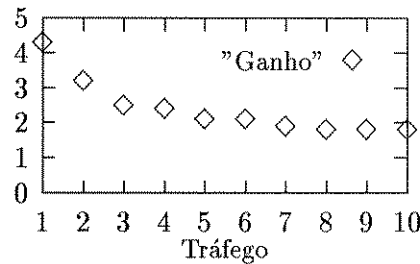


Figura 5: Ganho de Escala para um fator de concentração $f=100$.

poissonianos de intensidade A e f.A, onde f é um fator de concentração de tráfego. Para um determinado tráfego A (linha da tabela) é apresentado qual deve ser o fator de concentração necessário para que haja um determinado ganho de escala em relação ao custo (número de circuitos) por Erlang oferecido. Este ganho de escala é dado pela relação:

$$G = (N1/A)/(N2/f.A) = f.N1/N2. \quad (1)$$

A tabela 1 mostra que para, tráfegos acima de 6 Erlangs, é preciso um fator de concentração superior a 100 para que seja obtido um ganho de escala de 2. Para tráfegos acima de 6 Erlangs, caminhos indiretos são preferíveis a caminhos diretos se o tráfego nos caminhos indiretos for superior a 600 Erlangs. Caso contrário rotas diretas devem ser utilizadas.

A concentração de tráfego é possível em duas situações típicas: existência de alto interesse de tráfego entre duas centrais, o que atrai outros tráfegos, ou existência de um número consideravelmente grande de centrais com baixo tráfego cuja concentração é conseguida através de uma central tandem.

Para analisar a segunda alternativa, seja uma rede com $N+1$ centrais locais cada uma com T assinantes filiados e um tráfego médio de saída por assinante igual a "a". Supõe-se que todas as centrais são ligadas duas a duas por duas rotas unidirecionais. O tráfego oferecido por rota é dado por:

$$A = T.a/(2.N) \quad (2)$$

No caso das centrais serem interligadas apenas por um caminho indireto através de uma única central tandem, o tráfego oferecido por rota unidirecional é dado por:

$$A = T.a/2 \quad (3)$$

O fator de concentração, neste caso, é igual a N . A partir da tabela 1 é possível concluir que, supondo que o ganho de escala deve ser igual ou superior a dois, mesmo para redes de porte considerável (acima de 100 centrais), tráfegos oferecidos por rota acima de 6 Erlangs justificam economicamente a utilização de rotas diretas.

Para um bloqueio de 1%, um tráfego de 6 Erlangs necessita de 13 circuitos. A interligação direta de duas centrais, utilizando rotas unidirecionais, necessita 26 circuitos, viabilizando a implantação de um enlace PCM de 30 canais.

Portanto, para redes locais que utilizem meio de transmissão óptico com custo por circuito de junção aproximadamente igual para as diferentes centrais, pode-se estabelecer a seguinte regra de determinação do entroncamento, a ser aplicada como uma primeira aproximação: "As centrais que possuem um interesse de tráfego entre si superior a 6 Erlangs devem ser interligadas duas a duas por caminhos diretos e os demais tráfegos devem utilizar caminhos intermediários compostos de duas rotas".

Esta regra é válida se os custos de entroncamento forem aproximadamente iguais para todas as centrais. Em geral o custo de entroncamento é heterogêneo, uma vez que convivem na rede centrais de diferentes tecnologias, portes e fabricantes. Neste caso é preciso determinar, para cada rede, os tráfegos que devem utilizar caminhos diretos ou não com o emprego de técnicas de pesquisa operacional. Esta análise foi realizada excluindo fatores que, se levados em consideração, podem alterar significativamente o problema, tais como: Modularidade dos equipamentos, variação de carga de tráfego e a estrutura já instalada na rede.

3.2.2 Caminhos de Transbordo.

A análise de ganho de escala feita anteriormente adota como hipótese um plano de encaminhamento sem a utilização de caminhos alternativos (transbordo). A análise apresentada a seguir mostra que é possível conseguir ganhos de escala maiores que 2 para fatores de concentração menores que 10 quando se considera rotas alternativas. Seja uma rota com N circuitos na qual é oferecido um tráfego de intensidade A . O número de circuitos N é estabelecido para que a probabilidade de bloqueio seja de 1%. Seja o arranjo de uma rota direta com N_1 circuitos. A esta rota é oferecido um tráfego A , resultando numa probabilidade de bloqueio de 10%. O tráfego bloqueado transborda para uma rota com N_2 circuitos. O bloqueio final para o conjunto de rotas é de 1% do tráfego total. À rota N_2 é oferecido um tráfego de média $f.m$ e variância $f.v$, onde m e v são a média e a variância do tráfego transbordado da rota de N_1 circuitos e f é o fator de concentração na rota alternativa. N_2 é determinado de forma que a probabilidade de bloqueio na rota seja de 10%, resultando para o conjunto das duas rotas uma probabilidade de bloqueio de 1%. Neste caso o ganho de escala entre

TRÁF.	GANHO DE ESCALA				
	1.6	1.7	1.8	1.9	2.0
25	2.6	3.0	3.4	3.8	4.3
50	3.0	3.6	4.4	5.2	6.5
75	3.6	4.5	5.7	7.4	10.0
100	4.2	5.6	7.6	10.9	16.7
125	5.0	7.0	10.4	16.7	31.8

Tabela 2: Fatores de concentração de tráfego na rota alternativa e seus correspondentes ganhos de escala.

os dois arranjos pode ser estabelecido através da seguinte expressão:

$$G = (N - N1)/(N2/f) = f.(N - N1)/N2 \quad (4)$$

O valor $N - N1$ pode ser interpretado como o número de circuitos no primeiro arranjo que recebe o tráfego "transbordado" pelos $N1$ circuitos, enquanto que o valor $N2/f$ é a "parcela" dos $N2$ circuitos que recebe o tráfego transbordado pelos $N1$ circuitos do segundo arranjo.

A tabela 2 apresenta os fatores de concentração do tráfego transbordado para valores de intensidades de tráfego variando de 25 a 125 Erlangs.

A tabela 2 mostra que para tráfegos de até 75 Erlangs o fator de concentração é inferior a 10, para um ganho de escala igual a 2. Considerando-se conjuntamente os resultados das tabelas 1 e 3, e ainda a hipótese de custos homogêneos na rede tronco, é possível formular uma regra simples para um dimensionamento preliminar do entroncamento: "As centrais que possuem um interesse de tráfego entre si superior a 6 Erlangs devem ser interligadas duas a duas por rotas diretas. Os tráfegos entre 6 e 75 Erlangs devem ter uma rota direta dimensionada com bloqueio de 10% e devem transbordar para uma outra rota alternativa (via central tandem), escolhida de modo a se conseguir a maior concentração de tráfego possível. Se a rota alternativa for exclusiva para este fim, deve ser dimensionada com 10% de bloqueio, garantindo um bloqueio final de 1%. Os tráfegos maiores que 75 Erlangs devem ser encaminhados exclusivamente por rotas diretas." Conforme registrado anteriormente, esta regra somente é válida se os custos de entroncamento forem aproximadamente iguais para todas as centrais e se forem desprezados uma série de fatores que alteram significativamente este problema, tais como: a modularidade dos equipamentos e a estrutura já instalada na rede. A regra prática, consagrada no dimensionamento do entroncamento, considera uma probabilidade de bloqueio nas rotas finais de no máximo 1% e uma probabilidade de no máximo 10% para as rotas de alto uso. O principal motivo do bloqueio de 1% nas rotas finais é que, a estas rotas, são oferecidos tráfegos transbordados e também tráfegos de primeira escolha, que eventualmente não possuem rotas alternativas [1]. Esta questão pode ser tratada através da utilização de rota de proteção ou da técnica de reserva de circuitos, permitindo que o bloqueio na rota final seja de 1%, para os tráfegos de primeira escolha, e de 10%, para os tráfegos de transbordo.

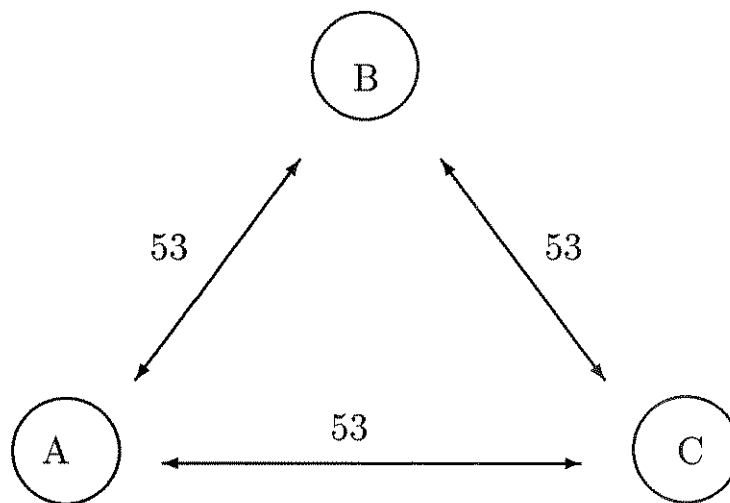


Figura 6: Rede lógica com rotas diretas interligando as centrais. Cada rota possui 53 circuitos, totalizando 159 circuitos para a rede.

3.2.3 Fraca Dependência Entre Roteamento e Entroncamento.

Para uma rede que utiliza comutação temporal e transmissão óptica pode-se adotar as seguintes hipóteses:

- Custo da transmissão associado à distância é desprezível .
- Custo de comutação é dominante em relação ao custo de transmissão.

Como consequência destas hipóteses, verifica-se que existe uma fraca relação entre a rede lógica (entroncamento) e a rede física (roteamento). O exemplo a seguir procura demonstrar este fato. Seja uma rede com 3 centrais e um interesse de tráfego entre elas de 20 Erlangs. Há duas topologias possíveis para a rede lógica:

- Centrais ligadas duas a duas por rotas diretas. Supondo-se todas as rotas bidirecionais, esta topologia resulta numa rede com três rotas, com 53 circuitos cada (bloqueio de 1%) (Figura 6).
- Duas das centrais utilizando a terceira como central de trânsito . Esta topologia resulta em duas rotas com 96 circuitos cada (Figura 7).

Supondo-se que o custo de um sistema de segunda hierarquia é inferior ao custo de três sistemas de primeira hierarquia, a solução para a rede física de custo mínimo é a implantação de dois sistemas de segunda hierarquia (120 canais), cada sistema interligando as centrais A e B e as centrais B e C, respectivamente. Portanto, a solução para a rede física é a mesma para os dois casos de redes lógicas (Figura 8).

O cenário puramente digital sugere ser possível resolver os problemas de roteamento e de entroncamento e encaminhamento separadamente. O objetivo de minimização do custo para os problemas

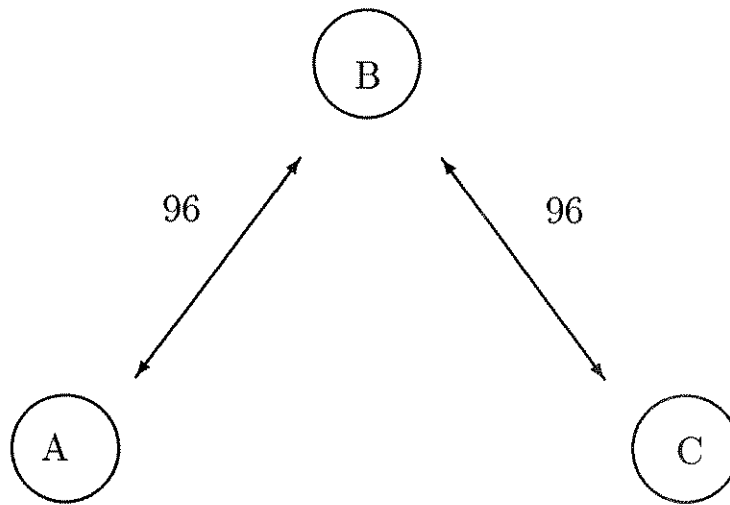


Figura 7: Rede lógica com centrais interligadas via central trânsito. Cada rota possui 96 circuitos, totalizando 192 circuitos para a rede.

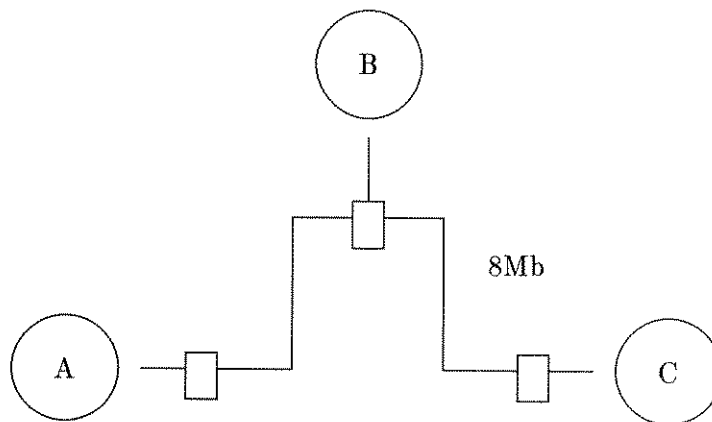


Figura 8: Rede física que interliga as três centrais.

de entroncamento e encaminhamento pode ser traduzido como minimizar o número de troncos de 30 canais utilizados, ponderados pelo custo dos troncos para cada tipo de central (uma vez que, por hipótese, o custo de transmissão é desprezível em relação ao custo de comutação).

3.2.4 As Modularidades dos Meios de Comutação e de Transmissão.

Os equipamentos digitais utilizam uma multiplexagem temporal de circuitos. Esta multiplexagem é feita sempre em módulos padronizados de circuitos.

Esta modularidade reduz a flexibilidade do planejamento, uma vez que o número de circuitos a serem implantados na rede deve ser um múltiplo dos módulos utilizados pelos diversos equipamentos. Este problema de modularidade pode ser separado em modularidade dos meios de comutação e de transmissão.

Define-se como folga de modularidade a diferença entre o número de circuitos instalados e o módulo do equipamento de transmissão ou comutação em uso efetivo. Por exemplo, se num sistema de transmissão de 480 canais forem utilizados 360 canais, a folga de modularidade é de 120 canais.

Os sistemas de transmissão digitais podem pertencer a duas tecnologias distintas: hierarquia digital síncrona (Synchronous Digital Hierarchy - SDH) e hierarquia digital plesiócrona (Plesiochronous Digital Hierarchy - PDH). Na PDH existem sistemas de diversas hierarquias, correspondendo cada hierarquia a um módulo de multiplexagem. Por exemplo, equipamentos de primeira hierarquia PDH possuem 30 canais de voz (64 Kbps), segunda hierarquia possuem 120 canais de voz, e assim por diante. Para que um canal de voz seja transportado num dado nível hierárquico é preciso passar por equipamentos de hierarquia inferior. Assim, para um canal de voz de um tronco digital de 30 canais de uma dada central possa ser multiplexado num sistema de 480 canais, é preciso que ele passe antes por um sistema de 120 canais. Os sistemas de hierarquia inferior multiplexados num dado sistema são chamados tributários do sistema.

O problema de modularidade de transmissão na hierarquia digital plesiócrona é tratado durante a fase do planejamento do roteamento. É um problema complexo, envolvendo várias hierarquias, cada uma com sua modularidade. A folga de modularidade neste caso, frequentemente não é aproveitável e é inútil se não existirem troncos digitais instalados e não utilizados nas centrais. Mesmo se existirem troncos disponíveis para aproveitamento da folga nos meios de transmissão, provavelmente é necessária a instalação de outros sistemas de transmissão, que podem por sua vez apresentar uma nova folga de modularidade. Na hierarquia digital síncrona, o problema da modularidade é atenuado, uma vez que tributários de baixa velocidade podem ser multiplexados em sistemas de alta capacidade, sem a necessidade de passar por hierarquias intermediárias.

As centrais telefônicas CPA-T atuais utilizam troncos digitais de 30 canais. Esta é portanto a modularidade da comutação. A figura 9 apresenta um gráfico do custo por Erlang cursado na rota, supondo-se uma modularidade de 30 canais. Pode-se verificar que o efeito da modularidade é forte para rotas com menos de 90 circuitos. Para rotas grandes, o problema da modularidade é desprezível.

Em muitos casos, esta modularidade pode ser desprezada totalmente devido a dois fatores:

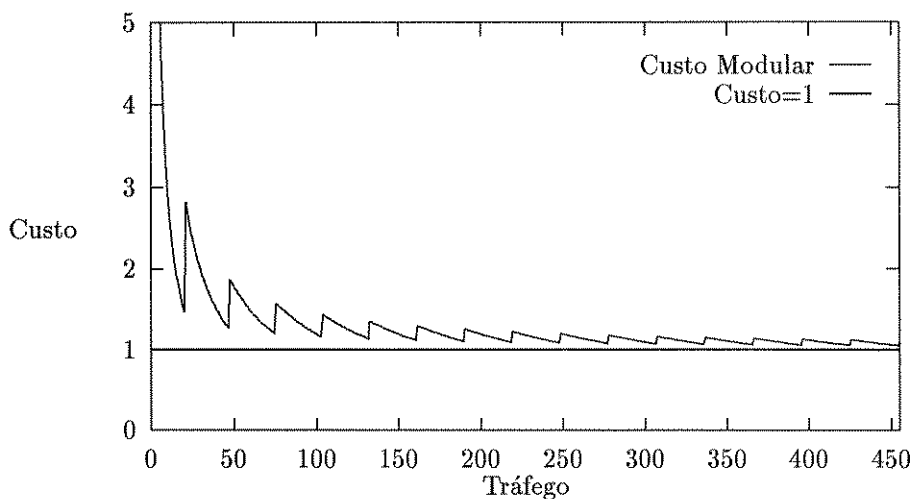


Figura 9: Gráfico do custo por Erlang cursado na rota, supondo-se uma modularidade de 30 canais.

- A folga de modularidade de comutação pode ser totalmente aproveitada, embora as rotas possam ficar superdimensionadas.
- A existência de interconectores digitais (digital cross connect) que são equipamentos de comutação semi-permanente. Estes equipamentos permitem utilizar um mesmo enlace de 30 canais para interligar mais de duas centrais, eliminando quase totalmente o problema da modularidade de comutação e atenuando o da modularidade de transmissão.

Por exemplo, seja uma rede com três centrais interligadas duas a duas por meio de rotas bidirecionais de 10 circuitos. A utilização de interconectores digitais permite que cada central possua um único tronco digital de 30 canais. Estes 30 canais podem ser distribuídos pelas 3 rotas e comutados pelo interconector num ponto central.

É importante observar que os interconectores digitais não competem diretamente com a função trânsito das centrais. Espera-se que o custo destes interconectores seja uma fração do custo de entroncamento de centrais. Por outro lado os interconectores não apresentam um ganho de escala no número de circuitos, uma vez que não concentram tráfegos.

Um aspecto relevante no acoplamento entre as fases de dimensionamento do entroncamento e do roteamento é devido ao fato que no cálculo do entroncamento entra em jogo uma série de parâmetros estatísticos, previsão de tráfego, probabilidade de bloqueio, etc., que permitem uma abordagem flexível na sua resolução, enquanto que no roteamento, o problema é traduzido em uma demanda precisa de circuitos em cada rota que deve ser atendida pelos meios de transmissão a serem instalados.

Alterações no entroncamento podem implicar em uma redução significativa de custos nos meios de transmissão, sem provocar uma degradação acentuada da qualidade de serviço, medida principalmente como bloqueio ponto a ponto de chamadas. Portanto, é conveniente que no processo de

planejamento haja uma interação entre a fase de determinação do encaminhamento e entroncamento e a fase de roteamento, se o custo da transmissão for comparável ao custo do entroncamento. No cenário atual, o custo de transmissão é, em geral, muito inferior ao do entroncamento. Desta forma, esta interação pode ser dispensada.

3.2.5 Distribuição de Tráfego e Dimensionamento Multi-horário.

O tráfego telefônico de uma rede apresenta padrões de variação ao longo do dia, dos dias da semana e do ano.

Esta variação não é uniforme em relação às diversas regiões atendidas pela rede telefônica. Por exemplo, para uma rede urbana pode-se esperar que no horário comercial o tráfego telefônico esteja concentrado nas regiões industriais e comerciais, para onde se desloca grande parcela da população. Nos demais horários o tráfego deve se concentrar nas regiões residenciais. Na rede interurbana verifica-se um comportamento do mesmo tipo, principalmente no caso de redes onde existam regiões com diferentes fusos horários.

O dimensionamento das redes pode levar em conta esta variação de intensidade de tráfego, utilizando circuitos de rotas ociosas num determinado período em auxílio de outras rotas sobrecarregadas. Este tipo de dimensionamento é chamado de multi-horário [14].

O dimensionamento multi-horário requer um tratamento muito mais complexo que o planejamento convencional. Na literatura são propostas várias estratégias para abordagem do problema, sem que nenhuma delas tenha sido universalmente adotada.

Um outro aspecto importante quanto à variação do tráfego é a robustez da rede em relação à panes e sobrecargas de tráfego.

Uma tendência na prestação de serviços, com a finalidade de vencer a concorrência pelo mercado, é o cuidado e atenção que se deve ter com a qualidade de atendimento ao cliente [2]. Nas empresas de telecomunicações, um dos aspectos a serem observados é a variação da qualidade de serviço ao longo do tempo. Pesquisas demonstram que a opinião dos clientes sobre a qualidade do serviço é mais influenciada pelos incidentes negativos do que pelos positivos. Assim a robustez da rede em presença de panes, sobrecargas e variações de tráfego devem ser consideradas para manutenção da qualidade de serviço.

3.2.6 História da Rede.

Um aspecto determinante no dimensionamento de redes telefônicas é a existência de equipamentos já implantados na rede. Estes equipamentos devem ser utilizados em primeiro lugar, antes que novos equipamentos sejam implantados.

Também existem aspectos operacionais que impedem modificações radicais na estrutura histórica da rede. Por exemplo sejam duas redes com topologias distintas. Na primeira rede as centrais são todas ligadas duas a duas. Se for feita uma ampliação do número de assinantes de uma central (por exemplo de 4.000 para 5.000 assinantes) será necessário um redimensionamento de todas as rotas

desta central. Isto exigirá uma modificação de todas as centrais da rede, o que pode ser bastante custoso. Na segunda rede, todas centrais estão ligadas a uma única central tandem. Neste caso a mesma ampliação do caso anterior pode ser feita modificando-se apenas duas centrais.

Estes aspectos tornam a solução de menor custo diferente daquela que seria obtida no caso de planejamento de uma rede totalmente nova. Assim a estrutura da rede fica marcada por sua história e o processo de planejamento deve incorporar este aspecto na sua formulação.

3.3 Gerência de Rede.

A determinação do encaminhamento e do entroncamento não envolve apenas a fase de planejamento das redes de comunicações. A operação destas redes também exige modificações visando uma adequação às diferenças entre as previsões da demanda e ao tráfego real solicitado pelos assinantes.

Atualmente, a gerência de rede, e em particular a gerência de tráfego, ocupa posição de destaque em muitas empresas de telecomunicações, pois permite supervisionar, monitorar e controlar diferentes contingências causadas por sobrecargas de tráfego ou panes [5], [19].

A flexibilidade de alterações do encaminhamento proporcionadas pelas centrais CPA/T, a existência da rede de sinalização de canal comum, que permite o conhecimento do estado quase que instantâneo da rede, e a existência da rede de gerência (Telecommunication Management Network - TMN) em um futuro próximo, permitirão ações de gerência de grande alcance. Neste cenário, o planejamento e a gerência de rede se confundem e o planejamento deve ser influenciado de forma que a rede resultante possa usufruir das facilidades de gerência.

As principais ações de gerência de tráfego são:

- Modificação do plano de encaminhamento.
- Modificação do entroncamento.
- Modificação do Roteamento.

Estas atividades são as mesmas da fase de planejamento colocadas num novo contexto, pois a gerência de redes, em geral, não conta com a implantação de novos equipamentos.

3.4 Conclusões.

As redes locais dos próximos anos utilizarão em sua maioria fibras ópticas fazendo com que os custos de transmissão sejam quase invariantes com a distância entre as centrais.

As centrais de comutação continuarão a exigir um esforço considerável no desenvolvimento de seu software aplicativo visando a incorporação contínua de novos serviços e o aumento de sua flexibilidade de ação em tempo real na rede. Portanto, é provável que o custo por canal de comunicação continue sendo maior na comutação do que na transmissão.

As novas tecnologias de gerência de rede indicam a possibilidade de ações cada vez de maior alcance na operação da rede, visando a melhoria de desempenho.

Neste contexto, a determinação do plano de encaminhamento e o dimensionamento das rotas devem ser feitas de forma integrada. As redes, particularmente as redes locais devem apresentar um ou dois níveis, sendo que a maior parte das chamadas deve ser estabelecida por rotas diretas ou caminhos formados por duas rotas. Caminhos com três ou mais rotas devem ser extraordinários. O problema de roteamento pode ser resolvido de forma separada, embora seja conveniente permitir uma interação com as demais etapas.

A modularidade de comutação não é um problema de grande importância do ponto de vista do dimensionamento das rotas.

Capítulo 4

Desenvolvimento de Ferramentas.

*Se você não espera encontrar o inesperado, não o encontrará;
porque é penoso procurá-lo, e difícil.
Heráclito (500 A.C.).*

Este capítulo aborda temas relacionados à forma das ferramentas de planejamento de redes telefônicas. De maneira geral, engenheiros de telecomunicações preocupavam-se quase exclusivamente com o conteúdo destes programas de computador, dedicando-se a aprimorar modelos matemáticos ou minimizar o tempo de execução. A forma dos programas de planejamento, como o formato de entrada e saída de dados ou como o processo de interação com o usuário eram desprezados, considerados temas triviais. Atualmente, há quase um consenso na área de telecomunicações de que a forma é tão importante quanto o conteúdo, que a forma e conteúdo são indissociáveis. A seguir são apresentados os motivos que levam a esta conclusão.

4.1 Quem Toma a Decisão?

Quem toma as decisões no processo de planejamento nas empresas de telecomunicações não são os engenheiros de planejamento; são os membros da área administrativa que avaliam os cenários de rede apresentados diante das finanças da empresa, da política de investimentos e outros fatores. Para estes decisores, tem pouca importância se os resultados foram obtidos por um brilhante e intrincado modelo matemático ou se foram obtidos por um eficiente procedimento heurístico. Basta que os resultados sejam confiáveis. É fundamental que os resultados sejam apresentados na forma de mapas e gráficos coloridos e não na forma de extensas tabelas numéricas. Em contrapartida, os engenheiros tendem a subestimar muitos dos fatores considerados pelos decisores na elaboração dos cenários de planejamento.

Durante o curso de doutorado, o autor desta tese tomou conhecimento do PLANITU, um pacote integrado de otimização e dimensionamento de redes de telecomunicações. Este pacote de planejamento foi desenvolvido pela União Internacional de Telecomunicações (International Telecommunications Union - ITU). Existem muito poucos pacotes atualmente tão abrangentes e tão

desenvolvidos graficamente. É possível iniciar o processo de planejamento de uma rede local, por exemplo, a partir do mapa da cidade, onde para cada quadra é possível realizar a previsão de tráfego independentemente. O PLANITU integra as fases de localização de centrais, entroncamento e roteamento, empregando a metodologia clássica de planejamento.

A utilização de um pacote integrado, em vez de um conjunto de ferramentas especializadas, aumenta o poder de planejamento. A integração das etapas de planejamento permite que os resultados obtidos em cada etapa possam ser utilizados de forma mais eficiente nas etapas seguintes. A comunicação com a esfera administrativa também é beneficiada, tornando a apresentação dos resultados mais completa e mais clara.

O PLANITU não tem sido empregado pelas empresas brasileiras, apesar de suas qualidades, entre elas o preço de aquisição que é gratuito. Três fatores contribuem para este fato:

- O ambiente computacional (Sistemas DIGITAL) é bastante incomum nas empresas brasileiras.
- O PLANITU exige uma extensa e detalhada base de dados de custos. Estes custos, em geral, não estão disponíveis atualmente para o planejador.
- O treinamento e o suporte ao usuário não é considerado adequado pelas empresas.

A diferença do custo de implantação de novos equipamentos numa rede muito bem planejada e uma rede não tão bem planejada é da ordem de 10% a 15% do custo de implantação. Uma rede bem planejada é resultado de um bom processo de otimização dos recursos da rede. Esta otimização emprega uma função de custo, tradicionalmente fundamentada numa base de dados dos custos dos equipamentos, custo de manutenção, custos dos serviços e numa estimativa de demanda de serviços de telecomunicações, traduzida numa previsão de tráfego dos assinantes. A imprecisão destes elementos, em geral, é maior que 10%. Fatores como instabilidade da economia, inflação, crescimento desordenado das cidades, exigência de licitações para compra de equipamentos contribuem para que as estimativas de custos e de tráfego sejam imprecisas ou mesmo impossíveis.

Em 1990 Stefanuto [38] publicou um interessante trabalho avaliando os resultados de um convênio entre o Centro de Pesquisa e Desenvolvimento da TELEBRÁS (CPqD) e a UNICAMP cujo objetivo foi desenvolver ferramentas computacionais para planejamento de redes telefônicas. Foi gerado um conjunto de ferramentas abrangendo todos os problemas de planejamento de redes telefônicas.

Stefanuto entrevistou pessoas envolvidas neste processo de desenvolvimento e todos consideraram o projeto muito bem sucedido. Entretanto, estas ferramentas foram sub-utilizadas pelo sistema TELEBRÁS. Muitos engenheiros da área de planejamento não conhecem estas ferramentas. Algumas empresas optaram por desenvolver suas próprias ferramentas.

Entre os motivos que levaram a esta situação está o fato de que as ferramentas possuíam uma comunicação homem-máquina pouco desenvolvida e o pacote não era integrado, exigindo assim um grande esforço do usuário. Pode-se dizer que estas ferramentas pertencem à era FORTRAN, uma era em que os computadores utilizavam interfaces não gráficas e apresentavam limitações em

termos de velocidade de processamento e capacidade de memória. Por outro lado as ferramentas foram desenvolvidas para atuar num cenário geral, ou seja para planejar uma rede genérica qualquer. Ferramentas desenvolvidas pela própria empresa podem incorporar características particulares desta ou daquela rede, tornado-as mais eficientes e satisfazer melhor as preferências e necessidades dos planejadores.

4.2 Conclusões.

Esta tese foi direcionada a partir de reflexões sobre os aspectos apresentados neste capítulo. Concluiu-se que o trabalho de doutorado não deveria resultar em ferramentas computacionais acabadas, uma vez que, para torná-las adequadas aos usuários, é exigido um grande esforço de elaboração de rotinas para banco de dados, entrada e saída de dados, comunicação homem-máquina, que estão fora do escopo do trabalho de doutorado. O trabalho buscou uma compatibilidade entre os modelos propostos e as ferramentas atualmente desenvolvidas e em uso no sistema TELEBRÁS.

Especificação de Ferramentas.

A partir dos elementos apresentados anteriormente, foi especificado um conjunto ideal de ferramentas computacionais integradas para a determinação do entroncamento e do plano de encaminhamento das redes de telecomunicações. O trabalho de doutorado foi realizado com o objetivo de atender esta especificação. Estas ferramentas são listadas a seguir, identificadas por sua função no processo de planejamento.

Otimização de Custo do Caminho de Primeira Escolha.

Os dados de entrada para esta ferramenta são:

- Matriz de tráfego ponto-a-ponto.
- Função custo de entroncamento entre centrais.

O resultado obtido é a determinação de um caminho para cada tráfego, isto é, do plano de encaminhamento considerando-se que cada chamada telefônica seja enviada por apenas um caminho de primeira escolha.

No caso onde o custo dos circuitos de junção for diferente para as diversas centrais da rede, esta ferramenta é complexa. Os métodos numéricos para a solução de problemas de grande porte com ganho de escala admitem apenas soluções aproximadas e exigem grande esforço computacional. No caso de redes com custos uniformes a determinação dos caminhos de primeira escolha pode ser executada por algoritmos simplificados e eficientes.

Otimização de Custo dos Caminhos Alternativos.

Os dados de entrada para esta ferramenta são:

- Matriz de tráfego ponto-a-ponto.
- Função custo de entroncamento entre centrais.
- Caminhos de primeira escolha.
- Critérios de qualidade de serviço.

O resultado desta ferramenta é a determinação do plano de encaminhamento e o dimensionamento das rotas do entroncamento. Esta ferramenta pode ser vista como a evolução das ferramentas tradicionais de dimensionamento baseadas nas concepções de Rapp.

Avaliação de Desempenho.

Os dados de entrada para esta ferramenta são:

- Matriz de tráfego ponto-a-ponto.
- Plano de encaminhamento.
- Dimensionamento das rotas.

O resultado desta ferramenta é uma análise de desempenho da rede que permite estudar o efeito de panes e sobrecargas de tráfego e que permite validar os resultados de dimensionamento obtidos pelas demais ferramentas.

Esta ferramenta também é usada durante a operação da rede e permite analisar propostas de reconfigurações da rede e modificações no seu plano de encaminhamento.

Otimização do Desempenho do Encaminhamento.

Os dados de entrada para esta ferramenta são:

- Matriz de tráfego ponto-a-ponto.
- Dimensionamento das rotas.
- Critérios de qualidade de serviço.

O resultado obtido é um plano de encaminhamento que proporciona a maximização do desempenho da rede. O objetivo desta ferramenta é a alteração do plano de encaminhamento de redes em operação. Visa adequar a rede ao perfil de tráfego estimado a partir da operação da rede e do conhecimento de eventos modificadores deste perfil como por exemplo, férias, dia das mães, horário comercial versus horário não comercial, etc.

Otimização de Desempenho de Rede.

Os dados de entrada para esta ferramenta são:

- Matriz de tráfego ponto-a-ponto.
- Plano de encaminhamento.
- Dimensionamento das rotas.
- Roteamento da rede.
- Critérios de qualidade de serviço.

O resultado obtido é uma rede modificada em relação ao encaminhamento, entroncamento e roteamento, que proporcione desempenho ótimo. Esta ferramenta opera num cenário onde não há implantação de novos equipamentos. Os equipamentos já existentes são logicamente reconfigurados.

Características.

Neste capítulo foram identificados os seguintes elementos fundamentais que devem ser observados no desenvolvimento de ferramentas de planejamento de redes de telecomunicações:

- A saída de dados das ferramentas computacionais devem ser compreensivas não somente para os engenheiros de planejamento, mas também para o pessoal do setor administrativo das empresas de telecomunicações.
- As ferramentas de planejamento devem possuir alta interatividade com o usuário, permitindo o aproveitamento da experiência do usuário bem como a aplicação de decisões estratégicas.
- O processo de planejamento deve considerar as imprecisões dos dados de entrada, que tipo de dados estão disponíveis para o usuário. A rede planejada deve ser robusta diante de situações de contingência ou variações de tráfego.
- As ferramentas devem ser desenvolvidas em ambientes universais (como DOS ou UNIX) empregando linguagens como C ou PASCAL.
- As ferramentas de planejamento e gerenciamento da rede devem ser integradas. Deve haver uma uniformidade na interface homem-máquina.

Capítulo 5

Modelos de Redes.

- *Gatinho sorridente - começou ela, bastante timidamente ... - Você pode me dizer, por favor, qual o caminho que devo tomar agora?*

- *Depende bastante para onde você quer ir - disse o gato.*

- *Não me importa muito - disse Alice.*

- *Então, não importa qual o caminho que você tome - disse o gato.*

Lewis Carrol.

O ponto de partida desta tese de doutorado foi a realização de um trabalho complementar ao apresentado na dissertação de mestrado “Encaminhamento de Chamadas Telefônicas: Modelamento e Simulação” [10], onde foram apresentados duas ferramentas para cálculo de bloqueio baseadas uma num modelo analítico (ANATRAF) e outra num modelo de simulação a eventos discretos (SIMTRAF). O primeiro objetivo proposto para o trabalho de doutorado foi incorporar a estes modelos a técnica de encaminhamento de reserva de circuitos e o encaminhamento com retorno (“crankback” ou reencaminhamento). A seguir são apresentadas as modificações propostas para implementação destas técnicas ao modelo analítico. Também é apresentada uma proposta de evolução do modelo analítico, em relação ao modelo originalmente proposto.

5.1 Reserva de Circuitos.

A estratégia de reserva de circuitos é um mecanismo de encaminhamento de chamadas, que permite que determinados tráfegos ocupem circuitos de uma rota apenas se o número de circuitos vagos for superior a um certo limiar.

A seguir é apresentada uma extensão do modelo de Erlang Invertido [27], [10] para o caso de reserva de circuitos. A incorporação no modelo do mecanismo de reserva de circuitos foi proposta por Le Gall e Bernussou [28], que apresentaram um modelo para uma célula básica de transbordo mútuo com reserva de circuitos.

A figura 10 apresenta um arranjo de duas rotas com N_0 e N_1 circuitos respectivamente. A cada uma das rotas é oferecido, em primeira escolha, um tráfego poissoniano (A_0 e A_1) que transbordam

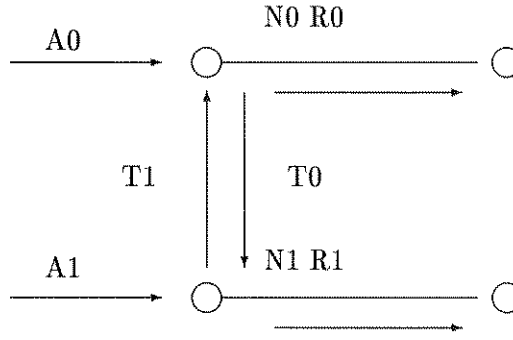


Figura 10: Caso elementar de encadeamento com transbordo mútuo e reserva de circuitos.

para a outra rota. O tráfego de transbordo é bloqueado se na rota de transbordo não houver pelo menos $R+1$ circuitos livres (R_0 e R_1 são os limiares das rotas 0 e 1).

Seja P_{N_0, N_1} a probabilidade de todos circuitos estarem ocupados. Sejam P_{01} e P_{10} as probabilidades de bloqueio dos tráfegos A_0 e A_1 respectivamente. Estas probabilidades podem ser calculadas aproximadamente através das seguintes expressões [28]:

$$P_{N_0, N_1} = \left[\frac{A_0 \cdot N_1}{A \cdot A_1^{N_1}} \cdot \sum_{j=N_1-R_1}^{N_1} \frac{A_1^j}{j! \cdot \text{Erl}(A, N_0 + j)} + \frac{A_1 \cdot N_0}{A \cdot A_0^{N_0}} \cdot \sum_{j=N_0-R_0}^{N_0} \frac{A_0^j}{j! \cdot \text{Erl}(A, N_1 + j)} \right]^{-1} \quad (5)$$

$$P_{01} = \frac{N_0!}{A_0^{N_0}} \cdot P_{N_0, N_1} \cdot \sum_{j=N_0-R_0}^{N_0} \frac{A_0^j}{j!} \quad (6)$$

$$P_{10} = \frac{N_1!}{A_1^{N_1}} \cdot P_{N_0, N_1} \cdot \sum_{j=N_1-R_1}^{N_1} \frac{A_1^j}{j!} \quad (7)$$

Onde:

$$A = A_0 + A_1$$

$$\text{Erl}(A, N) = \frac{A^N / N!}{\sum_{i=0}^N A^i / i!}$$

Estas fórmulas, em analogia à formula de Erlang, são representadas pela seguinte notação, em homenagem à Dra. Le Gall que propôs este modelo:

$$[P_{01}, P_{10}] = \text{LGall}(A_0, A_1, N_0, N_1, R_0, R_1). \quad (8)$$

Estas expressões envolvem as seguintes aproximações:

A probabilidade de haver N_1 circuitos ocupados na rota 1 e j circuitos ocupados na rota 0 é dada pela expressão:

$$P_{j, N_1} = \frac{N_0! \cdot A_0^j}{A_0^{N_0} \cdot j!} \cdot P_{N_0, N_1} \quad ; \quad N_0 - R_0 \leq j \leq N_0 \quad (9)$$

Analogamente para a rota 1:

$$P_{N_0,j} = \frac{N_1! \cdot A_1^j}{A_1^{N_1} \cdot j!} \cdot P_{N_0,N_1} \quad ; \quad N_1 - R_1 \leq j \leq N_1 \quad (10)$$

Estas aproximações são fundamentadas no fato de que, quando o número de circuitos ocupados na rota ultrapassa $N-R$ circuitos, o comportamento da rota depende apenas do tráfego protegido e das chamadas já estabelecidas. Novas chamadas transbordadas de outra rota não serão estabelecidas.

As expressões apresentadas anteriormente foram validadas numericamente em [28] e permitem desacoplar a dependência entre os tráfegos, resultando em expressões relativamente simples e apropriadas para um cálculo numérico eficiente.

No caso de $R_0 = R_1 = 0$ a expressão de P_{N_0,N_1} transforma-se na fórmula de Erlang para uma rota com $N_0 + N_1$ circuitos.

Para $N_1 = R_1 = 0$ tem-se um caso particular, onde uma rota com N_0 circuitos e um limiar de reserva igual a R_0 recebe dois tráfego poissonianos A_0 e A_1 . O tráfego A_0 ocupa os circuitos sem restrição enquanto que o tráfego A_1 ocupa circuitos apenas se houverem pelo menos $R_0 + 1$ circuitos livres [17].

A probabilidade de que todos circuitos estejam ocupados (Probabilidade de bloqueio do tráfego A_0) é dada pela seguinte expressão:

$$P_0 = \frac{A^{N_0-R_0} \cdot A_0^{R_0}}{N_0!} \cdot \left[\sum_{j=N_0-R_0+1}^{N_0} \frac{A_0^{j-N_0+R_0} \cdot A^{N_0-R_0}}{j!} + \sum_{j=0}^{N_0-R_0} \frac{A^j}{j!} \right]^{-1} \quad (11)$$

A probabilidade de que $N_0 - R_0$ circuitos estejam ocupados (probabilidade de bloqueio do tráfego A_1) é dada por:

$$P_{0R} = \frac{P_0 \cdot N_0}{A_0^{N_0}} \cdot \sum_{j=N_0-R_0}^{N_0} \frac{A_0^j}{j!} \quad (12)$$

É utilizada a notação $[P_0, P_{0R}] = RErl(A_0, A_1, N_0, R_0)$ para representar estas fórmulas.

A partir destes elementos de rede básicos, é desenvolvido o modelo de Erlang Invertido para rotas com reserva de circuitos. A metodologia utilizada é a mesma do modelo desenvolvido para rotas sem reserva e é uma variante do modelo sugerido por Le Gall [28].

A apresentação do modelo supõe, sem perda de generalidade, que os tráfegos de primeira escolha não sofrem restrição e o tráfego transbordado sempre sofre restrição. A cada rota são associadas as variáveis x e x_r que representam as parcelas de tráfego cursado devido aos tráfegos poissonianos oferecidos sem restrição e com restrição respectivamente. A partir destas variáveis é possível construir as equações do modelo.

Seja um conjunto de rotas $R_l, R_m, \dots, R_i$. A estas rotas é oferecido um conjunto de tráfegos, cada um numa sequência própria de transbordo. O tráfego cursado numa rota genérica i é obtido a partir da expressão:

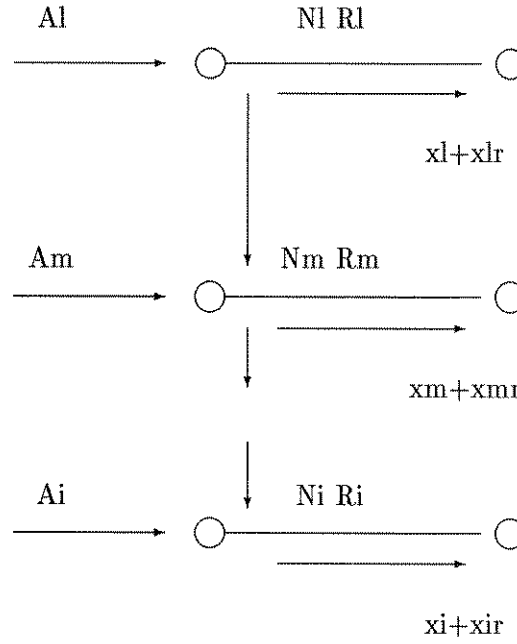


Figura 11: Arranjo básico de uma sequência de transbordo.

$$x_i = A_i \cdot (1 - P_i) \quad (13)$$

$$x_{ir} = \sum_{j \in \Phi} A_j \cdot (P_{lm...} - P_{lm...i}) \quad (14)$$

Onde:

x_i = Tráfego cursado na rota i que não sofre restrição de ocupação dos circuitos.

x_{ir} = Tráfego cursado na rota i que sofre restrição de ocupação dos circuitos devido a reserva de circuitos.

A_k = Tráfego oferecido k .

Φ = Conjunto dos tráfegos que transbordam para a rota i .

A_i = Tráfego oferecido em primeira escolha à rota i .

A_j = tráfego oferecido em primeira escolha à rota genérica l e que transborda por uma sequência de rotas $m, \dots$ até transbordar para a rota i .

P_i = Probabilidade de bloqueio na rota de primeira escolha.

$lm...i$ = Sequência de transbordo do tráfego j até transbordar para a rota i .

$P_{lm...}$ = Probabilidade conjunta de bloqueio do tráfego j nas rotas anteriores a i (probabilidade de transbordo para a rota i).

$P_{lm...i}$ = Probabilidade conjunta de bloqueio do tráfego j nas rotas anteriores e na rota i (probabilidade de transbordo de i).

As probabilidades de bloqueio são obtidas a partir de tráfegos fictícios, oferecidos a arranjos fictícios que se adaptem aos casos estudados anteriormente. Define-se tráfego fictício como sendo

um tráfego poissoniano que, quando oferecido a uma rota (ou conjunto de rotas), produz um certo tráfego escoado (conhecido).

O bloqueio em primeira escolha é obtido a partir das expressões:

$$Y_i = x_i / (1 - P_i)$$

$$Y_{ir} = x_{ir} / (1 - P_{ir})$$

$$[P_i, P_{ir}] = RErl(Y_i, Y_{ir}, N_i, R_i)$$

Onde Y_i e Y_{ir} são os tráfegos fictícios irrestrito e com restrição, respectivamente, da reserva de circuitos para a rota i .

O bloqueio conjunto para duas ou mais rotas é obtido a partir das seguintes expressões:

$$Y_{lm...i} = \frac{x_{ir}}{1 - P_{lm...i}} + \sum_{k \in \Psi} \frac{x_k}{1 - P_{lm...i}}$$

$$Y_{i...ml} = \frac{x_i}{1 - P_{i...ml}} + \sum_{k \in \Psi} \frac{x_{kr}}{1 - P_{i...ml}}$$

$$[P_{lm...i}, P_{i...ml}] = LGall(Y_{lm...i}, Y_{i...ml}, N_t, N_i, R_t, R_i)$$

Onde: $\Psi = \{ 1, m, \dots \}$ é o conjunto de todas rotas da sequência de transbordo.

$$N_t = \sum_{k \in \Phi} N_k$$

$$R_t = \sum_{k \in \Phi} R_k$$

No caso de não haver reserva de circuitos ($R_i = 0$), estas equações produzem o modelo original de Erlang Invertido.

Seja, por exemplo, o arranjo apresentado na figura 12. O tráfego A_0 seleciona a rota 0 em primeira escolha com transbordo para a rota 1. A rota 1 também recebe o tráfego A_1 em primeira escolha. O conjunto de equações para este caso é o seguinte:

$$x_0 = A_0 \cdot (1 - P_0) = 0$$

$$x_{0r} = A_1 \cdot (1 - P_{0r})$$

$$x_1 = A_1 \cdot (1 - P_1)$$

$$x_{1r} = A_0 \cdot (P_0 - P_{01})$$

A parcela do tráfego cursado na rota 0 correspondente ao tráfego A_0 é representada por x_0 . A variável x_{0r} neste caso particular é nula, uma vez que não há transbordo da rota 1 para a rota

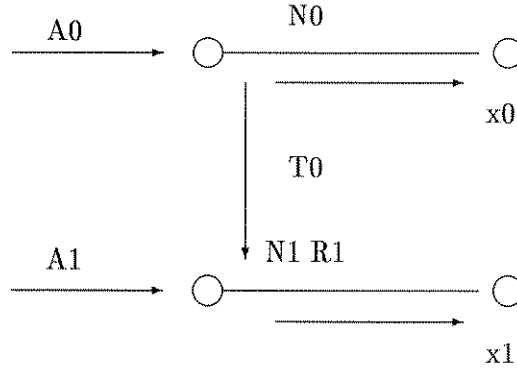


Figura 12: Arranjo de uma rota de primeira escolha e uma rota alternativa.

0. As variáveis x_1 e x_{1r} correspondem às parcelas do tráfego cursado devido ao tráfego A_1 e do transbordo da rota 0, respectivamente.

$$Y_0 = x_0 / (1 - P_0)$$

$$Y_{0r} = x_{0r} / (1 - P_{0r}) = 0$$

$$Y_1 = x_1 / (1 - P_1)$$

$$Y_{1r} = x_{1r} / (1 - P_{1r})$$

$$Y_{01} = (x_0 + x_{1r}) / (1 - P_{01})$$

$$Y_{10} = (x_1 + x_{0r}) / (1 - P_{10})$$

Os tráfegos fictícios são definidos como tráfegos poissonianos que, oferecidos às rotas ou conjunto de rotas, produzem os tráfegos cursados, calculados anteriormente.

$$[P_0, P_{0r}] = RErI(Y_0, Y_{0r}, N_0, R_0)$$

Equivalentemente, uma vez que Y_{0r} é nulo neste caso:

$$P_0 = ErI(Y_0, N_0)$$

$$[P_1, P_{1r}] = RErI(Y_1, Y_{1r}, N_1, R_1)$$

$$[P_{01}, P_{10}] = LGall(Y_{01}, Y_{10}, N_0, N_1, N_0, R_1)$$

Como não há transbordo da rota 1 para a rota 0, adota-se $R_0 = N_0$. Isto deve ser feito sempre que a parcela x_{ir} for igual a 0. As tabelas 3 e 4 apresentam resultados numéricos para este conjunto de equações. Os resultados são comparados aos obtidos em simulação.

N0	N1	R1	A0	A1	x0 ANA	x1 ANA	x0 SIM	x1 SIM
5	5	2	6	2	3.84	2.65	3.83	2.60
5	5	2	5	3	3.58	3.01	3.56	2.99
5	5	2	4	4	3.2	3.33	3.21	3.32
5	10	3	6	6	3.84	6.55	3.85	6.49
5	10	3	4	8	3.2	7.22	3.22	7.21
15	15	3	18	6	13.07	9.28	13.05	9.24
15	15	3	15	7	12.3	8.88	12.29	8.86
50	50	5	60	30	47.03	39.32	47.03	39.24
50	50	5	50	40	44.76	41.78	44.74	41.66
50	50	3	60	30	47.03	39.99	47.03	39.83

Tabela 3: Resultados numéricos para o sistema. São apresentados os parâmetros número de circuitos nas rotas (N_0 e N_1), o limiar de reserva da rota 1 (R_1), Os tráfegos oferecidos (A_0 e A_1). São apresentados os tráfego cursados nas rotas 0 e 1 obtidos a partir do modelo analítico e de simulação.

N0	N1	R1	A0	A1	DIF x1	BLOQ RT 1	P01 ANA	P01 SIM	DIF P01
5	5	2	6	2	0.02	0.38	0.23	0.24	0.02
5	5	2	5	3	0.01	0.33	0.20	0.21	0.04
5	5	2	4	4	0.00	0.31	0.15	0.16	0.04
5	10	3	6	6	0.01	0.20	0.21	0.21	0.04
5	10	3	4	8	0.00	0.18	0.13	0.15	0.08
15	15	3	18	6	0.00	0.16	0.09	0.09	0.02
15	15	3	15	7	0.00	0.09	0.05	0.05	0.06
50	50	5	60	30	0.00	0.09	0.06	0.06	0.01
50	50	5	50	40	0.00	0.08	0.04	0.05	0.06
50	50	3	60	30	0.00	0.07	0.04	0.04	0.01

Tabela 4: Resultados numéricos para o sistema. São apresentadas a diferença normalizada entre os resultados do modelo de simulação e analítico para o tráfego cursado na rota 1, a probabilidade de bloqueio na rota 1 e a probabilidade conjunta de bloqueio para o tráfego A_0 . São apresentadas a probabilidade de bloqueio da rota 1 e a diferença normalizada entre os resultados dos modelos para o bloqueio conjunto para o tráfego A_0 .

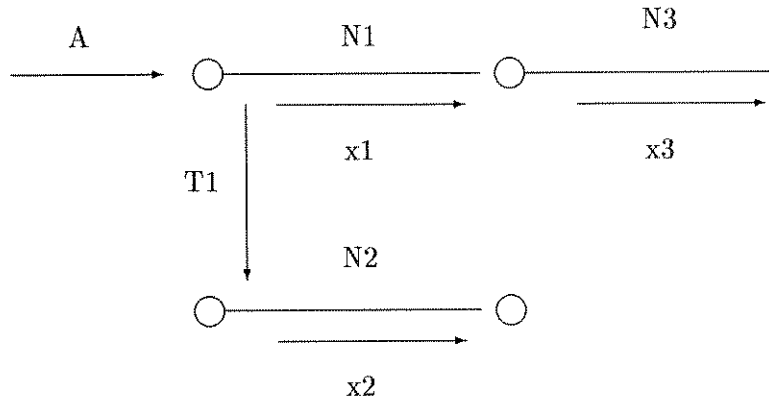


Figura 13: Caso elementar com "crankback".

5.2 Encaminhamento com Retorno.

O encaminhamento (ou reencaminhamento) com retorno ("crankback") consiste em permitir a escolha de rotas alternativas, numa dada central, para chamadas bloqueadas em rotas de saída de centrais à frente.

A figura 13 apresenta um exemplo de encaminhamento com "crankback". Neste arranjo existem as rotas 1 e 3 em série e a rota 2, que é alternativa à rota 1 com "crankback". Isto significa que as chamadas bloqueadas na rota 3 serão encaminhadas alternativamente para a rota 2.

O modelamento analítico para esta situação deve tratar o tráfego transbordado pela rota 3 da mesma forma como o tráfego transbordado pela rota 1.

As equações do modelo Erlang invertido para este caso são:

$$x_1 = x_3 = x = A.[1 - E(Y_1, N_1)].[1 - E(Y_3, N_3)]$$

$$Y_1 = x_1/[1 - E(Y_1, N_1)]$$

$$Y_3 = x_3/[1 - E(Y_3, N_3)]$$

$$x_2 = A.[E(Y_1, N_1) - E(Y_{12}, N_1 + N_2)] + A.[1 - E(Y_1, N_1)].[E(Y_3, N_3) - E(Y_{32}, N_3 + N_2)]$$

$$Y_{12} = x_1 + x_2/[1 - E(Y_{12}, N_1 + N_2)]$$

$$Y_{32} = x_3 + x_2/[1 - E(Y_{32}, N_3 + N_2)]$$

Esta modificação do modelo originalmente proposto em [28] não altera significativamente sua estrutura, uma vez que trata o "crankback" de forma semelhante a um transbordo comum. A introdução desta modificação, portanto, não altera a eficiência do modelo.

5.3 Conclusão.

No início do desenvolvimento destas ferramentas, em 1989, pretendia-se desenvolver uma ferramenta de análise de redes telefônicas, o ANATRAF, e validar seu modelo matemático por meio de simulações. Naquela época computador de pequeno porte mais poderoso disponível eram microcomputadores tipo PC-AT 286. Neste ambiente, a simulação de uma rede de médio porte (10.000 Erlangs), como a da cidade de Campinas, exigia muitas horas de tempo de CPU.

Após cinco anos o cenário mudou completamente. Atualmente estão disponíveis para realizar estas simulações máquinas tipo workstation SPARC 10 e PC-AT 486 de 60MHz. Neste novo ambiente, pode-se simular uma rede de médio porte em alguns minutos (da ordem de 10 minutos). Agora, o novo desafio é simular grandes redes, basicamente quatro redes: redes metropolitanas das cidades do Rio de Janeiro e São Paulo, a rede interurbana do Estado de São Paulo e a rede interurbana brasileira. Estas redes estão na faixa de 100.000 a 500.000 Erlangs comutados. Não há dúvidas de que os ambientes computacionais disponíveis às empresas de telecomunicações evoluirão nos próximos anos a ponto de permitir que estas grandes redes possam ser simuladas em tempos da ordem de minutos. Em contrapartida do tempo de simulação, os modelos de simulação apresentam vantagens sobre modelos analíticos. A introdução de novas características no modelo são feitas de forma muito mais imediata. Por exemplo, a introdução dos mecanismos de reserva de circuitos e reencaminhamento não exigiu nenhuma alteração extensa do SIMTRAF, e pode ser feita sem aproximações. Ao contrário do ANATRAF, que exigiu a formulação de um novo modelo matemático, e que é aproximado.

Estas características do modelo de simulação levaram à escolher esta técnica como ferramenta de análise. Ao longo do trabalho de doutorado, as avaliações das redes sempre utilizaram a ferramenta de simulação. Houve também um esforço para utilizar a técnica de simulação em processos de síntese.

Capítulo 6

Formulação Tipo Fluxo.

Todas as coisas fluem, nada permanece.

Heráclito (500 A.C.).

Este capítulo apresenta uma formulação tipo fluxo multiproduto de custo mínimo aplicada a dois problemas distintos:

O primeiro problema é a determinação do plano de encaminhamento e do dimensionamento das rotas de uma rede telefônica. Tradicionalmente o planejamento destas redes considera um plano de encaminhamento hierárquico, especificado arbitrariamente. A metodologia proposta obtém um plano de encaminhamento minimizando os custos de implantação de equipamentos na expansão das redes e define, indiretamente, o dimensionamento de suas rotas de entroncamento.

O segundo problema consiste em determinar para uma dada rede, qual o plano de encaminhamento que maximiza o tráfego escoado.

6.1 Encaminhamento de Mínimo Custo.

Um dos aspectos do planejamento pouco abordado, no caso de redes telefônicas, é a otimização do plano de encaminhamento. Diversos artigos tem tratado desta questão nas redes de dados, mas no caso de redes telefônicas, têm sido estudados apenas os problemas relativos ao encaminhamento dinâmico e às novas técnicas de encaminhamento.

Em geral, as ferramentas de planejamento que envolvem o plano de encaminhamento, adotam um plano pré-estabelecido pelo planejador. A abordagem clássica para o problema de dimensionamento do entroncamento consiste em arbitrar um plano de encaminhamento e determinar o entroncamento a partir de minimizações parciais para cada tráfego. Esta atitude está associada ao fato de que as centrais telefônicas anteriores à tecnologia CPA (Controle a Programa Armazenado) não apresentavam flexibilidade de encaminhamento e os seus custos de comutação de chamadas de trânsito eram elevados limitando consideravelmente as possibilidades de encaminhamento.

Atualmente, as centrais CPA possuem sofisticados recursos de encaminhamento e o seu custo para chamadas de trânsito é relativamente baixo. O planejamento da expansão das redes telefônicas

utilizando um plano de encaminhamento não otimizado resulta numa sub-utilização de seu potencial com consequente aumento de custo de implantação de equipamentos.

Os meios de transmissão mais utilizados são fibras ópticas, rádio digital e enlaces de satélite. A utilização de cabos subsiste apenas na rede de assinantes. Existe uma tendência de utilização de hierarquias de transmissão cada vez mais elevadas, ao mesmo tempo em que a qualidade das fibras permite longos percursos a custos cada vez mais baixos. Pode-se, portanto, admitir que o custo de transmissão possui apenas duas componentes: custo associado ao número de canais e custo associado à distância percorrida.

A tendência para este cenário é de que o custo associado ao número de canais seja dominante em relação ao custo associado à distância. Este fato traz importantes conseqüências para o planejamento telefônico. Assim numa rede onde os novos equipamentos implantados sejam totalmente digitais o custo dos meios de transmissão é quase independente da distância percorrida e a determinação da rede com mínimo custo de implantação é obtida pela minimização do custo de implantação das portas digitais.

Esta minimização depende da determinação do número de circuitos a serem implantados e da sua alocação nas rotas, de forma a minimizar a ociosidade dos equipamentos digitais devido às suas modularidades.

Com o desenvolvimento de novas tecnologias e do aprimoramento dos equipamentos e serviços, verifica-se uma tendência de redução relativa do custo de transmissão em relação ao custo de comutação. As centrais cada vez mais incorporam novas facilidades e novos serviços, aumentando seu custo. A fabricação de fibras e outros equipamentos tem seu custo reduzido. Esta tendência, e a complexidade dos problemas associados, justificam a partição do problema em determinação do encaminhamento e entroncamento e determinação do roteamento [15].

O problema de determinação do plano de encaminhamento e entroncamento, minimizando-se o custo de implantação não é um problema convexo. Dois fatores contribuem para isto: a modularidade dos equipamentos e o ganho de escala devido à concentração de tráfegos.

A estratégia proposta por este trabalho para resolução do problema de encaminhamento e de entroncamento é um processo em duas etapas.

A primeira etapa é a determinação do encaminhamento de mínimo custo, considerando que o plano de encaminhamento admite apenas um caminho para cada tráfego e que a probabilidade de bloqueio nas rotas é pequena.

A segunda etapa consiste na aplicação dos métodos clássicos de dimensionamento de rotas, tal como o proposto por Rapp [34], [35], utilizando os resultados obtidos na primeira etapa (custos por circuito nas rotas e o plano de encaminhamento) como dados de entrada.

A determinação do encaminhamento e do entroncamento de uma rede telefônica é resultado da resolução do seguinte problema:

Determinar o número de circuitos nas rotas e o plano de encaminhamento que minimize o custo de implantação dos novos circuitos, admitindo-se que a probabilidade de bloqueio nas rotas é desprezível e supondo-se conhecidos:

- As centrais existentes na rede e dados associados (central local ou tandem e capacidade de comutação).
- A matriz de tráfego oferecido.
- A capacidade das rotas já existentes.
- O custo de implantação de circuitos nas rotas.

A rede obtida a partir da resolução deste problema constitui um ponto inicial para as próximas etapas que busquem um refinamento da solução.

6.1.1 Formulação do Problema de Otimização.

O problema proposto é formulado como um problema de fluxo multiproduto de custo mínimo, onde o custo de transporte do fluxo (tráfego telefônico) representa o custo de implantação de circuitos numa rede.

$$\text{Min}_{x_{ij}^k} \sum_{i=1}^N \sum_{j=1}^N f_{ij}(\sum_{k=1}^T x_{ij}^k)$$

sujeito a:

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = +d_k ; \forall k, \text{ se } j \text{ é destino de } k (j = D(k))$$

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = -d_k ; \forall k, \text{ se } j \text{ é origem de } k (j = O(k))$$

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = 0 ; \forall k, \text{ caso contrário}$$

$$0 \leq x_{ij}^k \leq d_k$$

Onde:

x_{ij}^k é a intensidade do tráfego k no arco que une os nós i e j .

d_k é a intensidade do tráfego k .

N é o número de nós.

T é o número total de produtos (tráfegos).

$f_{ij}(\sum_{k=1}^T x_{ij}^k)$ é a função custo de implantação de circuitos no arco ij , que leva em conta o ganho de escala devido à concentração de tráfego.

Nesta formulação cada arco ij representa uma rota lógica da rede, ou seja, um conjunto de canais de voz entre duas centrais. O produto k representa o tráfego telefônico originado numa central $O(k)$ com destino a outra central $D(k)$.

Este tipo de formulação gera problemas de grandes dimensões, mesmo para pequenas redes. Assim, numa rede de 30 centrais interligadas duas a duas por rotas unidirecionais, onde cada central é origem de tráfegos com destino às demais centrais, a rede exibe 870 arcos e 870 tráfegos distintos. Isto dá origem a um problema com 756.900 variáveis. O número de variáveis evolui aproximadamente com a quarta potência do número de centrais.

Diversos trabalhos que tratam problemas desta mesma natureza, lançam mão de uma formulação tipo arco-caminho [16]. Neste caso as variáveis do problema representam os fluxos (tráfego)

que utilizam os caminhos possíveis na rede. Em geral, para cada tráfego, existem da ordem de 10 caminhos possíveis. Desta forma o número de variáveis é da ordem de dez vezes o número de tráfegos origem-destino e, portanto, muito menor que na formulação original.

A desvantagem deste tipo de formulação está na determinação dos caminhos a serem escolhidos, que deve ser feita pelo usuário, exigindo uma grande experiência do planejador e que pode condicionar significativamente a solução final do problema.

O tipo de formulação escolhida, arco-produto, foi baseada no fato de que se encontram disponíveis no mercado de informática equipamentos de pequeno porte, mas de grande capacidade computacional, com perspectivas de um aumento acelerado desta capacidade para o futuro. Nestes equipamentos é possível solucionar problemas de grande porte, da ordem de um milhão de variáveis.

As restrições matemáticas do problema determinam a conservação de tráfego para cada nó. Todo tráfego que entra no nó, incluindo o tráfego originado, deve ser igual ao tráfego que sai do nó, incluindo o tráfego terminado. Este tipo de restrição leva em conta a hipótese de que a probabilidade de bloqueio nas rotas é desprezível. Portanto não são considerados transbordos.

O custo de implantação de circuitos é uma função composta de duas outras: Uma que determina para um dado tráfego o número de circuitos necessários para uma dada probabilidade de bloqueio, outra que determina o custo de implantação para um dado número de circuitos. A segunda função depende da estrutura de custos das centrais interligadas pelos arcos e dos custos relacionados aos equipamentos de transmissão.

A hipótese de que a probabilidade de bloqueio é desprezível permite definir, como função que relacione a intensidade de tráfego ao número de circuitos necessários, a inversa da fórmula de Erlang para uma probabilidade de bloqueio pequena, arbitrariamente escolhida. Neste trabalho escolheu-se uma probabilidade de 0.5%.

A inversa da formula de Erlang é uma função côncava e que apresenta uma forte não linearidade para tráfegos pequenos (entre 0 e 10 Erlangs) e se aproxima assintoticamente a uma reta com derivada 1 para tráfegos grandes (acima de 100 Erlangs). Conseqüentemente espera-se que a função de custo para este problema apresente uma característica de ganho de escala, originando um problema não convexo. Por outro lado, a função custo de implantação de circuitos é complexa e heterogênea, dependendo do tipo de centrais envolvidas e outros fatores, gerando funções de custo diferentes para cada rota. Pode-se prever que estes custos serão modulares, uma vez que o entroncamento das novas centrais implantadas na rede será digital. Portanto, em geral, o problema não é concavo ou convexo.

O método escolhido para resolução do problema proposto exige que a função objetivo seja do tipo linear por partes, como definido a seguir:

$$f(x) = \begin{cases} c^1.x + F^1; & \text{Se } 0 \leq x \leq M^1 \\ c^2.x + F^2; & \text{Se } M^1 \leq x \leq M^2 \\ \dots & \\ c^R.x + F^R; & \text{Se } M^{R-1} \leq x \leq M^R \end{cases}$$

Onde a variável x representa um tráfego cursado.

6.1.2 Algoritmo de Resolução.

A minimização de problemas não convexos é um campo que envolve uma certa dificuldade. Soluções exatas para esta classe de problemas exigem um alto grau de especialização do algoritmo em relação ao tipo de função objetivo e do conjunto de restrições. Tais algoritmos, em geral, não se aplicam a problemas de grande porte, dada sua complexidade computacional.

Uma alternativa para resolução destes problemas é lançar mão de algoritmos que busquem uma boa solução. Tradicionalmente é realizada uma otimização local com este propósito. Entretanto, no caso de problemas não convexos, uma minimização local não leva a um ponto de mínimo global e, em geral, é fortemente dependente da condição inicial do algoritmo.

O algoritmo utilizado neste trabalho é uma adaptação do algoritmo apresentado por Bala-krishnan [6], originalmente desenvolvido para problemas de fretamento de cargas marítimas e que consiste de um procedimento dual de resolução associado a um procedimento heurístico para determinação de uma solução primal factível.

Uma vez que a função objetivo não é convexa, a solução do problema dual não é ótima, dado que há um "gap" de dualidade. Entretanto esta solução é um limitante inferior para a solução do problema primal, o que pode ser um dado valioso na avaliação da qualidade da solução determinada. A utilização de um método dual permite desacoplar o problema, de forma que a otimização possa ser realizada para cada arco individualmente. Isto confere ao algoritmo uma eficiência que permite o tratamento de problemas de grande porte.

Descrição Matemática.

Seja $P(x)$ o problema de programação matemática, onde a função objetivo é uma somatória de funções lineares por partes, representando o custo de implantação de circuitos em cada arco ij da rede:

$$P(x) : \text{Min} \sum_{i=1}^N \sum_{j=1}^N f_{ij}(\sum_{k=1}^T x_{ij}^k)$$

sujeito a:

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = +d_k ; \forall k, \text{ se } j \text{ é destino de } k (j = D(k))$$

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = -d_k ; \forall k, \text{ se } j \text{ é origem de } k (j = O(k))$$

$$\sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k = 0 ; \forall k, \text{ caso contrário}$$

$$0 \leq x_{ij}^k \leq d_k$$

Este problema é do tipo:

$$\text{Min}_x f(x)$$

Sujeito a:

$$\begin{aligned} g(x) &= 0 \\ 0 &\leq x \leq d \end{aligned}$$

Esta formulação pode sofrer uma relaxação do tipo Lagrangeana para gerar o seguinte problema dual [7].

$$\begin{aligned} &Max_v Min_x [f(x) + v'.g(x)] \\ &\text{sujeito a:} \\ &0 \leq x \leq d \end{aligned}$$

A relaxação do problema P(x) resulta no problema dual D(v).

$$\begin{aligned} D(v) : &Max_v Min_x \sum_{i=1}^N \sum_{j=1}^N f_{ij}(\sum_{k=1}^T x_{ij}^k) + \sum_{k=1}^T (v_i^k - v_j^k).x_{ij}^k + \sum_{k=1}^T v_{D_k}^k.d_k \\ &\text{Sujeito a:} \\ &0 \leq x_{ij}^k \leq d_k \\ &\text{Onde:} \\ &v_{O_k} = 0. \end{aligned}$$

A formulação do problema dual D(v) teve seus termos rearranjados. As restrições do problema P(x) são redundantes, pois para cada tráfego k considerado existe uma restrição de conservação de fluxo que deve ser eliminada. Uma vez que os multiplicadores tem uma interpretação de potenciais de custos, o potencial do nó de origem de um determinado tráfego deve ser nulo, ($v_{O_k} = 0$), e portanto, se o multiplicador da equação de conservação de fluxo do nó de origem de cada tráfego for nulo, a redundância é eliminada.

Subproblemas Lagrangeanos.

Para um determinado conjunto de multiplicadores de Lagrange, o problema D(v) pode ser separado por arco, gerando subproblemas $Da_{ij}(x)$ que podem ser solucionados independentemente:

$$\begin{aligned} Da_{ij}(x) : &Min f_{ij}(\sum_{k=1}^T x_{ij}^k) + \sum_{k=1}^T (v_i^k - v_j^k).x_{ij}^k \\ &\text{sujeito a:} \\ &0 \leq x_{ij}^k \leq d_k \\ &\text{Onde: } v_i^k - v_j^k = \Delta V. \end{aligned}$$

Para um dado conjunto v de multiplicadores de Lagrange, o problema $Da_{ij}(x)$ é resolvido por um método do tipo "guloso" (greedy). Os tráfegos são organizados em ordem crescente de ΔV e somados um a um. A cada tráfego somado, ou sempre que o tráfego no arco alcançar o extremo de um segmento da função linear por partes, calcula-se o valor da função para determinação do ponto de mínimo.

Resolução do Problema Dual.

A teoria de programação matemática mostra que, no caso de problemas não convexos, a solução de $D(v)$ para qualquer valor v arbitrado para o conjunto de multiplicadores é um limitante inferior para o valor da solução de $P(x)$.

Portanto a solução do problema a seguir leva ao melhor limitante inferior de $P(x)$.

$$\text{Max}_v D(v)$$

Esta maximização é feita por meio de dois algoritmos independentes: Um procedimento dual ascendente e um método de subgradiente.

Procedimento Dual Ascendente.

O procedimento dual ascendente tem como objetivo maximizar os termos $v_{D(k)}^k \cdot d_k$ do problema $D(v)$ sem modificar os termos restantes, garantindo o aumento da função dual. Esta maximização é feita incrementando-se o valor dos multiplicadores de Lagrange associados aos nós da rede iterativamente para cada tráfego, até que o ganho do procedimento não seja significativo. O incremento dos multiplicadores para um dado tráfego é obtido por meio da resolução do seguinte problema:

$$\text{Max } \delta_{D_k}$$

Sujeito a:

$$\delta_j^k - \delta_i^k \leq u_{ij}^k$$

$$\delta_i^k \geq 0$$

$$\delta_{O_k} = 0$$

Onde u_{ij}^k é o maior incremento possível na diferença dos multiplicadores, sem que as variáveis x_{ij}^k do problema $Da_{ij}(x)$ sofram modificações em seus valores.

Os multiplicadores de Lagrange são, então, atualizados para:

$$v_i^k = v_i^k + \delta_i^k$$

Este problema pode ser interpretado como a determinação do caminho de mínima distância numa rede, onde a distância do arco é dada pelos limitantes u_{ij}^k . A determinação dos limitantes u_{ij}^k é obtida da seguinte forma: Dado um certo tráfego, verifica-se, para cada arco, qual o maior incremento possível na diferença entre os multiplicadores associados aos nós interligados pelo arco.

Nos arcos em que o tráfego cursado não é nulo, qualquer variação dos multiplicadores gera uma alteração dos demais termos em $D(v)$, o que pode levar a uma diminuição da função dual. Portanto, faz-se $u_{ij}^k = 0$. Para os demais arcos (tráfego nulo) aplica-se o seguinte procedimento:

Inicialmente atribui-se à u_{ij}^k o valor:

$$u_{ij}^k = c_{ij}^r + v_i^k - v_j^k$$

Onde c_{ij}^r é o custo linear relativo ao segmento r da função de custo, determinado pelo tráfego cursado no arco ij da solução dual.

Caso a função não seja contínua e o tráfego cursado se encontre exatamente na interseção de dois segmentos, então o valor do degrau entre os dois segmentos dividido pela intensidade do tráfego k deve ser acrescido ao custo. Este procedimento é uma adaptação ao algoritmo proposto por Balakrishnan, uma vez que em seu algoritmo a função objetivo é considerada contínua.

Para a determinação do valor definitivo de u_{ij}^k em cada iteração do Dual Ascendente, deve-se obter os valores dos tráfegos cursados nos subproblemas lagrangeanos, nos pontos duais $(v_i^k - v_j^k)$ e $(v_i^k - v_j^k + u_{ij}^k)$.

Seja z_{ij} o custo da solução de $Da_{ij}(x)$ para o conjunto original de multiplicadores $(v_i^k - v_j^k)$ e seja n_{ij} o custo da solução para os multiplicadores atualizados, isto é, $(v_i^k - v_j^k + u_{ij}^k)$. O valor definitivo do limitante u_{ij}^k é obtido de acordo com a expressão:

$$u_{ij}^k = \max(0, c_{ij}^r + v_i^k - v_j^k - (z_{ij} - n_{ij})/m_{ij}^k)$$

Onde m_{ij}^k é a intensidade de tráfego k cursada no arco, na solução do novo problema, ou d_k se o tráfego não é cursado.

A escolha do valor de m_{ij}^k entre x_{ij}^k e d_k é uma variação em relação ao algoritmo originalmente proposto por Balakrishnan, uma vez que em seu artigo eram tratadas apenas funções côncavas e portanto $m_{ij}^k = d_k$.

A forma de escolha dos limitantes u_{ij}^k garante que a solução do problema $Da_{ij}(x)$ mantém inalterada as variáveis x_{ij}^k inicialmente nulas, garantindo também um aumento da função dual. Este esquema para escolha é conservador mas eficiente computacionalmente.

Método de Subgradiente.

O método de subgradiente utiliza um vetor (w_i^k) cujas componentes representam, para cada nó da rede e para cada tráfego, o quanto a intensidade de um determinado tráfego que entra no nó ultrapassa a intensidade de tráfego que sai do nó (ou o tráfego terminado no nó). Quando o nó é origem de um determinado tráfego, a componente correspondente é nula.

Cada iteração do método calcula o subgradiente, atualiza os multiplicadores de Lagrange, resolve o problema dual determinando um novo ponto (v, x) . O método itera até atingir um critério de parada.

O método não utiliza o subgradiente diretamente, mas um subgradiente modificado $\hat{w}$. Obtido a partir do primeiro. Na primeira iteração do método de subgradiente o vetor $\hat{w}$ é igual ao subgradiente. Nas iterações seguintes é atualizado segundo a expressão:

$$\hat{w} = w + \alpha \cdot \hat{w}$$

Onde α é um parâmetro pré-estabelecido ($\alpha=0.2$).

Os multiplicadores são atualizados pela expressão:

$$v_i^k = v_i^k + \epsilon \cdot \hat{w}_i^k$$

O parâmetro ϵ é dado pela expressão:

$$\epsilon = \lambda \cdot (Z - z) / \|\hat{w}\|^2$$

Onde:

Z é um limitante superior do valor ótimo da função objetivo obtido da melhor solução factível do problema primal.

z é um limitante inferior do valor ótimo da função objetivo obtido da solução do problema dual com o valor atual dos multiplicadores de Lagrange.

$\|\hat{w}\|^2$ é a norma euclidiana do vetor $\hat{w}$.

O parâmetro λ é escolhido numa busca unidimensional de forma a resultar num aumento do resultado do problema dual. Esta busca é feita atribuindo-se um valor inicial igual a 2. O valor de λ é então dividido por 2 até que seja obtido um incremento no resultado do problema dual.

Procedimento Heurístico para o Problema Primal.

Uma vez obtido um novo ponto (v, x) do problema dual é possível determinar um novo ponto (x) do problema primal aplicando-se um procedimento heurístico. O procedimento é dividido em duas etapas: construção de uma solução factível e otimização local.

Etapas 1 - Construção de uma Solução Factível. Quando o procedimento heurístico é executado imediatamente após o procedimento dual ascendente, esta etapa consiste na determinação do caminho de distância mínima para cada tráfego, utilizando como distância dos arcos valor do parâmetro u_{ij}^k . Em outro momento qualquer utiliza-se como solução inicial a última solução factível obtida.

Etapas 2 - Otimização Local. Este procedimento consiste na obtenção, para cada tráfego, de um novo caminho com menor custo de implantação, através da aplicação de um algoritmo de caminho de custo mínimo. Os tráfegos são iterativamente selecionados um a um, até que não seja mais possível reduzir o custo da rede. Este algoritmo admite três variantes quanto à função de custo empregada:

- **Custo Dual**, esta é a variante originalmente proposta por Balakrishnan, onde o custo por arco é o custo relativo utilizado na função dual, ou seja, é o custo atual do arco mais a parcela devido aos multiplicadores de Lagrange dos nós interligados pelo arco.

- **Custo por Unidade de Tráfego (Erlang)**, que é o custo do arco dividido pelo seu tráfego cursado. No caso de custos modulares este procedimento procura reduzir a folga de modularidade.

- **Custo do Tráfego**, que utiliza o custo para escoar um determinado tráfego no arco e é obtido pela diferença do custo do arco com e sem este determinado tráfego. Este procedimento procura reduzir o custo total de implantação.

Resumo do Algoritmo Proposto.

A figura 14 apresenta o algoritmo global de resolução do problema.

O primeiro passo é a inicialização dos multiplicadores de Lagrange. Para cada tráfego é resolvido um problema de caminho de mínima distância entre a origem e os demais nós da rede. Este problema utiliza como distância dos arcos o menor custo por Erlang cursado, obtido a partir da função de custo do problema original, ou seja, a distância é a inclinação da reta que passa pela origem e tangencia por baixo a função de custo. Conseqüentemente, os custos dos subproblemas lagrangeanos serão sempre positivos e os tráfegos serão todos nulos.

Em seguida é executado o procedimento dual ascendente, que itera até que não seja mais possível incrementar a função dual. Então é executado o procedimento heurístico, que fornecerá um limitante superior do custo de implantação a ser utilizado no método de subgradiente. Ao final do procedimento heurístico, é avaliado o incremento do custo dual de um ciclo. Se o incremento for inferior a um valor pré-determinado, ou se o número de iterações ultrapassar um certo limite, ou ainda, se a diferença entre o custos primal e dual for inferior a um critério pré-determinado (2%) então a execução é finalizada.

O método de subgradiente é aplicado em associação com o procedimento heurístico. Periodicamente (a cada 6 iterações do método de subgradiente) é executada uma iteração da heurística com o objetivo de melhorar (reduzir) o limitante superior do custo de implantação. O algoritmo itera dentro do método de subgradiente até que a contribuição das iterações para o incremento do custo dual seja inferior a um limiar estabelecido ou que o número de iterações ultrapasse um determinado valor.

6.1.3 Experimento Computacional.

Foram realizados alguns experimentos computacionais para testar as modificações introduzidas no método proposto por Balakrishnan e para uma avaliação da complexidade computacional.

Foram feitos experimentos com redes elementares, com uma rede artificial e com uma rede realista. As redes elementares foram utilizadas para teste de validação do programa e para avaliação de sua complexidade computacional. A rede artificial foi desenvolvida para permitir uma análise do funcionamento do método e de suas deficiências. A rede realista foi utilizada para verificar a utilidade da ferramenta proposta neste trabalho para o planejamento de redes telefônicas e para confrontar as hipóteses adotadas com uma situação prática.

Função de Custo.

Nos experimentos computacionais foram adotadas duas funções de custo: Uma contínua e outra modular. Nos dois casos a função de custo é linear por partes. Os segmentos das funções foram escolhidos para coincidir com os módulos de 30 canais.

O custo de uma rota é proporcional ao número de circuitos necessários para que tenha uma probabilidade de bloqueio de 0.5%. Este número de circuitos é calculado por meio da fórmula

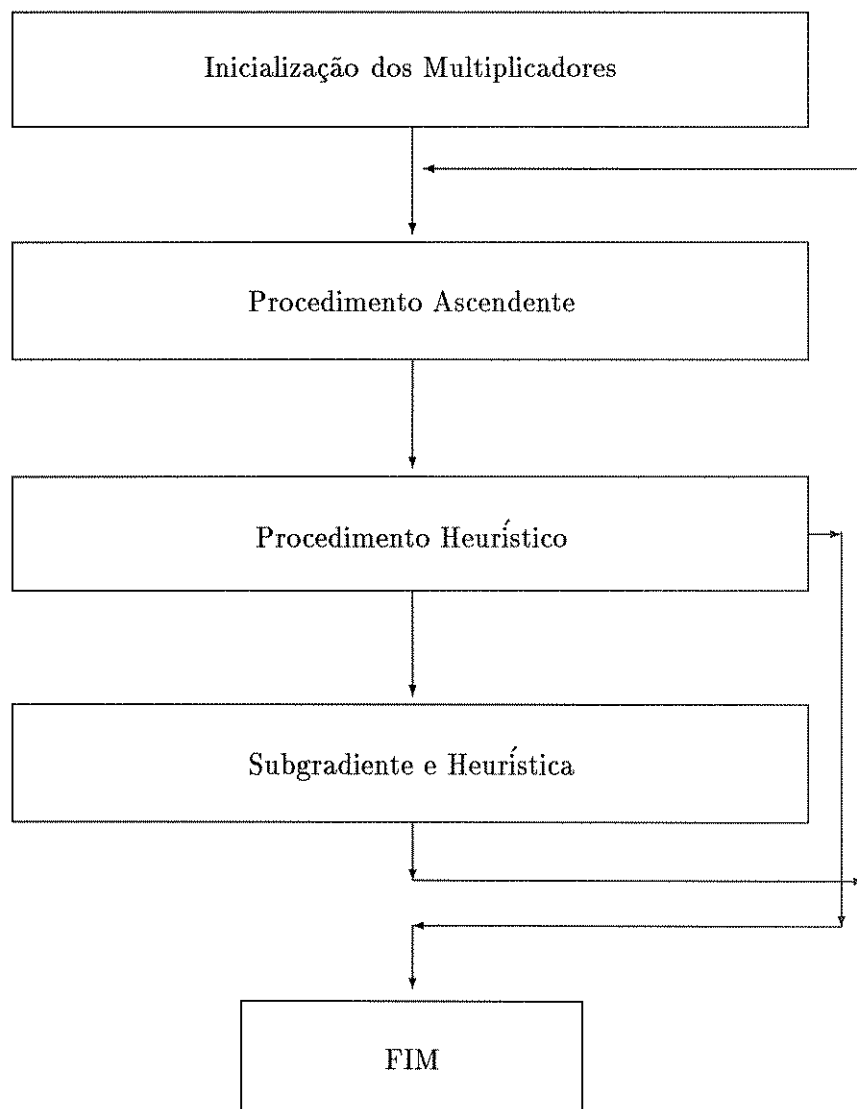


Figura 14: Algoritmo para Resolução do Problema de Encaminhamento de Custo Mínimo de Implantação.

TRÁFEGO	CUSTO LINEAR	CUSTO FIXO
7.00000	2.14290	0.00000
19.03390	1.25000	6.25030
44.75660	1.16630	7.84344
71.75510	1.11120	10.30953
99.38160	1.00859	17.67232
127.39560	1.00709	17.82139
155.67731	1.00608	17.95005
184.15720	1.00534	18.06526
212.79111	1.00477	18.17022
241.54849	1.00432	18.26598
270.40771	1.00395	18.35536
299.35269	1.00364	18.43917
328.37091	1.00338	18.51703
357.45270	1.00316	18.58925
386.59009	1.00296	18.66076
415.77689	1.00279	18.72647
445.00760	1.00263	18.79299
100000.00000	1.00000	19.96336

Tabela 5: Esta tabela apresenta, para cada segmento da função custo de tipo contínuo, o seu limite superior, o custo linear por Erlang e o seu custo fixo.

inversa de Erlang. Para cada rota, o custo é obtido multiplicando-se o número de circuitos por um fator correspondente ao custo por circuito em cada rota, que é dependente das centrais interligadas pela rota.

Na função contínua, os segmentos são contínuos e a derivada dos segmentos é decrescente em relação ao tráfego oferecido à rota. Este tipo de custo é aplicável à redes analógicas, onde o custo da rota varia linearmente com o número de circuitos.

Para o custo modular, a função é do tipo escada até 445 Erlangs (480 circuitos). Cada degrau correspondente a um número inteiro de sistemas de primeira hierarquia (30 canais) necessários para que a rota possa ser dimensionada com um bloqueio inferior a 0.5%. Este tipo de custo é aplicável à redes digitais. Para tráfegos acima de 445 Erlangs aplica-se um custo linear, por considerar-se, ao menos para os experimentos computacionais, que acima deste número de circuitos o efeito da modularidade pode ser desprezado.

As tabelas 5 e 6 apresentam, para cada segmento das funções de custo, o limite superior do segmento, o custo linear por Erlang e o custo fixo do segmento.

Tempo de Execução.

Foram criadas quatro redes semelhantes, onde todas as centrais admitem interligação por rota direta e originam tráfegos para as demais centrais da rede. O custo por circuito em todas as rotas é unitário e o resultado para as quatro redes é o estabelecimento de rotas diretas entre todas as centrais. Esta solução é encontrada pelo algoritmo em apenas uma iteração e a diferença entre os

TRÁFEGO	CUSTO LINEAR	CUSTO FIXO
19.03390	0.00000	30.00000
44.75660	0.00000	60.00000
71.75510	0.00000	90.00000
99.38160	0.00000	120.00000
127.39560	0.00000	150.00000
155.67731	0.00000	180.00000
184.15720	0.00000	210.00000
212.79111	0.00000	240.00000
241.54849	0.00000	270.00000
270.40771	0.00000	300.00000
299.35269	0.00000	330.00000
328.37091	0.00000	360.00000
357.45270	0.00000	390.00000
386.59009	0.00000	420.00000
415.77689	0.00000	450.00000
445.00760	0.00000	480.00000
916.00000	1.01910	26.49277
1391.00000	1.01050	34.37037
1868.00000	1.00630	40.21254
100000.00000	1.00000	51.98000

Tabela 6: Esta tabela apresenta, para cada segmento da função custo tipo Modular, o seu limite superior, o custo linear por Erlang e o seu custo fixo.

TOTAL DE NÓS	TEMPO DE CPU (seg.)
10	24
15	221
20	1.131
30	12.888

Tabela 7: Tempo de execução para diferentes portes de redes.

custos primal e dual é inferior à 2%, que é um dos critérios para término de execução do algoritmo. Estas redes foram utilizadas para avaliar o crescimento do tempo de execução computacional com o aumento do porte das redes. A tabela 7 apresenta os tempos de execução para redes com 10, 15, 20 e 30 nós de comutação. Estes resultados foram obtidos para uma estação de trabalho SPARC2.

Rede Artificial

A rede artificial possui 19 centrais, sendo 3 centrais do tipo A (centrais 0, 1 e 2) e 16 centrais do tipo B (centrais 3 a 18). A matriz de tráfego e a matriz de custo por circuito nas rotas estão apresentadas na tabela 8.

Há uma exceção à tabela 8 para o custo por circuito nas rotas entre as centrais 1 (tipo A) e 3 (tipo B). No caso das rotas 1-3 e 3-1 o custo é de 0.8. Esta exceção tem como objetivo tornar um pouco mais atrativo o encaminhamento utilizando a central 1 como trânsito. O modelo de custo aplicado na determinação da solução para esta rede é do tipo contínuo.

Tipos de Centrais	A-A	A-B	B-A	B-B
Custo por Circuito	0.9	0.9	0.9	1
Interesse de Tráfego	240	30	30	3

Tabela 8: Parâmetros da rede artificial para cada tipo de arco.

A solução ótima para esta rede, que pode ser determinada por inspeção graças a sua simetria, consiste em interligar as centrais 0, 1 e 2 com todas as demais centrais. Todo tráfego originado ou terminado nestas centrais é encaminhado diretamente. O tráfego entre as centrais tipo B deve ser encaminhado por meio da central 1, uma vez que para os tráfegos entre centrais tipo B e central 3 o encaminhamento pela central 1 é mais econômico pelo custo inferior por circuitos entre as demais centrais tipo B por um efeito de ganho de escala .

A solução encontrada no experimento computacional foi interligar as centrais 0, 1 e 2 diretamente com todas as demais centrais. O tráfego originado e terminado nestas centrais foi encaminhado diretamente. O tráfego entre a central 3 e as demais centrais tipo B foi encaminhado pela central 1. Os demais tráfegos entre centrais tipo B foi encaminhado pela central 0. O algoritmo foi capaz de perceber o alto ganho de escala em encaminhar os tráfegos entre centrais tipo B via central tandem, mas não foi capaz de perceber o ligeiro ganho de escala em utilizar a central 1 em lugar da central 0, escolhida simplesmente porque o número 0 vem antes de 1. Este efeito decorre da linearização por partes da função objetivo. O algoritmo somente é capaz de reconhecer ganhos de escala quando se passa de um segmento para outro da função objetivo, quando na realidade qualquer aumento de tráfego nos arcos produz estes ganhos. Para reduzir este efeito da linearização pode-se aumentar o número de segmentos. Entretanto, este procedimento reduz a velocidade do algoritmo. O número de segmentos da função objetivo deve ser escolhido de forma a equilibrar o tempo de execução e a precisão dos resultados.

A diferença entre o custo ótimo e o custo da solução encontrada pode ser avaliada em termo de número de circuitos. A solução ótima exige 7320 circuitos e a solução encontrada 7350 circuitos. A diferença entre o custo do problema dual e o custo primal ("gap" de dualidade) neste caso foi de 12% em relação ao custo da função dual. Neste caso verificou-se que a solução ótima encontra-se muito mais próxima da solução primal do que a solução dual. Nos experimentos computacionais foram verificados diferenças percentuais de 0% a 35% entre os custos primal e dual, em relação a custo dual.

Rede Realista

O experimento com uma rede realista foi feita com uma rede local de 29 centrais. Esta rede corresponde a um cenário proposto para a rede de Campinas para o ano de 1990.

Embora a rede possua um número relativamente alto de centrais, o porte do problema está aquém do limite da ferramenta desenvolvida para o experimento. Muitos nós de comutação da rede executam apenas a função de trânsito. Isto torna a matriz de tráfego relativamente incompleta. Por outro lado, por motivos operacionais e tecnológicos, muitas centrais possuem uma autonomia

	CPU	Custo Total	Troncos	Rotas
Modular	381 s.	9.9×10^6	20.068	119
Linear	500 s.	9.0×10^6	18.523	123
OTELO		13.1×10^6	17.765	140

Tabela 9: Resultados obtidos utilizando custos do tipo linear e modular comparados aos resultados obtidos por uma ferramenta clássica de planejamento. Não se conhece precisamente o tempo de execução para o programa OTELO, mas sabe-se que é da ordem de poucos minutos.

de encaminhamento restrita, eliminando a possibilidade de estabelecimento de algumas rotas. Isto reduz significativamente o número de variáveis do problema.

A rede foi estudada para um cenário utilizando custos lineares para todas rotas, e outro com custos modulares. Os resultados foram comparados aos obtidos com uma ferramenta de planejamento para dimensionamento de redes locais (OTELO). O programa OTELO foi desenvolvido pela TELESP para o dimensionamento da sua rede de troncos. Esta ferramenta utiliza um plano de encaminhamento previamente estabelecido. Esta comparação demonstra o potencial da metodologia proposta, embora não permita uma avaliação mais precisa a respeito do ganho obtido com sua aplicação, uma vez que o encaminhamento otimizado viola restrições técnicas da rede. A tabela 9 apresenta um resumo dos resultados dos experimentos.

No experimento foram definidos três tipos diferentes de rotas, com custos por circuito diferentes: rotas entre duas centrais digitais, entre duas centrais analógicas e entre uma central digital e outra analógica. O algoritmo foi aplicado duas vezes, utilizando uma função objetivo linear por partes contínua e uma função modular, como apresentado anteriormente.

A rede dimensionada pelo OTELO resulta numa rede com menos troncos e custo mais elevado. Isto significa que o algoritmo proposto escolheu caminhos com maior número de rotas, mas escolheu rotas com menor custo por circuito.

6.1.4 Aplicação em Redes de Grande Porte.

No experimento computacional verificou-se que, para o porte dos computadores utilizados, o algoritmo descrito é capaz de tratar o problema formulado para redes em torno de 30 centrais com um tempo de processamento da ordem de algumas horas. Este custo computacional aumenta rapidamente com o número de centrais. Entretanto, uma ligeira modificação na formulação do problema permite que o mesmo algoritmo trate redes com mais de 100 centrais. Esta variação consiste em considerar como um produto todo tráfego com destino a um dado nó. Neste caso o número de tráfegos não varia com aproximadamente o quadrado do número de centrais, mas linearmente. Isto reduz significativamente as dimensões do problema.

Esta variação da formulação do problema exige uma alteração do algoritmo no cálculo dos caminhos de mínimo custo. Uma vez que para cada tráfego existem diversas origens, apenas uma delas deve ser escolhida como referência.

6.2 Encaminhamento de Desempenho Ótimo.

A seguir é apresentada uma formulação para determinação de encaminhamento ótimo para uma dada rede. Os dados de entrada do problema são o número de circuitos nas rotas e o tráfego oferecido ponto-a-ponto. O resultado é o plano de encaminhamento, utilizando a técnica de repartição de carga, que minimiza o tráfego bloqueado na rede.

O algoritmo de otimização empregado neste caso é uma adaptação do algoritmo apresentado nas seções anteriores. A principal motivação para a utilização deste método foi testar modificações introduzidas no algoritmo proposto por Balakrishnan que o tornam capaz de resolver problemas não côncavos. O problema de determinação do encaminhamento abordado nesta seção, é semelhante ao problema de determinação de encaminhamento de custo mínimo de implantação, exceto que neste segundo o problema é não convexo. Para problemas convexos, como o proposto neste trabalho, existem métodos clássicos muito mais eficientes, particularmente porque o tipo da função objetivo é conhecida [37]. Portanto a escolha do método de solução não teve como objetivo desenvolver um método eficiente de resolução, mas aprimorar o método proposto por Balakrishnan.

O objetivo do problema de otimização do encaminhamento é minimizar o tráfego bloqueado nas rotas, ou seja, maximizar o tráfego cursado na rede. A técnica de encaminhamento considerada é a técnica de repartição de carga sem transbordo. A técnica de repartição de carga consiste em estabelecer para cada tráfego um certo número de caminhos que interliguem a origem e o destino das chamadas. As chamadas são distribuídas pelos caminhos de acordo com probabilidades estabelecidas. A solução do problema leva à determinação dos caminhos e as probabilidades de escolha dos caminhos associadas a cada tráfego.

São dados de entrada do problema:

- A topologia da rede.
- O número de circuitos nas rotas.
- A matriz de tráfego oferecido ponto-a-ponto.

6.2.1 Formulação do Problema de Otimização.

O problema proposto é formulado como um problema de fluxo multiproduto de custo mínimo, onde o custo de transporte do fluxo (tráfego telefônico) representa o tráfego bloqueado em cada rota, que é a mesma formulação do problema de custo mínimo.

$$\text{Min}_{x_{ij}^k} \sum_{i=1}^N \sum_{j=1}^N f_{ij} \left(\sum_{k=1}^T x_{ij}^k \right)$$

sujeito a:

$$\begin{aligned} \sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k &= +d_k ; \forall k, \text{ se } j \text{ é destino de } k (j = D(k)) \\ \sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k &= -d_k ; \forall k, \text{ se } j \text{ é origem de } k (j = O(k)) \\ \sum_{i=1}^N x_{ij}^k - \sum_{i=1}^N x_{ji}^k &= 0 ; \forall k, \text{ caso contrário} \end{aligned}$$

$$0 \leq x_{ij}^k \leq d_k$$

Onde:

x_{ij}^k é a intensidade do tráfego k no arco que une os nós i e j.

d_k é a intensidade do tráfego k.

N é o número de nós.

T é o número total de produtos (tráfegos).

$f_{ij}(\sum_{k=1}^T x_{ij}^k)$ é a função tráfego bloqueado no arco ij.

Nesta formulação cada arco ij representa uma rota lógica da rede, ou seja, um conjunto de canais de voz entre duas centrais. O produto k representa o tráfego telefônico originado numa central O(k) com destino a outra central D(k).

As restrições do problema determinam a conservação de tráfego para cada nó. Todo tráfego que entra no nó, incluindo o tráfego originado, deve ser igual ao tráfego que sai do nó, incluindo o tráfego terminado. Este tipo de restrição leva em conta a hipótese de que a probabilidade de bloqueio nas rotas é pequena. Portanto não são considerados transbordos.

Supõe-se, como aproximação, que o tráfego oferecido às rotas é poissoniano. Portanto a probabilidade de bloqueio na rota é dada pela fórmula B de Erlang e o tráfego bloqueado numa dada rota é dado pela expressão:

$$f_{ij} = \sum_{k=1}^T x_{ij}^k \cdot \text{Erl}(\sum_{k=1}^T x_{ij}^k, N_{ij}) \quad (15)$$

Onde:

f_{ij} é o tráfego bloqueado na rota ij.

x_{ij}^k é o tráfego oferecido à rota ij.

$\text{Erl}(A, N)$ é a fórmula B de Erlang para um tráfego A e um número de circuitos N.

N_{ij} é o número de circuitos na rota ij.

T é o número de tráfegos existentes na rede.

A função objetivo do problema representa a soma de todo tráfego bloqueado nas rotas da rede. Supondo-se que o tráfego oferecido a cada rota é poissoniano e independente das demais rotas, o tráfego bloqueado na rede será a soma do tráfego bloqueado em cada rota, dado pela expressão:

$$x_b = x \cdot \text{Erl}(x, n) \quad (16)$$

Onde:

x_b é o tráfego bloqueado na rota.

x é o tráfego oferecido à rota.

$\text{Erl}(x, n)$ é a probabilidade de bloqueio dada pela fórmula B de Erlang dado que a rota possui n circuitos.

O método escolhido para resolução do problema proposto exige que a função objetivo seja do tipo linear por partes, como definido a seguir:

$$f(x) = \begin{cases} c^1.x + F^1; & Se 0 \leq x \leq M^1 \\ c^2.x + F^2; & Se M^1 \leq x \leq M^2 \\ \dots & \\ c^R.x + F^R; & Se M^{R-1} \leq x \leq M^R \end{cases}$$

Onde a variável x representa um tráfego oferecido.

Portanto é preciso realizar uma linearização da função probabilidade de bloqueio nas rotas dada pela equação 15. Esta expressão pode ser linearizada com precisão utilizando-se relativamente poucos segmentos. Para valores de tráfego pequenos a probabilidade de bloqueio é aproximadamente nula. Para valores elevados de tráfego ela apresenta um comportamento quase linear. Estas duas faixas de tráfego podem ser aproximadas por um segmento cada. A faixa de tráfego que deve ser representada por múltiplos segmentos é aquela que corresponde às probabilidades de bloqueio entre 0.5% e 20%.

A função da equação 16 é linearizada, para cada rota, de acordo com o seguinte procedimento. Inicialmente são determinados os valores de tráfego oferecido que definem os segmentos lineares da função ($M^1, M^2, \dots$). São escolhidos doze valores de acordo com as expressões:

$$M^i = \arg_x [b_i = Erl(x, n)]; i = 1, 2, \dots, 11$$

$$M^{12} = 100000$$

Onde:

b_i é a i -ésima componente do vetor $b = [0.001, 0.0025, 0.005, 0.0075, 0.01, 0.02, 0.03, 0.04, 0.05, 0.075, 0.10]$

$Erl(x, n)$ é a probabilidade de bloqueio dada pela fórmula B de Erlang dado que a rota possui n circuitos.

Os parâmetros c^i e F^i são obtidos pelas equações:

$$F^1 = 0$$

$$F^i = Erl(M^i, n); 1 < i \leq 12$$

$$c^{12} = 1$$

$$c^1 = \frac{M^1 \cdot Erl(M^1, n)}{M^1}$$

$$c^i = \frac{M^i \cdot Erl(M^i, n) - M^{i-1} \cdot Erl(M^{i-1}, n)}{M^i - M^{i-1}}; 1 < i < 12$$

A técnica de repartição de carga torna a hipótese de que o tráfego é poissoniano bastante precisa se as probabilidades de bloqueio na rede forem pequenas. Quando o tráfego é repartido aleatoriamente, os sub-tráfegos resultantes também são poissonianos e oferecidos a um único caminho sem transbordo. A fórmula proposta para cálculo do tráfego Bloqueado (f_{ij}) é convexa em relação ao tráfego oferecido à rota.

6.2.2 Algoritmo de Resolução.

O algoritmo utilizado nesta seção é uma adaptação do algoritmo empregado para o problema de encaminhamento de custo mínimo apresentado anteriormente. As fases do algoritmo e seu embasamento teórico é o mesmo já apresentado anteriormente. A seguir são enfocadas os pontos do algoritmo que foram modificados.

Procedimento Dual Ascendente.

O procedimento dual ascendente tem como objetivo maximizar os termos $v_{D(k)}^k \cdot d_k$ do problema $D(v)$ sem modificar os termos restantes, garantindo o aumento da função dual. Os multiplicadores de Lagrange são incrementados até que não haja mais ganho entre interações do procedimento. O incremento dos multiplicadores para um dado tráfego é obtido por meio da resolução do seguinte problema:

$$\text{Max } \delta_{D_k}$$

Sujeito a:

$$\delta_j^k - \delta_i^k \leq u_{ij}^k$$

$$\delta_l^k \geq 0$$

$$\delta_{O_k} = 0$$

Onde u_{ij}^k é o maior incremento possível na diferença dos multiplicadores, sem que as variáveis x_{ij}^k do problema $Da_{ij}(x)$ sofram modificações em seus valores.

Os multiplicadores de Lagrange são, então, atualizados para:

$$v_i^k = v_i^k + \delta_i^k$$

A determinação dos limitantes u_{ij}^k é obtida da seguinte forma:

Dado um certo tráfego, verifica-se, para cada arco, qual o maior incremento possível na diferença entre os multiplicadores associados aos nós interligados pelo arco.

Nos arcos em que o tráfego cursado não é nulo, qualquer variação dos multiplicadores gera uma alteração dos demais termos em $D(v)$, o que pode levar a uma diminuição da função dual. Portanto, faz-se $u_{ij}^k = 0$. Para os demais arcos (tráfego nulo) aplica-se o seguinte procedimento:

Inicialmente atribui-se à u_{ij}^k o valor:

$$u_{ij}^k = c_{ij}^r + v_i^k - v_j^k$$

Onde c_{ij}^r é o bloqueio linear relativo ao segmento r da função objetivo, determinado pelo tráfego oferecido no arco ij da solução dual.

Para a determinação do valor definitivo de u_{ij}^k em cada iteração do Dual Ascendente, deve-se obter os valores dos tráfegos cursados nos subproblemas lagrangeanos, nos pontos duais $(v_i^k - v_j^k)$

e $(v_i^k - v_j^k + u_{ij}^k)$.

Seja z_{ij} o tráfego bloqueado da solução de $Da_{ij}(x)$ para o conjunto original de multiplicadores $(v_i^k - v_j^k)$ e seja n_{ij} o tráfego bloqueado da solução para os multiplicadores atualizados, isto é, $(v_i^k - v_j^k + u_{ij}^k)$.

O valor definitivo do limitante u_{ij}^k é obtido de acordo com a expressão:

$$u_{ij}^k = \max(0, c_{ij}^r + v_i^k - v_j^k - (z_{ij} - n_{ij})/m_{ij}^k)$$

Onde m_{ij}^k é a intensidade de tráfego k cursada no arco, na solução do novo problema, ou d_k se o tráfego não é cursado.

Método de Subgradiente.

O método de subgradiente utiliza um vetor $(\hat{w}_i^k)$. No caso dos nós que originam tráfegos, as componentes associadas são nulas. Nos demais casos as componentes do vetor representam, para cada nó da rede e para cada tráfego, de quanto o multiplicador associado deve ser incrementado (ou decrementado), supondo que os demais multiplicadores permaneçam constantes, de forma que o custo lagrangeano se iguale ao custo do segmento imediatamente superior (ou inferior) da função de custo. Para cada nó é calculado um fator de factibilidade, que é o quanto a intensidade de um determinado tráfego que entra no nó ultrapassa a intensidade de tráfego que sai do nó (ou o tráfego terminado no nó). Se o fator é positivo o multiplicador associado deve ser decrementado, caso contrário incrementado. Se o fator for nulo então a solução é factível para aquele determinado nó e tráfego. Portanto o multiplicador não deve ser modificado.

A cada iteração, o método calcula o subgradiente, atualiza os multiplicadores de Lagrange, resolve o problema dual determinando um novo ponto (v, x) . O método itera até atingir um critério de parada.

Os multiplicadores são atualizados pela expressão:

$$v_i^k = v_i^k + \lambda \cdot \hat{w}_i^k$$

O parâmetro λ assume inicialmente o valor 1. Se a iteração do algoritmo leva a uma solução inferior à obtida na iteração anterior então a iteração é anulada e uma nova tentativa é realizada com o valor de λ dividido por 2. Este procedimento é repetido até que seja obtido uma iteração que resulte numa nova solução do problema dual superior à anterior.

Resumo do Algoritmo Proposto.

A figura 15 apresenta o algoritmo global de resolução do problema.

O primeiro passo é a inicialização dos multiplicadores de Lagrange. Para cada tráfego é resolvido um problema de caminho de mínima distância entre a origem e os demais nós da rede. Este problema utiliza como distância dos arcos o menor bloqueio por Erlang oferecido, obtido a partir da função

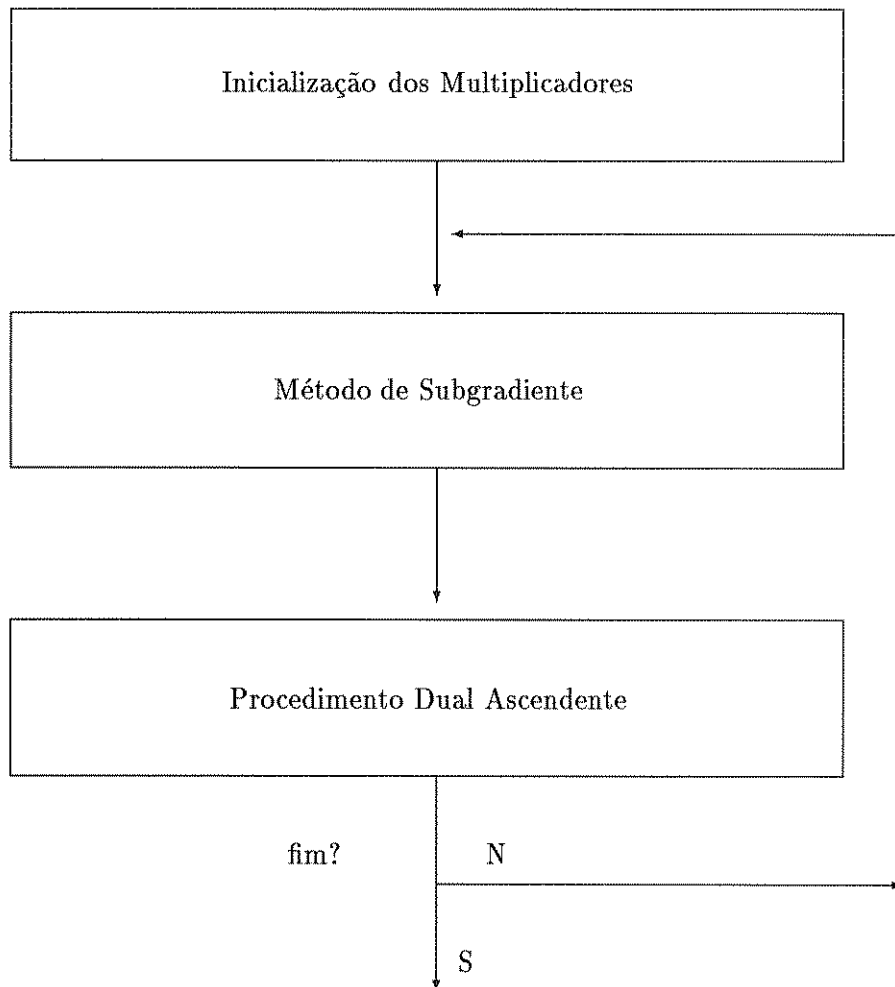


Figura 15: Algoritmo para Resolução do Problema de Encaminhamento de Ótimo Desempenho.

objetivo do problema original, ou seja, a distância é a inclinação da reta que passa pela origem e tangencia por baixo a função objetivo. Conseqüentemente, a função objetivo dos subproblemas lagrangeanos serão sempre positivos e os tráfegos serão todos nulos.

Após a inicialização, o algoritmo itera dentro do método de subgradiente até que seja alcançada uma solução factível ou que o número de iterações ultrapasse um determinado valor.

Em seguida é executado o procedimento dual ascendente, que itera até o número de iterações ultrapassar um certo limite, caso em que o procedimento de subgradiente é novamente executado, ou se for encontrada uma solução factível, quando a execução do algoritmo é finalizada.

Como o problema proposto é convexo, não é preciso executar o procedimento heurístico para obter uma solução factível. Esta solução é obtida ao final do procedimento dual ascendente.

Foi desenvolvido um programa, implementando este algoritmo. Neste caso não cabe um discussão similar ao do problema de custo mínimo. A dimensão do problema de encaminhamento de desempenho ótimo é muito inferior, dado que se conhece as rotas existentes, de forma que o tempo de execução é da ordem de segundos. Por outro lado, devido à convexidade do problema não há a questão da precisão do resultado, que é sempre ótimo. O programa teve como objetivo apenas validar as modificações propostas no algoritmo original.

6.3 Conclusões

Neste capítulo foi apresentada a formulação de um problema de determinação do plano de encaminhamento para redes telefônicas, de forma a minimizar o custo de implantação de equipamentos.

Foi proposto um método de solução para esta formulação e foram realizados experimentos computacionais que atestaram a viabilidade prática do método. Uma série de adaptações e extensões foram feitos no algoritmo original utilizado na resolução do problema visando a sua adaptação à formulação modular e ao aumento de sua eficiência numérica. Uma aplicação preliminar do algoritmo em um rede de porte significativo, de 30 centrais, apresentou uma economia considerável no custo de implantação, demonstrando o potencial desta técnica.

Esta formulação estabelece uma nova metodologia para a determinação do plano de encaminhamento e do dimensionamento das rotas de uma rede telefônica, que leva em consideração o ganho de escala proporcionado pela concentração de tráfegos e os custos heterogêneos de ligação entre cada par de centrais existente na rede.

Esta metodologia é composta de duas etapas: a primeira determina um plano de encaminhamento considerando apenas caminhos de primeira escolha para cada tráfego, e a segunda consiste no processo clássico de dimensionamento de rotas, utilizando os resultados da primeira etapa.

Este capítulo também apresentou o problema de determinação do plano de encaminhamento de ótimo desempenho para uma rede telefônica que utilize a técnica de encaminhamento por repartição de carga.

O algoritmo utilizado neste caso é uma adaptação do algoritmo apresentado por Balakrishnan [6]. Foi introduzida uma modificação no método de subgradiente originalmente proposto. Foram

introduzidas alterações no método de resolução dos subproblemas lagrangeanos de forma a tratar qualquer tipo de função objetivo linear por partes (originalmente era possível apenas tratar funções côncavas).

Capítulo 7

Formulação Tipo Cadeia de Markov.

Temos, por isso, que encarar o presente estado do universo como resultante do estado anterior e como causa do que virá. Se tivéssemos, por um instante, uma inteligência que pudesse compreender todas as forças que animam a natureza e a respectiva situação dos seres que a compõem - uma inteligência suficientemente vasta para poder submeter esses dados a uma análise, - ela iria abarcar em uma mesma fórmula os movimentos dos maiores corpos do universo e os dos menores átomos; para ela, nada seria incerto, e o futuro, assim como o passado, estaria presente a seus olhos.

Laplace.

Este capítulo apresenta um novo método de determinação do plano de encaminhamento para uma rede que empregue a técnica de encaminhamento dependente do estado. Este método está fundamentado na teoria de processo de decisão markoviano. O método pode ser empregado tanto na determinação de um plano de encaminhamento estático como pode ser integrado a um sistema de encaminhamento dinâmico dependente do estado. O processo de otimização é implementado por meio de uma simulação da rede.

O encaminhamento dependente do estado é uma técnica de encaminhamento que tem sido intensamente estudada recentemente [36], [12]. Os equipamentos de comutação da rede telefônica já são capazes de encaminhar chamadas individualmente, de acordo com o estado da rede. Várias empresas desenvolvem projetos, na tentativa de viabilizar o emprego prático deste tipo de encaminhamento.

Seu princípio de funcionamento supõe que as centrais telefônicas podem, em tempo real, determinar o estado da rede, ou seja, quantos circuitos estão ocupados em cada rota da rede. A rede pode então encaminhar o tráfego de forma distinta, para diferentes estados da rede.

Pode-se classificar como encaminhamento dependente do estado um grande número de técnicas. Por exemplo, a técnica de minimização da capacidade residual é uma técnica de encaminhamento dependente do estado. As chamadas são estabelecidas nos caminhos que possuem o maior número de circuitos desocupados. Dependendo do número de circuitos desocupados a chamada é encaminhada por um ou outro caminho. No contexto deste trabalho, o encaminhamento dependente do estado é definido na sua forma mais geral. Para cada estado da rede, há um caminho estabelecido para

cada tráfego (ou não há caminho e o tráfego é rejeitado naquele estado). O encaminhamento com minimização da capacidade residual é portanto um caso particular (ou uma solução possível) de encaminhamento dependente do estado.

7.1 Cadeia Semi-Markoviana

O método de determinação do plano de encaminhamento, proposto neste trabalho, está fundamentado na teoria de processo de decisão markoviano [20], [39]. O processo de estabelecimento das chamadas na rede é representado por meio de uma cadeia semi-markoviana com recompensa. Esta representação pode ser aplicada a uma rota isoladamente ou a rede como um todo. A seguir são apresentados alguns conceitos sobre cadeias semi-markovianas e convenções que serão utilizadas ao longo do capítulo.

Uma cadeia semi-markoviana com recompensa é um processo que apresenta um certo número de estados. O processo passa de um estado para outro ao longo do tempo. Entre duas transições de estado, o processo permanece num dado estado durante um certo tempo. O tempo de permanência é uma variável aleatória. Cada vez que o processo visita um dado estado, há um custo de permanência no estado, que é também uma variável aleatória. Por exemplo, seja uma rota com um dado número de circuitos. Sempre que a rota atinge o estado onde todos os circuitos estão ocupados, chamadas são bloqueadas. O número médio de chamadas bloqueadas cada vez que a rota atinge este estado é o custo médio associado ao estado. Dado que o processo se encontra num estado, pode ocorrer uma transição para qualquer outro estado. A escolha do próximo estado é aleatória, segundo um vetor de probabilidade constante.

Uma cadeia semi-markoviana com recompensa é definida pelos seguintes elementos:

- Estados do processo.
- P_{ij} = Probabilidade do sistema alcançar o estado j numa transição, dado que está no estado i .
- τ_i = Esperança do tempo de permanência no estado i .
- c_i = Esperança do custo de permanência no estado i .

São adotadas duas premissas básicas:

- A cadeia não possui estados transientes ou absorventes (cadeia homogênea). Isto significa que o número médio de transições por unidade de tempo para qualquer estado é constante ao longo do tempo.
- Os parâmetros que definem a cadeia (P_{ij} , τ_i e c_i) são constantes e independentes da história da rede e dependem apenas do estado i .

Seja $P_{ij}^{(n)}$ a probabilidade do processo alcançar o estado j em n transições, dado que esta inicialmente no estado i .

Conhecendo-se P_{ij} pode-se calcular $P_{ij}^{(n)}$ através das equações:

$$P_{ij}^{(n+1)} = \sum_k P_{ik}^{(n)} \cdot P_{kj} ; n \geq 0 \quad (17)$$

$$P_{ij}^{(0)} = \begin{cases} 0 ; i \neq j \\ 1 ; i = j \end{cases} \quad (18)$$

Seja π_j a esperança da fração do tempo em que o processo permanece no estado j ao longo do tempo (ou seja a probabilidade de encontrar o processo no estado j para $t \rightarrow \infty$).

$$\pi_j = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m P_{ij}^{(n)} \quad (19)$$

Seja a variável $\tau_i^{(m)}$ a esperança do tempo utilizado para realizar as m primeiras transições partindo-se do estado i , definida pela seguinte equação recorrente:

$$\tau_i^{(1)} = \tau_i$$

$$\tau_i^{(m+1)} = \tau_i^{(m)} + \sum_j P_{ij}^{(m)} \cdot \tau_j ; m > 1$$

Seja $V_m(i)$ a esperança do custo total do processo dado que ocorrem m transições partindo inicialmente do estado i . Portanto:

$$V_m(i) = \sum_{k=0}^{m-1} \sum_j P_{ij}^{(k)} \cdot c_j$$

Seja g o custo médio estacionário por unidade de tempo (ou a taxa de custo de longo prazo) associado ao processo definido por:

$$g = \lim_{m \rightarrow \infty} \frac{V_m(i)}{\tau_i^{(m)}} ; \quad \forall i$$

É possível demonstrar [39] que se v_i e g são valores que satisfazem o sistema de equações:

$$v_i = c_i - g \cdot \tau_i + \sum_j P_{ij} \cdot v_j ; i \in I \quad (20)$$

Então:

$$v_i - v_j = \lim_{m \rightarrow \infty} (V_m(i) - V_m(j))$$

A expressão $v_i - v_j$ representa a diferença do custo de se iniciar o processo no estado i em vez do estado j . No caso de uma rota telefônica, $v_{k+1} - v_k$ representa o número de futuras chamadas bloqueadas devido ao estabelecimento de uma chamada quando a rota se encontra no estado k .

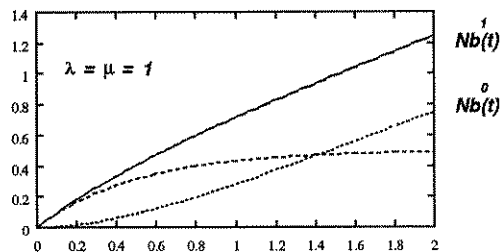


Figura 16: Número médio acumulado de clientes que encontram uma fila M/M/1/1 ocupada ao longo do tempo. O limite para tempo infinito da diferença entre as duas curvas é o custo relativo do estado ocupado em relação ao estado livre.

A figura 16 apresenta o número médio acumulado de clientes que encontram uma fila M/M/1/1 ocupada ao longo do tempo, que neste caso é o custo total do processo, quando o sistema inicialmente está livre ou ocupado. O limite para tempo infinito da diferença entre as duas curvas é o custo relativo do estado ocupado em relação ao estado livre.

7.2 Políticas de Encaminhamento

A teoria de processos markovianos com recompensa permite abordar o problema de determinação de políticas ótimas. Seja uma cadeia semi-markoviana com um determinado número de estados. A cada estado (i) é possível estabelecer um conjunto de alternativas de ação $(a_i \in A_i)$. Cada alternativa (a_i) resulta num determinado tempo de permanência no estado $(\tau_i(a_i))$ e em determinadas probabilidades de transição para outros estados $(P_{ij}(a_i))$. Escolhendo-se para cada estado uma das alternativas possíveis, define-se um política (R) . Pode-se afirmar que cada política resulta numa cadeia semi-markoviana distinta. Por hipótese-se, admite-se que qualquer política R resulta numa cadeia homogênea.

Seja, por exemplo, uma máquina capaz de produzir dois tipos de peças. Existem duas linhas de produção distintas que consomem um tipo de peça cada. É possível estocar no máximo N_1 e N_2 peças de cada tipo, respectivamente.

O tempo de produção da máquina e o tempo entre o consumo de duas peças pela linha de produção são variáveis aleatórias exponencialmente distribuídas. O número de peças estocadas define os estados do processo. A cada estado é preciso definir qual o próximo tipo de peça a ser produzido na máquina (exceto no estado (N_1, N_2) em que nenhuma peça será produzida). Esta decisão envolverá os custos de interromper uma ou outra linha de produção. Assim a cada estado existem duas ações possíveis: produzir peça para a linha 1 ou para a linha 2. Escolhendo-se uma ação em cada estado obtem-se uma política.

O modelamento de uma rede telefônica por meio de uma cadeia semi-markoviana permite estabelecer o plano de encaminhamento ótimo para uma dada rede. O plano de encaminhamento pode ser definido como um conjunto de regras que determina a forma como os circuitos das diversas rotas são alocados para o estabelecimento de uma chamada. No caso de um encaminhamento dependente do estado da rede, para cada estado existe uma regra para o estabelecimento das chamadas. Assim,

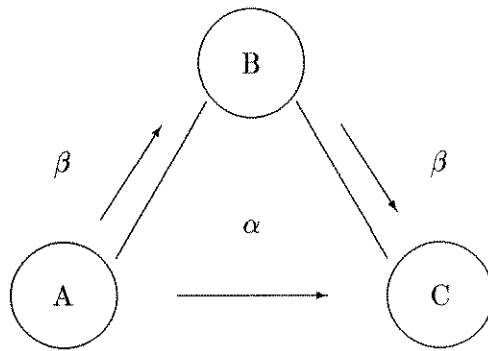


Figura 17: As centrais A e B e C são interligadas por duas rotas. Existem tráfegos com origem em A e destino em B e C, e com origem em B com destino em C. O tráfego entre A e C utiliza a central B como trânsito e possui uma taxa de geração de chamadas α . Os demais tráfegos possuem uma taxa β .

neste caso, o plano de encaminhamento corresponde a uma política.

Seja a rede apresentada pela figura 17. A rede possui três centrais e duas rotas. Existem tráfegos com origem na central A e destino nas centrais B e C, e com origem na central B com destino na central C. O tráfego entre A e C utiliza a central B como trânsito e possui uma taxa de geração de chamadas α . Os demais tráfegos possuem uma taxa β . Deseja-se estabelecer um plano de encaminhamento que permita reduzir o bloqueio de chamadas ao mínimo (ou maximizar o tráfego cursado). Se cada rota da rede possuir apenas um circuito é possível estabelecer duas políticas:

- O tráfego α é sempre aceito.
- O tráfego α é sempre bloqueado.

É possível perceber intuitivamente que se a intensidade dos tráfegos β for alta, resultando numa alta ocupação dos circuitos, o tráfego α deve ser bloqueado. Cada chamada entre as centrais A e C ocupa dois circuitos. As demais chamadas ocupam apenas um circuito. Desta forma é possível estabelecer perto do dobro de chamadas se o tráfego α for descartado. Por outro lado se o tráfego β for nulo (ou muito pequeno) então o tráfego α deve ser admitido. A rede admite outras políticas, como bloquear os tráfego β , mas que são inferiores às duas políticas propostas.

7.3 Política Ótima

A partir do modelamento da rede como um cadeia semi-markoviana é possível determinar a política de encaminhamento ótima. Existem diversos métodos de otimização de políticas. A seguir é apresentado o algoritmo de iteração de política [39].

Teorema 1 *Sejam $g(R)$ e $v_i(R)$, os custos associados a uma cadeia semi-markoviana para uma dada política R . Se para uma outra política $\bar{R}$ é válida a expressão:*

$$c_i(\bar{R}) - g(R) \cdot \tau_i(\bar{R}) + \sum_j p_{ij}(\bar{R}) \cdot v_j(R) \leq v_i(R) \quad (21)$$

Então é possível demonstrar que [39]:

$$g(\bar{R}) \leq g(R)$$

Portanto a política $\bar{R}$ é igual ou superior à política R , pois o custo $g(\bar{R})$ é igual ou inferior a $g(R)$.

Seja um processo com um determinado número de estados e um conjunto de ações possíveis definido para cada estado. A partir deste resultado pode-se estabelecer o seguinte algoritmo para determinar a política ótima:

1. Determinar uma política inicial R arbitrária, escolhendo-se uma ação para cada estado.
2. Para a política R , calcular a solução $\{g(R), v_i(R)\}$ do sistema de equações:

$$\begin{aligned} v_i &= c_i(R) - g(R) \cdot \tau_i(R) + \sum_j p_{ij}(R) \cdot v_j \\ v_s &= 0 \end{aligned}$$

Onde s é um estado escolhido arbitrariamente.

3. Para cada estado i , determinar uma ação que satisfaça a equação:

$$a_i = \operatorname{argmin}_{a_i \in A_i} c_i(a_i) - g(R) \cdot \tau_i(a_i) + \sum_j p_{ij}(a_i) \cdot v_j(R)$$

Definir uma nova política $\bar{R}$ como o conjunto de ações $\{a_i\}$.

4. Se $R = \bar{R}$ então R é a política ótima. Caso contrário, voltar ao passo 2 com a política $\bar{R}$ no lugar da política R .

7.4 Modelamento de Redes Telefônicas por meio de Cadeias Semi-Markovianas.

O processo de estabelecimento de chamadas numa rede telefônica admite diversas formas de representação por cadeias semi-markovianas. Uma rede é caracterizada por sua topologia, pelo tráfego oferecido e pelo plano de encaminhamento. No caso geral, um estado da rede é representado por uma combinação possível de estabelecimento de chamadas em diferentes caminhos. Seja um estado i arbitrário. Este estado é caracterizado por uma matriz X_i , onde cada componente da matriz

indica o número de chamadas, correspondente a um determinado tráfego, estabelecidas num dado caminho da rede [8].

Seja a rede da figura 17. A rede apresenta três caminhos e três tráfegos distintos. Os caminhos são AB ($p=1$), BC ($p=2$) e ABC ($p=3$). Os tráfegos são AB ($t=1$), BC ($t=2$) e AC ($t=3$). A cada estado, portanto, corresponde uma matriz com nove elementos:

$$X_i = \begin{bmatrix} x_{1,1}^i & x_{1,2}^i & x_{1,3}^i \\ x_{2,1}^i & x_{2,2}^i & x_{2,3}^i \\ x_{3,1}^i & x_{3,2}^i & x_{3,3}^i \end{bmatrix}$$

Onde $x_{t,p}^i$ representa o número de chamadas do tráfego t estabelecidas no caminho p no estado i . Se, por exemplo, no caso em que cada rota possua apenas um circuito, esta rede é representada por cinco estados:

$$X_1 = \begin{bmatrix} 000 \\ 000 \\ 000 \end{bmatrix} \quad X_2 = \begin{bmatrix} 100 \\ 000 \\ 000 \end{bmatrix} \quad X_3 = \begin{bmatrix} 000 \\ 010 \\ 000 \end{bmatrix} \quad X_4 = \begin{bmatrix} 100 \\ 010 \\ 000 \end{bmatrix} \quad X_5 = \begin{bmatrix} 000 \\ 000 \\ 001 \end{bmatrix}$$

O custo de permanência nos estados admite diversas possibilidades de definição. Para o método de otimização proposto neste trabalho, o custo representa o número médio de chamadas bloqueadas durante a permanência do estado. É o mesmo custo empregado por Ott e Krishnan [33], [25]. Outros parâmetros possíveis são a tarifação média auferida, o número médio de chamadas estabelecidas ou o próprio número de circuitos ocupados no estado. Na seção 7.10 é apresentado um exemplo onde o custo de permanência no estado é a tarifação média auferida.

Os parâmetros da cadeia semi-markoviana derivam das taxas de nascimento e morte dos tráfegos da rede. No caso do exemplo, supondo-se que a taxa de nascimento para os três tráfegos é λ e o tempo médio de duração das chamadas é $\frac{1}{\mu}$, as probabilidades de transição, o tempo médio e o custo médio de permanência do estado 4 são dados por:

$$\tau_4 = 1/2 \cdot \mu \quad c_4 = 3 \cdot \lambda \cdot \tau_4 = 3 \cdot \lambda / 2 \cdot \mu \quad P_{4,3} = P_{4,2} = \mu / 2 \cdot \mu = 0.5$$

Outras formas de representação de estados são possíveis. Por exemplo, um estado pode ser uma combinação de circuitos ocupados nas rotas da rede. Neste caso, para cada estado há uma matriz correspondente de seis elementos:

$$X_i = \begin{bmatrix} x_{1,1}^i & x_{1,2}^i \\ x_{2,1}^i & x_{2,2}^i \\ x_{3,1}^i & x_{3,2}^i \end{bmatrix}$$

Onde $x_{t,r}^i$ representa o número de chamadas estabelecida na rota r no estado i . Supondo que cada rota possua apenas um circuito, esta rede é representada por cinco estados:

$$X_1 = \begin{bmatrix} 00 \\ 00 \\ 00 \end{bmatrix} \quad X_2 = \begin{bmatrix} 10 \\ 00 \\ 00 \end{bmatrix} \quad X_3 = \begin{bmatrix} 00 \\ 01 \\ 00 \end{bmatrix} \quad X_4 = \begin{bmatrix} 10 \\ 01 \\ 00 \end{bmatrix} \quad X_5 = \begin{bmatrix} 00 \\ 00 \\ 11 \end{bmatrix}$$

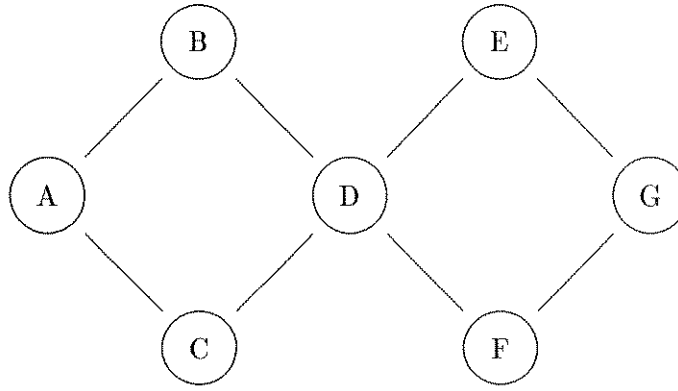


Figura 18: Rede com seis centrais. A definição dos estados deve ser preferivelmente do tipo tráfego caminho.

Esta segunda formulação, no caso da rede exemplo e na maioria das redes, é equivalente à primeira formulação. Entretanto ela apresenta alguns problemas na representação de determinadas redes, como, por exemplo no caso da rede da figura 18. Neste caso uma chamada entre os nós A e G pode ser estabelecida em quatro caminhos diferentes: ABDEG, ABDFG, ACDEG e ACDFG. Seja, por exemplo, a situação em que uma chamada está estabelecida no caminho ABDEG e outra no caminho ACDFG. O estado que representa esta situação é o mesmo que representa uma outra situação: uma chamada estabelecida no caminho ABDFG e outra no caminho ACDEG. A consequência é que a determinação dos parâmetros relacionados ao estado (probabilidade de transição, custo e tempo de permanência) pode ser muito mais difícil que no primeiro método de representação dos estados.

7.5 Algoritmo de Otimização de Encaminhamento

Seja uma rede telefônica com tráfego oferecido poissoniano e dois planos de encaminhamento (políticas R e $\bar{R}$) que determinam, para cada estado, um caminho para estabelecimento ou o bloqueio das chamadas. A partir do tráfego oferecido, da topologia da rede e do plano de encaminhamento define-se uma cadeia semi-markoviana.

Escolhida uma política, o tempo de permanência num estado pode ser calculado a partir das taxas de nascimento e de morte das chamadas:

$$\tau_i(R) = \frac{1}{\sum_{t \in A_i} \lambda_t + \sum_{t \in E_i} n(i, t) \cdot \mu_t} \quad (22)$$

Onde:

$\tau_i(R)$ ($\tau_i(\bar{R})$) é o tempo médio de permanência no estado i para a política R ($\bar{R}$).

λ_t é a taxa média de geração de chamadas do tráfego t .

μ_t é a taxa média de morte de uma chamada do tráfego t .

A_i ($\bar{A}_i$) é o conjunto de tráfegos que não são bloqueados no estado i , sob a política R ($\bar{R}$), e que, portanto, provocam uma transição para outro estado.

E_i é o conjunto de tráfegos que tem uma ou mais chamadas estabelecidas no estado i .

$n(i, t)$ é o número de chamadas do tráfego t estabelecidas no estado i .

Para cada estado em que há bloqueio de tráfegos, associa-se um custo médio de permanência no estado:

$$c_i(R) = \tau_i(R) \cdot \sum_{t \in B_i} \lambda_t \quad (23)$$

Onde B_i ($\bar{B}_i$) é o conjunto de tráfegos cujas chamadas são bloqueadas no estado i para a política R ($\bar{R}$).

Se, sob a política R , o estado $j = J(t, i, R)$ é alcançado a partir do estado i com o nascimento de uma chamada de um dado tráfego t , a probabilidade de transição de estado é calculada a partir da expressão:

$$P_{iJ(t, i, R)}(R) = \frac{\lambda_t}{\sum_{t \in A_i} \lambda_t + \sum_{t \in E_i} n(i, t) \cdot \mu_t} \quad (24)$$

Onde $J(t, i, R)$ é uma função que fornece o estado j alcançado numa transição causada pelo nascimento de uma chamada do tráfego t no estado i .

Se, sob a política R , o estado $j = K(t, i)$ é alcançado a partir do estado i com o término de uma chamada de um dado tráfego t , a probabilidade de transição de estado é calculada a partir da expressão:

$$P_{iK(i, t)}(R) = \frac{\mu_t}{\sum_{t \in A_i} \lambda_t + \sum_{t \in E_i} n(i, t) \cdot \mu_t} \quad (25)$$

Onde $K(i, t)$ é uma função que fornece o estado j alcançado numa transição causada pela morte de uma chamada do tráfego t no estado i .

Teorema 2 *Seja uma política $\bar{R}$ construída a partir da política R da seguinte forma:*

- *Se num dado estado i , sob a política R , uma chamada é estabelecida por um certo caminho, levando o processo do estado i para o estado j , mas existe um caminho alternativo, que leva a um estado k , tal que $v_k(R) - v_i(R) < v_j(R) - v_i(R)$, então sob política $\bar{R}$ é utilizado o caminho alternativo.*
- *Se num dado estado i , sob a política R , uma chamada é rejeitada, mas há um caminho para seu estabelecimento, levando o processo do estado i para o estado k , tal que $v_k(R) - v_i(R) < 1$, então sob a política $\bar{R}$ a chamada é estabelecida neste caminho.*
- *Se num dado estado i , sob a política R , uma chamada é estabelecida num determinado caminho, levando o processo do estado i para o estado j , tal que $v_j(R) - v_i(R) > 1$, então sob a política $\bar{R}$ a chamada é rejeitada neste estado.*

Então a taxa de custo médio do processo sob a política $\bar{R}$ é inferior do que sob a política R , isto é:

$$g(\bar{R}) \leq g(R)$$

Prova:

Da equação 20 obtem-se as equações para os custos relativos:

$$v_i(R) = c_i(R) - g(R) \cdot \tau_i(R) + \sum_j p_{ij}(R) \cdot v_j(R)$$

Substituindo-se as expressões 22, 23, 24 e 25 obtem-se:

$$v_i(R) = \tau_i(R) \cdot \left[\sum_{t \in B_i} \lambda_t - g(R) + \sum_{t \in A_i} \lambda_t \cdot v_{J(t,i,R)}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R) \right]$$

Multiplicando-se ambos os lados da expressão por $\frac{\tau_i(\bar{R})}{\tau_i(R)}$ tem-se:

$$\frac{\tau_i(\bar{R}) \cdot v_i(R)}{\tau_i(R)} = \tau_i(\bar{R}) \cdot \left[\sum_{t \in B_i} \lambda_t - g(R) + \sum_{t \in A_i} \lambda_t \cdot v_{J(t,i,R)}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R) \right] \quad (26)$$

Sejam os conjuntos:

AB_i = conjunto de tráfegos admitidos na rede sob a política R e rejeitados sob a política $\bar{R}$ no estado i.

BA_i = conjunto de tráfegos rejeitados sob a política R e admitidos sob a política $\bar{R}$ no estado i.

M_i = conjunto de tráfegos admitidos sob a política R e admitidos em caminhos diferentes sob a política $\bar{R}$ no estado i.

A partir da equação 22 tem-se:

$$\frac{1}{\tau_i(\bar{R})} = \frac{1}{\tau_i(R)} + \sum_{t \in BA_i} \lambda_t - \sum_{t \in AB_i} \lambda_t$$

$$\frac{\tau_i(\bar{R})}{\tau_i(R)} = 1 + \tau_i(\bar{R}) \cdot \left(\sum_{t \in AB_i} \lambda_t - \sum_{t \in BA_i} \lambda_t \right)$$

Substituindo-se na equação 26 tem-se:

$$v_i(R) = \tau_i(\bar{R}) \cdot \left(\sum_{t \in B_i} \lambda_t - g(R) + \sum_{t \in A_i} \lambda_t \cdot v_{J(t,i,R)}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R) - \sum_{t \in AB_i} \lambda_t \cdot v_i(R) + \sum_{t \in BA_i} \lambda_t \cdot v_i(R) \right)$$

Pela definição da política $\bar{R}$ tem-se:

$$\begin{aligned} v_i(R) &< v_{J(t,i,\bar{R})}(R) - 1 & ; t \in AB_i \\ v_i(R) &> v_{J(t,i,\bar{R})}(R) - 1 & ; t \in BA_i \\ v_{J(t,i,\bar{R})}(R) &< v_{J(t,i,R)}(R) & ; t \in M_i \end{aligned}$$

Portanto:

$$v_i(R) > \tau_i(\bar{R}).[\sum_{t \in B_i} \lambda_t - g(R) + \sum_{t \in A_i} \lambda_t \cdot v_{J(t,i,\bar{R})}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R) - \sum_{t \in AB_i} \lambda_t \cdot (v_{J(t,i,\bar{R})}(R) - 1) + \sum_{t \in BA_i} \lambda_t \cdot (v_{J(t,i,\bar{R})}(R) - 1)]$$

Mas $\bar{A}_i = A_i + BA_i - AB_i$.

Então:

$$v_i(R) > \tau_i(\bar{R}).[\sum_{t \in B_i} \lambda_t - g(R) + \sum_{t \in \bar{A}_i} \lambda_t \cdot v_{J(t,i,\bar{R})}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R) + \sum_{t \in AB_i} \lambda_t - \sum_{t \in BA_i} \lambda_t]$$

Mas $\bar{B}_i = B_i + AB_i - BA_i$, então:

$$v_i(R) > \tau_i(\bar{R}).(\sum_{t \in \bar{B}_i} \lambda_t - g(R) + \sum_{t \in \bar{A}_i} \lambda_t \cdot v_{J(t,i,\bar{R})}(R) + \sum_{t \in E_i} n(i,t) \cdot \mu \cdot v_{K(t,i)}(R))$$

Substituindo-se as expressões 22, 23, 24 e 25 obtém-se:

$$v_i(R) > c_i(\bar{R}) - g(R) \cdot \tau_i(\bar{R}) + \sum_j p_{ij}(\bar{R}) \cdot v_j(R) \quad (27)$$

De acordo com o teorema 1, $g(\bar{R}) \leq g(R)$. Portanto, a política $\bar{R}$ possui uma taxa de custo médio inferior à política R . É possível demonstrar, de forma análoga, que os teoremas 1 e 2 são válidos se as inequações forem invertidas e consequentemente que, se forem escolhidas sucessivamente políticas melhores, a política ótima é alcançada num número finito de passos.

Este teorema mostra que, conhecendo-se os parâmetros de uma cadeia semi-markoviana que represente uma rede telefônica, onde os tráfegos oferecidos são poissonianos e se aplique um dado plano de encaminhamento não ótimo, é possível estabelecer um novo plano de encaminhamento mais eficiente, mesmo sem conhecer nenhum parâmetro da cadeia sob a nova política de encaminhamento. No caso geral, as cadeias semi-markovianas não apresentam esta propriedade. É por esta razão que o método de determinação da política ótima, apresentado na seção 7.3, exige que, na determinação da ação a ser escolhida para um dado estado, sejam conhecidos os parâmetros da cadeia sob a nova política.

Esta propriedade permite estabelecer (ao menos em teoria) um procedimento para otimizar o plano de encaminhamento com um número finito de passos, através de medições de tráfego na rede. Dado que a rede telefônica opera sob uma dada política de encaminhamento, é possível determinar, via medições, o tempo e o custo de permanência nos estados e as probabilidades de transição de estado. Assim é possível calcular os custos relativos e determinar uma nova política, melhor que a anterior.

7.6 Infactibilidade de Cadeias Representando Redes Telefônicas.

Em geral, o número de estados de uma cadeia necessários para representar uma rede telefônica é imenso. O número de estados é tão grande que seria um erro chamá-lo de número astronômico.

Número de estrelas visíveis a olho nú	Entre 3000 e 7000
Maior número em finanças	$496.585.346 \times 10^{12}$
Distância do quasar PKS 2000-300	1.23×10^{20} km
Massa da Lua	735×10^{20} kg
Número de Avogrado	6×10^{23}
Gugol	10^{100}
Maior número para uma calculadora de 12 dígitos	10^{109}
Número possível de jogadas no xadrez	$10^{10^{50}}$
Gugolplexo	$10^{Gugol} = 10^{10^{100}}$
Número de Skewes	$10^{10^{10^{34}}}$

Tabela 10: Grandes Números.

Seja uma rede com N nós (centrais telefônicas) interligados dois a dois por rotas unidirecionais. Em geral, as redes telefônicas possuem entre duas (redes locais de pequenas cidades) e uma centena de centrais. As redes que ultrapassam uma centena de centrais em geral são subdivididas em redes menores. Cada uma das rotas pode possuir até cerca de um milhar de circuitos. Seja uma rede onde as chamadas são estabelecidas exclusivamente por rotas diretas unidirecionais entre as centrais. O número de rotas (R) e de estados (I) na rede pode ser calculado a partir do número de nós da rede (N) e do número de circuitos nas rotas (C):

$$R = N.(N - 1)$$
$$I = \prod_{j=1}^R (C + 1) = (C + 1)^R$$

Uma rede de médio porte, em que todas as rotas contenham nove circuitos cada e a rede possua 50 centrais, todas interligadas duas a duas, pode ser representada, portanto, por uma cadeia de 10^{2450} estados. A tabela 10 apresenta alguns números grandes, que auxiliam a imaginar o que representa tal número de estados [22].

Portanto, em geral não é factível modelar uma rede telefônica como uma única cadeia semi-markoviana, na tentativa de aplicar os métodos clássicos de otimização. Uma alternativa é tratar o problema de forma separada por rotas.

7.7 Modelos Separados por Rota.

Ott e Krishnan [25] apresentaram uma técnica de encaminhamento dependente do estado da rede baseado em decisões de Markov. O processo de decisão consiste em verificar o custo de estabelecimento de cada chamada. Este custo é o número adicional de futuras chamadas que serão bloqueadas, caso a chamada presente seja estabelecida na rede. Este custo é avaliado para vários caminhos possíveis na rede. A chamada é encaminhada pelo caminho de menor custo, desde que este custo seja inferior a uma chamada. Se o custo for superior a este valor, a chamada deve ser bloqueada, permitindo assim que no futuro mais de uma chamada seja estabelecida, em média.

O cálculo do custo de estabelecimento das chamadas pode ser calculado separadamente por rota, se o processo de ocupação for independente para cada rota.

Seja um tráfego poissoniano, com taxas de nascimento e morte das chamadas igual a λ e μ , respectivamente, oferecido a uma rota com C circuitos. No instante em que é oferecida uma chamada à rede, k circuitos estão ocupados ($k < C$). Se a nova chamada for estabelecida, haverá um certo número de chamadas bloqueadas no futuro. Se a chamada não for estabelecida haverá um número menor de chamadas bloqueadas no futuro. A diferença do custo das duas situações é o custo de estabelecimento de uma chamada ($D(k)$):

$$D(k) = Nb_{k+1} - Nb_k \quad (28)$$

Onde Nb_k é o número de chamadas bloqueadas ao longo do tempo, dado que no instante inicial a rota se encontra no estado k .

Embora o número de chamadas bloqueadas Nb_k seja infinito, a diferença de chamadas bloqueadas, dado que se parte de dois estados iniciais distintos ($Nb_{k+1} - Nb_k$), é finita e pode ser calculada modelando-se a rota como uma cadeia semi-markoviana com recompensa.

O processo de ocupação dos circuitos pode ser modelado por uma cadeia semi-markoviana com $C+1$ estados. Cada estado n representa os circuitos ocupados. O custo de permanecer em qualquer estado é nulo, exceto para o estado C onde o custo médio c_C é dado pela taxa de nascimento de chamadas multiplicado pelo tempo médio de permanência no estado C :

$$c_C = \frac{\lambda}{C \cdot \mu}$$

As probabilidades de transição de cada estado são:

$$\begin{aligned} P_{C,C-1} &= 1 \\ P_{0,1} &= 1 \\ P_{i,i-1} &= i \cdot \mu / (i \cdot \mu + \lambda) \quad ; \quad i \neq C \\ P_{i,i+1} &= \lambda / (i \cdot \mu + \lambda) \quad ; \quad i \neq 0 \end{aligned}$$

Os tempos médios de transição são:

$$\begin{aligned} \tau_C &= 1 / C \cdot \mu \\ \tau_i &= 1 / (\lambda + i \cdot \mu) \end{aligned}$$

Portanto o sistema de equações é dado por:

$$\begin{aligned} v_0 &= -\frac{g}{\lambda} + v_1 \\ v_i &= (-g + \lambda \cdot v_{i+1} + i \cdot \mu \cdot v_{i-1}) / (\lambda + i \cdot \mu) \quad ; \quad 0 < i < C \\ v_C &= (\lambda - g) / C \cdot \mu + v_{C-1} \end{aligned}$$

Sabe-se, a partir da teoria de filas, que a taxa de chamadas bloqueadas numa rota é dada pela taxa de nascimento de chamadas multiplicada pela probabilidade de bloqueio na rota, embora seja possível obter este resultado também a partir da resolução do sistema de equações:

$$g = \lambda \cdot \text{Erl}\left(\frac{\lambda}{\mu}, C\right)$$

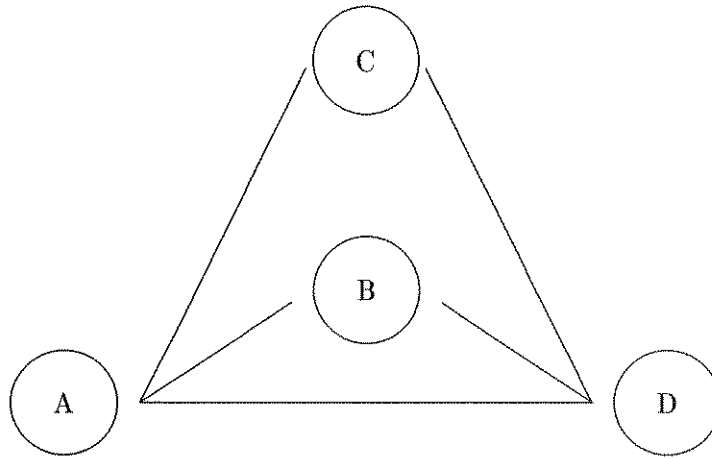


Figura 19: Rede com duas centrais de trânsito. O tráfego com origem em A e destino em D possui três caminhos possíveis. A escolha de qual caminho utilizar depende do algoritmo de encaminhamento dependente do estado da rede.

O custo de estabelecer uma chamada na rota, dado que a rota se encontra com k circuitos ocupados, em termos de futuras chamadas bloqueadas é:

$$D(k) = v_{k+1} - v_k$$

Portanto, a partir do sistema de equações, chega-se a:

$$D(k) = \frac{Erl(\frac{\lambda}{\mu}, C)}{Erl(\frac{\lambda}{\mu}, k)} \quad (29)$$

7.8 Algoritmo de Encaminhamento.

O algoritmo de encaminhamento proposto por Ott e Krishnan consiste em calcular, a cada chamada e para cada caminho, o número de futuras chamadas bloqueadas se a presente chamada for estabelecida no caminho. O caminho escolhido será aquele cuja soma de custos das rotas for mínimo. Se o menor custo for superior a uma chamada, então a chamada não deve ser estabelecida.

Seja o critério de decisão D_i a esperança do número adicional de futuras chamadas bloqueadas no caminho i devido ao estabelecimento da chamada em questão. Para encaminhar uma chamada, cada caminho possível é avaliado segundo este critério e seja D_{min} o menor valor entre todos os caminhos. Se o valor de D_{min} for superior a 1, então a chamada é bloqueada. Caso contrário a chamada é estabelecida no caminho em que o critério é mínimo.

Seja a rede apresentada na figura 19. O tráfego entre as centrais A e D possuem três alternativas possíveis, os caminhos AD, ABD e ACD.

O custo de estabelecer uma chamada em cada um deles é dado pelas expressões:

$$D_{AD}(k_{AD}) = Erl(A_{AD}, C_{AD}) / Erl(A_{AD}, k_{AD})$$

$$D_{ABD}(k_{AB}, k_{BD}) = \text{Erl}(A_{AB}, C_{AB}) / \text{Erl}(A_{AB}, k_{AB}) + \text{Erl}(A_{BD}, C_{BD}) / \text{Erl}(A_{BD}, k_{BD})$$

$$D_{ACD}(k_{AC}, k_{CD}) = \text{Erl}(A_{AC}, C_{AC}) / \text{Erl}(A_{AC}, k_{AC}) + \text{Erl}(A_{CD}, C_{CD}) / \text{Erl}(A_{CD}, k_{CD})$$

$$D = \min\{D_{AD}, D_{ABD}, D_{ACD}\}$$

Onde:

$A_{AB}, A_{AC}, A_{AD}, A_{BD}$ e A_{CD} são os tráfegos oferecidos às rotas.

$k_{AB}, k_{AC}, k_{AD}, k_{BD}$ e k_{CD} são os estados das rotas (número de circuitos ocupados em cada rota) no instante de estabelecimento das chamadas.

Portanto o algoritmo deve estabelecer a chamada no caminho de menor custo, se $D \leq 1$; caso contrário a chamada deve ser bloqueada.

O algoritmo parte da premissa de que o tráfego oferecido às rotas é poissoniano. Esta é uma boa aproximação para redes onde a probabilidade de bloqueio nos caminhos de primeira escolha é baixo e o tráfego de primeira escolha em todas as rotas é majoritário. No caso de rotas diretas, o custo de estabelecimento de uma chamada é sempre inferior à unidade. A equação 29 fornece um resultado sempre no intervalo $[0,1]$. Portanto o algoritmo sempre leva ao estabelecimento de uma chamada num caminho composto por uma rota direta, se houver circuitos livres. Caso contrário um caminho alternativo é escolhido.

7.9 Modelamento de Uma Fila GI/M/C/C.

No modelamento de uma rede telefônica por meio de uma cadeia semi-markoviana, o tráfego e o plano de encaminhamento determinam as probabilidades de transição entre estados, o custo e o tempo de permanência nos estados. Para que o modelo seja válido, é preciso que estes três parâmetros sejam dependentes apenas do estado correspondente; portanto devem ser independentes do modo como foi alcançado aquele estado.

No caso de filas markovianas a condição de independência é satisfeita. Uma vez que o tempo entre geração e o tempo de duração das chamadas são variáveis aleatórias que apresentam distribuição exponencial negativa, a cada transição o processo é renovado, uma vez que estas variáveis são "sem memória". As probabilidades de transição, o custo e o tempo de permanência num estado são calculados diretamente a partir das taxas de geração e término de chamadas.

Entretanto, o potencial de utilização das cadeia semi-markovianas vai além de fila markovianas. Seja uma fila do tipo GI/M/C/C, isto é, um processo de nascimento e morte com $C+1$ estados; o tempo entre nascimentos é uma variável aleatória com distribuição arbitrária independente e o tempo de duração das chamadas é uma variável aleatória com distribuição exponencial negativa.

Quando ocorre uma morte de uma chamada, não há renovação do processo, uma vez que o tempo entre chegadas não é exponencialmente distribuído. Por exemplo, seja o tempo entre nascimentos

uma variável uniformemente distribuída no intervalo $[0, 2\mu]$. Se um dado estado i ($0 < i < C$) é alcançado com o nascimento de uma chamada, o tempo de permanência médio é maior que no caso em que o mesmo estado é alcançado com uma morte. Portanto não é possível representar a fila GI/M/C/C de forma imediata, como é possível para uma fila markoviana. Os parâmetros para um dado estado seriam dependentes da trajetória percorrida pelo processo até atingir o estado, o que viola a definição de cadeias semi-markoviana. Entretanto, é possível estabelecer a seguinte definição abstrata de estado: um estado é caracterizado pelo número de circuitos ocupados imediatamente após o estabelecimento (ou bloqueio) de uma chamada.

Neste caso, os estados da cadeia semi-markoviana não correspondem aos estados de ocupação da fila. Define-se como uma transição da cadeia semi-markoviana o instante do estabelecimento (ou bloqueio) de uma chamada. Uma vez que, por hipótese, o tempo entre nascimento é uma variável independente (GI), a cada transição o processo é renovado. A transição da cadeia semi-markoviana não corresponde a uma transição de estado na fila GI/M/C/C.

Dado que a cadeia se encontra no estado i , as probabilidades de transição estão ligadas ao número de mortes que ocorrem entre dois nascimentos. Se não ocorrer nenhuma morte, o processo passa ao estado $i+1$ (exceto no caso do estado C , que permanece em C mesmo). Se ocorrer uma morte, o processo permanece no estado i ; se ocorrer duas mortes, passa ao estado $i-1$; e assim por diante. Portanto:

$$P_{ij} = \begin{cases} \int_0^\infty Pm_{i-j+1}(t) \cdot p_t(t) \cdot dt & ; j \leq i+1, i < C \\ 0 & ; j > i+1, i < C \\ \int_0^\infty Pm_{C-j+1}(t) \cdot p_t(t) \cdot dt & ; j < C, i = C \\ \int_0^\infty Pm_0(t) \cdot p_t(t) \cdot dt + \\ \int_0^\infty Pm_1(t) \cdot p_t(t) \cdot dt & ; j = C, i = C \end{cases}$$

Onde:

$Pm_k(t)$ é a probabilidade de ocorrer k mortes no intervalo de tempo $[0, t]$.

$p_t(t)$ é a função densidade de probabilidade do tempo entre dois nascimentos.

É importante observar que, neste modelo, há a possibilidade do processo realizar uma "transição" de um estado para o mesmo estado. O tempo entre transições é dado por:

$$\tau_i = \int_0^\infty t \cdot p_t(t) \cdot dt$$

O custo de permanência no estado C é a probabilidade de nascimento de uma chamada antes que ocorra uma morte:

$$c_C = \int_0^\infty Pm_0(t) \cdot p_t(t) \cdot dt$$

7.10 Modelamento Individual das Rotas.

Uma rota que pertence a uma rede onde os tráfegos oferecidos são poisonianos, pode ser modelada, aproximadamente, como uma cadeia semi-markoviana, de forma similar ao caso de uma fila

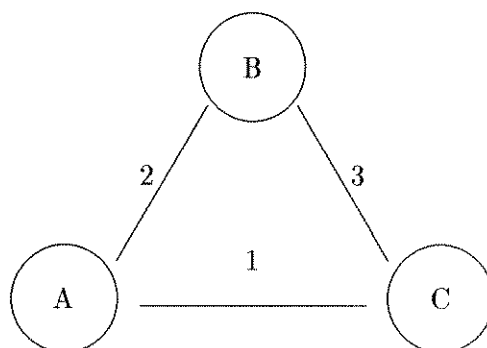


Figura 20: Rede elementar. O tráfego AC utiliza a rota 1 em primeira escolha e as rotas 2 e 3 como caminho de transbordo.

GI/M/C/C. Os estados da cadeia podem ser simplesmente definidos pelo número de circuitos ocupados na rota, ou de forma mais complexa, pelo número de circuitos ocupados e pelo tráfego que originou a chamada que levou a rota aquele estado.

Se o tempo médio de duração das chamadas não for igual para todos os tráfegos oferecidos à rota, então o tempo entre mortes não possui distribuição exponencial negativa como na fila GI/M/C/C. Entretanto, o processo de morte continua a ser um processo "sem memória". No caso particular de uma rede onde o tempo médio de duração das chamadas é igual para todos os tráfegos, a rota é uma fila GI/M/C/C.

O tráfego oferecido à rota é do tipo GI porque depende do estado da rede. Mesmo que a rota permaneça num mesmo estado, outras rotas podem mudar de estado na rede, alterando o tráfego oferecido à rota. Por exemplo, seja um tráfego encaminhado por uma rota direta entre duas centrais. Quando todos os circuitos da rota direta estão ocupados, há uma rota alternativa (rota de proteção). O tráfego oferecido a esta rota alternativa não é poissoniano.

O cálculo da distribuição do tempo entre nascimento de chamadas numa rota pode ser infactível. O processo de nascimento de chamadas é dependente do estado da rede. Portanto é preciso considerar todas as combinações de estados por que a rede pode passar no intervalo de tempo entre duas chamadas oferecidas à rota considerada. Este número de combinações pode ser bastante elevado e o cálculo analítico infactível.

Se for adotada a hipótese de que o processo de ocupação de uma rota é independente das demais rotas, então pode-se calcular o custo de estabelecimento de uma chamada na rede de forma separada, como apresentado na seção 7.8.

7.10.1 Rede Elementar.

Esta seção apresenta um exemplo de modelamento de uma rede, de forma a ilustrar o que foi apresentado anteriormente.

Seja a rede elementar apresentada na figura 20. As rotas 1, 2 e 3 possuem 1 circuito cada.

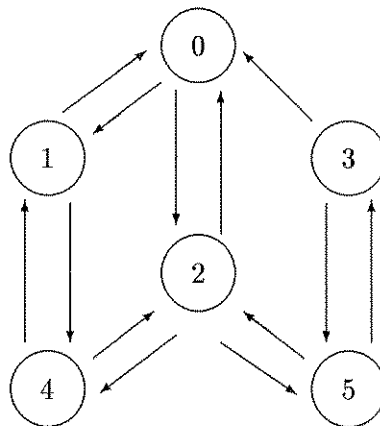


Figura 21: Diagrama dos estados e transições de uma cadeia semi-markoviana que representa a rede elementar.

Estado	0	1	2	3	4	5
Caminho AB	0	1	0	0	1	0
Caminho AC	0	0	1	0	1	1
Caminho ABC	0	0	0	1	0	1

Tabela 11: Mapeamento da ocupação das rotas para cada estado da rede.

O tráfego AC (0.3 Erlangs), utiliza a rota 1 como caminho de primeira escolha e as rotas 2 e 3 como caminho alternativo. O tráfego AB (0.3 Erlangs) utiliza unicamente a rota 2. Os tráfegos oferecidos são poissonianos ($\lambda = 0.3$ e $\mu = 1$).

A seguir são apresentados 3 modelos para esta rede. O primeiro modelo é uma representação da rede como cadeia semi-markoviana com recompensa. O segundo modelo é uma representação separada das rotas, obtida a partir da agregação de estados da cadeia semi-markovianas que representa a rede como um todo. O terceiro modelo é uma representação das rotas como filas tipo GI/M/C/C.

7.10.2 Modelo semi-markoviano.

A figura 21 apresenta os estados da cadeia semi-markoviana que representam a rede elementar da figura 20.

Os estados da cadeia são definidos pelo número de chamadas pertencentes a cada tráfego, estabelecidas nos três caminhos da rede (AB, AC e ABC). Uma transição de estado ocorre quando uma chamada é estabelecida ou terminada.

O custo (que na verdade é uma recompensa) de permanência num dado estado é definido como a receita gerada pelas chamadas estabelecidas. Cada chamada gera uma receita de uma unidade monetária por segundo. As probabilidades dos estados são obtidas resolvendo-se o seguinte sistema de equações:

$$\begin{aligned}
2.\lambda.\pi_0 - \mu.(\pi_1 + \pi_2 + \pi_3) &= 0 \\
(\lambda + \mu).\pi_1 - \lambda.\pi_0 - \mu.\pi_4 &= 0 \\
(2.\lambda + \mu).\pi_2 - \lambda.\pi_0 - \mu.(\pi_4 + \pi_5) &= 0 \\
(\lambda + \mu).\pi_3 - \mu.\pi_5 &= 0 \\
(2.\mu).\pi_4 - \lambda.(\pi_1 + \pi_2) &= 0 \\
\pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 &= 1
\end{aligned} \tag{30}$$

$$\begin{aligned}
\pi_0 &= 0.578 \\
\pi_1 &= 0.171 \\
\pi_2 &= 0.155 \\
\pi_3 &= 0.020 \\
\pi_4 &= 0.049 \\
\pi_5 &= 0.026
\end{aligned} \tag{31}$$

O tempo médio de permanência nos estados é obtido diretamente pelo inverso da soma das taxas de transição de cada estado:

$$\begin{aligned}
t_0 &= 1/2.\lambda = 1.667 \\
t_1 &= 1/(\lambda + \mu) = 0.769 \\
t_2 &= 1/(2.\lambda + \mu) = 0.625 \\
t_3 &= 1/(\lambda + \mu) = 0.769 \\
t_4 &= 1/2.\mu = 0.5 \\
t_5 &= 1/2.\mu = 0.5
\end{aligned} \tag{32}$$

A partir da definição dos elementos da cadeia, pode-se montar o sistema de equações 20 para cálculo dos custos relativos e da taxa de custo média, de acordo com as expressões 23, 24 e 25:

$$\begin{aligned}
v_0 &= 0 \\
v_0 &= t_0.(\lambda.v_1 + \lambda.v_2 - g) = 0 \\
v_1 &= t_1.(1 + \lambda.v_4 + \mu.v_0 - g) = 0.751 \\
v_2 &= t_2.(1 + \lambda.(v_4 + v_5) + \mu.v_0 - g) = 0.907 \\
v_3 &= t_3.(1 + \lambda.v_5 + \mu.v_0 - g) = 0.751 \\
v_4 &= t_4.(2 + \mu.(v_1 + v_2) - g) = 1.580 \\
v_5 &= t_5.(2 + \mu.(v_2 + v_3) - g) = 1.580 \\
g &= 0.497
\end{aligned} \tag{33}$$

Note que a receita média da rede é 0.497 unidades monetárias por segundo. Caso não houvesse bloqueio, a receita seria de 0.6 unidades monetárias por segundo.

7.10.3 Modelo agregado.

No modelo agregado para a rota 1, deseja-se representar o comportamento da rede apenas em termos da ocupação de seus circuitos. Isto significa que há apenas dois estados possíveis: rota

desocupada (0) e rota ocupada pelo tráfego AC (1). O estado 0 da rota 1, agrega três estados da rede (0, 1 e 3). O estado 1 da rota 1 agrega os outros três estados da rede (2, 4 e 5).

Seja o sistema de equações:

$$\begin{aligned}
 v_0 &= 0 \\
 v_0 &= t_0.(\lambda.v_1 + \lambda.v_2 - g) = 0 \\
 v_1 &= t_1.(0 + \lambda.v_4 + \mu.v_0 - g) = 0 \\
 v_2 &= t_2.(1 + \lambda.(v_4 + v_5) + \mu.v_0 - g) = 0.769 \\
 v_3 &= t_3.(0 + \lambda.v_5 + \mu.v_0 - g) = 0 \\
 v_4 &= t_4.(1 + \mu.(v_1 + v_2) - g) = 0.769 \\
 v_5 &= t_5.(1 + \mu.(v_2 + v_3) - g) = 0.769 \\
 g &= 0.231
 \end{aligned} \tag{34}$$

Este sistema é o mesmo sistema da rede como um todo (33), exceto que o custo de permanência em cada estado apenas computa as chamadas estabelecidas na rota 1. Considera-se que chamadas estabelecidas nas rotas 2 e 3 não geram receita.

O custo relativo dos estados da rota 1, é uma média ponderada dos custos relativos dos estados agregados. O peso dos custos relativos de um estado agregado é a probabilidade de que o processo esteja no estado agregado, dado que está no estado 0 (ou 1) da rota 1.

$$\begin{aligned}
 va_0^1 &= (\pi_0.v_0 + \pi_1.v_1 + \pi_3.v_3)/(\pi_0 + \pi_1 + \pi_3) = 0 \\
 va_1^1 &= (\pi_0.v_2 + \pi_1.v_4 + \pi_3.v_5)/(\pi_0 + \pi_1 + \pi_3) = 0.769
 \end{aligned} \tag{35}$$

O modelo agregado para a rota 2 é obtido de forma análoga ao da rota 1. No caso da rota 2 há três estados: rota 2 desocupada (0), rota 2 ocupada pelo tráfego AC (10) e ocupada pelo tráfego AB (01).

$$\begin{aligned}
 v_0 &= 0 \\
 v_0 &= t_0.(\lambda.v_1 + \lambda.v_2 - g) = 0 \\
 v_1 &= t_1.(1 + \lambda.v_4 + \mu.v_0 - g) = 0.751 \\
 v_2 &= t_2.(0 + \lambda.(v_4 + v_5) + \mu.v_0 - g) = 0.138 \\
 v_3 &= t_3.(1 + \lambda.v_5 + \mu.v_0 - g) = 0.751 \\
 v_4 &= t_4.(1 + \mu.(v_1 + v_2) - g) = 0.811 \\
 v_5 &= t_5.(1 + \mu.(v_2 + v_3) - g) = 0.811 \\
 g &= 0.267
 \end{aligned} \tag{36}$$

$$\begin{aligned}
 va_{00}^2 &= (\pi_0.v_0 + \pi_2.v_2)/(\pi_0 + \pi_2) = 0.029 \\
 va_{00}^2 &= 0 \\
 va_{01}^2 &= (\pi_0.v_1 + \pi_2.v_4)/(\pi_0 + \pi_2) - v_{00} = 0.735 \\
 va_{10}^2 &= v_5 - v_{00} = 0.782
 \end{aligned} \tag{37}$$

Como neste caso o custo relativo médio do estado 00 é não nulo (va_{00}^2), é preciso subtrair uma constante de todos os custos para que o custo referencial (va_{00}^2) seja nulo.

7.10.4 Modelo GI/M/C/C.

Neste modelo, as rotas devem ser modeladas de forma similar à descrita na seção 7.10. Neste exemplo, considera-se que o custo (na verdade recompensa) de permanência no estado é a receita gerada pela chamada e considera-se uma transição apenas quando uma chamada é estabelecida (chamadas bloqueadas não caracterizam uma transição. O tráfego oferecido à rota 1 é poissoniano e portanto pode ser modelada como uma fila M/M/1/1. Os tempos de permanência nos estados são dados por:

$$\begin{aligned} t_0 &= 1/\lambda = 3.333 \\ t_1 &= 1/\mu = 1 \end{aligned} \quad (38)$$

Dado que o tráfego oferecido é poissoniano, o sistema de equações para cálculo dos custos relativos é exato:

$$\begin{aligned} vg_0^1 &= 0 \\ vg_0^1 &= -gg1.t_0 + v_1 = 0 \\ vg_1^1 &= (1 - gg1).t_1 + v_0 = 0.769 \\ gg1 &= 0.231 \end{aligned} \quad (39)$$

A rota 2 exige maior trabalho na definição dos parâmetros. Neste caso, um estado é definido como o número de circuitos ocupados por um dado tráfego, imediatamente após o estabelecimento de uma chamada. Existem, portanto, dois estados possíveis: rota 2 ocupada pelo tráfego AC (10) e ocupada pelo tráfego AB (01). O estado 00 não existe a partir desta definição, uma vez que não é possível que, após estabelecer uma chamada, não haja nenhum circuito ocupado. Entretanto, é possível criar o estado 00 de forma artificial. Seja um certo tráfego poissoniano, com taxa de geração de chamadas e tempo médio de retenção infinitesimal. Quando uma chamada deste tráfego é estabelecida, o que ocorre com probabilidade infinitesimal, o processo atinge o estado 00.

O primeiro passo para definição do modelo é obter os tempos e custos médios de permanência nos estados e as probabilidades de transição.

Seja pe_i^k a probabilidade de haver uma transição do estado i , definido no modelo da rede como um todo, para o estado k . Pode-se expressar os valores destas probabilidades por meio dos seguintes sistemas de equações:

$$\begin{aligned} pe_1^{01} &= (\mu.pe_0^{01} + \lambda.pe_4^{01})/(\lambda + \mu) \\ pe_4^{01} &= (\mu.pe_1^{01} + \mu.pe_2^{01})/(2.\mu) \\ pe_2^{01} &= (\lambda + \mu.pe_0^{01})/(2.\lambda + \mu) \\ pe_0^{01} &= (\lambda + \lambda.pe_2^{01})/(2.\lambda) \end{aligned} \quad (40)$$

Estas equações recorrentes foram obtidas da seguinte maneira: seja, por exemplo, a probabilidade pe_2^{01} . Quando o sistema está no estado 2, pode-se atingir o estado 01 de diversas maneiras. A maneira direta consiste no nascimento de uma chamada AB. Porém pode ocorrer a morte de um chamada AC (com o sistema indo para o estado 0) e então pode-se alcançar o estado 01 de

diversas maneiras. A probabilidade pe_2^{01} é a probabilidade de atingir diretamente o estado 01 mais a probabilidade de atingí-lo passando pelo estado 0, que é a probabilidade de atingir o estado 0 multiplicada pela probabilidade de alcançar o estado 01 a partir do estado 0. De forma análoga se procede para os demais estados.

$$\begin{aligned} pe_5^{10} &= (\mu \cdot pe_3^{10} + \mu \cdot pe_2^{10}) / (2 \cdot \mu) \\ pe_3^{10} &= (\mu \cdot pe_0^{10} + \lambda \cdot pe_5^{10}) / (\lambda + \mu) \\ pe_0^{10} &= (\lambda \cdot pe_2^{10}) / (2 \cdot \lambda) \\ pe_2^{10} &= (\lambda + \mu \cdot pe_0^{10}) / (2 \cdot \lambda + \mu) \end{aligned} \quad (41)$$

As probabilidades de transição entre estados são obtidas através de uma média ponderada das probabilidades anteriormente calculadas:

$$\begin{aligned} pg_{01}^{01} &= (\pi_0 \cdot pe_1^{01} + \pi_2 \cdot pe_4^{01}) / (\pi_0 + \pi_2) \\ pg_{01}^{10} &= 1 - pg_{01}^{01} \end{aligned} \quad (42)$$

$$\begin{aligned} pg_{10}^{10} &= pe_5^{10} \\ pg_{10}^{01} &= 1 - pg_{10}^{10} \end{aligned} \quad (43)$$

$$\begin{aligned} pg_{00}^{01} &= (\pi_0 \cdot pe_0^{01} + \pi_2 \cdot pe_2^{01}) / (\pi_0 + \pi_2) \\ pg_{00}^{10} &= (\pi_0 \cdot pe_0^{10} + \pi_2 \cdot pe_2^{10}) / (\pi_0 + \pi_2) \end{aligned} \quad (44)$$

Resultando em:

$$\begin{aligned} pg_{00}^{01} &= 0.835 \\ pg_{00}^{10} &= 0.165 \\ pg_{01}^{01} &= 0.833 \\ pg_{01}^{10} &= 0.167 \\ pg_{10}^{01} &= 0.787 \\ pg_{10}^{10} &= 0.213 \end{aligned} \quad (45)$$

De forma análoga, o tempo e o custo de permanência nos estados é dado pelo seguinte sistema de equações:

$$\begin{aligned} te_1^{01} &= (1 + \mu \cdot te_0^{01} + \lambda \cdot te_4^{01}) / (\lambda + \mu) \\ te_4^{01} &= (1 + \mu \cdot te_2^{01} + \mu \cdot te_1^{01}) / (2 \cdot \mu) \\ te_0^{01} &= (1 + \lambda \cdot te_2^{01}) / (2 \cdot \lambda) \\ te_2^{01} &= (1 + \mu \cdot te_0^{01}) / (2 \cdot \lambda + \mu) \end{aligned} \quad (46)$$

$$\begin{aligned} te_3^{10} &= (1 + \mu \cdot te_0^{10} + \lambda \cdot te_5^{10}) / (\lambda + \mu) \\ te_5^{10} &= (1 + \mu \cdot te_2^{10} + \mu \cdot te_3^{10}) / (2 \cdot \mu) \\ te_0^{10} &= (1 + \lambda \cdot te_2^{10}) / (2 \cdot \lambda) \\ te_2^{10} &= (1 + \mu \cdot te_0^{10}) / (2 \cdot \lambda + \mu) \end{aligned} \quad (47)$$

ESTADO	ROTA 1	ROTA 2	REDE	GI/M/C/C	ERRO
0	0	00	0.000	0.000	0%
1	0	01	0.751	0.735	2%
2	1	00	0.907	0.769	15%
3	0	10	0.751	0.778	4%
4	1	01	1.580	1.504	5%
5	1	10	1.580	1.547	2%

Tabela 12: Esta tabela apresenta para cada estado da rede quais os estados correspondentes nas rotas 1 e 2, o custo calculado para a rede como um todo, a soma dos custos calculados para as rotas separadamente pelo modelo GI/M/C/C e o erro em relação ao modelo da rede como um todo.

$$\begin{aligned}
tg_{01} &= (p_0.te_1^{01} + p_2.te_4^{01})/(p_0 + p_2) = 2.782 \\
tg_{10} &= te_5^{10} = 3.778 \\
tg_{00} &= (p_0.te_0^{10} + p_2.te_2^{10})/(p_0 + p_2) = 6.622
\end{aligned} \tag{48}$$

$$\begin{aligned}
ce_1 &= (1 + \lambda.ce_4)/(\lambda + \mu) \\
ce_4 &= (1 + \mu.ce_1)/(2.\mu) \\
ce_3 &= (1 + \lambda.ce_5)/(\lambda + \mu) \\
ce_5 &= (1 + \mu.ce_3)/(2.\mu)
\end{aligned} \tag{49}$$

$$\begin{aligned}
cg_{01} &= (p_0.ce_1 + p_2.ce_4)/(p_0 + p_2) = 1 \\
cg_{10} &= ce_5 = 1 \\
cg_{00} &= 0
\end{aligned} \tag{50}$$

Obtidos os parâmetros da cadeia semi-markoviana, determinam-se os custos relativos:

$$\begin{aligned}
vg_{00}^2 &= 0 \\
vg_{00}^2 &= c_{00} - gg2.tg_{00} + p_{00}^{01}.vg_{201} + p_{00}^{10}.vg_{210} = 0 \\
vg_{01}^2 &= c_{01} - gg2.tg_{01} + p_{01}^{01}.vg_{201} + p_{01}^{10}.vg_{210} = 0.735 \\
vg_{10}^2 &= c_{10} - gg2.tg_{10} + p_{10}^{01}.vg_{201} + p_{10}^{10}.vg_{210} = 0.778 \\
gg2 &= 0.267
\end{aligned} \tag{51}$$

7.11 Comparação dos Resultados.

A seguir são comparados os resultados obtidos para os três modelos. O modelo para a rede como um todo é exato. No modelo agregado, os custos relativos separados por rota são exatos. No modelo GI/M/C/C são aproximados.

O objetivo dos modelos separados é permitir calcular o custo relativo de um estado da rede a partir dos estados agregados correspondentes. Por exemplo, deseja-se calcular o custo relativo do estado 0, somando-se o custo relativo obtido para o estado 0 da rota 1 e o estado 00 da rota 2. Como o processo de ocupação das rotas não é independente, há um erro cometido tanto no modelo agregado como no modelo GI/M/C/C.

Comparando-se os resultados do modelo GI/M/C/C e do modelo agregado para a rota 2, pode-se concluir que o modelo GI/M/C/C apresenta resultados para os custos separados por rota quase exatos. Isto significa que o modelamento das rotas como filas GI/M/C/C é muito bom. Assim a imprecisão do modelo separado por rotas depende apenas da validade da hipótese de considerar o processo de ocupação de circuitos independente para cada rota.

A tabela 12 apresenta, para cada estado da rede, os custos relativos calculados a partir da cadeia semi-markoviana aplicada à rede como um todo (exato) e os custos calculados como a soma dos custos calculados separadamente, pelo modelo geral. A tabela permite verificar a validade da hipótese de que o processo de ocupação das rotas é aproximadamente independente. Com exceção do estado 2, os erros cometido são bastante reduzidos. Dado que as rotas possuem apenas 1 circuito e que 19% do tráfego oferecido à rota 2 é o transbordo da rota 1, pode-se considerar o exemplo como um estudo de um caso severo.

7.12 Proposta de Encaminhamento Dependente do Estado Separado por Rotas.

A seção anterior demonstra a complexidade de modelar o processo de ocupação de uma rota como uma fila GI/M/C/C a partir do tráfego oferecido à rede. Entretanto este tipo de modelamento permite que se estime o valor dos custos relativos separados por rota a partir de medições da rede ou de simulação. A seguir é apresentado um método fundamentado neste tipo de abordagem.

A partir dos conceitos sobre processos de decisão markoviana, apresentados anteriormente, é possível propor o seguinte procedimento heurístico de encaminhamento dependente do estado:

Seja uma rede telefônica e um plano (política) inicial de encaminhamento ($R=0$). O plano de encaminhamento estabelece um conjunto de um ou mais caminhos possíveis para o estabelecimento de cada chamada e um critério para escolha do caminho. Para cada tráfego, há um caminho de primeira escolha. Se este caminho estiver bloqueado, um caminho alternativo é escolhido, de forma a minimizar o bloqueio de chamadas total na rede.

Se, num dado caminho, uma das rotas estiver com todos os circuitos ocupados ou se tiver sido atingido o limiar de proteção da rota (técnica de reserva de circuitos), para aquela chamada, então a rota e o caminho estão bloqueados. Apenas caminhos não bloqueados podem ser escolhidos para o estabelecimento das chamadas. Uma chamada somente pode ser rejeitada se todos os caminhos previstos no plano de encaminhamento estiverem bloqueados. Sempre que uma chamada for rejeitada pela rede, contabiliza-se um insucesso para cada rota bloqueada, em cada um dos caminhos. Se num dado caminho houver mais de uma rota bloqueada, computa-se um insucesso apenas na rota mais próxima à origem da chamada.

Considere-se cada rota como uma cadeia semi-markoviana, a exemplo de uma fila G/M/C/C. Seja o estado i definido como o estado em que i circuitos estão ocupados na rota, no instante em que uma chamada do tráfego t é estabelecida na rota. O custo associado ao estado i e ao tráfego t é definido como o número médio de insucessos contabilizados na rota, durante a permanência no

estado. A taxa de custo da rota r ($g_r(R)$) é, portanto, o número médio de insucessos computados na rota por unidade de tempo.

Sejam $V_{r,t}(i)$ e $V_{r,t}(i+1)$ os custos relativos correspondentes aos estados i e $i+1$. Portanto, a diferença $D_{r,t}(i) = V_{r,t}(i) - V_{r,t}(i+1)$ representa o número de chamadas bloqueadas a mais quando a rota inicialmente está no estado $i+1$ em relação a quando a rota está no estado i .

Após um determinado período, estabelece-se um novo plano de encaminhamento ($R=1$), baseado nos custos relativos obtidos anteriormente. Para cada chamada, determina-se para os caminhos não bloqueados um custo de estabelecimento da chamada definido por:

$$C_{p,t} = \sum_{r \in R(p)} D_r(i(r), t)$$

Onde:

- p é o caminho considerado.
- $C_{p,t}$ é o custo de estabelecimento da chamada do tráfego t no caminho p .
- $R(p)$ é o conjunto das rotas que pertencem ao caminho p .
- $i(r)$ é o estado da rota r antes do estabelecimento da chamada.
- $D_r(i(r), t)$ é a diferença dos custos relativos para os estados $i(r)+1$ e $i(r)$ quando uma chamada do tráfego t é estabelecida.

O caminho escolhido deve ser aquele que apresentar o menor custo de estabelecimento.

Após um novo período de medições pode-se determinar um novo plano de encaminhamento ($R=R+1$) e o procedimento se repete indefinidamente, até que um critério de parada seja satisfeito.

7.12.1 Estimativa dos Custos Relativos

Para aplicação do procedimento de encaminhamento é necessário obter o valor dos custos relativos para cada estado de cada rota. Estes valores são obtidos a partir de um estimador, calculado a partir de medições de tráfego. Cada rota deve possuir um contador de insucessos. Sempre que uma chamada é estabelecida na rota, o valor do contador é memorizado, assim como o estado da rota imediatamente após o estabelecimento da chamada (número de circuitos ocupados). Ao final de cada chamada computa-se o número de insucessos na rota durante a vida da chamada. O valor médio de insucessos durante uma chamada estabelecida em cada estado consiste no estimador do custo relativo.

Para uma rota com C circuitos e tráfego oferecido poissoniano de intensidade A (taxa de geração de chamadas λ e tempo médio de duração das chamadas $1/\mu$), pode-se calcular o valor da diferença entre custos relativos de dois estados consecutivos ($D_r(i) = V_r(i) - V_r(i+1)$) por meio expressão:

$$D(k) = \text{Erl}\left(\frac{\lambda}{\mu}, C\right) / \text{Erl}\left(\frac{\lambda}{\mu}, k\right) \quad (52)$$

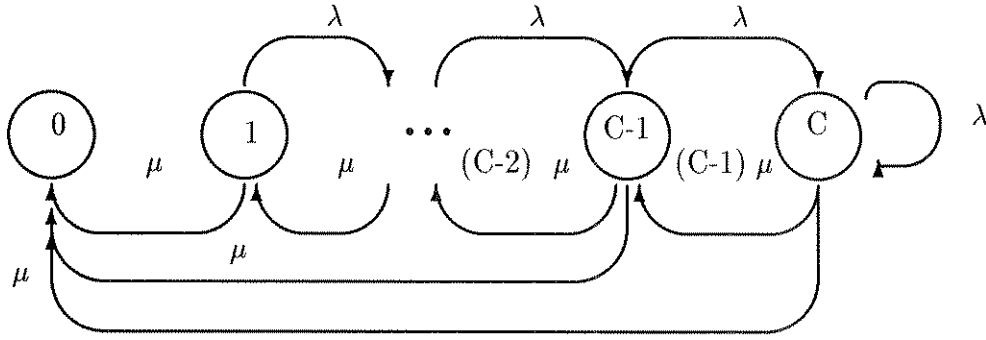


Figura 22: Diagrama de transição de estados de uma cadeia de Markov para calcular o número médio de chamadas bloqueadas durante uma chamada, que é o número de chamadas bloqueadas até que o processo alcance o estado 0. O estado 0 representa o fim da chamada.

Neste caso particular o estimador do custo relativo para cada estado pode ser obtido a partir da solução do seguinte sistema de equações:

$$\begin{aligned} V(C) &= [\lambda.(1 + V(C)) + (C - 1).\mu.V(C - 1)]/[\lambda + C.\mu] \\ V(i) &= [\lambda.V(i + 1) + (i - 1).\mu.V(i - 1)]/[\lambda + i.\mu] ; 1 < i < C \\ V(1) &= \lambda.V(2)/[\lambda + \mu] \end{aligned} \quad (53)$$

Este sistema de equações recorrentes pode ser obtido a partir do diagrama de transição de estados da cadeia de Markov apresentado na figura 22. O estimador é dado pelo número de chamadas bloqueadas até que o processo alcance o estado 0. O estado 0 representa o fim da chamada. As equações podem ser obtidas da seguinte maneira. Suponha-se que inicialmente o sistema está no estado 1 (após o nascimento de uma dada chamada considerada). Podem ocorrer dois eventos: a chamada pode terminar, e o número de chamadas bloqueadas durante sua vida é nulo ou nasce uma nova chamada. Neste caso, o número de chamadas bloqueadas é igual ao número médio de chamadas bloqueadas quando o estado inicial é 2. Fazendo a média ponderada pela probabilidade de ocorrência dos dois eventos, chega-se a equação recorrente para $V(1)$. Analogamente, obtêm-se as equações para os demais estados. Nestes casos existe a probabilidade de nascer uma nova chamada, morrer a chamada considerada ou morrer uma outra chamada qualquer. No estado C , quando nasce uma chamada, esta é bloqueada e o sistema volta ao estado C .

Embora a precisão do estimador não seja acurada, o resultado prático é satisfatório. Se for escolhido um dado estado, verificado o valor da diferença teórica dos custos relativos para aquele estado e encontrada a diferença, baseada nos estimadores, que mais se aproxima do valor teórico, verifica-se que o erro corresponde a dois ou três circuitos (para rotas de até 90 circuitos e bloqueios de até 9,5%). Por exemplo, para uma rota de 30 circuitos e bloqueio de 2%, a diferença teórica de custos para os estados 22 e 23 é 0.133. A diferença baseada nos estimadores que mais se aproxima deste valor corresponde aos estados 24 e 25. Neste caso o erro cometido nos custos, corresponde a um erro de 2 circuitos ocupados.

Além do erro relacionado à definição do estimador, existe o problema estatístico. Para que o estimador seja confiável, é preciso obter um número conveniente de amostras. Uma rota pode

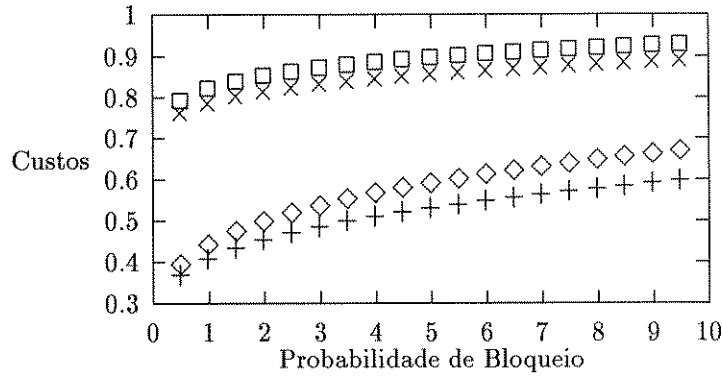


Figura 23: O gráfico apresenta, para os estados C-1 e C (onde C é o número de circuitos), as diferenças de custos relativos exatas (◇) e estimadas (+) para uma rota de 10 circuitos; exatas (□) e estimadas (×) para uma rota com 90 circuitos.

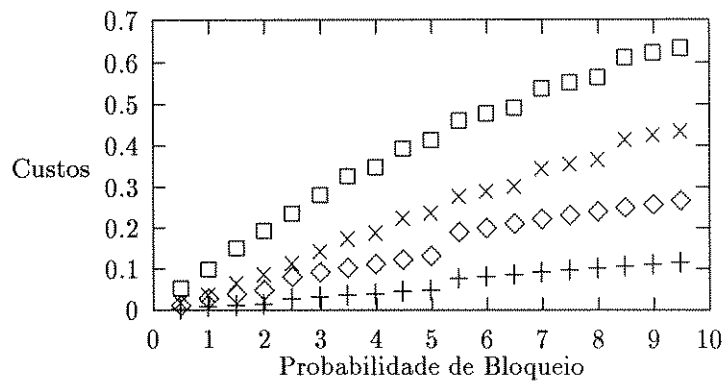


Figura 24: O gráfico apresenta, para os estados M e M+1 (onde M é dado por $M=C.(10-B)$, B é a probabilidade de bloqueio), as diferenças de custos relativos exatas (◇) e estimadas (+) para uma rota de 10 circuitos; exatas (□) e estimadas (×) para uma rota com 90 circuitos.

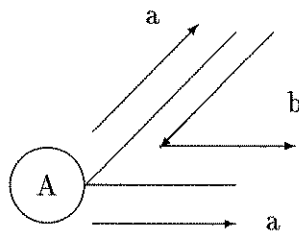


Figura 25: O tráfego originado em A (a) pode escolher entre dois caminhos alternativos. O caminho do tráfego b contém duas rotas que pertencem a estes caminhos alternativos.

apresentar um alto número de estados pouco prováveis. Neste caso pode ser necessário um longo período de medições para que estimadores confiáveis sejam obtidos.

7.12.2 Discussão do Método de Encaminhamento

O método de contagem de insucessos exige mais sofisticação quando são empregadas rotas bidirecionais, para que seja evitada a múltipla contagem de insucessos. Isto ocorre se um determinado caminho utiliza duas rotas que pertencem a dois caminhos alternativos para algum outro tráfego. Seja o exemplo apresentado na figura 25. O tráfego a, originado na central A, possui dois caminhos alternativos. O tráfego b emprega um caminho que inclui as duas rotas pertencentes aos caminhos de a. Quando as duas rotas estão bloqueadas e uma chamada do tráfego a é rejeitada, é atribuído um insucesso à cada rota. Do ponto de vista do tráfego b foram rejeitadas duas chamadas naquele caminho. Desta forma, ocorre uma superestimativa dos custos neste caminho do tráfego b. Portanto a contagem neste tipo de rede não deve ser por rota. Por exemplo, uma solução possível é ter um contador de insucessos para chamadas originadas e outro para cada par de rota de entrada e saída para chamadas trânsito. Outra alternativa é empregar um contador único de chamadas de saída para cada central. Também pode-se considerar implementar o encaminhamento mesmo com esta distorção em alguns casos. O grau de seu efeito depende da topologia, do plano de encaminhamento e do tráfego da rede.

Dentro da metodologia de processos de decisão markovianos, uma chamada deveria ser rejeitada num caminho se a soma das diferenças de custos relativos para todas as rotas do caminho fosse superior a 1, dado que o estabelecimento da chamada causaria a perda de mais de uma chamada no futuro. Uma das características incorporadas a este método de encaminhamento é de que as chamadas apenas são rejeitadas pela técnica de reserva de circuitos. Uma vez que há uma imprecisão do estimador de custos relativos, se o método de otimização empregasse exclusivamente o critério de custo mínimo, poderiam ocorrer consequências indesejáveis. Chamadas que deveriam ser rejeitadas (ou aceitas) seriam aceitas (ou rejeitadas) na rede. Por outro lado chamadas que deveriam ser estabelecidas em caminhos de primeira escolha poderiam ser estabelecidas em caminhos alternativos. A técnica de reserva de circuitos ameniza o efeito destas distorções do encaminhamento.

O encaminhamento dependente do estado procura reduzir o número total de chamadas bloqueadas na rede. Entretanto a solução de mínimo bloqueio pode não ser a melhor solução. Por

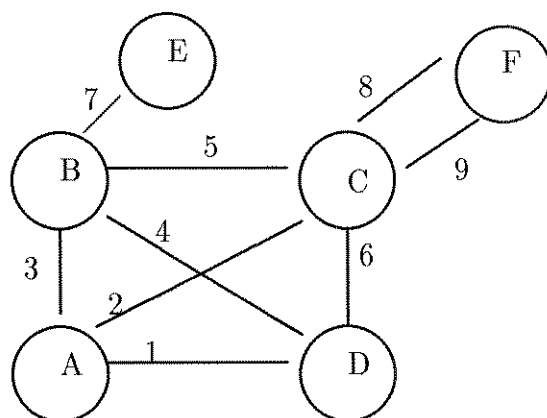


Figura 26: Rede hierárquica de dois níveis.

exemplo, a situação de mínimo bloqueio numa rede pode apresentar uma probabilidade de bloqueio total de 0.5% das chamadas, mas uma probabilidade de 10% para um determinado tráfego. Seria preferível, por exemplo, uma solução que apresentasse 0.7% de bloqueio total e bloqueio inferior a 1% para todos os tráfegos. O mecanismo de reserva de circuitos permite obter esta solução.

O emprego de uma rota de primeira escolha também contribui para uma sensível melhora no desempenho da técnica de encaminhamento proposta. Além de contribuir para reduzir os efeitos das imprecisões também reduz a necessidade de determinar o caminho de custo mínimo a cada chamada. Apenas chamadas que encontram o caminho de primeira escolha bloqueado exigem este procedimento.

De maneira geral, o custo de estabelecimento de uma chamada num dado estado é dependente do tráfego a que pertence a chamada. A probabilidade de bloqueio numa rota pode ser diferente para cada tráfego. Por exemplo, no caso de uma rota que recebe um tráfego de primeira escolha e outro de transbordo, o tráfego de transbordo é mais bloqueado que o tráfego de primeira escolha. O mesmo tipo de efeito pode ocorrer com o estimador. Por outro lado, se for empregado um estimador de custo relativo para cada tráfego, pode ser necessário um grande período de tempo para que se obtenha um valor confiável. Neste caso é mais apropriado que se utilize um único estimador para todos tráfegos, ou alguns estimadores para grupos de tráfegos. Por exemplo, um estimador para o tráfego de primeira escolha e outro para o tráfego de transbordo.

7.12.3 Resultados Experimentais

A seguir são apresentados os resultados de simulações de duas redes, que permitem avaliar o desempenho do encaminhamento dependente do estado.

Foram implementadas duas versões de encaminhamento dependente do estado. Uma utiliza um caminho de primeira escolha (SDA). Na segunda versão a escolha dos caminhos é feita unicamente pelo critério de custo mínimo (SDB).

A rede apresentada na figura 26 é uma rede hierárquica de dois níveis. O número de circuitos nas rotas foi calculado pelos métodos clássicos de dimensionamento [34], [35] de forma a permitir

ROTA	CIRCUITOS	SDA	SDB	FAR
1	21	17.6	17.4	17.4
2	10	5.7	5.6	6.5
3	11	1.9	2.0	1.2
4	56	48.3	48.8	48.0
5	18	2.8	3.1	2.8
6	71	53.1	53.3	53.7
7	66	49.2	49.8	49.6
8	66	49.6	49.6	49.3
9	12	5.0	4.9	4.9

Tabela 13: A tabela apresenta para cada rota o número de circuitos e o tráfego cursado para cada simulação.

TRÁFEGO	CAMINHOS
AD	1, 2-6, 3-4, 3-5-6
AF	2-9, 3-5-9
ED	7-4, 7-5-6
FD	8-6

Tabela 14: A tabela apresenta o plano de encaminhamento hierárquico, para o qual a rede foi dimensionada (FAR). Para cada tráfego são listados, em ordem de preferência, os caminhos para cada tráfego.

um bloqueio ponto-a-ponto inferior a 1% para todos os tráfego. O encaminhamento originalmente utilizado é do tipo transbordo, sem reserva de circuitos (FAR). Os caminhos permitidos para o encaminhamento dependente do estado são os mesmos do encaminhamento original da rede. Também não foi utilizada reserva de circuitos.

As simulações foram subdivididas em 100 subintervalos com 3600 segundos. O tempo médio de duração das chamadas é 240 segundos. A cada subintervalo são recalculados os custos relativos, a partir das medições feitas desde o início da simulação. O encaminhamento inicial é um encaminhamento arbitrário, semelhante ao plano de encaminhamento fixo.

No caso de encaminhamento dinâmico, as redes são simuladas duas vezes. Na primeira vez, a cada subintervalo os custos relativos são recalculados e, conseqüentemente, o plano de encaminhamento alterado. Na segunda simulação os custos são fixos, de acordo com os valores finais obtidos na primeira simulação. Na primeira simulação, portanto, calcula-se os custos relativos para obter o encaminhamento dinâmico e na segunda simulação verifica-se o desempenho do plano de

ORIG-DEST	TRÁFEGO	FAR	SDA	SDB
AD	50	0.12	0.08	0.03
AF	50	0.31	0.63	0.35
ED	20	0.47	0.58	0.51
FD	5	0.71	0.65	0.67
TOTAL	125	0.50	0.53	0.49

Tabela 15: A tabela apresenta os tráfegos oferecidos e o bloqueio ponto-a-ponto (%).

ROTA	CIRCUITOS	TRÁFEGO OFERECIDO
1-2	36	41.2
1-3	24	10.5
1-4	324	326.7 (286)
1-5	48	30.7
1-6	48	43.7
2-3	96	37.7
2-4	96	102.5
2-5	108	115.2
2-6	96	123.8
3-4	12	17.9
3-5	48	10.3
3-6	24	19.9
4-5	192 (72)	119.1
4-6	84	124.5 (74.5)
5-6	336	190.7

Tabela 16: A tabela apresenta o número de circuitos para cada rota e o tráfego oferecido entre cada par de centrais da rede. Existem três cenários de simulação. Um cenário de sobrecarga (S1), um cenário onde simula-se uma pane na rota 4-5, que passa de 192 para 72 circuitos (S2), e um cenário onde os tráfegos 1-4 e 4-6 são menores que na condição de sobrecarga (S3).

encaminhamento obtido.

A segunda rede simulada é uma rede não hierárquica com seis centrais, todas conectadas duas a duas por rotas bidirecionais. O encaminhamento originalmente utilizado é do tipo transbordo, com emprego das técnicas de "crankback" e reserva de circuitos. Originalmente, o número de caminhos alternativos permitidos era limitado a no máximo 2 caminhos (FAR-2). Para fins de comparação, a rede foi simulada também com um encaminhamento fixo sem limitações para transbordo (FAR-4). Esta rede foi apresentada por Mitra [30] num estudo comparativo de técnicas de encaminhamento. Todas as simulações reservam 5 circuitos para o tráfego de rota direta.

As simulações desta rede utilizam um contador de insucessos por rota (portanto sujeito à múltipla contagem de insucessos), e um único estimador de custo para todos os tráfegos em cada rota.

Foram feitas simulações utilizando dois planos distintos de encaminhamento dependente do estado. Um plano permite que as chamadas sejam estabelecidas em apenas alguns caminhos, os mesmos utilizados com o plano de encaminhamento FAR2 (SDA2 e SDB2). O outro plano permite estabelecer chamadas em qualquer caminho (SDA4 e SDB4), a exemplo do plano FAR4

Esta rede permite demonstrar o potencial da técnica de encaminhamento proposta. Os tráfegos da rede estão fora da condição nominal, para a qual as rotas foram dimensionadas. A rede foi simulada em três condições: rede em sobrecarga, tal como apresentado por Mitra, com pequenas mudanças no tráfego oferecido (S1), rede com pane, onde foi simulada uma perda do número de circuitos na rota entre as centrais 4 e 5 (S2), e a mesma rede de S1 com alguns tráfegos reduzidos (S3).

As simulações foram executadas em 100 subintervalos de 4800 segundos cada. O tempo médio

TRÁFEGO	CAMINHOS ALTERNATIVOS
1-2	3 - 4 - (5 - 6)
1-3	5 - 2 - (5 - 6)
1-4	(2 - 3 - 5 - 6)
1-5	3 - 6 - (2 - 4)
1-6	5 - (2 - 3 - 4)
2-3	(1 - 4 - 5 - 6)
2-4	3 - (1 - 5 - 6)
2-5	3 - 6 - (1 - 4)
2-6	3 - (1 - 4 - 5)
3-4	1 - 2 - (5 - 6)
3-5	6 - (1 - 2 - 4)
3-6	2 - 5 - (1 - 4)
4-5	1 - 6 - (2 - 3)
4-6	5 - 1 - (2 - 3)
5-6	(1 - 2 - 3 - 4)

Tabela 17: A tabela apresenta a sequência de escolha dos caminhos para os encaminhamentos fixos FAR2 (FAR4). As simulações foram realizadas com dois tipos de sequências. Uma que permite apenas dois caminhos alternativos no máximo e outra que permite sempre quatro caminhos. Por exemplo, uma chamada entre as centrais 1 e 6, no encaminhamento FAR2 tenta inicialmente a rota direta 1-6. Se ela estiver congestionada, então é tentado o caminho 1-5-6. Se congestionado, a chamada é bloqueada. Já no encaminhamento FAR4 tenta-se ainda os demais caminhos (pela ordem 1-2-6, 1-3-6, 1-4-6). As mesmas seqüências são utilizadas inicialmente para o encaminhamento dependente do estado, quando ainda não há custos estimados.

SIMULAÇÃO	FAR2	FAR4	SDA2	SDA4	SDB2	SDB4
S1	5.0	4.2	5.2	4.3	5.3	4.8
S2	11.6	12.1	11.3	11.8	11.5	11.9
S3	3.7	3.2	4.1	3.1	4.1	3.1

Tabela 18: Resultados das simulações para a rede não hierárquica. A tabela apresenta a probabilidade de bloqueio total da rede (%).

SIMULAÇÃO	SDA2	SDA4	SDB2	SDB4
S2	11.6	11.7	11.5	11.8
S3	4.0	3.2	4.0	3.1

Tabela 19: Resultados das simulações para a rede não hierárquica, com os custos calculados para S1. A tabela apresenta a probabilidade de bloqueio total da rede (%).

Estimativas de Custos.

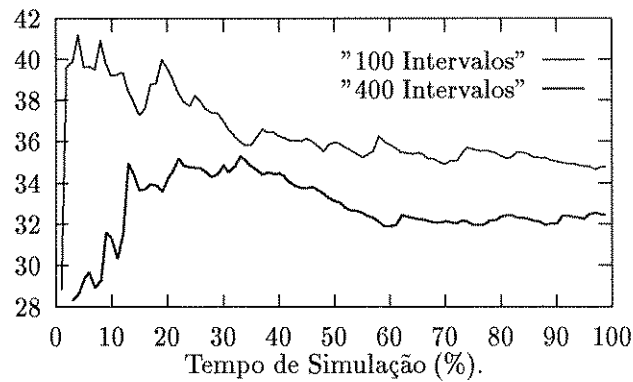


Figura 27: A figura apresenta os valores dos custos relativos do estado 324 circuitos ocupados da rota 1-4, isto é, o número médio de chamadas bloqueadas na rota 1-4 durante a vida de uma chamada que, no instante de nascimento, levou a rota para o estado de 324 circuitos ocupados.

Estimativas de Custos.

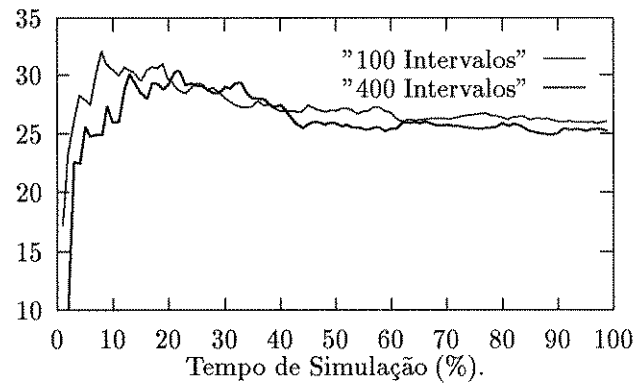


Figura 28: A figura apresenta os valores dos custos relativos do estado 314 circuitos ocupados da rota 1-4.

Estimativas de Custos.

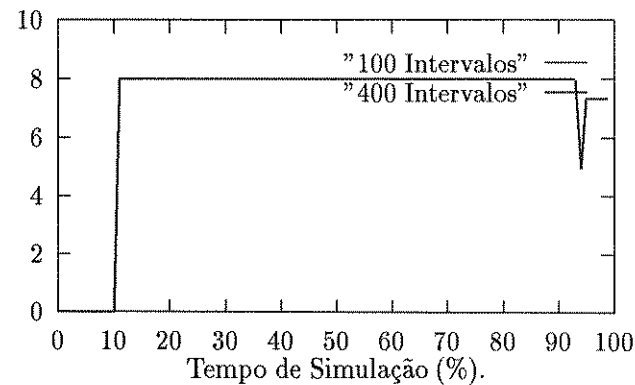


Figura 29: A figura apresenta os valores dos custos relativos do estado 96 circuitos ocupados da rota 2-3.

Estimativas de Custos.

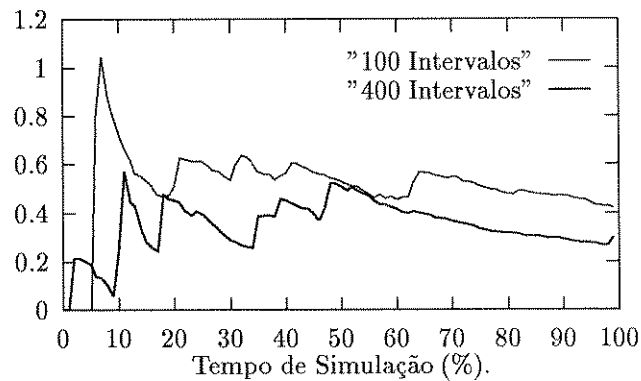


Figura 30: A figura apresenta os valores dos custos relativos do estado 71 circuitos ocupados da rota 2-3.

de retenção das chamadas é de 240 segundos. A cada subintervalo são recalculados os custos relativos, a partir das medições feitas desde o início da simulação. O encaminhamento inicial é um encaminhamento arbitrário, semelhante ao plano de encaminhamento fixo.

No caso de encaminhamento dinâmico, as redes são simuladas duas vezes. Na primeira vez, a cada subintervalo os custos relativos são recalculados e, conseqüentemente o plano de encaminhamento alterado. Na segunda simulação os custos são fixos, de acordo com os valores finais obtidos na primeira simulação.

As simulações mostram que, como já havia observado Mitra, o encaminhamento fixo com "crank-back" e reserva de circuitos é bastante robusto quando são empregadas um ou dois caminhos alternativos. Quando o número de caminhos alternativos cresce seu desempenho pode ser comprometido em algumas situações. O encaminhamento dependente do estado se aproximou do desempenho do encaminhamento fixo, mostrando-se competitivo, particularmente no tratamento de encaminhamento com grande número de caminhos alternativos. A tabela 19 apresenta os resultados de simulações para a rede nas condições S2 e S3, utilizando-se o plano de encaminhamento obtido para a condição S1. Os resultados mostram que o encaminhamento dependente do estado obtido é bastante robusto no caso de panes ou variação de tráfego.

Para cada rota e para cada estado da rota á um estimador de custo relativo. Escolheu-se quatro estimadores representativos, para que fossem coletados seus valores ao longo das simulações. As figuras 27, 28, 29 e 30 apresentam a evolução dos custos relativos ao longo da simulação para duas rotas (1-4 e 2-3). Para cada rota escolheu-se um estado muito provável e o estado em que todos os circuitos da rota estão ocupados. A rota 1-4 é uma rota muito congestionada (314 Erlangs de tráfego escoado em 324 circuitos) com um fator de pico de 0.2 (variância dividida pela média média do tráfego escoado). A rota 2-3 é uma rota descongestionada (74 Erlangs para 96 circuitos) com um fator de pico de 0.8. Neste caso pode-se perceber que há um pequeno número de amostras. Os gráficos comparam a evolução dos custos para duas simulações com o mesmo tempo de duração. Uma é dividida em 100 subintervalos e outra em 400 subintervalos.

7.13 Conclusão

Este capítulo apresentou um novo método para determinação do encaminhamento dependente do estado para uma dada rede. O método é uma evolução do encaminhamento dependente do estado, proposto por Ott e Krishna. A metodologia aplicada resultou num método de determinação do encaminhamento ótimo para uma dada rede, aplicando-se a teoria de processos de decisão markovianos.

A forma proposta de implementação do método de encaminhamento mostrou-se competitiva, em termos de minimização do bloqueio da rede, em relação ao encaminhamento fixo, com reserva de circuitos e "crackback".

O método de otimização do encaminhamento apresenta algum potencial de melhoria de desempenho. É possível melhorar a precisão do estimador, de forma a obter redes com probabilidade de bloqueio ainda inferiores, relaxando-se os mecanismos de reserva de circuitos e caminho de primeira escolha.

Potencialmente, o método proposto poderia ser aplicado diretamente como forma de encaminhamento adaptativo, de forma integrada à rede. As centrais poderiam realizar as medições necessárias e periodicamente alterar seus custos relativos, conforme as variações de tráfego. Entretanto o mecanismo de controle parece muito complicado e a convergência do algoritmo mostrou-se demasiadamente lenta. Para a implementação de um encaminhamento adaptativo, é necessário que as centrais façam medições de tráfego, determinando as taxas de nascimento e morte. Periodicamente, estes dados podem ser coletados, a rede simulada com auxílio do método proposto, e o novo encaminhamento pode ser implementado, passando-se os custos relativos para as centrais da rede.

O resultados obtidos mostraram que o método pode ser aplicado como uma excelente ferramenta computacional para planejamento e operação. Por exemplo, através da simulação de uma rede, pode-se determinar o encaminhamento dependente do estado fixo ótimo para a rede.

Capítulo 8

Formulação Tipo Rentabilidade.

O vencedor é aquele que permanecer no jogo por último sem falir ou aquele que tiver acumulado o maior capital ao final de um período pré-estabelecido.

Principal regra do jogo “Monopólio”.

Este capítulo aborda o problema de dimensionamento do entroncamento de uma rede telefônica utilizando o princípio de Moe [18]. Esta abordagem emprega um critério de lucratividade para dimensionar os equipamentos de uma rede telefônica. Esta metodologia não tem sido largamente empregada, principalmente devido a dois fatores: o grau de serviço da rede dimensionada não é garantido e o método exige o cálculo do bloqueio ponto-a-ponto. Por outro lado, o método apresenta uma simplicidade do algoritmo de resolução e permite tratar de forma simples o dimensionamento multi-horário.

8.1 Dimensionamento clássico.

Na abordagem clássica, o problema de entroncamento é formulado da seguinte forma (supondo-se, sem perda de generalidade, que o custo de implantação de circuitos é o mesmo para todas as rotas):

Minimizar o número de circuitos utilizados na rede de forma que a probabilidade de bloqueio dos tráfegos satisfaça um determinado grau de serviço. São dados o plano de encaminhamento e o vetor de tráfego oferecido ponto-a-ponto.

O grau de serviço pode ser especificado de maneiras diversas. A forma mais simples (e de acordo com o plano de encaminhamento da rede nacional, apresentado anteriormente) é estabelecer uma probabilidade de bloqueio máxima para cada rota. As rotas são classificadas em rotas de alto uso e rotas finais. Rotas de alto uso podem ter uma probabilidade de bloqueio típicas de até 10%, e rotas finais de até 1%. Esta formulação pode ser descrita matematicamente como:

$$\text{Min}_N \quad E.N \quad (54)$$

Sujeito a:

$$\mathcal{B}(N, A) \leq B_{MAX}$$

Onde:

E é um vetor linha unitário de dimensão igual ao número de rotas da rede.

$N = [N_1, N_2, \dots]'$ é o vetor coluna do número de circuitos nas rotas.

A é o vetor tráfego oferecido ponto-a-ponto.

$\mathcal{B}(N, A) = [B_1(A, N), \dots]'$ é o vetor cujos elementos são as probabilidades de bloqueio das rotas. Cada probabilidade de bloqueio é dada por uma função do número de circuitos nas rotas, do vetor de tráfego oferecido e do plano de encaminhamento da rede.

B_{MAX} é o vetor das probabilidades de bloqueio máximas admitidas para cada rota.

O número de circuitos numa rota é um número inteiro. Entretanto, nos problemas de otimização de entroncamento, usualmente considera-se o número de circuitos um número real. Ao final do processo de otimização o número de circuitos é arredondado para um número inteiro.

Este tipo de formulação apresenta características que tornam o problema de difícil solução. De maneira geral, trata-se de uma minimização não convexa e com restrições. Este tipo de problema exige algoritmos de resolução aproximados ou que exigem grande esforço computacional. Esta formulação também torna penosa a abordagem do problema de planejamento multi-horário. Cada novo horário considerado introduz um novo conjunto de restrições, o que torna a solução do problema mais difícil.

8.2 Rentabilidade.

Quando uma empresa de telecomunicações instala um tronco numa rede telefônica (ou qualquer outro equipamento), ela está investindo um certo capital que deve ser recuperado, com lucro, ao longo do tempo de vida útil do tronco.

Conhecidos o tempo médio de vida útil e o custo de um tronco, pode-se calcular uma taxa de custo, ou seja, quanto custa por unidade de tempo a implantação de um tronco na rede. Também pode-se calcular a taxa de receita gerada pelo tronco, ou seja, a receita gerada por unidade de tempo. Define-se a taxa de rentabilidade (R) como a taxa de receita menos a taxa de custo. Quanto maior a taxa de rentabilidade dos equipamentos, mais a empresa será lucrativa.

Deste ponto de vista, pode-se colocar o problema de dimensionamento da seguinte forma:

Determinar o número de circuitos em cada rota da rede, maximizando a rentabilidade global da rede.

$$Max_N \quad R(A, N) \tag{55}$$

Onde:

R é o vetor taxa de rentabilidade da rede.

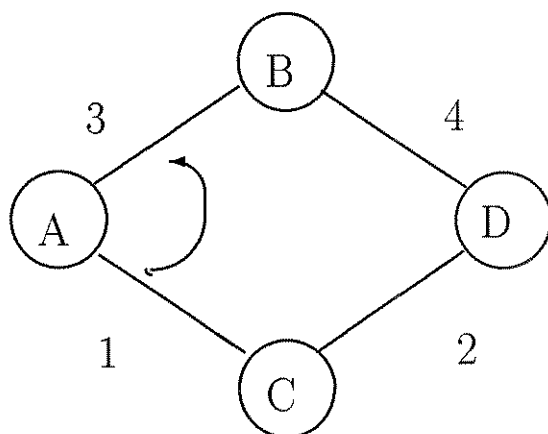


Figura 31: Rede exemplo para discussão da função rentabilidade.

Este problema é do tipo irrestrito. Se a taxa de rentabilidade, em função do número de circuitos nas rotas, for pseudo-côncava então o problema pode ser resolvido muito eficientemente por um algoritmo de subgradiente.

Por exemplo, seja a rede apresentada na figura 31. Suponha que exista um único tráfego AD oferecido á rede. Existem dois caminhos possíveis para este tráfego. O caminho ACD, composto pelas rotas 1 e 2, e um caminho alternativo ABD, composto pelas rotas 3 e 4. Supondo-se o número de circuitos nas rotas 2, 3 e 4 constantes, assim como a intensidade do tráfego AD, pode-se obter um gráfico unidimensional da rentabilidade da rede em função do número de circuitos na rota 1, apresentado na figura 32. Quando a rota 1 não existe, o tráfego escoar pelo caminho ABD, gerando uma certa receita. A rentabilidade é dada pela receita menos o custo dos troncos implantados nas rotas 2, 3 e 4. A medida que se instalam circuitos na rota 1, o tráfego escoado na rede aumenta, e o custo da rede aumenta. Entretanto a receita adicional gerada supera o custo de instalação de troncos na rota 1. Quando o número de circuitos na rota 1 atinge um valor superior ao número de circuitos na rota 2, o tráfego cursado para de aumentar em função do aumento de troncos na rota 1. De fato, no caso do exemplo, quando há bloqueio na rota 2, não há transbordo para o caminho ABD, provocando uma diminuição abrupta do tráfego escoado na rede e conseqüente redução de receita. Isto se deve ao fato que as chamadas passam a ser bloqueadas na rota 2 em vez da rota 1. Desta forma não há mais tráfego escoado no caminho ABD. A partir deste ponto a rentabilidade cai, devido ao aumento linear do custo de circuitos instalados na rota 1.

8.3 Dimensionamento de uma Rota.

Seja, por exemplo, o problema de dimensionamento de uma rota com um tráfego poissoniano oferecido de intensidade A . São dados do problema a taxa de receita das chamadas (1 unidade monetária cobrada por cada segundo de duração das chamadas) e a taxa de custo de implantação de um circuito na rota (15 centavos de unidade monetária por segundo).

A taxa de rentabilidade é dada pela seguinte expressão:

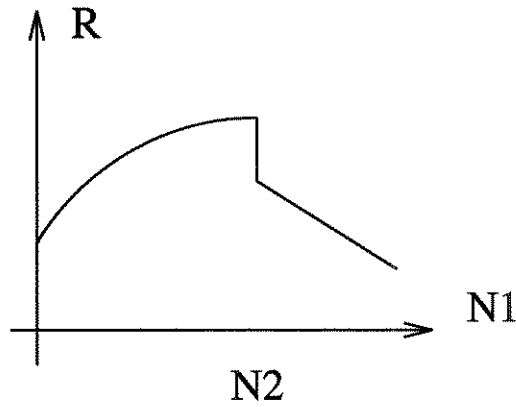


Figura 32: Gráfico da rentabilidade da rede em função do número de circuitos na rota 1 da rede exemplo.

$$R = T.x(A, N) - Cr.N \quad (56)$$

Onde:

T é a taxa de receita.

$x(A, N)$ é o tráfego cursado na rota.

Cr é a taxa de custo por circuito.

N é o número de circuitos na rota.

A condição suficiente para que N seja ótimo é dada por:

$$\frac{d(T.x(A, N) - Cr.N)}{dN} = 0 \quad (57)$$

Portanto, supondo-se que o custo é constante em função do número de circuitos:

$$\frac{dx}{dN} = \frac{Cr}{T} \quad (58)$$

Se o número de circuitos é uma variável inteira não é possível falar em diferenciação. Neste caso é possível estabelecer uma condição de otimalidade análoga. Seja a função $\frac{\Delta x}{\Delta N}(A, N)$ definida como sendo a diferença do tráfego cursado ($x_{N+1} - x_N$) na rota quando o número de circuitos é $N+1$ e N . Supondo-se que a rota é dividida num grupo de circuitos e um circuito, de forma que este circuito (último circuito) é ocupado somente se o grupo está com todos circuitos ocupados, então esta função fornece o tráfego cursado no último circuito. A rentabilidade é máxima se, e somente se, as seguintes equações são satisfeitas:

$$\frac{\Delta x^+}{\Delta N}(N) \leq \frac{Cr}{T} \quad (59)$$

$$\frac{\Delta x^-}{\Delta N}(N) \geq \frac{Cr}{T} \quad (60)$$

Onde:

$\frac{\Delta x^+}{\Delta N}$ é o acréscimo de tráfego na rota quando o número de circuitos passa de N para $N + 1$.

$\frac{\Delta x}{\Delta N}$ é o acréscimo de tráfego na rota quando o número de circuitos passa de $N - 1$ para N .

A função $\frac{\Delta x}{\Delta N}$ para uma rota que recebe um tráfego oferecido poissoniano é dada pela expressão:

$$\frac{\Delta x}{\Delta N} = a.[Erl(a, N) - Erl(a, N + 1)] \quad (61)$$

Onde $Erl(a, N)$ é a fórmula B de Erlang.

Pode-se interpretar esta condição da seguinte maneira: para cada circuito adicionado à rota, obtém-se um aumento do tráfego cursado e, conseqüentemente, da receita. Também há um aumento de despesa, uma vez que o circuito adicional representa um custo adicional. Se o número de circuitos na rota for inferior ao número ótimo, o aumento da receita é superior ao aumento do custo de implantação de mais um circuito. Portanto, existe um tráfego oferecido à rota que não é atendido, conseqüentemente há um lucro potencial sendo desprezado.

Se o número de circuitos for superior ao número ótimo então a receita adicional não compensa o custo adicional. Neste caso, os circuitos acima do ótimo são deficitários e portanto se forem eliminados, há um aumento da rentabilidade.

Quando o ponto ótimo é alcançado, um aumento de circuitos leva a admitir parcelas de tráfego não rentáveis e uma diminuição de circuitos leva a desprezar parcela do lucro potencial.

8.4 Dimensionamento de Rede.

No caso do dimensionamento de uma rede, a metodologia é similar. O problema de dimensionamento pode ser formulado da seguinte maneira:

$$Min_N \{T.x - Cr.N\} \quad (62)$$

Onde

T é o vetor de taxa de receita para cada tráfego oferecido.

x é vetor tráfego ponto-a-ponto cursado na rede.

Cr é o vetor taxa de custo por circuito nas rotas.

N é o vetor número de circuitos nas rotas.

Adotando-se por hipótese que a função tráfego escoado ponto-a-ponto é pseudo-côncava em relação ao número de circuitos nas rotas, a condição suficiente de otimalidade é:

$$\nabla R(A, N) = \nabla \{T.x - Cr.N\} = 0 \quad (63)$$

Ou seja, para uma dada rota i :

$$\frac{\partial \{T.x - Cr.N\}}{\partial N_i} = 0 \quad (64)$$

$$\frac{\partial T.x}{\partial N_i} = Cr_i \quad (65)$$

Portanto:

$$\sum_{j=1}^{\beta} \frac{\partial t_j \cdot x_j}{\partial N_i} = Cr_i \quad (66)$$

Onde β é o número de tráfegos oferecidos à rede.

No caso de dimensionamento multi-horário, existe uma derivada parcial do tráfego cursado em relação ao número de circuitos nas rotas diferente para cada horário, uma vez que pode haver um plano de encaminhamento diferente para cada horário. Dado que a rentabilidade da rede é a média da rentabilidade em cada horário, ponderada pela sua duração, a derivada da função rentabilidade será a média ponderada das derivadas de rentabilidade nos diversos horários.

Sem perda de generalidade, considere-se que a rentabilidade de todos os tráfegos são iguais a uma unidade. Pode-se definir uma função $\frac{\Delta x}{\Delta N_i}(N)$ definida como sendo a diferença da receita do tráfego cursado na rede quando o número de circuitos da rota i é $N+1$ e N , mantido constante o número de circuitos nas demais rotas. De posse de um modelo para cálculo de bloqueio (por exemplo, o modelo de Erlang invertido ou o modelo de simulação) é possível calcular, dado o número de circuitos em cada rota, o tráfego escoado em toda a rede. Portanto pode-se calcular o valor da função $\frac{\Delta x}{\Delta N_i}(N)$. Então o problema pode ser resolvido por meio do seguinte algoritmo:

1. Inicializa-se o vetor N arbitrariamente.

2. Para cada rota i calcula-se o valor de $\frac{\Delta x}{\Delta N_i}^+(N)$ e $\frac{\Delta x}{\Delta N_i}^-(N)$.

3. Calcula-se o vetor Gr da seguinte maneira:

Se $\frac{\Delta x}{\Delta N_i}^-(N) \geq Cr_i$ e $\frac{\Delta x}{\Delta N_i}^+(N) \leq Cr_i$ então $Gr_i = 0$

Se $\frac{\Delta x}{\Delta N_i}^-(N) \geq Cr_i$ e $\frac{\Delta x}{\Delta N_i}^+(N) \geq Cr_i$ então $Gr_i = +1$

Se $\frac{\Delta x}{\Delta N_i}^-(N) \leq Cr_i$ e $\frac{\Delta x}{\Delta N_i}^+(N) \leq Cr_i$ então $Gr_i = -1$

4. Se $Gr=0$ então N é ótimo. Caso contrário, atualiza-se o número de circuitos (N) nas rotas para $N + Gr$ e volta-se ao passo 2.

8.5 Garantia do Grau de Serviço.

O principal questão desta abordagem é que não há uma garantia de que o grau de serviço será satisfatório para todos os tráfegos. Existe, inclusive a possibilidade de que tráfegos muito deficitários simplesmente não sejam atendidos. O grau de serviço dependerá da rentabilidade de cada tráfego. Na formulação clássica, os tráfegos mais rentáveis podem subsidiar os tráfegos menos rentáveis.

Na formulação tipo rentabilidade, também é possível introduzir um mecanismo de subsídio, se a rentabilidade dos tráfegos for artificialmente modificada. Isto exige que se considere uma taxa de receita diferenciada para cada tráfego. Este mecanismo não será abordado neste trabalho.

Pequenas variações de bloqueio, causadas por pequenas variações no número de circuitos, produzem grandes variações de rentabilidade. Por exemplo, seja uma rota com 30 circuitos e um

N	N+1	$\frac{\Delta x}{\Delta N}$	N	N+1	$\frac{\Delta x}{\Delta N}$
520	521	0.383	30	31	0.136
515	516	0.472	29	30	0.177
510	511	0.509	28	29	0.225

Tabela 20: Comparação da rentabilidade de duas rotas, com tráfegos oferecidos de 500 e 22 Erlangs respectivamente, com pequenas variações do número de circuitos nas rotas. Pequenas variações no número de circuitos provocam grandes alterações na rentabilidade das rotas.

tráfego oferecido de 22 Erlangs e uma probabilidade de bloqueio de cerca de 2%. Se forem retirados dois circuitos, a probabilidade de bloqueio aumenta para 4%, e a função $\frac{\Delta x}{\Delta N}$ passa de 0.136 para 0.225, o que representa um aumento de 65% na rentabilidade do último circuito.

A seguir é feita uma estimativa pessimista da taxa de custo e da taxa de receita para um circuito. Atualmente um impulso tarifário custa 3 centavos de dolar aproximadamente. Considerando-se que em média as chamadas são tarifadas por um pulso a cada 180 segundos de duração, que a receita se concentra num período de 10 horas por dia, 5 dias por semana. Supondo-se que 10% da receita bruta é utilizada para amortizar o investimento num período de 10 anos.

Portanto, a taxa de receita (por segundo) por Erlang cursado na rede é de:

$$T = \frac{0.03 \cdot 10 \cdot 5}{180 \cdot 24 \cdot 7 \cdot 10}$$

O custo de implantação de um tronco numa central é da ordem de 300 dólares por tronco. Portanto a taxa de custo de um tronco ao longo de dez anos é de:

$$Cr = \frac{300}{10 \cdot 365 \cdot 24 \cdot 3600}$$

Portanto:

$$\frac{Cr}{T} = 0.19$$

Este valor corresponde a uma probabilidade de bloqueio de cerca de 3% para um tráfego de 22 Erlangs cursado numa rota direta com 29 circuitos. Esta rentabilidade é obtida, por exemplo, pelo último circuito de uma rota com 29 circuitos, para um tráfego oferecido de 22 Erlangs, o que resulta num bloqueio de cerca de 3% das chamadas.

Ressalte-se que esta é uma estimativa pessimista. Espera-se que a receita seja bastante superior a esta na realidade. Portanto, as redes dimensionadas por este método tenderão a apresentar baixas probabilidades de bloqueio.

8.6 Método Heurístico.

Desenvolveu-se um método heurístico, baseado em simulação da rede, para dimensionamento de redes telefônicas. A simulação é dividida em subintervalos. Ao final de cada sub-intervalo é feita uma estimativa das derivadas parciais do tráfego cursado na rede, em relação ao número de circuitos

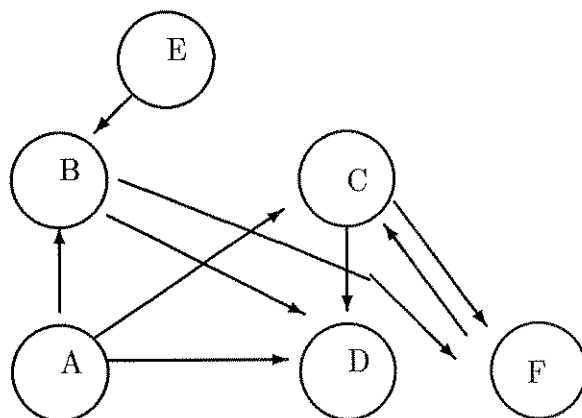


Figura 33: Rede hierárquica de dois níveis, dimensionada pelo método heurístico.

nas rotas. Um algoritmo semelhante ao apresentado na seção 8.4 é então utilizado para determinar se deve-se ou não adicionar ou subtrair um circuito de cada uma das rotas.

Durante a simulação, as rotas são subdivididas num grupo de circuitos e o último circuito da rota. Este circuito é ocupado apenas se todos os demais circuitos estiverem ocupados. Então o tráfego cursado, dividido pelo número de rotas ocupadas desde a origem até o destino da chamada, neste último circuito é o estimador $Es1$. Sempre que uma chamada é estabelecida no último circuito de uma rota e os caminhos alternativos a esta rota estão todos congestionados, então é computada uma chamada quase-bloqueada na rota do último circuito e nas rotas congestionadas dos caminhos alternativos. O estimador $Es2$ é o tráfego quase-bloqueado na dada rota.

A cada subintervalo de simulação decide-se se o número de circuitos de uma rota deve ser incrementado, decrementado ou mantido de acordo com o seguinte algoritmo:

1. Se $0.9Es2 < Cr/T$ então retira-se um circuito da rota.
2. Se $1.1Es2 > Cr/TeEs2 > Cr/T$ então adiciona-se um circuito na rota.

8.7 Experimento Computacional.

A seguir são apresentados os resultados do dimensionamento de uma rede hierárquica, apresentada na figura 33. Foram dimensionadas duas redes com a mesma topologia, mas com o vetor de tráfego distinto. Para este dimensionamento, arbitrou-se $Cr/T = 0.15$, para todos os tráfegos e rotas, e empregou-se encaminhamento fixo clássico (transbordo). O tempo de dimensionamento foi de 2 minutos numa workstation tipo Sparc 1. A rede também foi dimensionada pelo programa Otelos, que emprega a metodologia clássica de dimensionamento, com uma probabilidade de bloqueio de 0.5% nas rotas.

A rede dimensionada pelo programa OTELO, resultou num bloqueio global de 0.6% em ambos os casos. O dimensionamento pelo critério de rentabilidade resultou num bloqueio global de 1.5% na versão A e 2.2% na versão B. Em contrapartida, o custo de implantação da rede foi 8% e 15% inferior ao da solução obtida pela metodologia clássica.

Rota	Versão A			Versão B		
	Orig.	Dimens.	Bloq. (%)	Orig.	Dimens.	Bloq. (%)
AD	21	17	26	82	87	4
AC	10	16	5	14	15	6
AB	11	2	5	19	5	28
BD	67	62	0	33	26	1
CD	69	69	5	37	31	3
FC	66	64	1	32	27	2
EB	66	62	1	32	26	4
CF	11	8	1	11	7	11
BF	5	1	21	5	2	11
TOTAL	326	301		265	226	

Tabela 21: Número de circuitos originalmente planejado e redimensionado a partir da abordagem tipo rentabilidade e probabilidade de bloqueio nas rotas no segundo caso.

Tráfego	Versão A		Versão B	
	Intens. (Erl)	Bloq. (%)	Intens. (Erl)	Bloq. (%)
AD	20	1.2	80	0.6
AF	5	6.6	5	11.4
ED	50	1.6	20	4.3
FD	50	1.2	20	3.9

Tabela 22: Probabilidade de bloqueio ponto-a-ponto.

8.8 Conclusão

O princípio de Moe merece ser reconsiderado para o dimensionamento do entroncamento e determinação do plano de encaminhamento. No atual contexto de competitividade e eficiência das empresas de telecomunicações, parece oportuno discutir a substituição do critério clássico (dimensionamento do entroncamento sob o critério de minimização do investimento com restrições de grau de serviço) pelo princípio de Moe (dimensionamento da rede sob o ponto de vista de maximização da rentabilidade).

As ferramentas obtidas a partir desta abordagem são simples, precisas e computacionalmente eficientes. Apresentam também um grande potencial para iteração com o usuário, se enquadrando muito bem nas características desejadas para as ferramentas de planejamento.

A abordagem é muito adequada ao dimensionamento de redes que utilize a técnica de agendamento do plano de encaminhamento e permite que o problema de dimensionamento multi-horário seja tratado de forma imediata.

Capítulo 9

Conclusão.

Podemos ensaiar a organização de nossa desordem segundo esquemas sábios e de provada virtude, mas há de restar um mundo de essências mais íntimas que, esse, permanecerá sempre intato, irreduzível e desdenhoso das invenções humanas.

Sérgio Buarque de Holanda.

Os estudos apresentados nesta dissertação demonstram que a otimização do plano de encaminhamento, de forma integrada ao dimensionamento das rotas, pode contribuir para reduzir significativamente o custo de implantação de redes telefônicas. Mostrou-se que o principal problema que deve ser resolvido no dimensionamento das rotas é decidir quando implantar ou não rotas diretas entre centrais e, quando não há rotas diretas, escolher qual o caminho indireto a ser utilizado. Apresentou-se um algoritmo eficiente para resolver este problema. Foi demonstrado como a introdução de cabos ópticos, como meio de transmissão, contribuem para que as redes utilizem intensivamente rotas diretas e para que tenham uma estrutura em um ou dois níveis de comutação.

As centrais telefônicas tipo CPA-T permitem introduzir novas técnicas de encaminhamento. Algumas destas técnicas, particularmente as técnicas de encaminhamento dinâmico, permitem melhorar significativamente o desempenho das redes. O encaminhamento dinâmico é a forma mais eficiente de aumentar a robustez da rede em relação a panes ou sobrecarga.

O encaminhamento das redes brasileiras ainda utiliza a estrutura hierárquica com a técnica de transbordo. Em muitas redes, o número de centrais tipo CPA instaladas já é significativo. Nestas redes já é possível implantar novos tipos de plano de encaminhamento. A maioria destas centrais é capaz de executar os mecanismos de agendamento, reserva de circuitos e reencaminhamento. Estes elementos permitem estabelecer um plano de encaminhamento de desempenho significativamente melhor que o atual.

Com a aplicação do mecanismo de reserva de circuitos, não há razão para que qualquer rota entre centrais CPA não seja utilizada como rota alternativa. Os tráfegos de primeira escolha não são prejudicados uma vez que estão protegidos. O custo de utilizar a função tandem nestas centrais é pequeno. Portanto todas as centrais locais são candidatas a realizar a função tandem. Consequentemente a estrutura recomendada para uma rede como esta é interligar as centrais conforme

a viabilidade econômica, determinar os caminhos de primeira escolha e utilizar os demais caminhos possíveis como rotas alternativas. Para dimensionar uma rede com estes mecanismos pode-se adotar o seguinte procedimento:

1. Dimensionar as rotas da rede e determinar os caminhos de primeira escolha, de forma a minimizar o custo de implantação da rede.
2. Determinar o limiar de reserva de circuitos para cada rota.

Este trabalho demonstrou que a aplicação da teoria de decisão de Markov ao problema de otimização de redes telefônicas é viável, sem que seja preciso lançar mão de hipóteses demasiadamente simplificadoras. A partir desta teoria, foi apresentado um algoritmo para determinação do plano de encaminhamento dependente do estado de desempenho ótimo.

No decorrer da elaboração deste trabalho, ficou evidente que o emprego da técnica de simulação a eventos discretos é uma alternativa promissora, não apenas como ferramenta de análise de redes mas também aplicada aos processos de síntese otimizada. Atualmente já é possível simular redes de médio porte (1.000 a 30.000 Erlangs) em tempos de CPU da ordem de minutos. Certamente, em pouco tempo, será possível simular também redes de grande porte (30.000 a 1.000.000 Erlangs).

Bibliografia

- [1] *Encaminhamento de Chamadas Telefônicas Automáticas e Semi-automáticas*. Sistema de Práticas Telebrás, Série Engenharia 210-110-700, 01/02/78./78.
- [2] R. G. Ackerley. Overall grade-of-service models for the british telecom network. In *ITC 12*, page 1.1B.1, Junho 1988.
- [3] G. R. Ash, R. H. Cardwell, and R. P. Murray. Design and optimization of network with dynamic routing. *The Bell System Technical Journal*, pages 1787–1820, Outubro 1981.
- [4] G. R. Ash, A. K. Hafker, and K. R. Krishnam. Servicing an real-time control of networks with dynamic routing. *The Bell System Technical Journal*, pages 1821–1845, Outubro 1981.
- [5] Gerald R. Ash. Design and control of networks with dynamic nonhierarchical routing. *IEEE Communications Magazine*, pages 34–40, October 1990.
- [6] Anantaram Balakrishnan and Stephen C. Graves. A composite algorithm for a concave-cost network flow problem. *Networks*, 19:175–202, 1989.
- [7] Mokhtar S. Bazaraa and C. M. Shetty. *Nonlinear Programming - Theory and Algorithms*. John Wiley & Sons, 1979.
- [8] G. Bel, P. Chemouil, J. M. Garcia, and F. Le Gall. Adaptative traffic routing in telephone networks. *Large Scale Systems Journal*, 8(3), 1985.
- [9] Ivanil Sebatião Bonatti. *Gestion de Réseaux de Service: Application au Réseau Téléphonique Interurbain*. PhD thesis, Universidade Paul Sabatier de Toulouse, 1981. No 767.
- [10] Amaury Kruel Budri. Encaminhamento de chamadas telefônicas: Modelamento e simulação. Master's thesis, FEE/UNICAMP, Jul 1991.
- [11] B. Camoin and A. Masson. Wtude de la robustesse des structures de rwseaux - Le modrle malepeste. NT/RCC/RTR 415, CNET, 1978.
- [12] Prosper Chemouil, Janusz Filipiak, and Paul Gauthier. Performance issues in the design of dynamically controlled circuit-switched networks. *IEEE Communications Magazine*, pages 90–95, October 1990.

- [13] Edson R. de Pieri. Encaminhamento de chamadas telefônicas: Otimização de parâmetros. Master's thesis, FEE/UNICAMP, 1987.
- [14] M. Eisemberg. Engineering traffic networks for more than one busy hour. *The Bell System Technical Journal*, 56(1):1–20, January 1977.
- [15] Paulo M. França, Jurandir F. R. Fernandes, and Hermano M. F. Tavares. Expansão de redes telefônicas. *SBA: Controle e Automação*, 1(3):207–215, 1987.
- [16] Anilton S. Garcia. *Planejamento de Entroncamento em Redes Telefônicas Urbanas em Processo de Digitalização*. PhD thesis, FEE UNICAMP, 1987.
- [17] Andre Girard. Blocking probability of noninteger groups with trunk reservation. *IEEE Transactions on Communications*, 33(2):113–120, Fevereiro 1985.
- [18] Andre Girard. *Routing and Dimensioning in Circuit-Switched Networks*. Addison Wesley Publishing Company, 1990.
- [19] Detelev G. Haenschke, David A. Kettler, and Eric Oberer. Network management and congestion in the U.S. telecommunications network. *IEEE Transactions on Communications*, 29(4), 1981.
- [20] Ronald A. Howard. *Dynamic Programming and Markov Processes*. The MIT Press, 1969.
- [21] Masaichi Kajiwara, Masato Eguchi, and Keniuchi Mase. Future local telephone network configurations. *IEEE Transactions on Communications*, COM-28(1):102–111, January 1980.
- [22] Edward Kasner and James Newman. *Matemática e Imaginação*. Zahar Editores, Rio de Janeiro, 1968.
- [23] Martin Kerner, Howard L. Lemberg, and David M. Simmons. An analysis of alternative architectures for the interoffice network. *IEEE Journal on Selected Areas in Communications*, SAC-4(9):1404–1413, December 1986.
- [24] Peter B. Key and Graham A. Cope. Distributed dynamic routing schemes. *IEEE Communications Magazine*, pages 54–64, October 1990.
- [25] K. R. Krishnan. Markov decision algorithms for dynamic routing. *IEEE Communications Magazine*, pages 66–67, October 1990.
- [26] Luiz Carlos Machado Lage. Planejamento de redes de entroncamento em Áreas urbanas multi-centrais. Master's thesis, Departamento de Ciência da computação UFMG, 1989.
- [27] F. Le Gall and J. Bernossou. An analytical formulation for grade of service determination in telephone networks. *IEEE Transactions on Communications*, COM-31(3), March 1983.

- [28] F. Le Gall and J. Bernossou. Blocking probabilities for trunk reservation policy. *IEEE Transactions on Communications*, 35(3):313–318, Marco 1987.
- [29] Kenichi Mase and Hisao Yamamoto. Advanced traffic control methods for network management. *IEEE Communications Magazine*, pages 82–88, October 1990.
- [30] Debasis Mitra and Judith B. Seery. Comparative evaluations of randomized and dynamic strategies for circuit-switched networks. *IEEE Transactions on Communications*, 39(1):102–115, January 1991.
- [31] Shyam L. Moondra. Impact of emerging switching-transmission cost tradeoffs on future telecommunications networks architectures. *IEEE Journal on Selected Areas in Communications*, 7(8):1207–1218, October 1989.
- [32] A. O. Moreno. Planejamento de redes telefônicas - Determinação do circuito de junção. Master's thesis, ITA, Agosto 1974.
- [33] T. J. Ott and K. R. Krishnan. State dependent routing of telephone traffic and the use of separable routing schemes. In *Teletraffic Issues in Advanced Information Society*, pages 867–872. ITC 11, North-Holland, 1985.
- [34] Y. Rapp. Planing of junction network in a multi-exchange area - I. General principles. *Ericson Technics*, (1), 1964.
- [35] Y. Rapp. Planing of junction network in a multi-exchange area - II. Extensions of the principles and applications. *Ericson Technics*, (1), 1964.
- [36] Jean Regnier and W. Hugh Cameron. State-dependent dynamic traffic management for telephone networks. *IEEE Communications Magazine*, pages 42–53, October 1990.
- [37] Brunilde Sansò, Francois Soumis, and Michel Gendreau. On the evaluation of telecommunications network reliability using routing models. *IEEE Transactions on Communications*, 39(10):1494–1501, Outubro 1991.
- [38] Giancarlo N. Stefanuto. Os limites do sucesso da pesquisa acadêmica no ambiente industrial. In *XV Simpósio Nacional de Pesquisa de Administração em Ciência e Tecnologia.*, page G12, October 1990.
- [39] Henk C. Timjs. *Stochastic Modelling and Analysis: A Computational Approach*. John Wiley & Sons, 1986.
- [40] B. Wallström. Methods for optimizing alternative routing networks. *Ericsson Technics*, (1), 1969.
- [41] Roger I. Wilkinson. Theories for tool traffic engineering in the U.S.A. *The Bell System Technical Journal*, (35):421–514, 1956.

- [42] Richar B. Wolf. Advanced techniques for managing telecommunications networks. *IEEE Communications Magazine*, pages 76–81, October 1990.

Índice

- área de grupamento, 7
- área geográfica, 7
- agenda, 8
- algoritmo, 46, 51, 57, 61, 78, 79
- ambiente, 29
- ANATRAF, 33
- arco-produto, 45
- automatos aprendizes, 10
- Balakrishnan, 46, 57
- côncavo, 57
- cadeia homogênea, 66
- cadeia semi-markoviana, 66, 68, 70, 77, 79, 82, 88
- capacidade residual, 10
- central de trânsito, 3, 7
- central local, 3
- central tandem, 3
- classe, 7
- concentração de tráfego, 18, 20
- contagem, 92
- convexa, 46
- convexo, 43, 57
- CPA, 2, 8, 13, 26, 42
- CPqD, 29
- crankback, 12, 33, 40
- custo, 45, 51, 78, 82
- custo estacionário, 67
- custo relativo, 77
- custo relativo, 67, 89
- custos, 14, 21, 29
- custos relativos, 89
- DAR, 8
- decisão, 78
- decisão de Markov, 76, 92
- decisão markoviana, 65
- decisores, 28
- decomposição, 3
- digital cross conect, 24
- dimensionamento, 102
- DNHR, 8
- DTM, 8
- dual ascendente, 48, 60
- encaminhamento, 2, 4, 6, 57, 68, 78
- encaminhamento adaptativo, 2
- encaminhamento dependente do estado, 10, 65, 88
- encaminhamento dinâmico, 8
- encaminhamento hierárquico, 11
- entroncamento, 4, 15
- Erlang invertido, 33
- estágios de linha remotos, 15
- estado, 66, 70, 75, 82
- estimador, 89, 90
- estrutura física, 6
- estrutura lógica, 6
- eventos discretos, 1
- ferramentas, 28, 32
- fila GI/M/C/C, 79, 81, 85, 88
- fluxo multiproduto, 57
- fluxo multiproduto, 44
- função objetivo, 58
- ganho de escala, 15, 16, 19, 43

- gap de dualidade, 46
- gerência de rede, 26
- grau de serviço, 105
- guloso, 47
- heurístico, 50
- hierarquia, 2, 5, 6
- horizonte, 3
- insucesso, 89, 92
- integração, 29
- ITU, 28
- lagrangeano, 47, 48, 60
- linear por partes, 58
- linearização, 59
- localização de centrais, 3
- modelo agregado, 83
- modelo analítico, 1, 33
- modularidade, 23, 43
- Moe, princípio de, 100
- multi-horário, 2, 25, 100
- objetivos, 1
- OTELLO, 56, 107
- otimização, 29
- PDH, 23
- planejamento, 3
- PLANITU, 28
- poissoniano, 7
- política, 68, 69, 73
- pseudo-côncava, 102
- rede artificial, 54
- rede física, 21
- rede hierárquica, 93
- rede local, 3
- rede não hierárquica, 95
- rede digital, 13
- rede hierárquica, 6, 10
- rede interurbana, 3
- rede lógica, 21
- rede local, 3, 14
- rede não hierárquica, 6
- rede realista, 55
- reencaminhamento, 12, 33
- relaxação, 47
- rentabilidade, 101
- repartição de carga, 7, 57
- reserva de circuitos, 8, 15, 33
- restrições, 47
- rota de alto uso, 11
- rota de proteção, 11
- rota direta, 11
- rota final, 11
- roteamento, 4
- SDH, 23
- SIMTRAF, 33
- simulação, 1, 33, 93
- STAR, 8
- STR, 8
- subgradiente, 49, 61
- técnica de encaminhamento, 7
- tempo de execução, 54
- TMN, 26
- tráfego poissoniano, 77
- transbordo, 2, 7, 19
- UNICAMP, 29