

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**ANÁLISE DE DESEMPENHO DE COMPOSIÇÃO DE
REDES EM REDES DE AMBIENTE**

MAURÍCIO SUHET SPÍNOLA

ORIENTADOR: DJAMEL H. SADOK

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

PUBLICAÇÃO: PPGENE.DM – 050/08

BRASÍLIA/DF: MAIO - 2008

FICHA CATALOGRÁFICA

SPINOLA, MAURICIO SUHETT

Análise de Desempenho de Composição de Redes em Redes Ambiente [Distrito Federal] 2008.

xvii, 150p., 210x297 mm (ENE/FT/UnB, Mestre, Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Elétrica

1. Composição de Redes

2. Redes Ambiente

3. Desempenho de Redes

4. Simulação

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

CESSÃO DE DIREITOS

AUTOR: Mauricio Suhett Spínola.

TÍTULO: Análise de Desempenho de Composição de Redes em Redes de Ambiente.

GRAU: Mestre

ANO: 2008

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Mauricio Suhett Spínola

R. 9 Norte Lote 3 AP. 1104.

71908-540 – Águas Claras – DF – Brasil.

AGRADECIMENTOS

Gostaria de agradecer ao professor e orientador Djamel H. Sadok, pela sua sábia orientação e ajuda na elaboração desta dissertação. À professora Judith Kelner pelos valiosos conselhos.

Esta dissertação foi desenvolvida com a cooperação de algumas pessoas do GPRT (Grupo de Pesquisa em Redes e Telecomunicações) da Universidade Federal de Pernambuco (UFPE), às quais gostaria de registrar meu agradecimento: Carlos Kamienski e Stênio Fernandes. Em especial, gostaria de agradecer ao Ramide Dantas e a Jennifer Lima, pela ajuda, disponibilidade e colaboração.

À Brasil Telecom, pela oportunidade de fazer o mestrado e pelo financiamento do mesmo. Ao meu co-orientador na Brasil Telecom, Mauro T. Fukuda.

DEDICATÓRIA

A Jesus Cristo, razão e sentido maior da minha vida, por ter me dado a oportunidade de concluir mais uma etapa.

Aos meus pais Raimundo e Ermita pelo amor e dedicação, meu sogro Asclepíades e minha sogra Rosélia pelo encorajamento, meus irmãos Márcio e Charles, minha irmã Andrea, pelo apoio.

À Rosa Cleide, amiga e esposa, quem foi a motivação e força para concluir esta dissertação; você foi o porto que me abrigou nos dias de tempestade e o vento que me levou ao fim do percurso; Você, juntamente com nossos filhos Larissa e Matheus, faz tudo valer a pena.

RESUMO

ANÁLISE DE DESEMPENHO DE COMPOSIÇÃO DE REDES EM REDES DE AMBIENTE

Autor: Mauricio Suhett Spínola

Orientador: Djamel H. Sadok

Programa de Pós-graduação em Engenharia Elétrica

Brasília, março de 2008.

Na medida em que o cenário de telecomunicações se torna altamente diverso, por exemplo, devido ao surgimento de tecnologias de acesso heterogêneas (WLAN, WiMAX, UMTS, etc) e novos tipos de rede, surge a necessidade de um procedimento uniforme para a cooperação entre estas redes. A Composição de Redes consiste justamente neste procedimento uniforme e dinâmico para se alcançar a cooperação entre redes. Este procedimento criará novas oportunidades de negócio na medida em que facilita uma cooperação flexível e mais dinâmica entre redes heterogêneas no plano de controle.

O conceito de Composição de Redes é parte integrante do conceito maior de Redes de Ambiente (Ambient Networks) que foi desenvolvido para facilitar a cooperação entre redes heterogêneas. A Composição de Redes permite que serviços sejam provisionados e recursos sejam compartilhados através de um acordo de composição (Composition Agreement).

O desempenho do procedimento de composição é um fator crítico na determinação da viabilidade das Redes de Ambiente uma vez que se espera que este procedimento seja realizado extensivamente durante uma interação típica entre um usuário e a rede. Avaliar sua estabilidade e escalabilidade se torna uma questão importante, já que estas composições envolvem um custo adicional em termos de tempo e processamento para sua realização.

Uma vez que não estão disponíveis atualmente implementações reais de Redes de Ambiente, optou-se, neste trabalho, por utilizar um simulador, denominado ANCSim (Ambient Networks Composition Simulator). Para as simulações, o comportamento estatístico das demandas de serviço foi extraído de medidas realizadas na rede GPRS da Brasil Telecom.

ABSTRACT

PERFORMANCE ANALYSIS OF NETWORK COMPOSITION ON AMBIENT NETWORKS

Author: Mauricio Suhett Spínola

Supervisor: Djamel H. Sadok

Programa de Pós-graduação em Engenharia Elétrica

Brasília, March of 2008.

Since the telecommunications scenario become highly diverse, for instance, because of the development of heterogeneous access technologies (like WLAN, WiMAX and UMTS) and new types of networks, it is necessary a uniform procedure for cooperation between these networks. A network composition consists exactly of this uniform and dynamic procedure to achieve cooperation between networks. New business opportunities could be created with this procedure since it has the potential to facilitate a more dynamic and flexible cooperation between heterogeneous networks on the control plane.

The concept of network composition is part of the broader concept of Ambient Networks that was developed to facilitate cooperation between heterogeneous networks. The network composition permit that services be provisioned and resources are shared through composition agreements.

The performance of the composition procedure is a critical factor in determine a feasibility of Ambient Networks model since it is expected that this procedure will be executed extensively during a typical interaction between a user and the network. To evaluate its stability and scalability is an important issue because these compositions involve an additional cost in terms of time and power processing to execute it.

Since there is no real implementation of Ambient Networks available currently, we have decided in this work to use a simulator, called ANCSim (Ambient Networks Composition Simulator). For the simulations I used statistical behavior for service demands measured on GPRS network of Brasil Telecom.

SUMÁRIO

1. INTRODUÇÃO.....	1
1.1. MOTIVAÇÃO - A NECESSIDADE DE CONVERGÊNCIA.....	1
1.2. OBJETIVOS.....	4
2. AS REDES DE AMBIENTE.....	5
2.1. O PROJETO AMBIENT NETWORKS.....	5
2.1.1. Áreas de Trabalho e Pacotes de Trabalho.....	7
2.2. REQUISITOS GERAIS DE ARQUITETURA.....	9
2.2.1. Redes Heterogêneas.....	10
2.2.2. Mobilidade.....	10
2.2.3. Composição.....	11
2.2.4. Segurança e Privacidade.....	11
2.2.5. Compatibilidade Retroativa e Migração.....	12
2.2.6. Robustez e Tolerância a Falhas.....	13
2.2.7. Qualidade de Serviço.....	13
2.2.8. Suporte a Múltiplos Domínios.....	13
2.2.9. Tarifação.....	14
2.2.10. Conhecimento do Contexto.....	14
2.2.11. Extensibilidade Plug-and-Play dos Serviços de Rede.....	14
2.2.12. Inovação e Utilidade de Aplicação.....	15
2.3. O CONCEITO DE AMBIENT CONTROL SPACE (ACS).....	15
2.4. O MODELO AMBIENT LAYER.....	20
2.4.1. Abstração de Fluxo.....	21
2.4.2. Abstração de Bearer.....	22
2.4.3. Sessão de aplicação.....	23
2.4.4. Estrutura de nomes.....	23
3. O FRAMEWORK PBMAN.....	24
3.1. INTRODUÇÃO.....	25
3.2. DEFINIÇÃO.....	26
3.3. POLICY DECISION NETWORK.....	29
3.4. POLICY AGENTS.....	30
3.5. POLICY INFORMATION MODEL.....	31
3.6. COMPOSIÇÃO NO PBMAN.....	32

3.6.1.	<i>Classes de Composições</i>	33
3.6.1.1.	<i>Composição tipo PDN/PDN</i>	33
3.6.1.2.	<i>Composição tipo Agente/PDN</i>	35
3.6.1.3.	<i>Composição tipo Agente/Agente</i>	37
3.6.2.	<i>Acordo de Composição</i>	37
3.7.	<i>X-PBMAN</i>	39
3.7.1.	<i>X-PEER Middleware</i>	39
3.7.2.	<i>Arquitetura de Software</i>	40
3.8.	<i>CENÁRIO PBMAN</i>	42
3.8.1.	<i>Transação do Serviço de Vídeo</i>	43
3.8.2.	<i>Modelagem do Cenário</i>	45
4.	COMPOSIÇÕES EM REDES DE AMBIENTE	47
4.1.	<i>PROCEDIMENTOS DE COMPOSIÇÃO</i>	48
4.1.1.	<i>Detecção de Mídia</i>	49
4.1.2.	<i>Descoberta/ Anúncio</i>	50
4.1.3.	<i>Estabelecimento da Segurança e da Conectividade</i>	50
4.1.4.	<i>Negociação do Acordo de Composição</i>	51
4.1.5.	<i>Realização do Acordo de Composição</i>	52
4.1.6.	<i>Procedimento de Decomposição</i>	52
4.1.7.	<i>Procedimento de Atualização da Composição</i>	52
4.2.	<i>TIPOS DE COMPOSIÇÃO</i>	53
4.2.1.	<i>Tipos de Composição – Perspectiva da Gerência de Recursos</i>	53
4.2.2.	<i>Tipos de Composição – Perspectiva da Gerência de Rede</i>	56
4.2.2.1.	<i>Absorção</i>	58
4.2.2.2.	<i>Overlay do tipo Gatewaying</i>	58
4.2.2.3.	<i>Overlay do tipo Interworking</i>	60
4.2.3.	<i>Acordos de Composição Multilaterais</i>	60
4.2.4.	<i>Processo de Decisão e Relações</i>	60
5.	O ANCSIM (AMBIENT NETWORK COMPOSITION SIMULATOR) 61	
5.1.	<i>VISÃO GERAL</i>	61
5.1.1.	<i>Elementos</i>	62
5.1.1.1.	<i>Redes</i>	62
5.1.1.2.	<i>Serviços</i>	63
5.1.1.3.	<i>Funcionalidade</i>	64
5.1.1.4.	<i>Composição no ANCSim</i>	64
5.2.	<i>SELECIONANDO UMA REDE DESTINO</i>	65
5.3.	<i>SELEÇÃO DE CAMINHO</i>	67
5.4.	<i>O PROCESSO DE COMPOSIÇÃO</i>	70

5.4.1.	<i>ACS e Acordo de Composição</i>	72
5.4.2.	<i>Decomposição</i>	72
5.4.3.	<i>Estrutura do ANCSim</i>	73
5.4.4.	<i>Eventos</i>	75
6.	AVALIAÇÃO DAS COMPOSIÇÕES	76
6.1.	METODOLOGIA.....	77
6.2.	MEDIÇÕES NA REDE GPRS	77
6.3.	COMPOSIÇÃO AGENTE/PDN	79
6.4.	TRATAMENTO ESTATÍSTICO	84
6.4.1.	<i>Extração dos outliers</i>	84
6.4.2.	<i>Teste de aderência (Goodness-of-fit)</i>	88
6.4.2.1.	Teste de Auto-Correlação.....	88
6.4.2.2.	Estimativa de parâmetros da distribuição teórica	91
6.4.2.3.	Teste de aderência I^2	92
6.5.	COMPOSIÇÃO PDN/PDN	96
6.6.	DURAÇÃO DE SERVIÇO	99
6.6.1.	<i>Rede GPRS Brasil Telecom</i>	99
6.7.	TAXA DE SOLICITAÇÃO DE SERVIÇO	106
6.8.	AMBIENTE DE SIMULAÇÃO E CONFIGURAÇÃO DO EXPERIMENTO.....	106
6.8.1.	<i>Topologia de Rede</i>	106
6.8.2.	<i>Infra-estrutura de simulação</i>	107
6.8.3.	<i>Rede UMTS</i>	111
6.8.4.	<i>Métricas consideradas</i>	111
6.9.	RESULTADOS DA SIMULAÇÃO	112
6.9.1.	<i>Escalabilidade</i>	112
6.9.2.	<i>Tempo de Resposta</i>	117
6.9.3.	<i>Estabilidade da Composição</i>	121
7.	CONCLUSÕES E TRABALHOS FUTUROS	126
	REFERÊNCIAS BIBLIOGRÁFICAS	128
	APÊNDICE	130
A.	RESULTADOS DOS TESTES DE ADERÊNCIA	131
	COMPOSIÇÃO AGENTE/PDN	131
	<i>Distribuição exponencial</i>	131
	<i>Distribuição Gamma</i>	133
	<i>Distribuição Weibull</i>	136

<i>Distribuição Normal</i>	139
<i>Distribuição Lognormal</i>	141
<i>Distribuição Chi-Squared</i>	144
<i>Distribuição Poisson</i>	146

LISTA DE TABELAS

Tabela 3.1. Destinos selecionados para AN1.....	45
Tabela 3.2. Associações dos destinos selecionados na AN1.	46
Tabela 3.3. Políticas selecionadas para AN1.	47
Tabela 6.1. Média e Desvio Padrão para composição Agente/PDN.....	83
Tabela 6.2. Média e Desvio Padrão para composição Agente/PDN (sem outliers). 87	
Tabela 6.3. Valores para função de auto-correlação.	89
Tabela 6.4. Parâmetros utilizados nas medições.....	97
Tabela 6.5. Parâmetros.....	98
Tabela 6.6. Proporções.....	98
Tabela 6.7. Médias e Desvios Padrão para composição PDN/PDN.	98
Tabela 6.8. Distribuições.	99
Tabela 6.9. Média e Desvio Padrão para a duração do serviço.	101
Tabela 6.10. Média e Desvio Padrão para a duração do serviço com extração de outliers.....	106
Tabela 6.11. Infra-estrutura de simulação.....	107
Tabela 6.12. Parâmetros de simulação.....	108
Tabela 6.13. Fatores e Níveis de Simulação.	109
Tabela 6.14. População por Estado brasileiro e proporção de redes Agentes.	110

LISTA DE FIGURAS

Figura 2.1 Estrutura atual do WWI.	6
Figura 2.2. Representação esquemática do ACS.	16
Figura 2.3. Conexão de ACSs através da interface ANI.	16
Figura 2.4. Integração do ACS com IMS.	20
Figura 2.5. Representação das abstrações de conectividade.	22
Figura 3.1. Arquitetura geral do PBMAN.	27
Figura 3.2. Policy Decision Network.	29
Figura 3.3. Modelo de Informação do PBMAN.	31
Figura 3.4. Composição de PDN _A e PDN _B	33
Figura 3.5. Arquitetura X-Peer.	40
Figura 3.6. Arquitetura P-Node.	41
Figura 3.7. Transação do Serviço de Vídeo.	43
Figura 4.1. Fases do procedimento de composição de redes.	49
Figura 4.2. Resultado da integração de rede.	54
Figura 4.3. Resultado de compartilhamento de controle, direito não exclusivo para controle de recurso.	55
Figura 4.4. Resultado de compartilhamento de controle, direito exclusivo para controle de recurso (delegação de controle).	56
Figura 4.5. Resultado do interworking de rede.	56
Figura 4.6. Tipos de overlays.	57
Figura 4.7. Absorção.	58
Figura 4.8. Gatewaying.	59
Figura 5.1. PDNs e Agentes no ANCSim.	62
Figura 5.2. Uma topologia de Ambient Network.	66
Figura 5.3. Distribuição de Pesos.	66
Figura 5.4. Caminho considerando Serviços de Usuário.	67
Figura 5.5. Serviço disponível no Caminho.	68
Figura 5.6. Caminho mais curto entre as redes origem e destino.	69
Figura 5.7. Processo de Composição.	70
Figura 5.8. Processo de Composição.	71
Figura 5.9. Diagrama de Classe do ANCSim.	74
Figura 5.10. Fluxo de uma solicitação de serviço no ANCSim.	76

Figura 6.1. Metodologia adotada.	77
Figura 6.2. Rede GPRS.	78
Figura 6.3. Decodificação, Análise e Armazenamento das mensagens.	79
Figura 6.4. Procedimento de Attach.	80
Figura 6.5. Tempos de composição Agente/PDN.	82
Figura 6.6. GPRS Attach simplificado.	84
Figura 6.7. Tratamento estatístico.	84
Figura 6.8. Tempos de composição Agente/PDN com extração de outliers.	87
Figura 6.9. Função de auto-correlação.	90
Figura 6.10. Comparação PDFs para distribuição empírica e analítica.	95
Figura 6.11. QQ-Plot para distribuição Chi-squared.	96
Figura 6.12. Composição PDN/PDN.	97
Figura 6.13. Procedimento de Ativação de Contexto PDP.	100
Figura 6.14. Procedimento de Desativação de Contexto PDP.	100
Figura 6.15. Duração de Serviço.	103
Figura 6.16. Duração de Serviço com extração de outliers.	105
Figure 6.17. Topologia para simulação.	107
Figure 6.18. Número de composições por solicitação de serviço por tamanho de rede (a) e por overtime (b).	113
Figure 6.19. Composições Agente/PDN ociosas com overtime 1 (a) e Composições PDN/PDN ociosas sem overtime (b).	114
Figure 6.20. Composições ociosas sem controle de admissão (a) e Composições ociosas com overtime 0.1 (b).	116
Figure 6.21. Tempo de resposta para diferentes cargas e tamanhos de rede (a) e a probabilidade de composição dada uma solicitação de serviço (b).	118
Figure 6.22. Topologia da RNP com 27 PDNs.	119
Figure 6.23. Tempo de resposta para uma topologia com 12 PDNs [18].	119
Figure 6.24. Taxa de bloqueio de serviço.	120
Figure 6.25. Número de composições Agente/PDN (a) e composições PDN/PDN (b).	122
Figure 6.26. Duração das composições Agente/PDN sob diferentes overtimes (a) e Taxa de bloqueio de serviço (b).	123
Figure 6.27. Duração das composições PDN/PDN para cargas diferentes (a) e para diferentes overtimes (b).	125

LISTA DE SÍMBOLOS E ABREVIATURAS

3G - Terceira Geração
3GPP - Third Generation Partnership Project
ACS - Ambient Control Space
ADSL - Asymmetric Digital Subscriber Line
AN – Ambient Network
ANCSim – Ambient Network Composition Simulator
ANI – Ambient Network Interface
ARI – Ambient Resource Interface
ASI – Ambient Service Interface
BSC - Base Station Controller
CA – Composition Agreement
COPS – Common Open Policy Service
FMC - Fixed-Mobile Convergence
GPRS - General Packet Radio Service
GSM - Global System for Mobile Communications
HLR - Home Location Register
HSPA – High-Speed Packet Access
IMS – IP Multimedia Subsystem
IP – Internet Protocol
LTE – Long Term Evolution
MMS - Multimedia Messaging Service
MS - Mobile Station
NAT – Network Address Translation
PBM - Policy-Based Management
PBMAN - Policy-Based Management for Ambient Networks
PDN - Policy Decision Network
PEP - Policy Enforcement Point
PIB – Policy Information Base
PON – Passive Optical Network

QoS – Quality of Service
RAN - Radio Access Network
SGSN - Serving GPRS Support Node
SLA - Service Level Agreements
UMTS - Universal Mobile Telecommunication System
VLR - Visitor Location Register
VoIP – Voice Over IP
VPN – Virtual Private Network
WAP - Wireless Application Protocol
WCDMA Wide Code Multiple Division Access
WLAN – Wireless LAN
WWI - Wireless World Initiative
WWW - World Wide Web

1. INTRODUÇÃO

Este capítulo é uma introdução às Redes de Ambiente e aos problemas abordados por esta dissertação, com o objetivo de evidenciar a importância da análise de desempenho do procedimento de composição de redes. Também apresenta os principais objetivos a serem alcançados e a organização deste documento.

1.1. MOTIVAÇÃO - A NECESSIDADE DE CONVERGÊNCIA

O cenário de telecomunicações está passando por rápidas e profundas transformações. Neste cenário, as operadoras fixas e móveis estão enfrentando um novo desafio denominado de Convergência Fixo-Móvel (Fixed-Mobile Convergence – FMC), que representa a tendência pela busca da convergência de serviços e de redes. Significa a integração das tecnologias de acesso em um conjunto único de serviços.

A convergência pode ser considerada a partir de quatro pontos de vista: convergência de serviços, convergência de equipamentos, convergência de redes e convergência de negócios.

O principal paradigma da convergência de serviços é que os usuários devem poder permanecer agnósticos em relação à tecnologia e deve ter suas necessidades de comunicação satisfeitas em qualquer lugar usando os melhores meios possíveis. Isto é algumas vezes referido como “always best connected” (sempre melhor conectado) – é possível se comunicar tanto por acesso fixo como por acesso móvel usando a mesma interface de usuário.

A convergência de redes tem sido motivada pelo sucesso da Internet e sua utilidade para todos os tipos de serviços. O modelo do protocolo IP e da Internet estão sendo introduzidos em todas as áreas das comunicações, permitindo a evolução para as assim denominadas redes All-IP. Novas tecnologias de acesso (HSPA nos sistemas 3G e LTE, Wimax,

ADSL2+, VDSL, FTTH e redes PON) têm acelerado a convergência em direção às redes IP.

A convergência de negócios significa a oferta de um conjunto de serviços customizados usando qualquer canal de acesso. A convergência de negócios segue em linha com a mobilidade e a personalização. Tem a ver também com a necessidade de competir em um mercado cada vez mais aberto, onde as oportunidades precisam ser rapidamente e dinamicamente aproveitadas.

Para explorar todo o potencial da convergência será necessário implementar uma camada de controle que seja *dinâmica, cooperativa e configurável* em função das políticas de negócio. Esta é a idéia por detrás do paradigma das Redes de Ambiente.

A nova rede convergente deve ser construída levando-se em conta as variações nas demandas dos usuários e a criação de serviços espontâneos e adaptativos que possam ser disponibilizados a qualquer momento, em qualquer lugar e através de qualquer equipamento que o usuário preferir. Este objetivo somente pode ser alcançado se a rede suportar internamente a *reconfiguração dos seus recursos dinamicamente e sob demanda*.

Embora exista uma tendência ao domínio da tecnologia IP nas atuais redes fixas e móveis, tem-se que conviver com uma grande divergência na camada de controle. Diferentes ambientes de controle são estabelecidos para proporcionar serviços de rede como VPN, segurança, QoS, NAT, multicast, etc. O controle destes serviços estão se tornando cada vez mais fragmentados. Esta situação é de difícil administração com as atuais tecnologias.

A visão sustentada pelas Redes de Ambiente permite que as redes cooperem sob demanda, transparentemente e sem a necessidade de pré-configuração ou negociações off-line. Os usuários deixam de ser apenas donos de um terminal ou PC e passam a ser considerados donos e operadores de uma rede de equipamentos localizados em sua casa, seu escritório ou ao redor do seu corpo. Conseqüentemente, eles são incluídos na rede de cooperação e passam a ser tratados como operadores de redes especiais e de baixa complexidade. Esta abordagem é uma generalização que permite incluir os diferentes tipos de rede que estão surgindo, tal como redes veiculares e redes de sensores. Através da modelagem de terminais como redes, a rede passa a ser o bloco de construção básico desta nova arquitetura, permitindo que todos os tipos de rede sejam integrados a um sistema maior.

Nas Redes de Ambiente, a arquitetura de interfuncionamento está baseada em um plano de controle que coordena o estado da rede para funcionalidades como mobilidade, QoS e segurança.

O principal conceito inovador das Redes de Ambiente é a Composição de Rede. A composição de rede permite uma rápida adaptação da topologia de rede quando requerido por usuários móveis e redes em movimento, promovendo acesso a qualquer rede, incluindo as redes móveis pessoais, através de acordos instantâneos e sob demanda entre as redes, chamado de Acordo de Composição. Podemos imaginar a composição como um mecanismo para a negociação automática de roaming e/ou acordos de nível de serviço (Service Level Agreements – SLA), os quais são realizados manualmente hoje em dia.

Estas funcionalidades precisam ser gerenciadas de maneira integrada e flexível. A Gerência Baseada em Políticas (Policy-Based Management – PBM) atende a esta necessidade adequadamente, pois é uma abordagem para a administração de infraestruturas complexas de redes através do estabelecimento de políticas para tratar situações que são prováveis de acontecer, de maneira automática. Políticas podem ser utilizadas, por exemplo, para gerenciar e controlar o acesso aos recursos de uma rede pelo uso de regras e decisões de alto nível.

Além disso, é importante avaliar o desempenho da composição de rede, uma vez que a mesma representa um fator crítico na viabilidade das Redes de Ambiente devido à esperada alta demanda por composições em uma interação típica de um usuário com a rede. Adicionalmente, é necessário um melhor entendimento das principais características da composição e decomposição em Redes de Ambiente. Já que estes conceitos são inovadores, devem-se desenvolver novos modelos de gerenciamento e a análise de desempenho será de grande ajuda para orientar esta tarefa. Algumas das características que precisam ser bem entendidas é o estabelecimento, a duração, a eficiência, a estabilidade e a frequência das composições.

Este trabalho avalia a viabilidade do uso de composições dinâmicas de redes como um mecanismo capaz de fornecer acesso instantâneo a serviços para os usuários. O estudo utilizou o simulador de propósito específico desenvolvido pela UFPE, o ANCSim (Ambient Network Composition Simulator).

1.2. OBJETIVOS

O principal objetivo deste trabalho é a análise do mecanismo de composição de rede para Redes de Ambiente, utilizando para isso estatísticas de utilização de serviços coletadas numa rede real, a rede GPRS da Brasil Telecom. O mecanismo de composição de rede deve ser suficientemente escalável para satisfazer às altas demandas por serviços onipresentes para usuários móveis. As composições complexas devem ser mais estáveis e devem ocorrer com menos frequência, introduzindo um menor impacto não apenas no tempo de resposta percebido pelo usuário, mas também em termos da quantidade de recursos necessários para realizá-las. Por outro lado, as composições mais frequentes devem ser realizadas de maneira rápida e barata.

Para se avaliar o desempenho da composição, foram realizadas simulações considerando-se uma topologia evoluída da atual rede GPRS da Brasil Telecom, além de estatísticas de uso da rede.

O conjunto de experimentos foi elaborado para que os resultados finais pudessem oferecer um maior conhecimento da estabilidade e escalabilidade do mecanismo de composição. O comportamento da composição foi avaliado levando-se em consideração algumas métricas definidas com o propósito específico mencionado anteriormente. Exemplos de métricas consideradas neste trabalho são o número de solicitações de composição, a duração das composições e o número de composições ativas.

2. AS REDES DE AMBIENTE

Este capítulo descreve o Projeto Ambient Networks, apresentando os objetivos do projeto, seus principais impactos na evolução das redes de telecomunicações e suas principais fases de desenvolvimento. Depois, apresentam-se as inovações, princípios e características das Redes de Ambiente, enfatizando o conceito de Composição de Redes, descrevendo seus elementos funcionais, seus requisitos e procedimentos.

2.1. O PROJETO AMBIENT NETWORKS

O projeto Ambient Networks é parte de uma ação mais global chamada de Wireless World Initiative (WWI) [1],[2], que contém uma série de projetos cooperativos dentro do 6th Framework Project (FP6) da União Européia. Os diferentes projetos do WWI estão planejados para serem desenvolvidos num período de seis anos até 2010.

Ambient Networks é um projeto que busca criar soluções de rede para sistemas móveis além dos sistemas de 3ª Geração (3G). O WWI é um grande esforço conjunto entre indústria, academia e governo cujo principal objetivo é definir funções e sistemas móveis que forneçam aos usuários a melhor utilização com o menor custo.

A Figura 2.1 mostra os diferentes projetos que compõem o WWI. Cada projeto é coordenado por uma empresa específica. Por exemplo, o projeto Ambient Networks é coordenado pela Ericsson Research. Estes projetos são auto-contidos em cada área de pesquisa, mas complementam um ao outro. Em outras palavras, os resultados encontrados em cada projeto são combinados para formar uma solução completa [5].

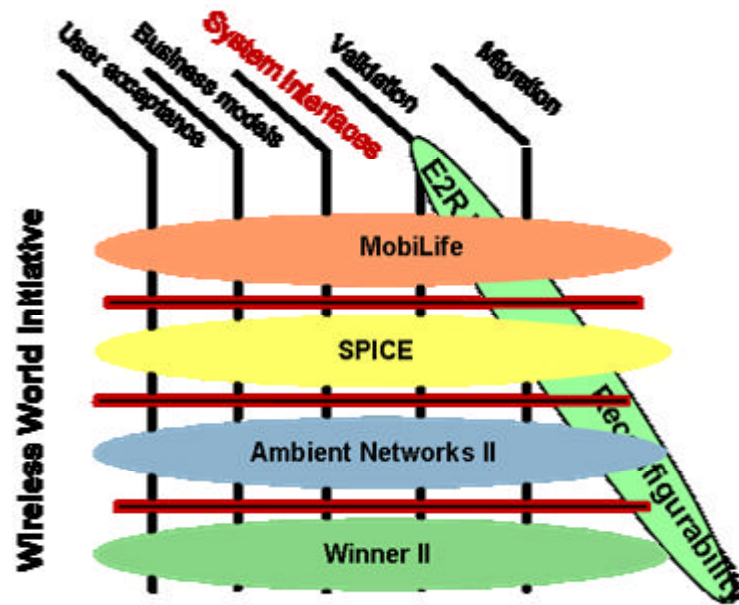


Figura 2.1 Estrutura atual do WWI.

O projeto WWI busca soluções para redes móveis além dos Sistemas 3G que sejam inovadoras e industrialmente exploráveis, que permitam a composição de redes através de fronteiras tecnológicas e de negócio. Para alcançar este objetivo, o projeto foca em conceitos como acesso multi-rádio, composição de redes, mobilidade e redes em movimento, roteamento e transporte inteligente de mídia, redes baseadas em contextos e gerência de segurança e de rede.

Os objetivos estratégicos do WWI são:

- Fornecer sistemas de comunicação ricos e de alto desempenho que ofereçam acesso onipresente a novos serviços, profissionais, indivíduos e multimídias.
- Oferecer sistemas sem fio a custos razoáveis que irão permitir que quatro bilhões de pessoas e quarenta bilhões de equipamentos se comuniquem no ano 2020. Isso permitirá que muito mais pessoas participem do processo econômico, reduzindo consideravelmente as desigualdades no mundo.

Para alcançar os objetivos do WWI, foi definido um cronograma que deve ser seguido por todos os projetos do WWI. Este cronograma tem três fases, com duração de dois anos cada. As fases estão assim especificadas:

- Fase 1: Exploração e análise tecnológica (2004 – 2005): durante esta fase, novos conceitos e novas tecnologias serão explorados e avaliados. Serão identificados requisitos genéricos de usuário. Os resultados esperados incluem (a) a identificação de um conjunto de tecnologias capaz de satisfazer aos objetivos do projeto, (b) os primeiros cenários de conceituação de sistemas e evolução, e (c) a definição de soluções flexíveis que satisfaçam as necessidades dos usuários.
- Fase 2: Desenvolvimento e especificação da tecnologia (2006 – 2007): Nesta fase, um trabalho colaborativo resultará em simulações e protótipos que demonstrarão a viabilidade tecnológica. Serão elaboradas as especificações para os fóruns de padronização. Os resultados esperados incluem simuladores, protótipos e submissão de padrões.
- Fase 3: Validação e Demonstração (2008 – 2009): Nesta fase, os conceitos desenvolvidos nos diferentes projetos integrados serão verificados em testes de larga escala, demonstrando a viabilidade dos sistemas propostos. Os resultados incluem (a) conceitos de sistemas totalmente acordados e (b) demonstradores que mostram aspectos do mundo móvel.

Atualmente os projetos estão na Fase 2. Os conceitos estudados e definidos na Fase 1 são agora implementados como simuladores e protótipos. A presente dissertação atende aos objetivos da Fase 2, consistindo num estudo de viabilidade técnica do conceito de Composição de Rede através de simulação.

Os resultados esperados para o projeto Ambient Networks é uma solução coerente e completa para redes sem fio; uma arquitetura construída a partir de componentes auto-gerenciados que reduzam os custos operacionais; e uma família de protocolos para composição de redes que garantam conectividade, gerência de recursos, segurança, gerenciabilidade, resolução de conflitos e tratamento de conteúdos.

2.1.1. Áreas de Trabalho e Pacotes de Trabalho

Para obter os resultados de acordo com os objetivos, e para estruturar o projeto, os trabalhos em Ambient Networks estão divididos em áreas de trabalho e fases. As Áreas de Trabalho combinam os trabalhos similares; as fases estruturam o projeto no tempo, permitindo uma pesquisa focada com um ciclo rápido de inovação e dando flexibilidade ao projeto enquanto mantém uma direção geral. Uma estrutura de Pacote de Trabalho (Work-Package) é entregue em cada fase.

O projeto Ambient Networks abrange várias áreas de pesquisa. Por exemplo, a definição de requisitos, desenvolvimento de conceitos e o desenho de uma arquitetura geral. Estas atividades são desenvolvidas na primeira área de trabalho denominada *Conceitos e Arquitetura*. Para garantir a relevância comercial e o impacto das Redes de Ambiente, existe uma segunda área de trabalho denominada *Interfaces de Negócio e Viabilidade Comercial*. A terceira área de trabalho, *Soluções Técnicas Chaves*, trata dos problemas técnicos. Finalmente, a quarta área está preocupada com *Prototipação e Validação*.

Considerando o tamanho destas áreas de trabalho e a atual fase do projeto, o trabalho é ainda dividido em tarefas ou pacotes [4]. Os principais pacotes de trabalho para as fases 1 e 2 são:

- Work-Package 1 (WP1) – Conceitos, Arquiteturas e Coordenação Técnica: O principal objetivo deste pacote de trabalho é desenvolver e disseminar o projeto baseado em cenários de usuários. É também responsável por harmonizar os mecanismos e funções de interação desenvolvidas pelo projeto.
- Work-Package 2 (WP2) – Acesso Multi-Rádio: O objetivo deste pacote é propor arquiteturas de acesso de multi-rádio que permitam o acesso sem fio como parte do conceito de Ambient Networks. Tenta encontrar e analisar alguns elementos críticos na rede de acesso para viabilizar técnica e economicamente a arquitetura proposta.
- Work-Package 3 (WP3) – Composição de Rede e Conectividade: Este pacote tem como objetivo oferecer mecanismos de conectividade que facilitem a composição entre Redes de Ambiente de tamanhos e capacidades variados, de forma simples e dinâmica. Descreve em detalhes o mecanismo de composição de redes.

- Work-Package 4 (WP4) – Mobilidade e Redes Móveis: Tem o objetivo de investigar, desenvolver e avaliar conceitos e mecanismos inovadores para a mobilidade e a gerência de redes em movimento. Define a arquitetura de mobilidade considerando os conceitos gerais estabelecidos no WP1.
- Work-Package 5 (WP5) – Transporte e Roteamento Multimídia Inteligente: O principal objetivo deste pacote é investigar novas abordagens para arquitetura de transporte e roteamento baseadas em funções de adaptação e roteamento, integrando os trabalhos desenvolvidos nos pacotes de trabalho WP6, WP3 e WP4.
- Work-Package 6 (WP6) – Context Aware Networks: Este pacote investiga, modela e estrutura as informações processadas por um gerente de contexto. Investiga esquemas flexíveis de descrição de dados para representação dos dados de contexto, facilitando a troca de informação entre as entidades que pertencem a uma Rede de Ambiente.
- Work-Package 7 (WP7) – Segurança: O principal objetivo é fornecer uma arquitetura de segurança que seja uniforme e completa, tornando possível a segurança das Redes de Ambiente em um ambiente caracterizado por fácil gerência, autenticação, privacidade e robustez sob ataque.
- Work-Package 8 (WP8) – Gerência de Rede: Explora, desenvolve e avalia soluções de gerência que sejam inovadoras para viabilizar a composição dinâmica de redes, reduzir os custos operacionais e aumentar a escalabilidade.

2.2. REQUISITOS GERAIS DE ARQUITETURA

A arquitetura de uma Rede de Ambiente é o resultado de um processo estruturado de projeto que começa com a identificação dos principais requisitos impostos sobre uma Rede de Ambiente e sua arquitetura intrínseca.

Os requisitos gerais estão relacionados diretamente com os problemas que as Redes de Ambiente precisam resolver. Alguns destes requisitos são similares aos requisitos já

encontrados em redes tipo GSM/3G ou a Internet. Contudo, os requisitos endereçam problemas não resolvidos nas atuais redes e, portanto, devem ser satisfeitos pelas Redes de Ambiente. Um conjunto dos doze requisitos descreve o que distingue as Redes de Ambiente de tecnologias anteriores.

2.2.1. Redes Heterogêneas

A arquitetura de rede deve suportar eficientemente diferentes tipos de tecnologias a fim de fornecer aos usuários uma conectividade ótima baseada em seus requisitos e preferências.

As Redes de Ambiente devem oferecer comunicação fim-a-fim independente de aplicação ou serviço. Os serviços de rede e os serviços de aplicação devem ser provisionados independentemente um do outro. Uma Rede de Ambiente deve esconder dos processos de aplicação a heterogeneidade de tecnologias, visíveis atualmente nas diferentes tecnologias de acesso. Além de tratar a heterogeneidade de acesso e de rede, as Redes de Ambiente devem tratar a heterogeneidade da cadeia de entrega de serviço, não impondo nenhuma restrição aos modelos de negócio que podem ser implementados em Redes de Ambiente.

2.2.2. Mobilidade

Uma Rede de Ambiente deve suportar a gerência de mobilidade para mobilidade de usuário, serviço, sessão, terminal e rede.

Os mecanismos de gerência de mobilidade devem ser capazes de localizar e atualizar a localização do usuário e suportar transparência de handover. Uma Rede de Ambiente deve ser capaz de suportar novos e existentes mecanismos de mobilidade que permitam que terminais e redes possam se mover sem estarem rigidamente ligados à sua “home network”. Deve ser suportada a transferência de informação de fluxo em tempo real.

A necessidade de mobilidade através de redes heterogêneas impõe algumas limitações no projeto da arquitetura de rede. Especificamente, as redes devem ser capazes de mover-se para diferentes localizações físicas ou lógicas a qualquer momento. Este requisito impacta as convenções de nomes e endereços dentro de uma Rede de Ambiente. É necessária uma função de controle específica para realizar a alocação, registro e se necessário a tradução e resolução de nomes para endereços físicos (locations e end-point IDs).

2.2.3. Composição

A arquitetura de rede deve suportar mecanismos que permitam negociação e acordos entre redes ou domínios administrativos durante o curso da comunicação.

O conceito de composição é uma nova função introduzida pelas Redes de Ambiente. O espaço de controle de uma rede deve ser responsável por gerar e transmitir a informação necessária para que outras redes possam compor com a rede atual. Com este requisito, introduz-se o conceito de negociação e acordos de serviços automáticos. Uma rede pode ser composta a partir de várias outras redes menores, que é vista externamente como uma “única” rede. Os componentes individuais da rede composta não perderão sua identidade ou capacidade de conectar-se a outras redes com sua identidade original.

2.2.4. Segurança e Privacidade

As Redes de Ambiente devem fornecer um esquema de segurança que seja completo e flexível, que opere consistentemente através de um ambiente em constante mudanças constituído por redes, entidades e serviços heterogêneos. Esta segurança de rede deve ser caracterizada por:

- Deve fornecer um ambiente amigável para o usuário, enquanto permanece invisível a este;

- Transição suave entre diferentes acessos e serviços;
- Operação confiável;
- Robustez e resiliência sob ataque;
- Fácil gerência de segurança;
- Proteção e privacidade dos ativos da rede e dos usuários;
- Proteção e privacidade da localização e da identidade;
- Auditoria.

É muito importante que seja possível o estabelecimento de uma relação de confiança entre as várias redes que interagem. O objetivo final é tornar uma Rede de Ambiente imune contra ameaças de segurança.

2.2.5. Compatibilidade Retroativa e Migração

As Redes de Ambiente devem suportar caminhos de migração a partir das redes e terminais existentes, por exemplo, permitindo o reuso da atual infra-estrutura dos sistemas 2G/3G incluindo suas evoluções bem como a Internet e sistemas de acesso não celulares, como xDSL e WLAN. As aplicações legadas devem poder rodar dentro do ambiente das Redes de Ambiente.

As Redes de Ambiente devem interagir com as redes legadas. Como estas não possuem os serviços esperados para as Redes de Ambiente, serviços avançados como composição de redes não seriam possíveis de implementação. Porém, as Redes de Ambiente podem interagir com as redes legadas através de funções de adaptação apropriadas que forneçam as funcionalidades que faltam.

2.2.6. Robustez e Tolerância a Falhas

A arquitetura e gerência de rede devem permitir a implementação de Redes de Ambiente que sejam escaláveis, custo-efetivas, robustas, confiáveis e com alta disponibilidade através de redes heterogêneas em ambientes dinâmicos.

Uma Rede de Ambiente deve evitar pontos únicos de falha. Seu projeto deve garantir o seu funcionamento mesmo quando isolada do resto da rede.

2.2.7. Qualidade de Serviço

Uma Rede de Ambiente deve fornecer múltiplas classes de QoS para serviços fim-a-fim quando necessário. Os mecanismos de QoS utilizados devem ser independentes de tecnologias específicas, mas deve fornecer uma coordenação de QoS consistente através das múltiplas tecnologias de acesso. Deve fornecer também a possibilidade de negociação de uma classe de QoS tanto do lado do usuário como do lado da rede.

O QoS acordado com uma aplicação ou com outra rede deve ser garantido pela Rede de Ambiente mesmo quando compondo com outra Rede de Ambiente ou quando interfaciando com redes legadas. As interfaces de uma Rede de Ambiente devem ser capazes de comunicar os parâmetros necessários para o estabelecimento e manutenção dos serviços baseados no QoS através das fronteiras da rede.

2.2.8. Suporte a Múltiplos Domínios

A arquitetura de rede deve suportar de maneira transparente as funcionalidades de rede entre múltiplos domínios administrativos, isto é, entre áreas operadas e gerenciadas por diferentes autoridades. Deve ser possível negociar com múltiplas Redes de Ambiente simultaneamente e formar acordos para atender uma ou várias aplicações.

2.2.9. Tarifação

A arquitetura de rede deve suportar mecanismos que permitam tarifar uma única entidade e posteriormente aplicar restrições quando apropriado.

As Redes de Ambiente devem fornecer maneiras seguras, confiáveis e eficientes para a coleta e gerência de informações de tarifação independente dos modelos de negócio em uso. Devem ser flexíveis para interagir com sistemas legados de tarifação e compensação.

2.2.10. Conhecimento do Contexto

As Redes de Ambiente devem suportar uma infra-estrutura comum para monitorar o contexto através de todas as funções presentes no Ambient Control Space (veja seção 2.3) para se adaptar automaticamente os serviços às variações do ambiente. Isso deve incluir o suporte à coleta, processamento, gerência e disseminação de informação de contexto.

Informações específicas do usuário tais como preferências pessoais, devem ser facilmente comunicadas através das redes para sempre alcançar a ação necessária com o menor envolvimento do usuário.

2.2.11. Extensibilidade Plug-and-Play dos Serviços de Rede

Deve suportar extensibilidade da camada de controle para que seja possível adaptar-se aos diferentes cenários de rede. A extensibilidade da camada de controle deve ser organizada de uma maneira plug-and-play que requeira pouco ou nenhum esforço de configuração.

Neste aspecto, as Redes de Ambiente fornecem uma evolução da atual arquitetura IMS (IP Multimedia Subsystem), que utiliza servidores de aplicação para a implementação de serviços, que estende o conjunto básico de serviços do IMS.

2.2.12. Inovação e Utilidade de Aplicação

As Redes de Ambiente devem suportar pequenos ciclos de inovação e implementação de aplicações. A arquitetura deve ser capaz de atrair desenvolvedores de aplicações pela oferta de uma API que seja de fácil migração, estável e completa. A utilidade desta API é de importância fundamental para o sucesso das Redes de Ambiente.

É importante oferecer aos desenvolvedores de aplicações o suporte e o desempenho dos serviços de rede que eles precisam, desde a simples conectividade até a entrega eficiente de mídia.

2.3. O CONCEITO DE AMBIENT CONTROL SPACE (ACS)

A noção de Ambient Control Space (ACS) é introduzida para abarcar todas as funções de controle em certo domínio de rede. Uma Rede de Ambiente é a soma do Ambient Control Space mais a infra-estrutura de conectividade.

Exemplos de funções desempenhadas pelo ACS incluem suporte a mobilidade, gerência de rede, composição de rede, controle de QoS, suporte à entrega de mídia, etc. A Figura 2.2 ilustra a representação de um ACS, com seus componentes e pontos de referência.

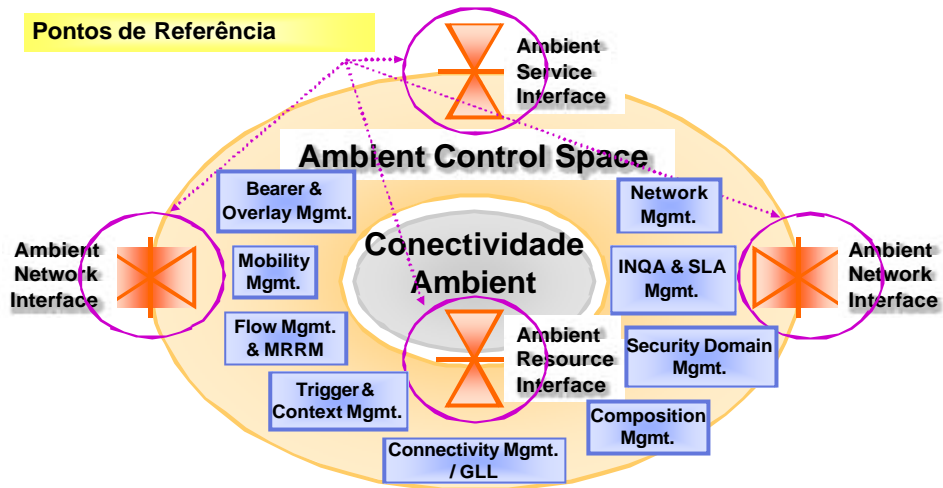


Figura 2.2. Representação esquemática do ACS.

A interface Ambient Network Interface (ANI) permite a cooperação através de diferentes ANs. A interface ANI é um ponto de referência que oferece uma maneira padronizada para conectar as funções de um ACS com funções de outro domínio, conforme Figura 2.3.

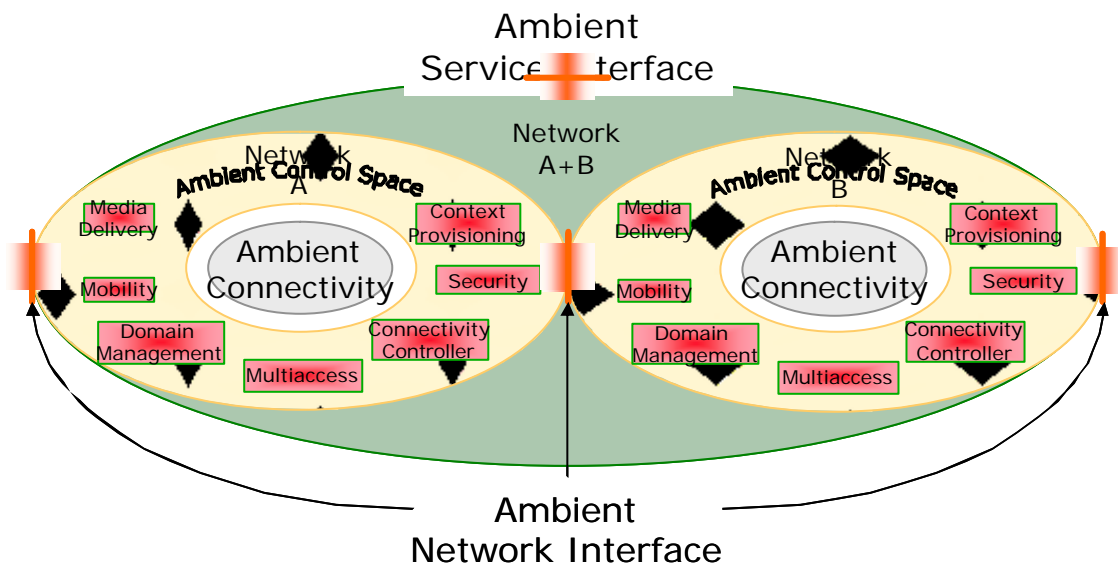


Figura 2.3. Conexão de ACSs através da interface ANI.

A interface ANI também é utilizada para anunciar a presença de funções de controle do ACS para domínios adjacentes.

O acesso aos serviços de uma AN acontece através da interface Ambient Service Interface (ASI). Mesmo em uma AN composta, este ponto de referência garante que apenas um ACS homogêneo seja visível para as entidades externas. Uma aplicação ou serviço, fazendo uso das funções no ACS, deve encontrar o mesmo ambiente, independentemente que qual interface ASI esteja conectada.

O ACS funciona baseado em uma abstração da infra-estrutura física de tal maneira que os recursos visíveis são apenas aqueles controlados pelo ACS. Os recursos têm acesso ao ACS através de um ponto de referência comum. Este ponto de referência é chamado de Ambient Resource Interface (ARI).

Uma AN é composta por dois componentes principais. O primeiro é o Ambient Connectivity, uma abstração da infra-estrutura de rede existente que é controlada pelo segundo componente, o Ambient Control Space. O ACS por sua vez é dividido em dois sub-componentes: as funções de controle (indicadas pelos retângulos na Figura 2.1) e as funções de suporte (framework functions) necessárias para que novas funções de controle consigam conectar-se e interagir com o ACS existente, executar suas tarefas de controle e coordenar com outras funções de controle.

As características chaves da arquitetura do ACS são a sua natureza distribuída e sua modularidade. Embora um pequeno conjunto de funções de controle precisa estar presente em qualquer ACS para que a rede seja uma rede “Ambient”, outras funções de controle podem ser adicionadas ou removidas do ACS durante a operação normal da rede.

Podemos imaginar um ACS como tendo duas metades. A primeira metade estabelece, mantém e movimentam a conectividade do plano de usuário e a segunda metade lida com a gerência de rede e de segurança, políticas e composição.

A parte do ACS que suporta e gerencia a conectividade do plano de usuário possui as seguintes funções:

- *Funções de gerência de comunicação e overlay* – oferece serviços de comunicação fim-a-fim para aplicações através da interface ASI. A gerência de overlay provê tanto comunicação básica fim-a-fim como serviços avançados de comunicação que

permitem a manipulação do fluxo de mídia ao longo da rede. Estes serviços avançados, chamados de redes overlay de serviços específicos, são construídos através de nós de processamento de mídia interconectados, chamados de Media Ports [4], que implementam transcodificação de mídia, caching e funções de sincronismo.

- *Funções de gerência de fluxo e de recursos de multi-rádio* – configura e estabelece fluxos. Multi-Radio Resource Management (MRRM) é a subclasse da gerência de fluxo responsável pelas decisões de handover para as partes de acesso via rádio. O procedimento de seleção de acesso decide (baseado em custo, QoS, políticas, etc.) qual dos fluxos de acesso disponíveis deve ser utilizado para fornecer um dado serviço de comunicação solicitado.
- *Funções de gerência de conectividade e de Generic Link Layer* – detecta novos enlaces e inicia o estabelecimento de conectividade entre as ANs. A generic link layer é uma subclasse da gerência de conectividade responsável por gerenciar os enlaces de acesso. Informação de topologia é enviada às entidades funcionais de gerência de comunicação e overlay para auxiliá-las em suas decisões. As solicitações de configuração e estabelecimento de fluxos são recebidas da gerência de fluxo e de recursos multi-rádio e os links são configurados correspondentemente via interface ARI. Os enlaces são monitorados através de medidas enviadas pela camada de conectividade.
- *Funções de gerência da mobilidade* – garante que o serviço de comunicação permanece independente de mudanças na conectividade ou de eventos de movimentos que ocorrem na camada de conectividade. O serviço de comunicação para as aplicações está amarrado a objetos de mais alto nível na estrutura de nomes que contém identificadores, tais como pontos de conexão para aplicações ou identificadores criptográficos para nós. A função chave da gerência de mobilidade é manter o mapeamento entre o localizador (locator) corrente e o identificador, permitindo a comunicação independente do atual ponto de conexão na rede. A entidade funcional de gerência de mobilidade toma decisões de handover baseada em triggers, informação de contexto e políticas.

- *Funções de gerência de triggering e contexto* – coordena o armazenamento, controle de acesso, disseminação e agregação de informação de contexto. As informações de contexto tais como triggers, preferências de usuário ou status da rede estão disponíveis na base distribuída de informação de contexto. Cada objeto é identificado por um identificador único e está registrado no coordenador de contexto. Os clientes de contexto podem recuperar informação de contexto diretamente da função de triggering e contexto ou são direcionados para a fonte do contexto que pode atender à solicitação. Também é possível receber atualizações contínuas sobre informações de contexto. As informações de contexto só podem ser acessadas mediante políticas explícitas para se garantir a proteção e privacidade das informações de negócio.

A segunda parte do ACS gerencia domínios administrativos. As principais áreas desta parte são:

- *Funções de gerência de domínio de segurança* – gerencia o registro e as políticas para domínios de segurança que são agrupamentos de recursos com um objeto de gerência comum de acordo com certas políticas de segurança. Os domínios de segurança permitem interações seguras dentro do grupo ou com outros objetos de uma maneira comum. A gerência de domínios de segurança fornece à gerência de composição informações sobre recursos disponíveis e políticas associadas e recebe informações sobre os acordos realizados.
- *Funções de gerência de composição* – governa o processo de composição através de coordenação com outras funções de controle, negocia com as ANs em composição e realiza os acordos de composição [2].
- *A gerência de rede* é realizada através de duas entidades. A gerência de plug-and-play lida com os processos relacionados à auto-configuração, como conectar um novo elemento, componente ou rede. A gerência de domínio acessa o banco de dados de políticas para suportar o processo de composição no plano de gerência. Ela lê o banco de políticas antes de um processo de composição e realiza uma negociação sobre o tipo de gerência de overlay. Finalmente, ela atualiza as políticas de rede em função dos resultados obtidos na composição.

- *Funções de SLA e INQA* (Inter-Network QoS Agreements) – estabelece e mantém acordos de nível de serviço (SLA) sobre recursos de conectividade.

Em termos de evolução de rede, a Figura 2.4 ilustra como o plano de controle representado pelo ACS poderia ser integrado à arquitetura IMS (IP Multimedia System) do 3GPP.

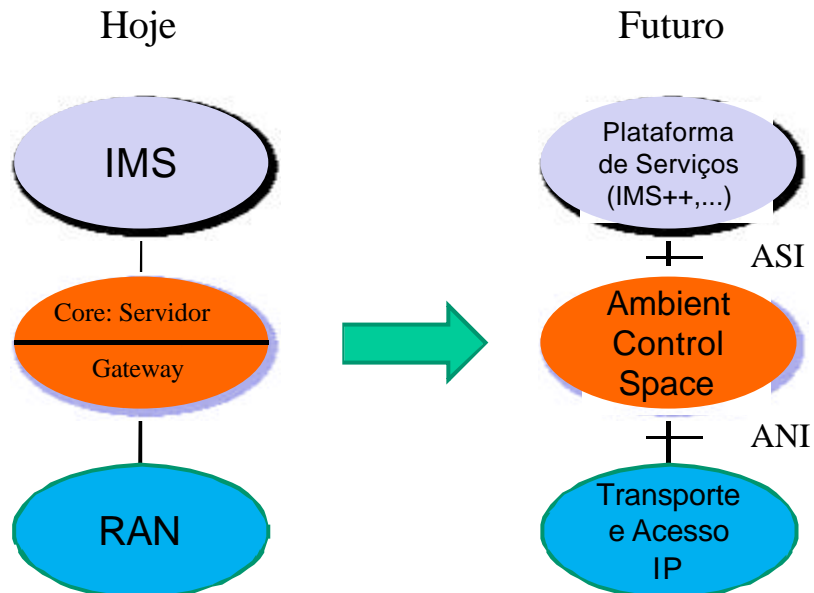


Figura 2.4. Integração do ACS com IMS.

2.4. O MODELO AMBIENT LAYER

O Ambient Control Space (ACS) precisa interagir com recursos do plano de usuário, que podem implementar várias tecnologias e serem acessíveis através de diversos protocolos. As funções de controle de uma AN devem permanecer tanto quanto possível independentes de tecnologias de rede específicas. Para cumprir com este objetivo, foi desenvolvido o Modelo de Camadas (Ambient Layer Model). O objetivo é permitir a interoperação com tecnologias legadas e uma fácil integração de tecnologias de rede futuras sem afetar os módulos de controle contidos no ACS.

O objetivo de permanecer independente de tecnologias específicas é alcançado através de uma abstração genérica do plano de usuário. As aplicações não devem conhecer sobre os

detalhes da camada de conectividade. As aplicações simplesmente criam uma entidade abstrata de conexão que atende suas necessidades de comunicação. As funções do ACS são responsáveis por estabelecer e manter esta conexão fim-a-fim através de qualquer tecnologia de conectividade. O modelo de abstração também garante que a implementação das funções de controle permaneça tão genérica quanto possível.

As abstrações de conectividade das ANs definem diferentes visões da conectividade física que fornece diferentes níveis de detalhamento. As funções de controle do ACS interagem com os recursos de rede através da interface ARI sem conhecer os detalhes de implementação ao configurarem os recursos de rede.

A transferência de dados propriamente dita através destes recursos é denominada de fluxo. Adicionalmente, o modelo de abstração fornece, para aplicações, serviços ou usuários interagindo com o ACS através da interface ASI, uma visão de mais alto nível da conectividade: o bearer é primitiva de comunicação deste nível. Ela esconde a implementação da conectividade e fornece apenas uma visão fim-a-fim.

2.4.1. Abstração de Fluxo

A seqüência de nós e enlaces que os pacotes de dados atravessam quando transmitidos entre dois nós é denominado de um caminho (path). Esta visão genérica serve de base para o modelo de abstração adotado pelas ANs.

As funções de controle do ACS, porém, não operam no nível de enlaces, nós e caminhos. A abstração de fluxo expõe apenas um subconjunto destas entidades. Como ilustrado na Figura 2.4, apenas algumas destas entidades do nível de conectividade são mostradas no nível de fluxo. As funções de controle do ACS operam sobre uma representação única da rede que é visível através da interface ARI.

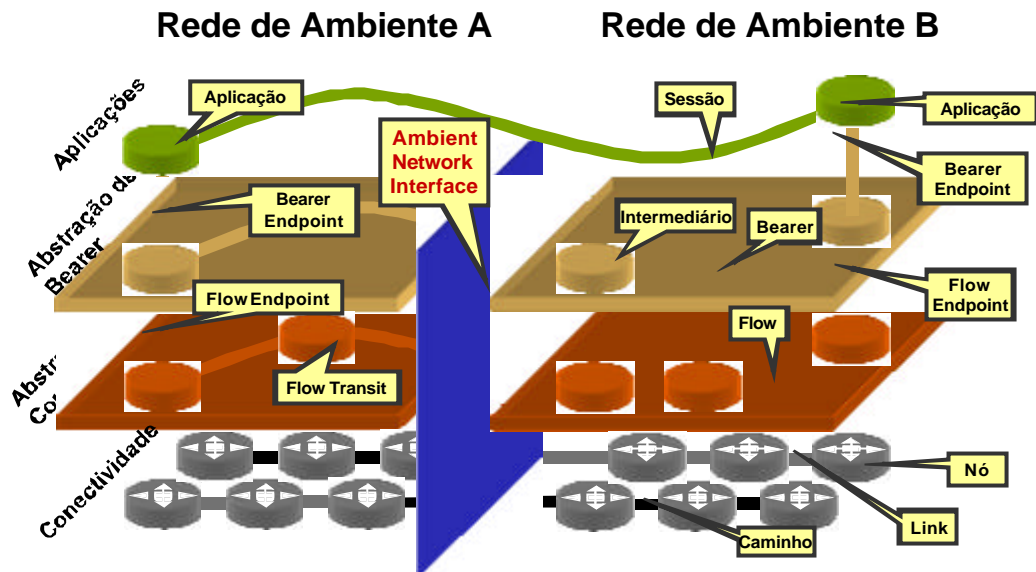


Figura 2.5. Representação das abstrações de conectividade.

Um fluxo é uma transferência de dados entre duas terminações de fluxo. Ele está delimitado a uma única tecnologia de rede e é definido como unidirecional. Um fluxo transfere dados transparentemente de acordo com certas características de desempenho. Um fluxo é, no mínimo, associado a um específico localizador de fonte e a um localizador de destino que estão ligados com as terminações do fluxo.

Um fluxo pode passar por recursos intermediários (flow transit) que não estão explicitamente ligados ao fluxo, mas que podem ser controlados através da interface ARI. O ACS pode usar as capacidades de controle e configuração da ARI para solicitar certos tratamentos do fluxo pelos recursos de conectividade.

A mobilidade de uma transferência de dados requer que um fluxo seja modificado ou que um novo fluxo seja criado e o antigo seja deletado.

2.4.2. Abstração de Bearer

As aplicações ou serviços operando acima do ACS não estão conscientes de detalhes como os fluxos que são usados para transportar os dados. Eles operam numa visão de mais alto

nível da conectividade denominada de *bearer*. O Bearer é um mecanismo de comunicação fim-a-fim que é oferecido na interface ASI.

As terminações do bearer, ao contrario dos fluxos, não estão ligadas a localizadores, mas a objetos de mais alto nível na estrutura de nomes. As funções do espaço de controle gerenciam o mapeamento de bearers em fluxos e eventuais atualizações neste mapeamento, isto é, em casos de eventos de mobilidade. Tais funções são capazes de construir um bearer a partir de um conjunto concatenado de fluxos usando a capacidade de resolução de nomes da estrutura de nomes do ACS. Múltiplos bearers podem ser mapeados para o mesmo fluxo. Atualmente o bearer é definido somente para unicast.

2.4.3. Sessão de aplicação

Acima da abstração de bearer, as aplicações podem definir sessões que combinem múltiplos bearers em entidades de transporte definidas pela aplicação. Por exemplo, uma aplicação poderia decidir solicitar diferentes bearers para tráfego de dados, voz e vídeo transferidos para uma aplicação peer de acordo com diferentes requisitos de cada tipo de tráfego. Para indicar que estes bearers estão relacionados, o termo sessão é usado.

2.4.4. Estrutura de nomes

A estrutura de nomes das ANs define mapeamentos dinâmicos entre os identificadores usados nos diferentes níveis das abstrações da conectividade. Com estes mapeamentos dinâmicos nos múltiplos níveis, os nomes dos objetos se tornam independente de sua localização e passa a ser possível o suporte à mobilidade em todos os níveis.

É tarefa do ACS gerenciar estes mapeamentos e fornecer mecanismos de resolução de nomes que permita mapear um nome de um objeto em uma camada superior a um identificador de uma camada mais baixa do objeto correspondente.

3. O FRAMEWORK PBMAN

Este capítulo apresenta um ambiente desenvolvido e implementado para tratar com a tarefa de gerenciamento em Redes de Ambiente, chamado Policy-Based Management for

Ambient Networks (PBMAN), descrevendo seus objetivos, terminologia e arquitetura. Mostra como a composição acontece no PBMAN considerando seus tipos, o Acordo de Composição e o processo de composição. Também detalha os middlewares XPeer e X-PBMAN, o protótipo de prova-de-conceito que implementa as principais funcionalidades do PBMAN.

3.1. INTRODUÇÃO

A complicada estrutura das redes sem fio, os atuais serviços Internet e os futuros serviços para os usuários móveis precisam ser gerenciados de uma maneira integrada e flexível. Eles necessitam de uma gerência distribuída escalável, que constitui um grande desafio.

A Gerência Baseada em Políticas (Policy-Based Management) [7] parece ser uma abordagem adequada para lidar com configurações dinâmicas e automáticas porque simplifica a administração de uma infra-estrutura de rede complexa de forma automática. As políticas descrevem como gerenciar e controlar o acesso à rede usando decisões e regras de alto nível. Formalmente, uma política é definida como um agregado de regras onde cada uma delas consiste de uma ou mais condições e ações.

O PBM foi desenvolvido pelo IETF [8] e é um modelo para gerenciamento baseado em políticas composto por quatro entidades: Policy Decision Points (PDPs), Policy Enforcement Points (PEPs), Policy Management Tool (PMT) e um repositório de políticas. O PDP é responsável por tratar as solicitações, interrogar o repositório de políticas, tomar decisões e distribuí-las aos PEPs. Os PEPs são entidades (por exemplo, roteadores) onde as ações são propriamente implementadas e/ou forçadas. O PMT suporta a especificação, edição e administração de políticas através de uma interface gráfica. Estas políticas são, então, armazenadas no repositório de políticas. Alguns protocolos são necessários dentro deste modelo, tais como COPS (Common Open Policy Service) para o interfuncionamento entre PDP e PEP, e LDAP (Lightweight Directory Access Protocol) para que o PDP possa acessar as políticas no repositório de políticas.

O modelo PBM do IETF não foi projetado para lidar com cenários heterogêneos onde frequentes mudanças acontecem, tais como lidar com usuários móveis com padrões de uso

altamente dinâmicos e necessidades de serviço imprevisíveis. A escalabilidade do sistema é uma limitação conhecida, uma vez que no modelo de duas camadas adotado pelo IETF um PDP somente pode controlar um número limitado de PEPs. Para resolver este problema, o modelo do IETF tem sido estendido para lidar com requisitos típicos de um cenário mais dinâmico. O Unified Policy-based Management (UPM) [9] propõe um modelo de três camadas, adicionando uma entidade intermediária entre os PDPs e os PEPs, conhecida como PEA (Policy Enforcement Agent). Este modelo hierárquico tem algumas limitações, uma vez que não abandona o modelo cliente/servidor.

Para lidar com os desafios das Redes de Ambiente, um novo modelo de gerenciamento baseado em políticas foi desenvolvido. Existem três principais motivações para propor um novo modelo. Primeiro, o modelo do IETF está focado em áreas específicas, tais como QoS e segurança, e em problemas mais simples de rede corporativas típicas. O escopo de aplicações para o gerenciamento baseado em políticas tem aumentado considerando o uso de serviços na Internet global. Segundo, o novo modelo considera cenários 3G/4G compostos por um imenso número de usuários móveis com mobilidade e padrões de uso de serviço altamente dinâmicos. O PBM tradicional não provê compartilhamento de informação dinâmica entre os domínios de política. Os requisitos especificados para Redes de Ambiente nos dá uma idéia da complexidade do novo ambiente. E terceiro, o novo modelo atende a requisitos importantes associados com estes novos ambientes, tais como escalabilidade, tolerância a falhas e balanceamento de carga, usando para isso a tecnologia Peer-to-Peer (P2P).

3.2. DEFINIÇÃO

O modelo PBMAN (Policy-Based Management for Ambient Networks) [11][10] tem o objetivo de desenvolver e implementar uma infra-estrutura para Redes de Ambiente. O PBMAN adota a técnica do PBM e a principal tecnologia é a tecnologia Peer-to-Peer (P2P).

De fato, o PBMAN é uma instância de um modelo mais abstrato chamado de P4MI (Peer-to-Peer Policy Management Infrastructure) [11]. Um princípio básico de projeto adotado

no PBMAN é manter a arquitetura genérica e simples. À medida em que novas experiências foram sendo realizadas com a implementação do modelo, novos recursos e funcionalidades serão adicionadas.

A Figura 3.1 ilustra a arquitetura básica do PBMAN que consiste de Policy Decision Network (PDN), Policy Enforcement Points (PEPs), usuários e suas interações. Os PEPs e usuários são chamados de agentes. Esta arquitetura é focada no papel e implementação do ACS nas várias redes em um cenário de Redes de Ambiente. Existem três tipos de implementação para o ACS: PDN ACS, User ACS e PEP ACS. O User ACS e o PEP ACS podem ser parte de um ACS Agente combinado.

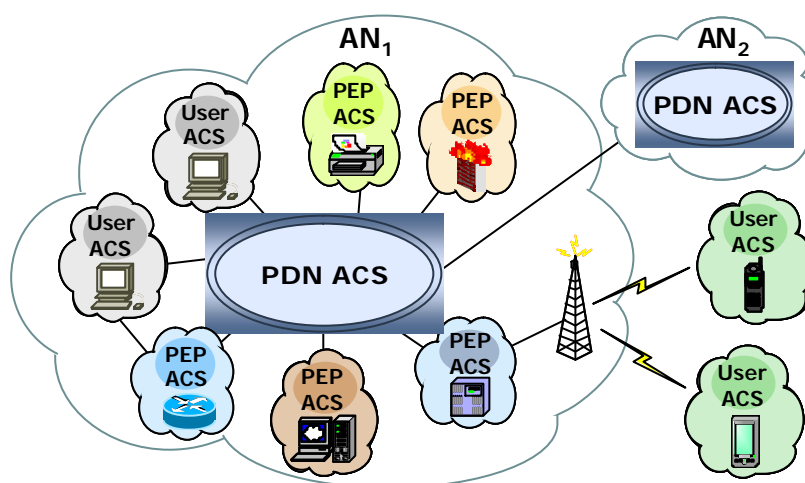


Figura 3.1. Arquitetura geral do PBMAN.

A figura mostra que os usuários e os PEPs são considerados casos especiais de redes pequenas. Existem muitos usuários e PEPs que estão conectados entre si e com a PDN por meio de uma rede fixa AN1. Por outro lado, AN1 tem alguns usuários móveis conectados à sua PDN, que podem ser usuários locais ou visitantes. A rede AN1 deve fornecer de forma transparente acesso e serviços para estes usuários de acordo com seus perfis, independentemente se eles são fixos ou móveis, locais ou visitantes. Para ser capaz disso, a rede AN1 deve trocar informações com outras ANs, e, portanto é necessária composição e decomposição.

O tipo de ACS necessário em qualquer situação é influenciado pelas diferentes funcionalidades de hosts, equipamentos ou redes que participam do PBMAN:

- Capacidade de processamento: uma baixa capacidade de processamento limita as funcionalidades do ACS, por exemplo, no caso de um PDA ou um telefone móvel.
- Memória e capacidade de armazenamento: aparelhos com memória e/ou disco pequenos, podem tornar difícil a implementação de um ACS com muitas funcionalidades.
- Padrão de mobilidade: terminais móveis que estão constantemente em movimento através de diferentes locais geográficos e que estão usando infra-estruturas de Redes de Ambiente não são adequados para assumir grandes responsabilidades no armazenamento e recuperação de políticas e informação de gerenciamento de outras redes.
- Conectividade e disponibilidade: este item está relacionado ao padrão de mobilidade, uma vez que usualmente uma alta mobilidade implica em conectividade instável e baixos níveis de disponibilidade.
- Capacidade de comunicação: um host com alta capacidade de processamento e com muita memória pode não ser apto a contribuir o processo de gerenciamento como um todo se ele estiver conectado a uma rede sem fio com baixa velocidade.
- Nível de dedicação: um servidor totalmente dedicado é preferível para assumir um papel importante no ACS do que um servidor que compartilhe os ciclos de processamento com outras tarefas.

É óbvio que um ACS de uma rede grande (um ACS tipo PDN) requer servidores com grande poder de processamento, muita memória e disco com grande capacidade de armazenamento, nenhuma mobilidade, alta disponibilidade, grande capacidade de comunicação e baixo nível de compartilhamento do processamento com outros serviços. Por outro lado, terminais móveis típicos são capazes de implementar apenas pequenas versões do ACS, gerenciando suas próprias necessidades e dependendo do PDN ACS para todo o resto.

3.3. POLICY DECISION NETWORK

A Policy Decision Network (PDN) é responsável por implementar a maioria dos conceitos e funcionalidades do modelo PBMAN, relacionados à gerência baseada em políticas e comunicação peer-to-peer, tais como armazenamento e recuperação de políticas e tomada de decisão sobre as solicitações recebidas. Em outras palavras, a PDN é uma implementação distribuída do ACS baseada no modelo Peer-to-Peer para o PBMAN. A Figura 3.2 ilustra a estrutura da PDN.

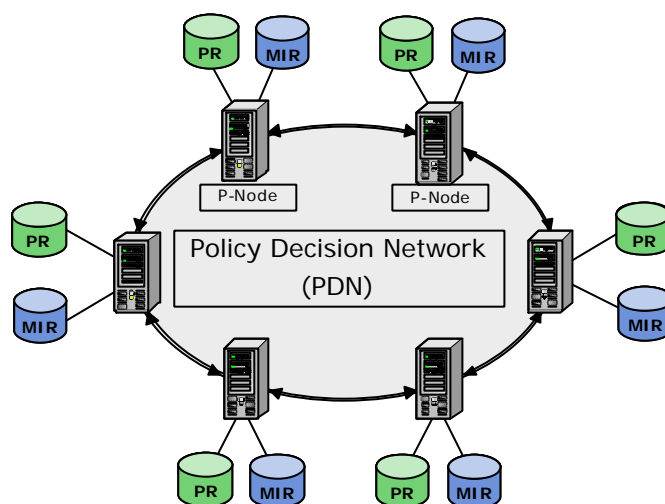


Figura 3.2. Policy Decision Network.

A PDN é composta por duas entidades principais – Pontos de Decisão e Repositórios. Um Ponto de Decisão, também chamado de PDN Node ou P-Node, é um servidor de política que assume parte do trabalho de uma PDN. A PDN é capaz de interoperar com outros P-Nodes por meio de uma rede Peer-to-Peer distribuída, baseada no padrão Distributed Hash Tables (DHT) [12], denominada de PDN-ring. A PDN-ring provê o PBMAN com as funcionalidades inerentes de sistemas P2P, tais como balanceamento de carga, tolerância a falhas e escalabilidade. Esta é a principal diferença entre uma PDN no modelo PBMAN e um PDP no modelo cliente/servidor PBM do IETF que foi projetado para interagir com um

conjunto de PEPs e não está intrinsecamente consciente da existência de outros PDPs, por exemplo, para o propósito de balanceamento de carga e tolerância a falhas.

A PDN tem dois repositórios de informação – o Policy Repository (PR) e o Management Information Repository (MIR). O PR armazena políticas de acordo com alguns requisitos, tais como facilitar o processo de pesquisa e recuperação de políticas. O PBMAN não especifica nenhuma tecnologia de armazenamento em particular, tais como LDAP, para que diferentes implementações possam interoperar. Além das políticas propriamente, existe uma necessidade de manter informações sobre entidades que são gerenciadas com políticas. Estas informações são armazenadas no MIR. Um exemplo de uma informação típica armazenada no MIR são informações de perfil (profile) para agentes e destinos, políticas para mapeamento e configuração de equipamentos e informações de controle e gerência da própria PDN.

3.4. POLICY AGENTS

Policy Agents (PA) são representados por hosts, equipamentos ou terminais utilizados por usuários ou pela rede para prover serviços e aplicar políticas. A interação entre agentes e a PDN é baseada na abordagem de uma rede P2P hierárquica do tipo DTH.

Existem dois tipos de agentes no PBMAN: PEPs e Usuários. Os PEPs são agentes dedicados a aplicar as políticas, tais como roteadores, firewalls e servidores de acesso remoto. Os agentes PEP são software e hardware para provimento de serviços, com a função de aplicar políticas de direito de uso, segurança, tarifação, etc. Exemplos deste tipo de PEP são servidores de impressão e de jogos. Os agentes Usuários representam aparelhos ou redes de aparelhos conectados que um dado usuário real está usando para acessar os serviços e recursos de uma AN.

Um ACS Agente roda em terminais de usuário e em PEPs. No PBMAN, User ACS e PEP ACS são considerados clientes do PDN ACS. O PEP ACS possui entidades que forçam a aplicação de políticas de uso, por exemplo, segurança, QoS e perfil de usuário.

A principal diferença entre um PDN ACS e um ACS Agente é que o PDN ACS é usado para armazenar e/ou prover políticas de toda uma AN, enquanto que o ACS Agente armazena e provê políticas apenas para si mesmo. Em outras palavras, ele pode solicitar políticas de diferentes destinos dos PDN ACSs ou Agente ACSs, mas não é capaz de provê-las para outros agentes.

3.5. POLICY INFORMATION MODEL

O modelo de informação apresenta a representação abstrata de entidades gerenciadas e como elas estão relacionadas umas às outras. O objetivo de um modelo de informação para PBMAN é permitir que o processamento de políticas e que as decisões sobre uso de serviços sejam realizadas de forma flexível e eficiente. A Figura 3.3 ilustra o modelo de informação.

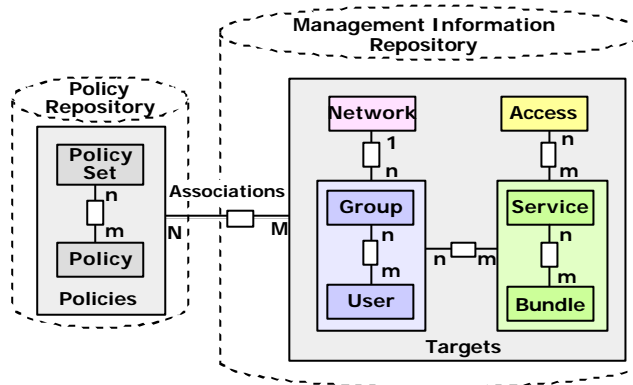


Figura 3.3. Modelo de Informação do PBMAN.

O modelo de informação do PBMAN consiste em três tipos principais de entidades de gerenciamento: políticas, destinos (targets) e associações. Os destinos são entidades, tais como serviços ou usuários (policy agents e targets), para as quais é possível associar políticas. Um passo importante em um sistema de gerenciamento baseado em políticas é a associação entre políticas e destinos. No modelo PBMAN, destinos são: usuário, grupo, rede (AN), serviço, grupo de serviços e classe de acesso. A classe de acesso é uma forma de classificar tecnologias de acesso por alguma funcionalidade comum, que é independente

de tecnologia. Um Policy Set é um grupo de políticas que serve para propósitos de organização.

Este modelo de informação utilizado no protótipo X-PBMAN e no cenário de vídeo apresentado na seção 3.8 é baseado num subconjunto do modelo de informação do PBMAN. Este modelo não é uma solução final e geral para Redes de Ambiente, uma vez que outros modelos podem ser sugeridos.

Existem dois tipos de serviços no PBMAN: serviços locais e Well Defines Services (WDS). Os serviços locais são usados para organizar usuários em serviços (por exemplo, assinantes em planos para um provedor) e é um conceito chave no PBMAN para permitir o entendimento e o provisionamento automático de serviços no momento de uma composição. Os serviços WDS são usados para garantir que um entendimento comum sobre as semânticas de um serviço seja alcançado por todas as redes que participam de uma entrega automática de serviço. Cada WDS necessita de uma especificação formal, similar a um SLA, onde um serviço é definido por diferentes parâmetros. Por exemplo, um serviço de QoS pode ser definido em termos de atraso, perda de pacotes e vazão.

3.6. COMPOSIÇÃO NO PBMAN

Composição de rede é a contribuição mais significativa da abordagem das Redes de Ambiente para a área de redes móveis e sem fio. Espera-se a composição de rede seja extensivamente utilizada em Redes de Ambiente. Os usuários terão acesso aos recursos e serviços da rede, a qualquer tempo, em qualquer lugar e através de qualquer tecnologia de acesso. Portanto, ela deve ser realizada tão rápida e eficiente quanto possível, de maneira a fornecer serviços sem interrupção aos usuários, bem como não impor um pesado custo de processamento e capacidade de comunicação ao sistema de gerenciamento.

A arquitetura do PBMAN é desenhada para facilitar operações como fusão de PDNs em uma única PDN e divisão e agrupamento de PDNs separadas.

3.6.1. Classes de Composições

Buscando objetivos de eficiência e desempenho, o modelo PBMAN identifica alguns tipos distintos de composições, para que seja possível otimizar cada uma delas. As composições são classificadas no PBMAN de acordo com o tipo de ACS envolvido (PDN e Agente). Três tipos diferentes de composição são possíveis considerando este critério: PDN/PDN, PDN/Agente e Agente/Agente.

3.6.1.1. Composição tipo PDN/PDN

Este tipo de composição requer a criação de um PDN ring. Alguns exemplos deste tipo são composições de redes de provedores (de qualquer tipo), redes privadas, redes caseiras e redes em movimento (carro, ônibus, trem e avião).

A Figura 3.4 ilustra uma típica composição PDN/PDN. A situação antes da composição é mostrada na Figura 3.4(a). Ambas PDN_A e PDN_B são PDNs simples (não composta com nenhuma outra PDN, mas provavelmente compostas com agentes), cada uma com quatro P-Nodes. Durante a composição, um novo PDN ring (PDN_{AB}) é criado e dois P-Nodes de cada PDN são escolhidos como membros da PDN_{AB} , como mostrado na Figura 3.4(b).

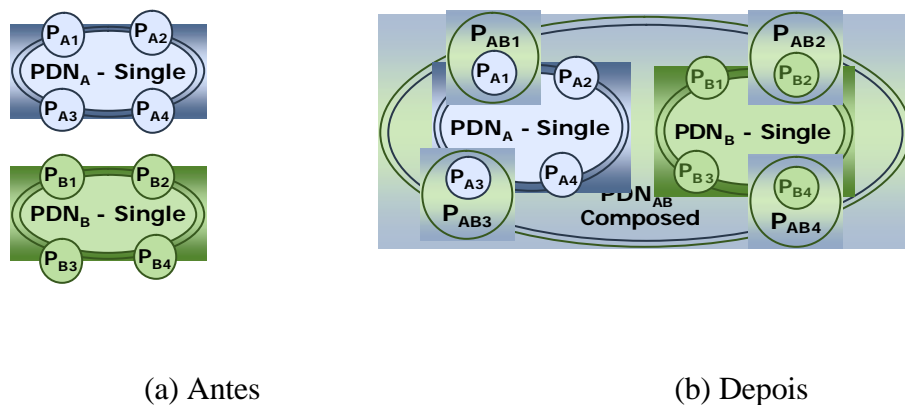


Figura 3.4. Composição de PDN_A e PDN_B .

O número de P-Nodes que farão parte da nova rede composta depende de uma política local. Uma vez que qualquer P-Node membro da nova composição pode atuar como uma gateway entre redes, tolerância a falhas e escalabilidade são obtidas de forma transparente, mesmo quando compondo múltiplas hierarquias de rede.

Um aspecto importante sobre composições do tipo PDN/PDN é que elas podem exigir um tempo considerável para serem realizadas, contudo elas tipicamente acontecerão apenas quando o primeiro usuário tenta acessar serviços de uma rede remota. Esse tempo é necessário para permitir a autenticação e autorização do P-Node de uma rede local baseada nas políticas da rede remota, a partir da solicitação do usuário. Para todos os acessos subseqüentes, o P-Node terá acesso instantâneo às informações da rede remota e a resposta, portanto, deve ser imediata. Por esta razão, espera-se que composições e decomposições deste tipo não aconteçam freqüentemente, uma vez que serão mais estáveis.

O ciclo de vida de uma composição entre PDNs é composto por seis fases: triggering (disparo), negociação, estabelecimento (setup), utilização, disparo de decomposição e execução da decomposição. O disparo da composição é o evento que motiva a nova composição, por exemplo, uma solicitação de serviço de um usuário para um PEP e deste para a PDN para processamento de política. Antes que se inicie a recuperação da política, o P-Node deve checar se é necessária uma composição. O disparo de decomposição pode ser um timeout quando a rede entende que a composição não está mais sendo usada.

Uma versão simplificada do processo de negociação é apresentada no seguinte algoritmo. Note que não estão incluídas situações não usuais, tais como erros, recusa da composição e renegociação.

- Uma negociação de composição inicia quando um P-Node da PDN de origem envia uma mensagem de solicitação de composição para um P-Node selecionado da PDN destino. Estes dois P-Nodes irão negociar a composição, cada um representando sua rede;
- A solicitação de composição é aceita, o P-Node de origem cria um novo PDN ring e publica a lista de P-Nodes que farão parte do PDN ring. Esta lista é obtida pela combinação da própria lista de P-Nodes de sua rede local com a lista da rede destino (remota) obtido com a mensagem de aceite;

- Os P-Nodes de origem e de destino enviam mensagens de join para todos os P-Nodes de suas PDNs que farão parte da nova PDN composta;
- Todos os P-Nodes aderem à nova PDN e republica suas políticas e outras informações relacionadas a políticas. Antes disso, os P-Nodes precisam certificar-se de que um determinado número de P-Nodes já aderiram à PDN, a fim de evitar redistribuição excessiva de chaves na DHT (isso acontece quando novos parceiros se juntam à rede DHT). A decisão de quando republicar a informação é baseada no percentual de P-Nodes que já se juntaram à rede. Se este percentual é atingido, a informação é republicada. Senão, o P-Node aguarda um tempo randômico e verifica novamente se o novo ring está pronto. Existe um mecanismo de temporização para lidar com a situação onde alguns P-Nodes levam muito tempo para se juntar à nova PDN.
- A composição é concluída quando todas as informações dos P-Nodes são republicadas na nova PDN, ou um timeout foi excedido. Quando a negociação da composição é disparada por uma solicitação de serviço de um usuário, o P-Node origem precisa saber quando a composição está pronta para ser usada.

Foi especificado um protocolo simples para composição no protótipo X-PBMAN e implementadas extensões a este algoritmo básico. Este protocolo estendido é usado no cenário apresentado na seção 3.8.

3.6.1.2. Composição tipo Agente/PDN

Este tipo de composição acontece para prover comunicação entre agentes e a PDN. Usos típicos são: a solicitação de serviços e recursos para usuários e a solicitação de políticas para PEPs, ambos da PDN.

Este tipo de composição é mais simples do que a composição PDN/PDN, uma vez que o agente conecta-se à PDN e não existe criação de PDN ring, eliminando-se o custo

associado da republicação de informação. Em sua forma mais básica, pode ser entendida como alguma forma de login, isto é, um processo de autenticação no qual o usuário ou PEP está apto a solicitar acesso a serviços. A solicitação pode ser enviada junto com a autenticação, isto é, a composição pode ser disparada por uma solicitação de serviço. A autenticação pode ser vista como o uso de uma senha, uma chave privada ou mesmo um token. A interação entre PDN e PEP só é permitida após composição do PEP com a PDN.

O exemplo básico de composição será a provisionamento do simples serviço de conectividade, tais como:

- Um usuário local compõe com a rede local através de DHCP.
- Um usuário remoto compõe com a rede remota através de PPP e protocolos de autenticação.

O acesso à Internet pode ser restringido pelo uso de um proxy. Os usuários podem acessar a rede externa (Internet ou intranet ou extranet) somente após serem autenticados com o proxy.

Usando a tecnologia IEEE 802.1x, somente usuários autenticados podem ter seus pacotes encaminhados através do ponto de autenticação, que pode ser um switch Ethernet para estações fixas ou um Access Point (AP) para aparelhos sem fio.

Um usuário pode ou não fazer o login na PDN (X-Peer) durante uma composição. Para ter acesso apenas a recursos e serviços (como conectividade básica), um usuário não precisa fazer login na PDN, mas apenas ser autenticado pelo PEP que provê aquele tipo de serviço. Por exemplo, em alguns casos, um PEP firewall pode informar ao usuário que se ele deseja acesso a qualquer serviço daquela rede, ele deve compor com a PDN. Em outros casos, apenas a autenticação com o PEP seja necessária. A questão é que quando um usuário faz um login numa PDN pode ser necessária a manutenção de alguma informação de estado e isso pode ser uma perda desnecessária de recursos. Um dado PEP pode fazer o login no lugar do usuário. Neste caso, o status do usuário na rede é “composto”, significando “online”, muito embora ele não mantenha uma conexão.

3.6.1.3. Composição tipo Agente/Agente

Este tipo de composição inclui as interações típicas de host para host e host para equipamentos existentes na atual Internet. Também são consideradas interações de agentes de forma ad-hoc.

A composição Agente/Agente é uma composição simples na qual nenhuma PDN é envolvida, não havendo, portanto, nenhum processo de login. Uma vez que os agentes podem ser representados por aparelhos de usuário de baixa capacidade, tais como PDAs e smartphones, não espera-se que eles implementem uma PDN. Além disso, podem existir diferenças entre composições do tipo Usuário/Usuário e Usuário/PEP. A primeira acontece quando usuários decidem colaborar de alguma maneira. A última, quando um usuário quer acessar algum serviço que é fornecido através de um PEP.

Pelos menos dois pontos importantes devem ser definidos para uma composição Agente/Agente:

- Descoberta de agente para negociação da composição: um protocolo de broadcast pode ser usado, como DHCP, para encontrar outros agentes. Um agente pode enviar mensagens de broadcast em qualquer interface de rede.
- Manutenção da composição: pode ser gerenciada por um mecanismo do tipo Gnutella, isto é, por meio de um protocolo de flooding. As composições Agente/Agente ainda estão num estágio inicial de desenvolvimento e necessitará de trabalhos futuros.

3.6.2. Acordo de Composição

Antes que uma composição aconteça deve haver uma negociação das políticas da composição e as redes envolvidas realizam um acordo (Composition Agreement – CA) representando a realização de uma composição.

Um Acordo de Composição no PBMAN é o conjunto de todas as políticas (usuário, suporte e composição), destinos e suas associações disponíveis em uma rede composta. As políticas controlam várias decisões sobre a composição como a construção do acordo de composição durante a negociação da composição entre duas ou mais redes.

As políticas controlam cada aspecto envolvido no compartilhamento de recursos e serviços dentro do PBMAN. Portanto, elas também controlam o processo de composição. Alguns aspectos relacionados à composição que podem ser controlados pelas políticas são:

- Indicar as condições para aceitar uma solicitação de composição baseado em diversos critérios, tais como critérios estratégicos, administrativos, financeiro ou técnico. Um exemplo de uma política pode ser: “Composição (com qualquer rede) somente é aceita se um mínimo de padrões de segurança estiver disponível”;
- Determinar o tipo de modelo de composição entre duas ou mais redes, tais como Integração de Rede, Controle Compartilhado ou Interfuncionamento de Rede;
- Decidir quais destinos (serviços, usuários, etc) e/ou recursos e políticas correspondentes estarão disponíveis na nova rede composta;
- Decidir como tratar com políticas conflitantes durante a negociação da composição. O resultado final tem que ser um estado consistente e aceitável para todas as redes envolvidas na composição;
- Gerar políticas de composição para a nova rede composta, decidindo como uma entidade compartilhada será propagada para outras composições onde uma rede composta estará envolvida;
- Decidir quais P-Nodes farão parte da nova PDN. Por exemplo, o número de P-Nodes pode ser proporcional (um percentual) ao tamanho das redes, as redes podem requerer um número mínimo de P-Nodes para aceitar a composição ou as redes podem escolher os P-Nodes pela proximidade aos PEPs, etc.;
- Mudanças no Acordo de Composição tais como o critério usado para definir se uma rede pode entrar ou deixar uma composição ou uma atualização no conjunto de serviços que uma rede provê na composição;

- Determinar regras para a decomposição. O evento de disparo é também definido por políticas e alguns disponíveis são: um temporizador após o último serviço ser usado, uma rede falha inesperadamente e isso representa uma decomposição, uma rede abandona uma composição e as redes restantes decidem que a composição não pode ser mantida por nenhuma razão ou uma rede detecta que outra rede não é mais confiável, porque está se comportando de maneira suspeita.

3.7. X-PBMAN

Um protótipo simplificado de prova-de-conceito do PBMAN, chamado de X-PBMAN, foi desenvolvido para explorar e entender a abordagem combinada de Policy-Based Management e Peer-to-Peer adotada no PBMAN. O X-PBMAN foi implementado usando o middleware peer-to-peer X-Peer [13], descrito na próxima seção.

3.7.1. X-PEER Middleware

X-Peer é um middleware projetado e implementado para suportar aplicações Peer-to-Peer (P2P), baseado numa arquitetura hierárquica onde super-peers (X-Peer nodes) comunicam-se entre si através de uma rede DHT. Esta característica permite aos nós X-Peer encontrar qualquer informação na rede em até $\log(n)$ saltos (hops), onde n é o número de super-peers. A principal vantagem desta proposta é a garantia da localização da informação numa rede distribuída e hierárquica. X-Peer tem o objetivo de suportar diferentes aplicações P2P em uma única plataforma de middleware, se diferenciando, desta forma, das atuais soluções que normalmente usam uma nova rede para cada nova aplicação. A atual implementação do X-Peer utiliza o FreePastry [14], uma implementação de código aberto do Pastry [15].

A Figura 3.5 ilustra a arquitetura X-Peer. Existem três níveis de comunicação: nós X-Peer se comunicam através de uma rede DHT; Aplicações se comunicam entre si através de um

protocolo P2P. Aplicações se comunicam aos X-Peers por meio de um protocolo X-Peer específico, através do qual o middleware X-Peer provê um conjunto de serviços às aplicações, tais como Join, Leave e Search.

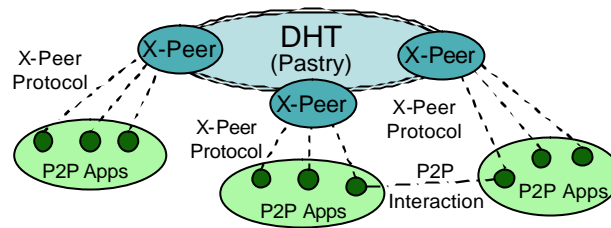


Figura 3.5. Arquitetura X-Peer.

Este middleware é suficientemente flexível para ser capaz de rodar como um sistema centralizado quando usando um único super node ou para rodar como um sistema P2P distribuído quando um nó X-Peer está co-localizado com cada aplicação P2P. Em outras palavras, é muito simples adicionar um novo nó X-Peer, de tal forma que se pode conseguir facilmente balanceamento de carga, tolerância a falhas e escalabilidade.

Esta característica da arquitetura X-Peer favorece seu uso na implementação do modelo PBMAN. Os agentes de política podem ser mapeados em aplicações, ao passo que os P-Nodes podem ser mapeados em nós X-Peer.

3.7.2. Arquitetura de Software

A arquitetura P-Node está ilustrada na Figura 3.6. Ela foi projetada e implementada para o protótipo X-PBMAN que realiza algumas das funções do PDN ACS. Algumas entidades externas com as quais o P-Node tem que interagir estão mostradas na figura. A versão atual da arquitetura é composta por seis módulos representando as Entidades Funcionais (Functional Entities – FE) do ACS. Um módulo específico de software do X-PBMAN implementa cada FE.

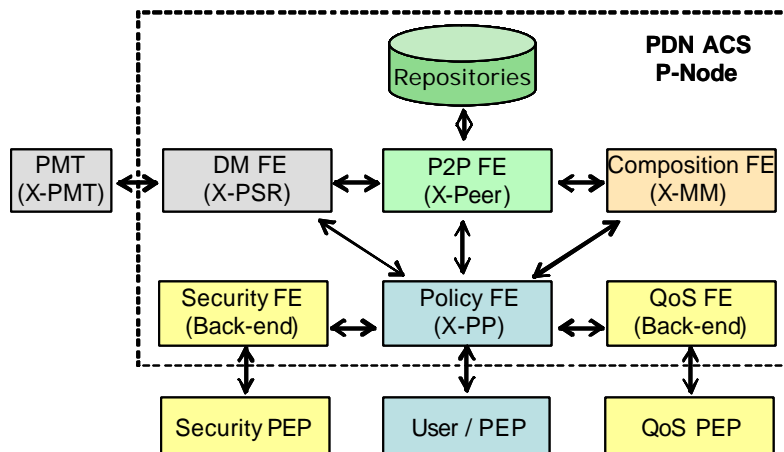


Figura 3.6. Arquitetura P-Node.

Alguns destes módulos rodam em espaços de endereçamento separados (processos) ao passo que outros rodam junto com o sistema principal. Uma descrição sucinta de cada um é dada a seguir.

- P2P-FE: é a função principal do P-Node, implementada diretamente pelo nó X-Peer. Os P-Nodes se comunicam entre si usando as facilidades fornecidas pelo middleware X-Peer.
- Composition FE: esta entidade funcional é implementada como um módulo de software agregado ao X-Peer, rodando na mesma máquina, mas em um espaço de endereçamento distinto. É desenvolvido como uma simples aplicação X-Peer. Este módulo é chamado de X-Peer Multi-ring Manager (X-MM), porque uma composição implica em criar um novo PDN ring (rede DHT) e gerenciar dois ou mais anéis simultaneamente.
- Policy FE: esta entidade funcional é responsável por lidar com as solicitações de serviço e processá-las. Ela é implementada como uma aplicação X-Peer, chamada de X-Peer Policy Processing Module (X-PP).
- DM FE (Data Management FE): é responsável por construir um modelo de dados mais sofisticado sobre modelo simples baseado no armazenamento e recuperação

de chave, fornecido pela funcionalidade X-Peer DHT. Ela funciona como uma API, oferecendo serviços de dados a outros módulos que precisam lidar com dados mais complexos. A DM FE também é uma aplicação X-Peer, chamada de X-Peer Policy Storage and Retrieval Module (X-PSR).

- QoS FE e Security FE: estas FEs são parte do back-end do sistema de políticas e são implementadas como classes do Policy FE, incluindo todos os métodos back-end dedicados a lidar com QoS e segurança, tais como configurar um QoS específico ou Security PEPs.
- Repositórios: o Policy Repository (PR) e o Management Information Repository (MIR) são implementados pelo método básico de armazenamento X-Peer storage.
- PMT: provê funcionalidades para edição de políticas e informação de gerenciamento (tais como targets) e seus relacionamentos. Sua implementação no X-PBMAN é denominada de X-PMT (X-Peer Policy Management Tool).
- QoS PEP e Security PEP: são módulos específicos responsáveis por serviços de configuração. Por exemplo, o QoS PEP pode marcar pacotes para a classe EF PHB do padrão DiffServ e o Security PEP pode abrir um túnel VPN, a partir de decisões tomadas pelos QoS FE e Security FE.
- Outros PEPs e Usuários: os usuários querem tipicamente compor com a PDN e os PEPs demandam decisões dela.

3.8. CENÁRIO PBMAN

Para analisar a efetividade da abordagem PBMAN, foi escolhido um serviço de Vídeo sob demanda (Video on Demand – VoD) . Este cenário foi modelado, implementado e testado usando o protótipo X-PBMAN. Ele consiste de duas redes, AN_1 e AN_2 , não diretamente conectadas, ambas provedores do serviço VoD. Usuários de uma rede podem assistir vídeos da outra rede, desde que as mesmas estejam compostas. Existem três Serviços Bem Definidos (Well-Defined Services – WDS): serviço de vídeo, de QoS e de criptografia.

Comercialmente, pode-se vender uma combinação destes serviços em quatro níveis de serviço distintos, isto é, vídeo com ou sem QoS e criptografia. Por exemplo, os serviços podem ser denominados de Vídeo Básico, Vídeo Premium (com QoS), Vídeo Seguro e Vídeo Seguro Premium. Os direitos de acesso ao serviço de vídeo e às funções de configuração de QoS e segurança são controlados por políticas.

3.8.1. Transação do Serviço de Vídeo

A Figura 3.7 ilustra uma transação típica para o serviço Vídeo Premium (com QoS), começando quando o usuário acessa a página web do provedor de vídeo até quando ele/ela obtém o stream de vídeo. As entidades responsáveis pelas funções de segurança (criptografia) não estão representadas na figura para não carregar muito a figura e facilitar o entendimento.

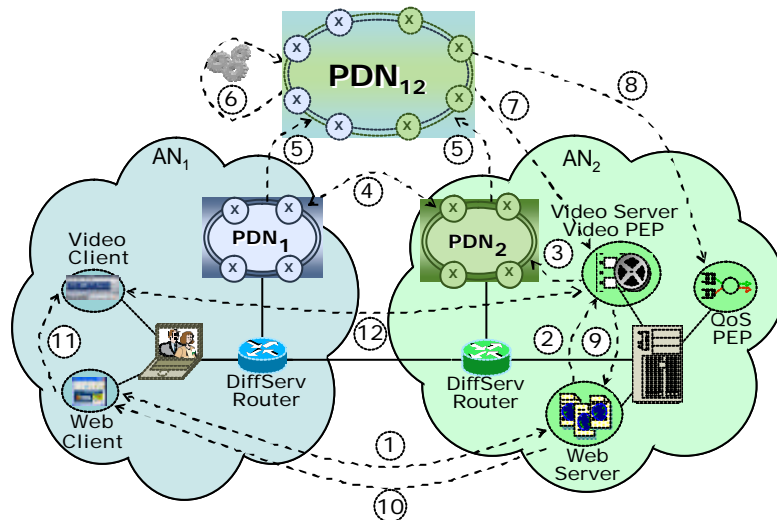


Figura 3.7. Transação do Serviço de Vídeo.

Uma transação de vídeo típica quando um usuário remoto da rede AN₁ acessa o servidor de vídeo na rede AN₂ é composta por doze passos:

1. O usuário inicia a transação ao acessar a página web do provedor de VoD;
2. O servidor web encaminha automaticamente a solicitação ao PEP de vídeo;
3. O PEP de vídeo envia uma solicitação à PDN;
4. A PDN2 detecta que o usuário é assinante da AN1 e que é necessária uma composição antes do processo de seleção de política. Então, a PDN2 contata a PDN1 e negocia os termos e condições da composição;
5. Caso a negociação da composição seja bem sucedida, as PDN2 e PDN1 criam um novo PDN ring e republicam todas as informações relevantes. Usando esta nova PDN composta, o P-Node é capaz de saber se o usuário contratou o serviço ou não;
6. A PDN inicia um processo de seleção de política. Ela decide se deve fornecer o serviço Básico ou Vídeo Premium ou se o serviço deve ser negado. Assumamos que o serviço Vídeo Premium seja autorizado;
7. A PDN envia uma resposta ao PEP de vídeo;
8. A PDN instrui o PEP de QoS para marcar os pacotes do stream de vídeo para DiffServ EF PHB. As garantias de QoS são aplicadas por roteadores DiffServ;
9. O PEP de vídeo responde ao servidor web, o qual dará continuidade à sua sessão pendente com o cliente web;
10. O servidor web apresenta ao usuário uma lista de vídeos;
11. Ao selecionar um canal de vídeo, o cliente web lança automaticamente o cliente de vídeo;
12. O cliente de vídeo contata o servidor de vídeo e a sessão de stream de vídeo é iniciada.

Este cenário foi modelado usando o modelo de informação de política do PBMAN e as políticas foram escritas na linguagem LPBMAN [2]. O modelo de informação define os destinos, políticas e suas associações. Para este cenário os destinos (targets) são usuários, grupos e serviços, que são associados uns aos outros para se obter o comportamento desejado. Também, as políticas estão associadas aos serviços.

3.8.2. Modelagem do Cenário

Como um exemplo, um subconjunto do modelo implementado no X-PBMAN é apresentado. A tabela 3.1 contém uma lista de destinos para a rede AN1, identificada por *eov*. Dois WDSs (Well-Defined Service) são necessários, vídeo e QoS, identificados como tais pelo sufixo @WDS. Na AN1, o target Group é usado para representar a assinatura do usuário a um serviço. Tanto o grupo Executive como o usuário José é local à AN1, identificados por *jose@eov*.

Tipo	Nome	Identificador
Network	Extreme On-line Video	EOV
Service	Video Service	video@WDS
	QoS Service	qos@WDS
Group	Executive	exec@eov
User	José	jose@eov

Tabela 3.1. Destinos selecionados para AN1.

A Tabela 32 mostra um conjunto de associações entre os destinos na AN1. O Executive Group está associado aos serviços de Vídeo e QoS, isto é, todos os usuários associados a este grupo serão assinantes do serviço Premium VoD (incluindo José). José poderá usar o serviço Vídeo Premium na AN2, desde que as redes estejam compostas e AN2 ofereça os serviços de Vídeo e QoS.

Type (A ® B)	Target A	Target B
---------------------	-----------------	-----------------

Group→Service	Executive	Video Service QoS Service
User→Group	José	Executive

Tabela 3.2. Associações dos destinos selecionados na AN1.

A Tabela 3.3 mostra três políticas selecionadas para AN1, envolvidas no provisionamento do serviço Premium VoD. As políticas p1@eov e p3@eov estão associadas ao serviço de vídeo, ao passo que a política p2 está associada ao serviço de QoS. A política p1@eov especifica que se o serviço é video@WDS e o usuário que solicita o serviço pertence ao Executive group, então duas ações são executadas. Primeiro, são recuperadas da PDN políticas adicionais associadas ao serviço qos@WDS, neste caso a política p2@eov. Em segundo lugar, o PEP de vídeo é informado que o usuário recebeu direito de acesso ao serviço Vídeo Premium. Neste exemplo, a string “service=premium_video” é enviada diretamente ao PEP de vídeo, por razões de simplicidade. Porém, lidar com mensagens de resposta de protocolo é um compromisso entre complexidade na especificação das políticas, na entidade FE parser ou no PEP.

A política p2@eov é válida quando o serviço é qos@WDS e a ação é “configure”. Neste caso, o endereço do QoS PEP associado ao PEP de vídeo é recuperado da PDN e um comando de backend para provisionamento de QoS é executado. A estrutura [map QOSPEP(req_pep_id)] significa que uma informação de mapeamento associada à constante QOSPEP (previamente armazenada) com o parâmetro req_pep_id é recuperada da PDN.

A política p3@eov é necessária para negar acesso a usuários não-assinantes. Em outras palavras, quando processando o algoritmo da política, as políticas p1@eov e p3@eov serão selecionadas porque ambas são associadas ao serviço de vídeo. No caso do usuário estar associado ao Executive group, a política p1 será executada devido a sua maior prioridade (conjunto 1). Caso contrário, a política p3 será executada.

policy p1@eov set 1 by User

<pre> if (service == video@WDS; usergroup ct exec@eov) then use qos.service=qos@WDS do return(“service=premium_video”) </pre>
<pre> policy p2@eov set 1 by QoS if (qos.service == qos@WDS) then set qos_pep=[map QOSPEP(req_pep_id)] do configure_qos_pep (qos_pep, client_addr) </pre>
<pre> policy p3@eov set 0 by User if (service == video@WDS) then do return (“service = deny”) </pre>

Tabela 3.3. Políticas selecionadas para AN1.

As políticas para AN2 são bastante similares àquelas definidas para AN1. Quando o usuário José da AN1 acessa o serviço de vídeo na AN2 a entidade Policy FE poderá encontrar todas as informações necessárias para avaliar e tomar decisões sobre aquela solicitação. Uma vez que ambos os serviços de vídeo e de QoS são do tipo WDS, eles serão compreendidos por ambas as redes.

Atualmente, o protótipo X-PBMAN implementa apenas a composição do tipo Integração de Rede, na qual todas as informações (políticas, destinos e associações) e todos os P-Nodes de uma dada PDN tomarão parte na nova rede composta.

4. COMPOSIÇÕES EM REDES DE AMBIENTE

Espera-se que as redes ANs operem em um ambiente de rede caracterizado por topologias bastante dinâmicas e redes heterogêneas que cooperam entre si. Acredita-se que novos

tipos de redes móveis ganhem importância, tais como as redes pessoais (Personal Area Networks – PAN), as Body Area Networks (BANs) e redes de sensores, todas interoperando entre si. O plano de controle destas redes precisa permitir, por exemplo, uma mobilidade sem interrupção (seamless mobility), Qos fim-a-fim, segurança e tarifação integradas. Um desafio é a heterogeneidade de tecnologias e soluções adotadas nos vários planos de controle.

O conceito de composição de rede é então introduzido para suportar interoperabilidade auto-organizada de planos de controle de redes heterogêneas. A composição de rede permite uma gerência consistente sobre as redes cooperantes e esconde dos elementos externos os detalhes das suas interconexões bem como as suas estruturas internas. Isso melhora a operação da rede e a eficiência de serviço.

Os detalhes da interoperação dos planos de controle entre as redes cooperantes são fixados em um acordo de composição. O estabelecimento de uma composição consiste na negociação e posterior realização de um acordo de composição. Tanto a negociação do acordo de composição como a realização da composição devem ser feitos de forma autônoma, isto é, eles são normalmente iniciados por processos internos e prosseguem com mínima interação com o usuário.

As políticas realizam um papel importante no processo de composição. A decisão de compor é baseada em políticas, a negociação do acordo de composição é baseada em políticas e o acordo de composição em si deve combinar de uma forma que faça sentido as políticas das redes ANs tal que a rede AN composta tenha suas próprias políticas que irão governar futuras composições.

4.1. PROCEDIMENTOS DE COMPOSIÇÃO

O processo de composição é descrito por um procedimento de composição, responsável pela estruturação da composição em uma seqüência de cinco fases distintas. Estas fases são Detecção de Mídia, Descoberta/Anúncio, Segurança e Estabelecimento de Conectividade, Negociação do Acordo de Composição e Realização do Acordo de Composição. A Figura 4.1 ilustra este procedimento.

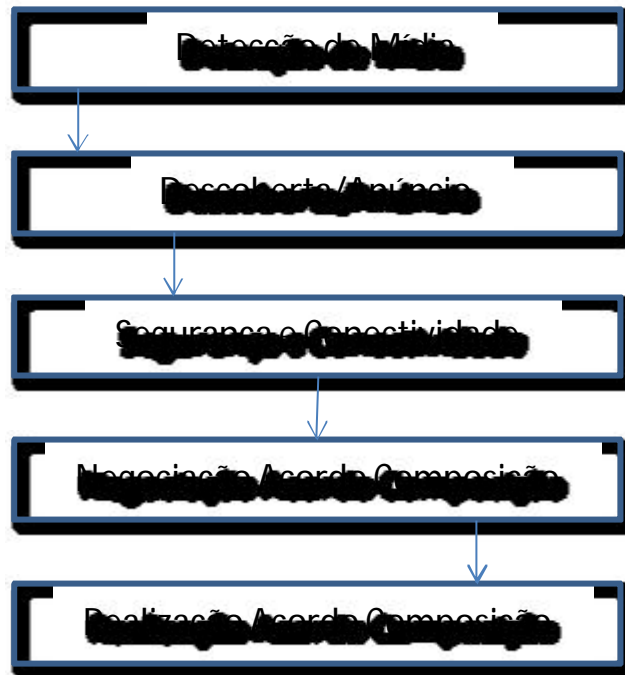


Figura 4.1. Fases do procedimento de composição de redes.

As fases não são necessariamente seguidas de forma unidirecional e linear. Por exemplo, é possível que futuros serviços sejam anunciados mesmo após o estabelecimento de uma associação segura ou que uma associação segura seja atualizada após determinar os detalhes do Acordo de Composição (Composition Agreement – CA) para permitir a necessária flexibilidade.

As cinco fases são discutidas em maior detalhe nas próximas seções.

4.1.1. Detecção de Mídia

Um primeiro passo é “sentir” o meio que permitirá a comunicação com uma rede AN vizinha. Este “sentir” o meio inclui a descoberta de um enlace para uma rede AN remota que não está diretamente conectada. Dependendo do cenário, diferentes tipos de detecção de mídia podem ser concebidos:

- Um operador conecta um novo ponto de acesso em sua rede.
- Duas redes operadas por diferentes operadoras se conectam pela primeira vez.
- Um aparelho de usuário é ligado e procura por redes em sua vizinhança.
- Um usuário inicia uma sessão que requer mais recursos do que foi alocado.
- Uma rede PAN necessita cooperar com uma rede AN remota (composição virtual).

4.1.2. Descoberta/ Anúncio

Dependendo da situação, depois da detecção da mídia segue-se uma fase de anúncio ou de descoberta. As mensagens trocadas na fase seriam tipicamente mensagens de broadcast. Numa rede nível 3, estas mensagens seriam enviadas como consultas direcionadas de composição a redes AN remotas ou enviadas usando-se técnicas de multicast.

Uma AN utiliza o processo de anúncio para oferecer recursos e serviços a outras redes AN. A mensagem de anúncio inclui o identificador criptográfico usado pela AN, que será utilizado para ligar um anúncio a uma AN particular e pode ser autenticado e autorizado numa fase posterior. Alternativamente, uma AN pode escutar os anúncios feitos por outras redes AN ou pode de forma ativa descobrir seus vizinhos. O procedimento de descoberta permite selecionar uma rede AN candidata à uma composição. Permite ainda a descoberta de identificadores, recursos, capacidades e serviços das outras redes.

O procedimento de composição termina nesta fase caso não seja identificada uma rede AN candidata a uma composição.

4.1.3. Estabelecimento da Segurança e da Conectividade

Quando a fase de descoberta resulta em uma rede AN candidata para a composição, as duas redes precisam estabelecer uma segurança e uma conectividade básicas. Uma forma eficiente de fazer isso é uma generalização do HIP (Host Identity Protocol) que inclui a geração de uma chave de sessão compartilhada usando o algoritmo Diffie-Hellmann. São usados identificadores criptográficos pertencentes às redes AN em comunicação e envolvidas na composição para mapear a chave estabelecida às ANs comunicantes e um desafio criptográfico é usado para protegê-las de ataques do tipo denial-of-service (DoS).

Podem ser utilizados protocolos existentes tais como HIP e IKEv2 para estabelecer a associação segura entre as duas redes AN. Porém, estes protocolos podem necessitar serem estendidos para que possam carregar payloads adicionais, tais como credenciais específicas para Redes de Ambiente, propósito da composição, etc.

Os identificadores das redes AN podem ser autenticados e autorizados por terceiros confiáveis ou o relacionamento pode ser baseado numa chave secreta pré-estabelecida.

A conectividade entre as duas redes AN é estabelecida em algum momento durante esta troca de mensagens ou imediatamente após a mesma.

4.1.4. Negociação do Acordo de Composição

O próximo passo do processo de composição é a negociação do atual acordo de composição (Composition Agreement). O CA inclui as políticas a serem seguidas na AN composta, o identificador da AN composta, como os recursos lógicos e físicos serão controlados e/ou compartilhados entre as redes AN, etc. Caso o CA inclua aspectos comerciais, ele deve ser assinado digitalmente por ambas as ANs para garantir o não repúdio.

Existem duas maneiras de se negociar um CA: centralizada e distribuída. Numa negociação centralizada, as entidades funcionais (Functional Entities – FE) das duas redes responsáveis pelo controle da composição negociam entre si, cada uma consultando outras FEs na mesma rede AN. Numa negociação descentralizada, cada FE negocia de forma independente a funcionalidade de controle sob sua responsabilidade, orquestrada pelo FE

de controle de composição. Neste caso, após o processo de negociação deve existir uma fase de consolidação interna entre os FEs em cada AN porque os acordos parciais negociados por cada FE pode ter certo grau de dependência entre eles.

4.1.5. Realização do Acordo de Composição

A fase de realização representa a conclusão da composição. Durante esta fase os elementos são configurados para refletir o acordo de composição. Portanto, cada AN deve realizar a configuração de seus próprios recursos, por exemplo, atualizando o registro de recursos, a base de dados de informações de contextos (Context Information Base – CIB) e o banco de dados de políticas e alocando recursos para conectividade e QoS através da interface ARI.

Dependendo do resultado da realização da composição pode ser necessário a realocação e reorganização dos endereços de rede. O resultado do processo de composição será ou uma nova rede AN ou uma rede AN aumentada (ou seja, uma AN é absorvida por outra), ou duas ANs que interoperam.

4.1.6. Procedimento de Decomposição

Uma ou mais ANs constituintes de uma composição pode decidir invocar o procedimento de decomposição para terminar a composição.

4.1.7. Procedimento de Atualização da Composição

Um acordo de composição pode ser alterado através do re-início da fase de negociação do acordo de composição. Existem várias razões para a atualização de uma composição tais

como novas necessidades de serviço, mobilidade, falha de enlaces, disponibilidade de novos enlaces ou novos serviços.

4.2. TIPOS DE COMPOSIÇÃO

Um processo de composição bem sucedido conduz a uma cooperação entre as redes ANs envolvidas que pode assumir várias formas. Existem duas abordagens para se analisar a cooperação resultante do processo de composição. A primeira considera a maneira como os recursos são compartilhados (gerência de recursos) entre as redes cooperantes e a segunda analisa sob o aspecto da gerência de rede.

4.2.1. Tipos de Composição – Perspectiva da Gerência de Recursos

Para caracterizar as formas como recursos são compartilhados, as seguintes premissas são assumidas em relação a uma rede AN:

- Cada AN implementa um ACS.
- Cada ACS controla recursos $R=\{r_1,r_2,\dots,r_n\}$, onde controle significa um direito exclusivo de alocar/desalocar/configurar/etc. recursos em R.
- O ACS é capaz de fornecer informação sobre o status dos seus recursos R a outras redes AN.
- O ACS tem a habilidade de alocar os recursos em R a partir de solicitações feitas por outras ANs, os quais podem assim ser utilizados por estas outras ANs.
- Cada ACS tem um conjunto de políticas $P=\{p_1,p_2,\dots,p_n\}$.
- O Acordo de Composição (CA) define a alocação, o controle, o status e a informação dos recursos.

- As políticas em P fornecem informações detalhadas sobre as regras em torno do procedimento de controle e alocação de recursos.
- Como pré-requisito, deve existir um relacionamento de confiança entre as redes ANs para que seja estabelecido um CA.

As Figuras 4.2 a 4.5 representam o status antes e depois de uma composição para cada tipo de composição previsto.

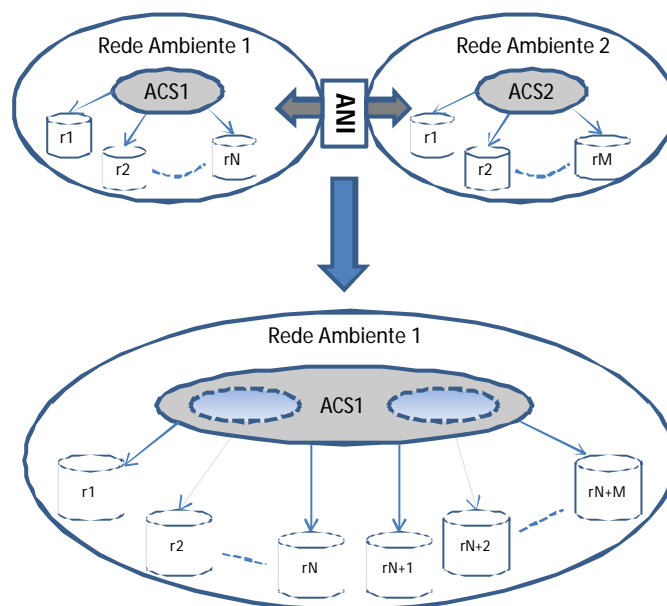


Figura 4.2. Resultado da integração de rede.

A Figura 4.2 ilustra o caso de integração de rede, onde o resultado da composição é uma nova rede AN que controla os recursos R, que corresponde à somatória dos recursos da AN1 e AN2.

Para o caso de compartilhamento de controle (control sharing), existem duas alternativas. A Figura 4.3 representa um direito não exclusivo de controlar recursos comuns e a Figura 4.4 representa o caso de concessão de direitos exclusivos de acesso de uma AN para outra (também conhecido como delegação de controle). O controle sobre um recurso compartilhado exige um ACS comum/virtual. Este ACS comum/virtual deve no mínimo controlar os recursos comuns, mas também controlar outros recursos não compartilhados.

Assim, um ACS comum/virtual é opcional para o caso de delegação de controle (Figura 4.4) e interworking de rede (Figura 4.5).

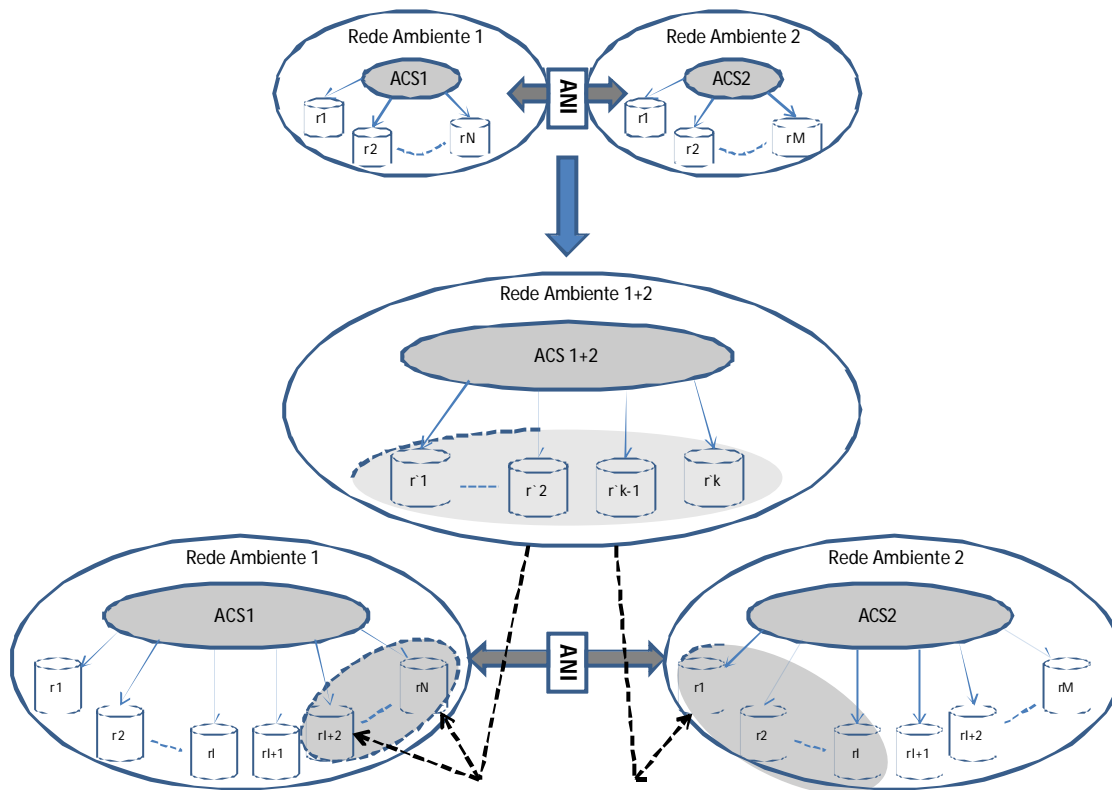


Figura 4.3. Resultado de compartilhamento de controle, direito não exclusivo para controle de recurso.

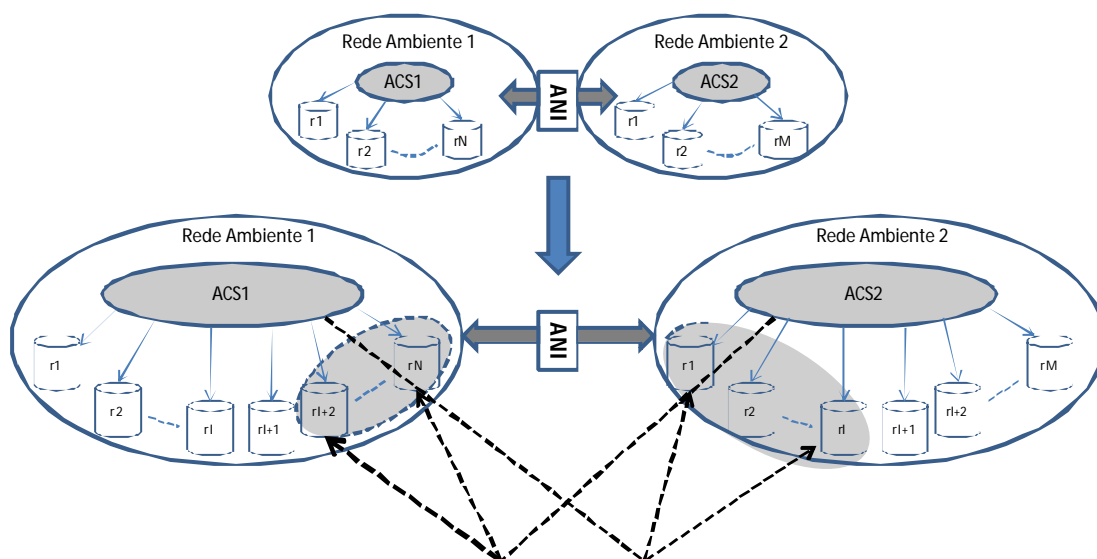


Figura 4.4. Resultado de compartilhamento de controle, direito exclusivo para controle de recurso (delegação de controle).

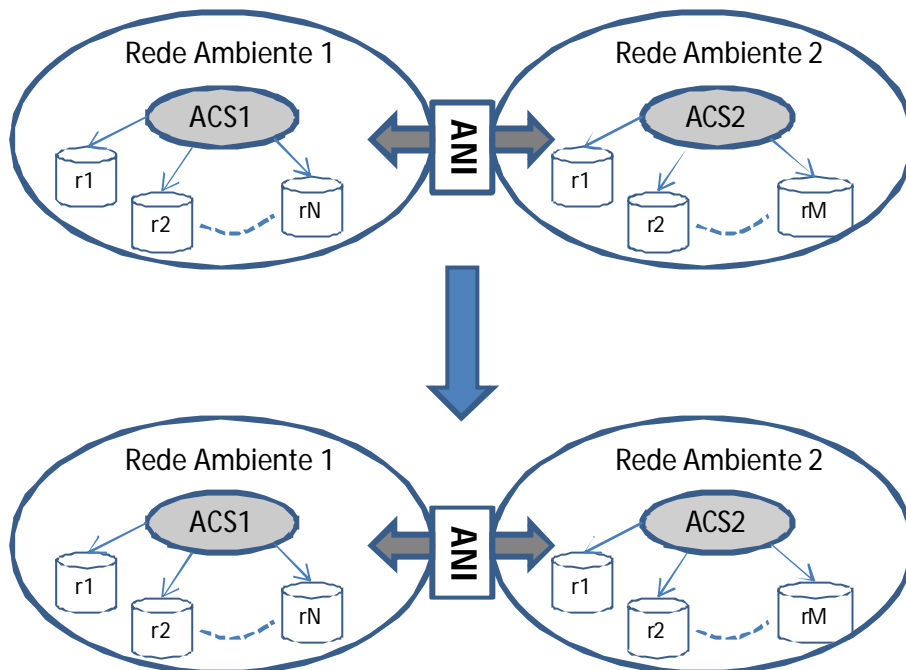


Figura 4.5. Resultado do interworking de rede.

4.2.2. Tipos de Composição – Perspectiva da Gerência de Rede

A gerência de rede utiliza uma arquitetura Peer-to-Peer (P2P) overlay para garantir uma arquitetura de rede escalável para as Redes de Ambiente. Este modelo mantém um overlay separado para cada ACS. Isto significa que cada AN está associada a um overlay que representa o ACS.

Os componentes básicos deste modelo overlay são peers, super-peers e overlays. Um overlay é um conjunto de peers pertencentes a um mesmo domínio de gerência e formando uma rede virtual. Cada overlay elege um super-peer para representar o overlay no mundo externo. É importante salientar que este super-peer é responsável somente pelas negociações com outros overlays e não tem nenhum privilegio especial dentro do seu overlay.

Os super-peers também podem formar overlays de maior hierarquia, criando assim uma estrutura de rede hierárquica. Outra característica desta estrutura hierárquica é que os níveis hierárquicos não são absolutos. Isto significa que não se pode definir um índice de nível hierárquico absoluto a um overlay.

Em Redes de Ambiente, os processos de composição de rede devem acontecer automaticamente sem a interação do usuário (tanto quanto possível) e devem ser baseados em preferências dos usuários pré-definidas.

Os usuários definem suas preferências por meio de regras de políticas. Estas regras podem ser regras estritas que expressam requisitos explícitos do usuário ou regras mais permissivas que refletem apenas o desejo do usuário. Os usuários formam redes com outros usuários que têm políticas similares (ou ao menos mutuamente aceitáveis). Quando uma AN é formada, a política geral reflete as preferências comuns da comunidade, que deve ser continuamente mantida de acordo com as políticas correntes dos atuais membros.

A política comum de rede define as preferências da comunidade que forma a rede. Esta política é usada quando distintas redes tentam se interconectar. Baseado nas políticas das duas redes, o processo de composição pode ter resultados diferentes. Os tipos distintos de overlay de gerenciamento resultantes deste processo são representados na Figura 4.6.

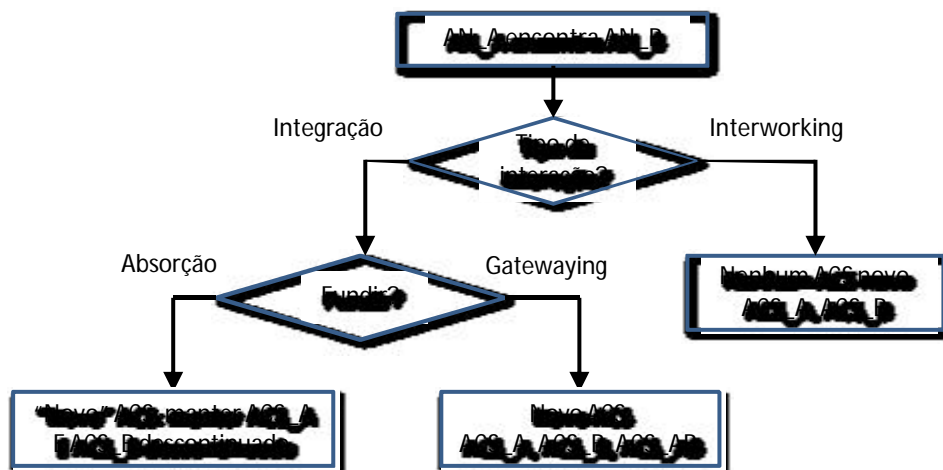


Figura 4.6. Tipos de overlays.

4.2.2.1. Absorção

Se os bancos de dados de políticas das duas redes são parecidos em conteúdo (não existe contradições nas preferências e as diferenças entre políticas podem ser aceitas mutuamente), as redes podem se juntar numa forma denominada de *absorção*. A absorção é a fusão completa das duas redes. Elas irão formar um único domínio de gerência com um super-peer comum, conforme ilustrado na Figura 4.7.

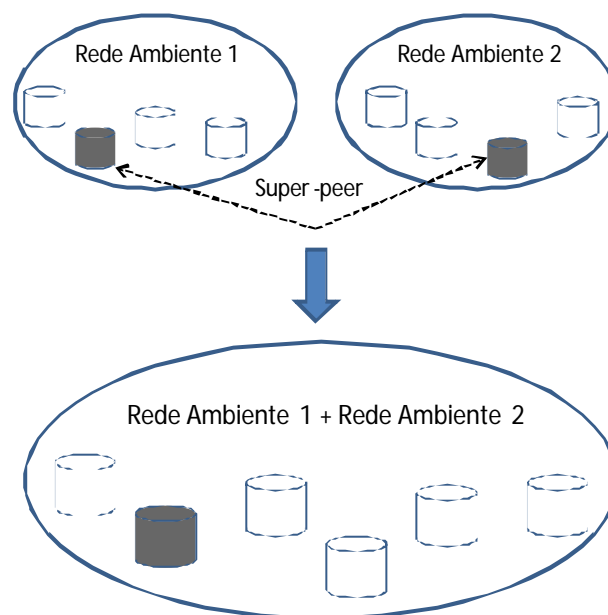


Figura 4.7. Absorção.

4.2.2.2. Overlay do tipo Gatewaying

Se a absorção das redes for impossível pelas regras das políticas, existe uma solução de compromisso entre absorção total e separação total. Esta forma de operação é chamada de *Gatewaying*.

É importante salientar que redes conectadas através de gatewaying não são redes separadas; elas são redes compostas mas elas não compartilham todos os recursos. Isto

significa que elas serão vistas como uma única rede, e portanto elas precisam ser representadas por um super-peer e um banco de dados de políticas consistente, como ilustrado na Figura 4.8. Portanto, o super-peer das redes combinadas criará um overlay de nível superior. Os membros desta rede elegerão um super-peer que irá representar as duas redes para o mundo exterior.

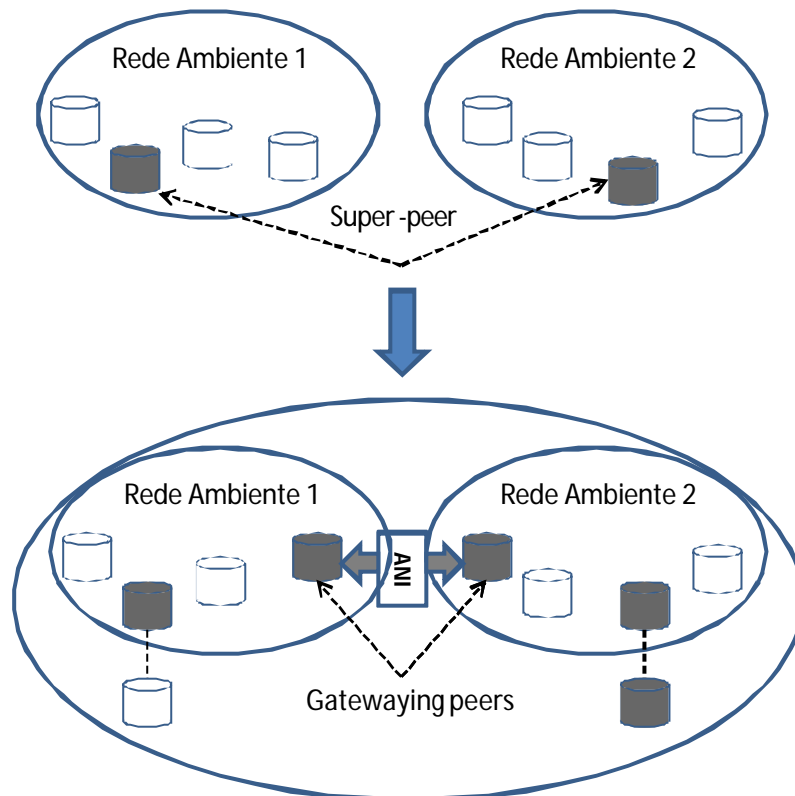


Figura 4.8. Gatewaying.

No gatewaying, a construção do banco de dados de políticas comum é mais difícil porque podem existir contradições entre as regras de políticas das duas redes. As políticas contraditórias devem ser eliminadas do conjunto comum de políticas. As redes compostas e toda a comunidade são descritas por suas políticas comuns e representadas pelo super-peer de maior nível. Os recursos não passados para o ACS overlay podem ser acessados através de overlays e super-peers das redes AN originais.

4.2.2.3. Overlay do tipo Interworking

O overlay do tipo interworking é selecionado se nenhum overlay novo é criado. Ele será um interworking independente entre as duas redes AN. Por isso é chamado de interworking (interoperação). As mesmas interfaces AN para o mundo exterior serão usadas.

4.2.3. Acordos de Composição Multilaterais

Quando duas redes AN compõem de acordo com a arquitetura de Redes de Ambiente, um novo ACS comum/virtual pode ser criado dependendo do Acordo de Composição (CA) negociado. A composição sempre ocorre entre duas redes AN num dado momento, portanto é uma operação bilateral de rede em sua forma básica. Contudo, esta operação bilateral pode ser aplicada múltiplas vezes para que possa ser suportada composições multilaterais consistindo de três ou mais redes. Neste caso, existe um CA que é acordado entre todas as redes AN constituintes. Assim, se a composição não resulta na criação de um ACS comum/virtual, então não existirá um ACS relacionado ao CA acordado, refletindo a composição estabelecida. Esta é a razão porque tais composições não podem ser entendidas de acordo com a arquitetura de Redes de Ambiente. Portanto, o ACS comum/virtual é um pré-requisito para composições multilaterais independentemente de como os recursos são compartilhados.

4.2.4. Processo de Decisão e Relações

O processo de decisão do tipo de composição envolve a análise de dois fatores distintos: (1) o controle comum dos recursos acordados e (2) a necessidade de suportar composições multilaterais. A Figura x6 representa este processo de decisão e as relações entre tipos de composições, suporte a acordos de composição multilaterais e tipos de overlay de gerência.

5. O ANCSIM (AMBIENT NETWORK COMPOSITION SIMULATOR)

Este capítulo apresenta o simulador Ambiente Network Composition Simulator (ANCSim), proposto e desenvolvido pelo GPRT (Grupo de Pesquisa em Redes e Telecomunicações) da Universidade Federal de Pernambuco (UFPE) para avaliar os conceitos introduzidos nos capítulos anteriores. São descritos os elementos que compõem o ANCSim e suas principais funcionalidades, incluindo o Acordo de Composição, a realização da composição, a decomposição e a estrutura do simulador.

5.1. VISÃO GERAL

Uma implementação adequada da composição de rede é um fator primordial para a viabilidade das Redes de Ambiente, uma vez que são esperadas grandes demandas por composição numa interação típica do usuário com a rede. Contudo, atualmente a composição encontra-se no campo teórico e abstrato porque o conceito foi recentemente introduzido.

Uma vez que a avaliação do conceito de composição de rede numa implementação real não é possível atualmente, optou-se por realizar uma análise através de simulação. O simulador ANCSim foi especificado e implementado com este objetivo. O ANCSim é um simulador de propósito específico para composição de Redes de Ambiente.

As ferramentas de simulação existentes, tais como ns-2, apresentam um nível de granularidade (avaliação através de pacotes) e/ou abstração (nenhum módulo para Ambient Networks) que não satisfazem as necessidades de análise de desempenho de composição de rede. Portanto, optou-se por desenvolver um simulador específico para este propósito.

5.1.1. Elementos

O simulador ANCSim é composto por dois elementos principais: Ambient Networks (ANs) e serviços. Estes elementos estão associados da seguinte maneira: cada AN tem um conjunto de serviços que são utilizados por seus usuários ou oferecidos a outras ANs.

5.1.1.1. Redes

O simulador ANCSim suporta os tipos de redes propostos no modelo PBMAN: PDN e Agentes. A figura 5.1 mostra um exemplo da associação entre estas duas redes.

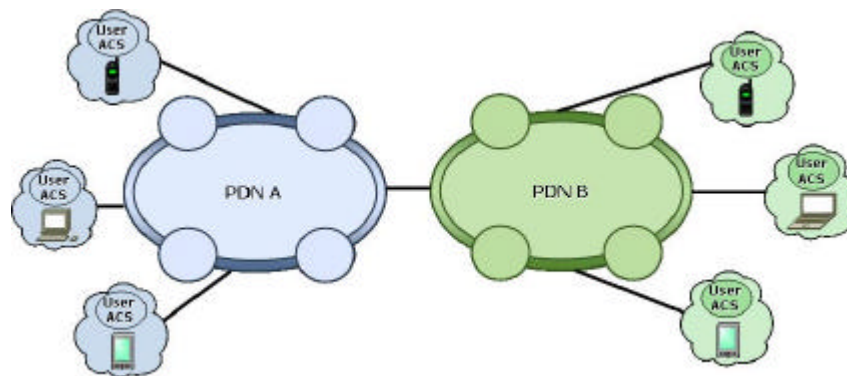


Figura 5.1. PDNs e Agentes no ANCSim.

As redes PDN são responsáveis por duas funções: tomar decisões sobre solicitações de serviço e manter informações de gerência como os serviços sendo utilizados pelas redes compostas ou o número de composições ativas em um instante particular. PDNs podem ser compostas uma com a outra.

As Redes Agente ou simplesmente Agentes são conectados a PDNs. No ANCSim, cada agente deve ser conectado a exatamente uma PDN. Este tipo de rede é representado por usuários ou por equipamentos de usuário. A diferença entre Agente e Rede Agente é que o Agente é um único usuário enquanto que a Rede Agente é uma rede composta por um usuário e alguns equipamentos. Eles são responsáveis por solicitar serviços a outros Agentes ou Redes Agente, que no ANCSim são chamados de Redes Destino. Estas solicitações são processadas pela PDN onde o Agente está conectado.

5.1.1.2. Serviços

No ANCSim, os serviços são classificados em dois grupos de acordo com suas aplicações: serviços que têm o propósito de suportar os serviços requisitados pelos usuários (chamados Serviços de Usuário) ou outros serviços (chamados Serviços de Suporte).

Os Serviços de Usuário são fornecidos aos usuários pertencendo à AN que provê o serviço e aos usuários pertencendo a outras ANs. Exemplos deste tipo de serviço são serviços de Vídeo, Voz sobre IP e Internet Banking. Os serviços de suporte são providos com o objetivo de fornecer níveis de garantia aos serviços de usuário, tais como boa qualidade de voz ou de vídeo e melhoria na segurança de transações importantes. Exemplos deste grupo são os serviços de QoS e VPN.

No ANCSim, cada serviço de usuário é oferecido pelas Redes Agente e tem um peso que representa sua importância na rede. O peso é usado para escolher uma Rede Destino para o serviço. Em outras palavras, a Rede Destino é aquela que oferecerá o serviço requisitado junto com a rede requisitante. Quanto maior o peso, maior a probabilidade de uma rede ser escolhida como destino.

A duração de cada serviço é configurada no início da simulação e varia de acordo com o tipo de serviço. Por exemplo, sessões de serviço de vídeo e VoIP normalmente duram mais do que um serviço de Banking e estas características são levadas em conta para definir o parâmetro “duração” no ANCSim.

Adicionalmente, cada serviço é configurado com um limite que especifica o número máximo de usuários que estão usando o serviço simultaneamente. Quando este limite é atingido, qualquer solicitação de serviço é negada aos usuários. O serviço está disponível novamente quando um dos usuários usando o serviço deixa de usá-lo, isto é, a sessão de serviço é terminada. Este evento libera o recurso para que outros usuários possam usá-lo.

5.1.1.3. Funcionalidade

O ANCSim implementa algumas funcionalidades importantes de uma AN em uma abstração de alto nível. É assumido que as redes são estáveis e que não ocorrem falhas como perda de pacotes e atrasos excessivos. A implementação inclui alguns aspectos do Ambient Control Space (ACS), Composition Agreement (CA) e procedimentos de composição e decomposição.

Para a simulação, os parâmetros de duração e taxa de solicitação de serviços serão derivados de medições realizadas na rede GPRS da Brasil Telecom para o serviço de dados.

5.1.1.4. Composição no ANCSim

A composição de rede no ANCSim é disparada a partir de solicitações de serviço, mas nem toda solicitação de serviço inicia uma solicitação de composição. Isso é determinado pela rede destino selecionada e sua disponibilidade. Em outras palavras, se a AN destino provê o serviço solicitado e se ela ainda aceita sessões de usuários, então para que este serviço possa ser fornecido do ponto de vista de gerenciamento uma ou mais composições podem

ser necessárias. Toda vez que uma solicitação de serviço é recebida, o provedor verifica todas as composições necessárias para o estabelecimento daquele serviço.

Existem dois tipos de solicitação de serviço no ANCSim: uniservice e multiservice. No uniservice, a solicitação contém apenas um tipo de serviço específico que o usuário está interessado naquele momento. Ou seja, o usuário pergunta por um determinado tipo de serviço. No multiservice, a solicitação contém não apenas um serviço, mas também outros serviços que o usuário pode estar interessado no futuro. Além do serviço solicitado, a solicitação inclui outros serviços que não são necessários naquele momento. O objetivo é evitar futuras requisições de composição. Como tal, futuras solicitações de serviço podem não ser disparadas, tentando diminuir o número de tentativas de composição.

Os serviços são classificados em dois tipos de acordo com a solicitação: serviço essencial ou opcional. Um serviço é essencial quando ele é necessário para uma composição bem sucedida. Se este serviço não pode ser provido, a composição irá falhar e uma mensagem de erro é retornada. Contudo, a composição não falha quando apenas serviços opcionais não estão disponíveis. Normalmente, solicitações de uniservice contêm serviços essenciais, ao passo que solicitações multiservice combinam serviços essenciais e opcionais.

5.2. SELECIONANDO UMA REDE DESTINO

Toda vez que um serviço é iniciado sua Rede Agente de destino é selecionada aleatoriamente, de acordo com um parâmetro administrativamente configurado, o peso, que representa a importância do serviço em cada rede. A Rede Agente com maior peso tem maior probabilidade de ser escolhida como rede destino. O seguinte procedimento é realizado pelo ANCSim para processar a rede destino e é ativado toda vez que uma nova solicitação de serviço é disparada:

1. Seleção das Redes Agente que provêm o serviço solicitado;
2. Recuperação do peso do serviço para cada par serviço/rede resultado do item 1;
3. Geração de um número randômico que segue uma distribuição uniforme;

4. Seleção da rede que contém o número gerado no item 3 considerando o conjunto de valores ao qual ele corresponde;
5. Se a rede destino selecionada é a rede fonte, a composição não será realizada; caso contrário, um caminho entre a fonte e o destino precisa ser encontrado.

A Figura 5.2 apresenta uma topologia que ilustra com mais detalhes o procedimento descrito acima. As PDNs são denominadas de PDN1, PDN2, PDN3, PDN4 e PDN5. Os Agentes são denominados de AG1, AG2, AG3, AG4, AG5, AG6 e AG7.

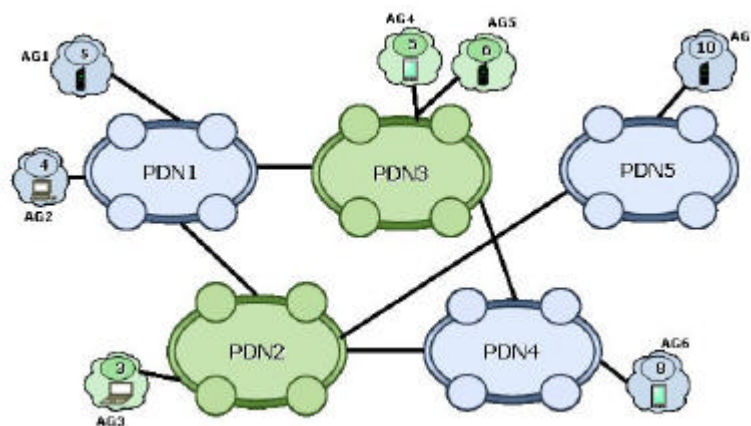


Figura 5.2. Uma topologia de Ambient Network.

Considerando que cada rede provê o mesmo serviço e este serviço tem o peso indicado na Figura 5.2, a seguinte distribuição poderia ser observada na Figura 5.3:



Figure 5.3. Distribuição de Pesos.

A diferença entre os intervalos corresponde ao peso do serviço em cada rede. Por exemplo, o peso do serviço na rede AG1 é 5 enquanto na rede AG5 ele é 6.

Após gerar a distribuição de pesos, um número é aleatoriamente escolhido. Por exemplo, se o número escolhido é 10, a rede destino será AG3.

5.3. SELEÇÃO DE CAMINHO

A seleção de um caminho apropriado entre as redes origem e destino depende do tipo do serviço envolvido na solicitação de serviço. Se o serviço é um serviço de usuário (aplicação, tal como VoIP), o caminho será formado pelo Agente e a PDN que origina a solicitação do serviço e o Agente e a PDN que são o destino desta solicitação. Este tipo de serviço precisa estar disponível somente nas terminações do caminho (serviço non-path-based). Este procedimento é ilustrado na Figura 5.4.

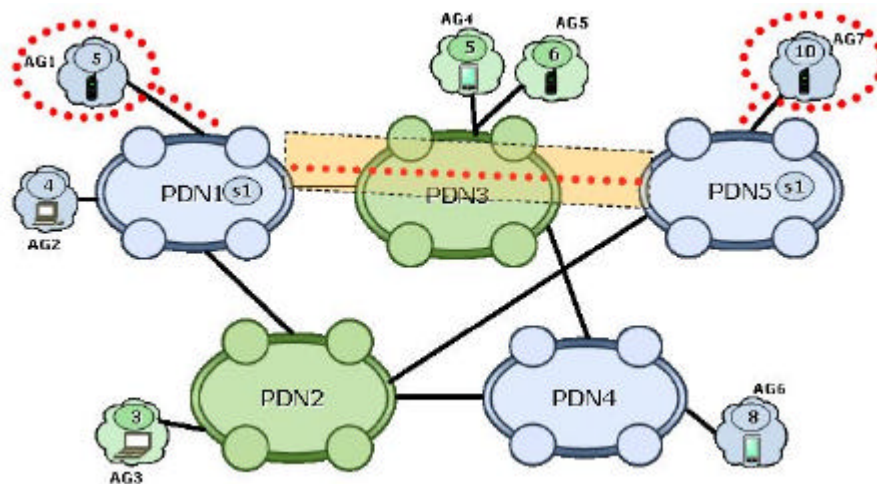


Figura 5.4. Caminho considerando Serviços de Usuário.

Assumindo que o serviço foi solicitado pela rede AG1 e que a rede destino seja a rede AG7, o caminho será composto por AG1 – PDN1 – PDN5 – AG7. A composição entre PDN1 e PDN5 será uma Composição Virtual. O procedimento de composição será explicado na seção 5.4.

Porém, se a solicitação é para algum serviço de suporte (por exemplo, QoS), os serviços solicitados devem estar disponíveis em todas as redes que compõem o caminho (serviço

path-based). Neste caso, foram consideradas duas abordagens para a seleção do caminho a partir dos candidatos. A primeira é descrita abaixo e ilustrada na Figura 5.5:

1. Todos os caminhos entre as redes origem e destino são encontrados;
2. Encontra-se o caminho mais curto onde o serviço está disponível entre todos os caminhos encontrados;
3. Uma falha de serviço ocorrerá caso não exista caminho que satisfaça estas condições.

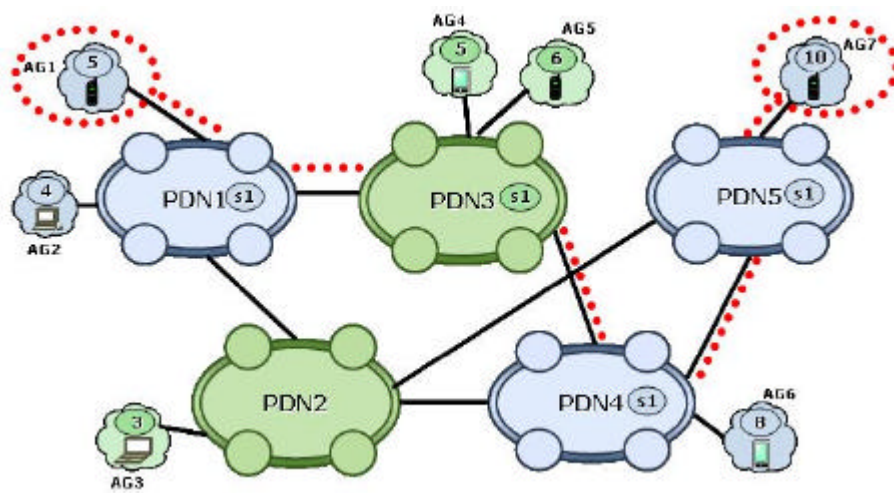


Figura 5.5. Serviço disponível no Caminho.

Assumindo que o serviço foi solicitado pela rede AG1 e que a o destino é a rede AG7, os caminhos encontrados entre estas duas redes são: AG1– PDN1 – PDN2 – PDN5 – AG7, AG1– PDN1 – PDN3 – PDN4 – PDN5 – AG7, AG1– PDN1 – PDN2 – PDN4 – PDN5 – AG7 and AG1– PDN1 – PDN3 – PDN4 – PDN2 – PDN5 – AG7. O caminho mais curto onde o serviço está disponível é: AG1– PDN1 – PDN3 – PDN4 – PDN5 – AG7, ilustrado na Figura 5.5. Quando mais de um caminho com o mesmo tamanho é encontrado, o primeiro caminho selecionado será o primeiro caminho encontrado.

A segunda abordagem é mais simples, descrita a seguir, está ilustrada na Figura 5.6:

1. Todos os caminhos entre as redes origem e destino são encontrados;
2. O caminho mais curto é selecionado;

3. São negociadas sequencialmente composições entre cada par de redes vizinhas.
4. Uma falha de serviço ocorrerá se o serviço não pode ser provido por alguma rede ao longo do caminho. Neste caso, as composições realizadas até este momento precisam ser desfeitas, diferentemente da primeira abordagem onde o serviço está disponível no caminho.

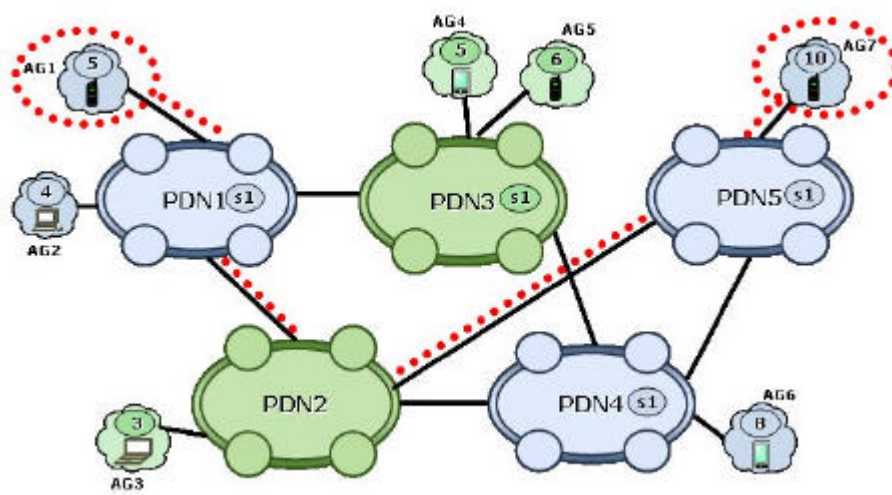


Figura 5.6. Caminho mais curto entre as redes origem e destino.

Foram encontrados os mesmos caminhos do exemplo anterior, mas o caminho selecionado é o mais curto: AG1 – PDN1 – PDN2 – PDN5 – AG7. Note que a rede PDN2 não provê o serviço. Por esta razão, ocorrerá uma falha de serviço.

Ambas as abordagens tem alguns problemas. A primeira abordagem procura fornecer um roteamento de composição que tenta descobrir o menor caminho onde o serviço solicitado está disponível em todas as redes. O roteamento da composição é um problema novo que tem sido estudado pelo Ambient Networks Project [16] e não faz parte do escopo desta dissertação. A segunda abordagem considera apenas o menor caminho, mesmo que exista um outro caminho que satisfaça as condições. Considerando as vantagens e desvantagens, optou-se pela segunda abordagem, porque em uma rede de acesso dinâmica os usuários precisam ter acesso instantâneo aos serviços e o roteamento é normalmente um procedimento lento.

5.4. O PROCESSO DE COMPOSIÇÃO

O processo de composição pode ser iniciado após a seleção do menor caminho. O processo de composição acontece entre todos os pares de redes adjacentes no caminho e são executados numa ordem sequencial.

Uma composição entre duas redes pode falhar por três razões: o serviço solicitado não está disponível na rede, o serviço já está sendo usado por outros usuários e não existem sessões disponíveis ou o número máximo de composições por rede atingiu seu limite.

Duas abordagens diferentes são usadas para compor duas redes e a escolha entre elas depende do tipo do serviço solicitado (serviços tipo path-based ou serviços tipo non-path-based).

Considerando o caminho selecionado na seção 5.3, quando o serviço solicitado é apenas para um serviço de usuário (non-path-based), composto por AG1 – PDN1 – PDN5 – AG7, a composição será realizada sequencialmente entre as redes AG1 e PDN1 (1- Agente/PDN), PDN1 e PDN5 (2 – PDN/PDN), PDN5 e AG7 (3 – Agente/PDN) como mostrado na Figura 5.7.

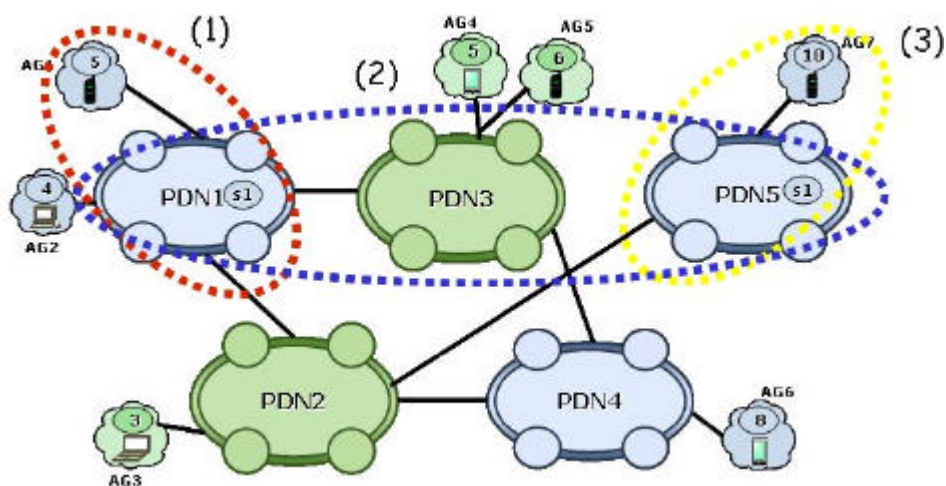


Figura 5.7. Processo de Composição.

Depois que as composições acontecem, a rede AG1 poderá usar o serviço fornecido pela rede AG7. Se já existe uma composição entre as redes envolvidas, o Acordo de Composição será modificado para refletir os novos serviços.

Quando s1 (service 1) é um Serviço de Vídeo, por exemplo, as Redes Agente precisam autenticar-se (composição Agente/PDN) nas PDNs onde estão conectadas. As PDNs envolvidas compõem umas com as outras para prover o serviço e nenhuma outra composição é necessária.

Quando o serviço solicitado é um serviço do tipo path-based a composição deve ser realizada entre todas as redes adjacentes no caminho. Considerando o mesmo caminho encontrado para este tipo de serviço na seção 5.3, AG1 – PDN1 – PDN2 – PDN5 – AG7, as composições serão: AG1 e PDN1 (1 - Agente/PDN), PDN1 e PDN2 (2 - PDN/PDN), PDN2 e PDN5 (3 - PDN/PDN), PDN5 e AG7 (4 - Agente/PDN) como indicado na Figura 5.8.

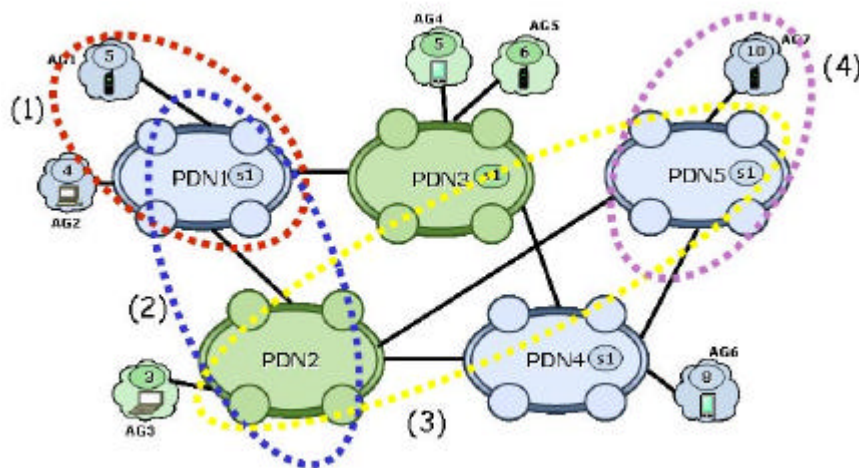


Figura 5.8. Processo de Composição.

Como pode ser notado, PDN1 e PDN2 não podem compor porque PDN2 não suporta o serviço solicitado. Assim, uma mensagem de erro será retornada para o requisitante e as composições feitas anteriormente (AG1 e PDN2) devem ser desfeitas.

Se o serviço s1 é um Serviço de Vídeo com QoS, as composições para QoS devem ser realizadas em cada rede que compõe o caminho porque o QoS, como um serviço do tipo

path-based, precisa ser provido ao longo de todo o caminho da origem até o destino. Uma composição para o serviço de vídeo também deve ser realizada, mas por questão de simplicidade apenas uma composição (Vídeo + QoS) é considerada.

5.4.1. ACS e Acordo de Composição

O simulador ANCSim implementa o conceito de ACS através da implementação das redes PDN e Agente, e é responsável por gerenciar o acordo de composição. A entidade funcional Composition-FE (C-FE) é representado e é responsável por tratar o processo de composição, decidindo quais redes necessitam compor umas com as outras.

No ANCSim, o Acordo de Composição é composto pelos serviços que as redes em composição concordam em compartilhar entre si. Este acordo de composição é armazenado pelo ACS de cada rede PDN. Em outras palavras, o ACS mantém as composições ativas das redes conectadas a ele. Desta forma, quando acontece uma nova composição ou uma atualização em serviços compartilhados de uma composição existente, o ACS é responsável por gerenciar e atualizar o acordo de composição armazenado nele.

5.4.2. Decomposição

Uma decomposição ocorrerá toda vez que não mais houver serviço sendo usado pelas redes compostas. Uma decisão importante a ser tomada é se a composição deve ser desfeita imediatamente depois de detectado que ela não é mais necessária ou se ela deve durar por algum tempo na esperança de que uma nova solicitação de serviço irá demandar que uma nova composição seja feita. Parâmetros definidos no início da simulação determinam quando a composição pode ocorrer.

No primeiro caso, toda vez que um serviço é terminado, é realizado um verificação para saber se as redes podem permanecer compostas ou não. Por exemplo, quando o serviço terminado envolve serviços de suporte, todas as redes ao longo do caminho precisam ser

verificadas para saber se as composições adjacentes podem ser desfeitas. As decomposições ocorrerão entre redes que não mais compartilham nenhum tipo de serviço. Neste caso, o acordo de composição é finalizado e os serviços não mais serão compartilhados.

No segundo caso, esta verificação ocorre depois de tempo especificado determinado através de fatores definidos no início da simulação. Esta abordagem procura evitar ter que recompor as redes logo após terem sido decompostas. Estas duas abordagens serão avaliadas.

5.4.3. Estrutura do ANCSim

O simulador ANCSim é um simulador de evento discreto desenvolvido com a linguagem de programação C++. O diagrama de classes do simulador está ilustrado na Figura 5.9, que mostra o relacionamento entre as classes que compõem o simulador. A principal classe é denominada Simulator e é responsável por coordenar seus componentes e a rotina de simulação. É constituída por variáveis de estado e um escalonador de eventos. As variáveis de estado representam o estado do simulador enquanto o escalonador determina qual evento executará em um ponto específico do tempo de simulação.

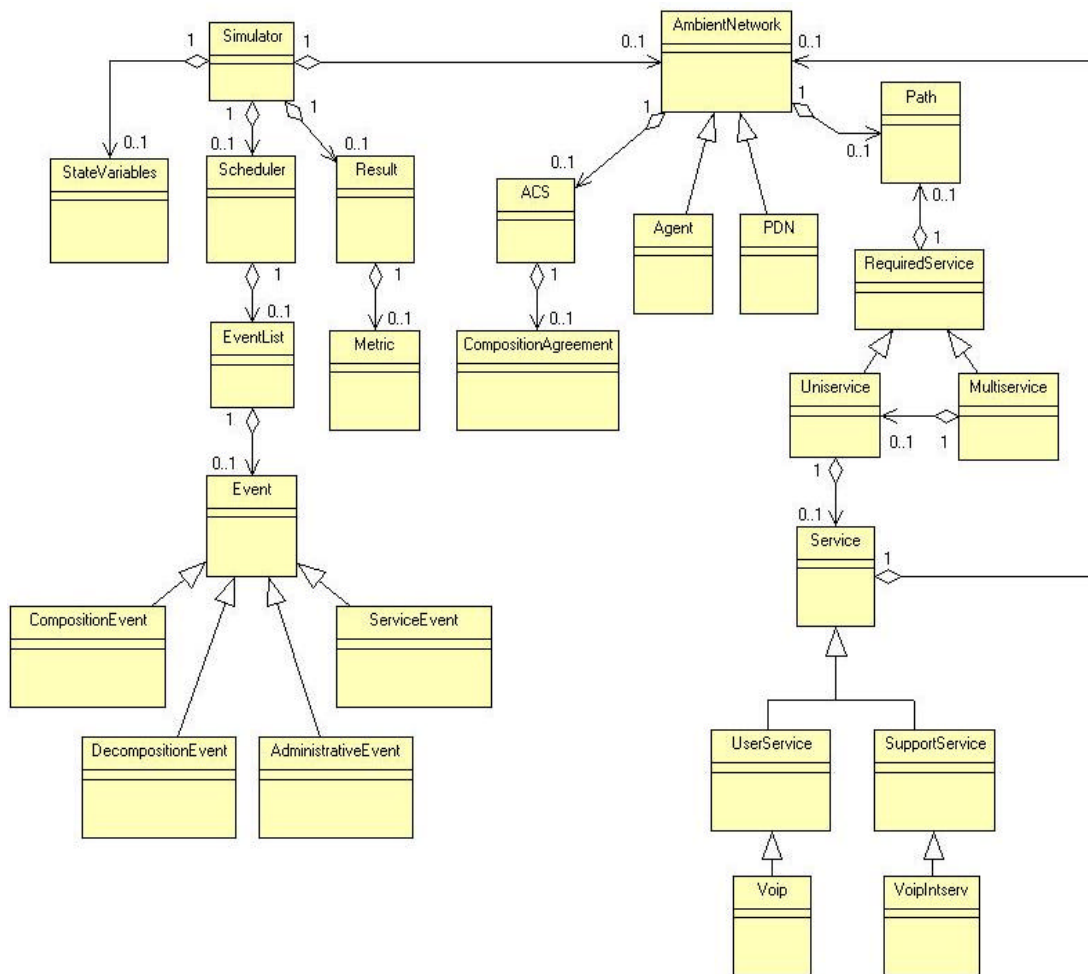


Figura 5.9. Diagrama de Classe do ANCSim.

O simulador também é composto da topologia sendo simulada, a qual consiste em um conjunto de Redes de Ambiente. ANs podem ser Agentes ou PDNs e cada uma contém um ACS responsável pelo Acordo de Composição. As redes Agente são capazes de solicitar serviços a outras redes. Uma solicitação de serviço pode ser Uniservice ou Multiservice e os serviços sendo solicitados podem ser UniService ou SupportService.

O número de replicações e o tempo de simulação são definidos antes que a simulação inicie. A simulação termina quando o tempo de simulação é atingido e os resultados (representado pela classe Result) são retornados.

5.4.4. Eventos

O ANCSim é um simulador orientado a eventos. Os eventos determinam as ações que serão executadas em um tempo específico. Os eventos são classificados em quatro tipos: Eventos de Serviço, Eventos de Composição, Eventos de Decomposição e Eventos Administrativos.

Os Eventos de Serviço são divididos em três: Solicitação de Serviço, Início de Serviço e Fim de Serviço. O evento Solicitação de Serviço é responsável por solicitar serviços a uma AN específica. O evento Início de Serviço sinaliza quando o serviço começa e o evento Fim de Serviço finaliza o serviço requisitado. O tempo entre estes dois últimos eventos (duração do serviço requisitado) e o intervalo entre as solicitações seguem uma distribuição exponencial. Cada serviço tem sua própria duração e taxa de solicitação determinada de acordo com seu tipo.

Os Eventos de Composição são divididos em dois tipos: Início de Composição e Fim de Composição. O Início de Composição é responsável por iniciar o processo de composição e o Fim de Composição completa o processo. A duração do processo de composição foi determinada através de medições realizadas na rede GPRS da Brasil Telecom para as composições do tipo Agente/PDN. Após coleta de amostras, foram derivadas distribuições estatísticas que foram utilizadas para os eventos de composição.

Existem dois tipos de Eventos de Decomposição: Início de Decomposição e Fim de Decomposição. O primeiro evento é responsável por iniciar o processo de decomposição e o segundo, por finalizá-lo. O intervalo de tempo entre estes dois eventos depende do tempo adicional (overtime) da composição: se não existe tempo adicional, o evento Fim de Decomposição é escalonado imediatamente. Caso contrário, a duração do tempo adicional é calculada e o evento Fim de Decomposição é escalonado considerando este resultado.

Os Eventos Administrativos refere-se a eventos pré-programados que têm o objetivo de recuperar algumas informações à medida que a simulação avança. Os Eventos de Polling podem ser citados como um exemplo desta categoria. Eles são responsáveis por obter o número de sessões e composições ativas num dado momento.

A Figura 5.10 ilustra o fluxo de uma solicitação de serviço típica no ANCSim.

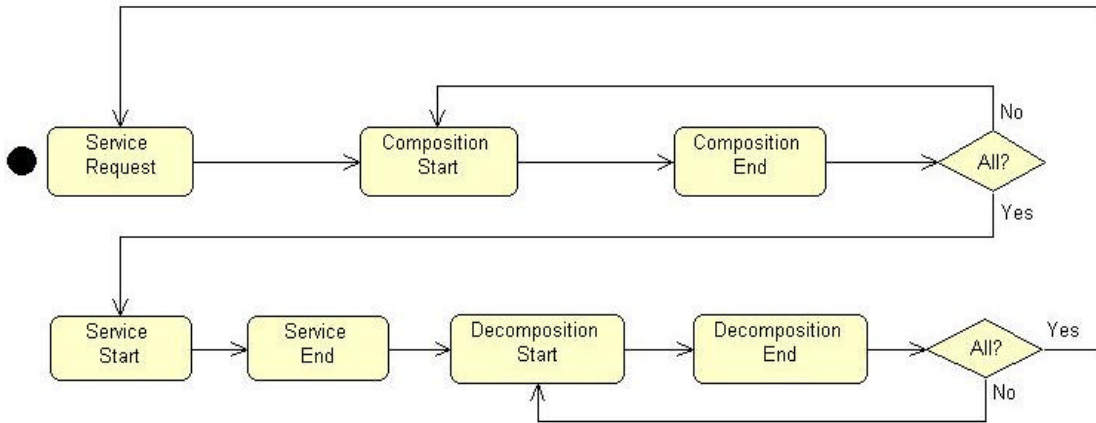


Figura 5.10. Fluxo de uma solicitação de serviço no ANCSim.

O processo inicia quando um serviço é solicitado por uma rede Agente. Após ser encontrado um caminho entre as redes de origem e de destino, o processo de composição entre as redes que compõem o caminho é iniciado. Quando o processo de composição é completado, o serviço é fornecido. A finalização do serviço dispara o processo de decomposição entre as redes ao longo do caminho. Finalmente, quando as decomposições terminam, uma outra solicitação para este serviço pode ser processada.

6. AVALIAÇÃO DAS COMPOSIÇÕES

Este capítulo apresenta os resultados dos processos de composição e decomposição em Redes de Ambiente utilizando o modelo implementado no simulador ANCSim. Primeiramente, são descritas as medições realizadas na rede GPRS da Brasil Telecom. Depois, apresenta-se o ambiente de simulação utilizado no experimento, bem como a configuração do cenário incluindo a topologia, os parâmetros e as métricas. Finalmente, são apresentados e analisados os resultados obtidos com a simulação.

6.1. METODOLOGIA

A metodologia utilizada neste trabalho para a avaliação da composição de redes está representada na Figura 6.1 a seguir.

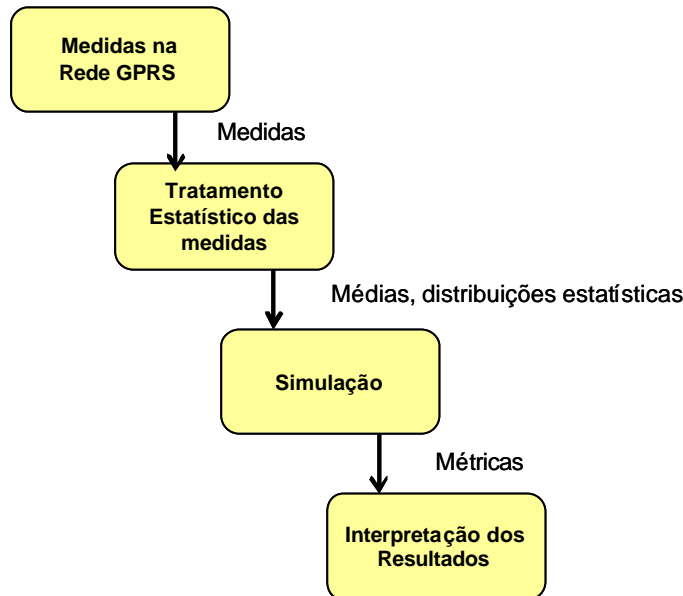
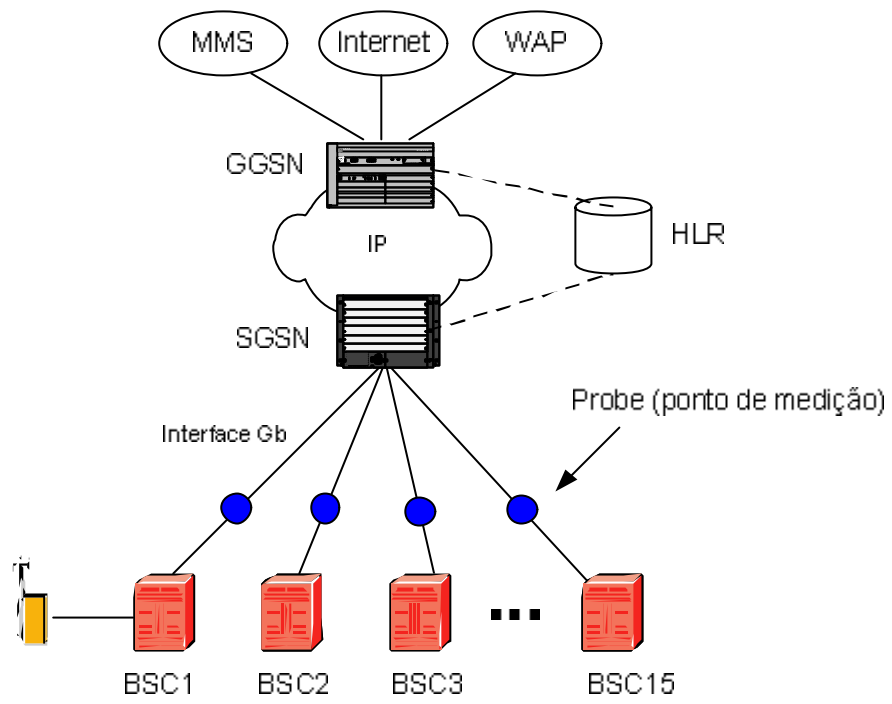


Figura 6.1. Metodologia adotada.

6.2. MEDIÇÕES NA REDE GPRS

Com o objetivo de fornecer resultados mais próximos da realidade, o modelo de simulação utiliza valores para duração das composições, taxa de solicitação de serviços e duração de serviço, obtidos através de medidas realizadas na rede GPRS da Brasil Telecom. Distribuições estatísticas foram derivadas a partir das medidas realizadas, as quais foram utilizadas para gerar os valores destes parâmetros durante a simulação.

A Figura 6.2 ilustra a topologia simplificada da rede GPRS utilizada durante as medições.



Para realizar as medidas, as mensagens de sinalização interceptadas na interface Gb foram enviadas a um Analisador de Protocolos que posteriormente realiza a filtragem, decodificação e armazenamento das mensagens, conforme ilustra a Figura 6.3.

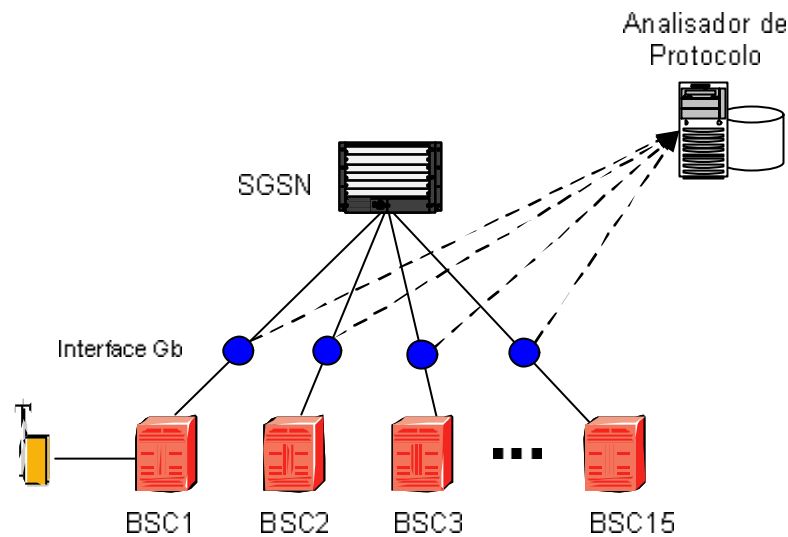


Figura 6.3. Decodificação, Análise e Armazenamento das mensagens.

Para as medidas de duração de Contexto PDP (sessão de dados), utilizamos o produto comercial Protocol Analysis da Agilent como analisador de protocolo. Para as medidas de duração do processo de Attach (registro na rede), utilizamos o produto comercial Session Trace da Agilent como analisador de protocolo.

6.3. COMPOSIÇÃO AGENTE/PDN

Uma composição do tipo Agente/PDN, tipicamente um processo de autenticação ou login, corresponde na rede GPRS ao processo de registro da estação móvel na rede, conhecido como processo de Attachment (ou simplesmente Attach). O processo de attach envolve a autenticação do usuário pela rede GPRS, através de consulta à base de assinantes (HLR).

Assim, utilizaremos as medidas de tempo de attach como medida para a duração da composição do tipo Agente/PDN.

O tempo de Attach (registro) é medido pela diferença temporal entre as mensagens de *Attach Request* e *Attach Accept*, conforme Figura 6.4.

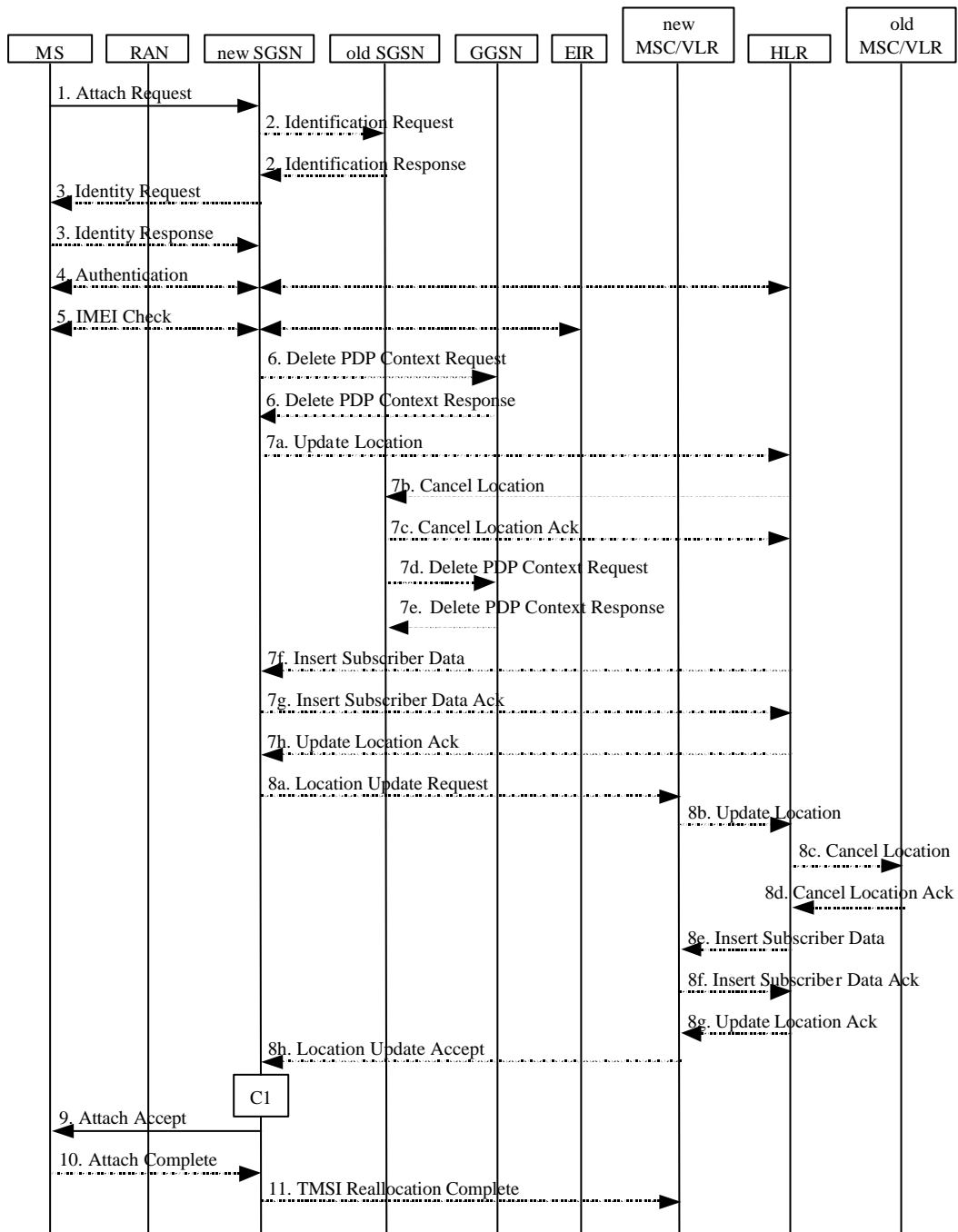
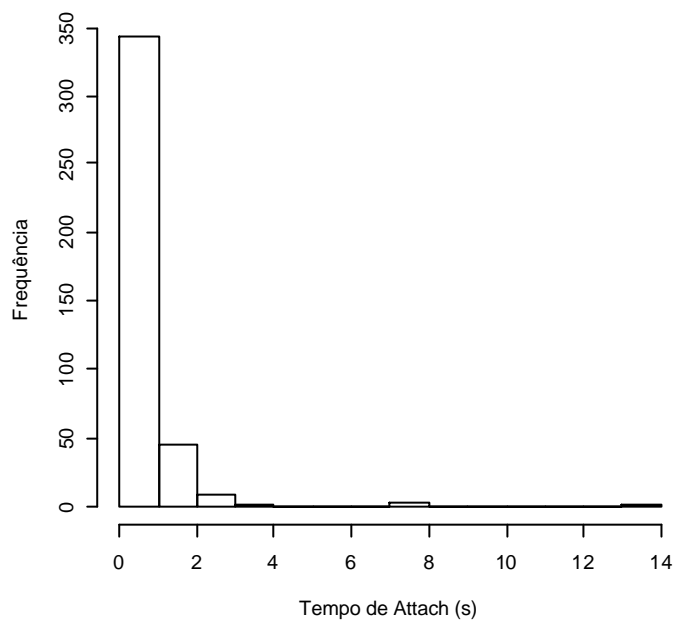


Figura 6.4. Procedimento de Attach.

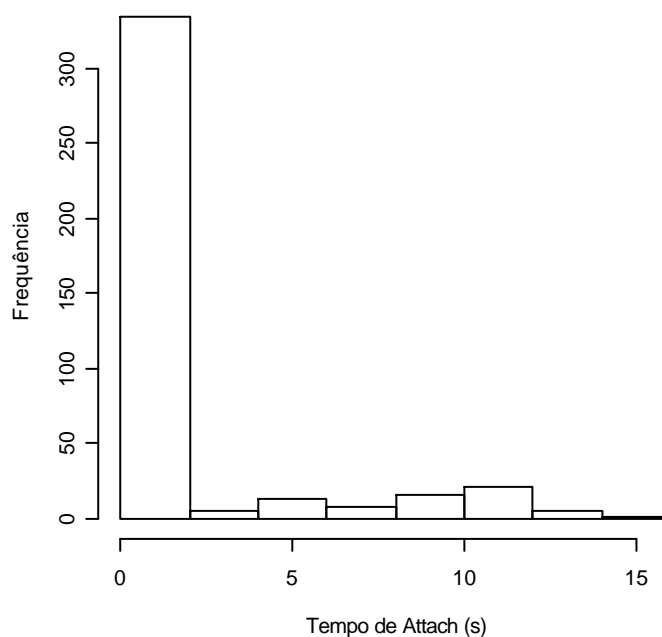
Foram interceptadas as mensagens de Attach Request e Attach Accept com suas marcas de tempos (timestamps) e realizada a diferença temporal para se obter o tempo transcorrido para realizar o attach na rede.

Foram realizadas medidas de tempo de Attach em quatro dias distintos. Os resultados obtidos estão representados nos histogramas a seguir.

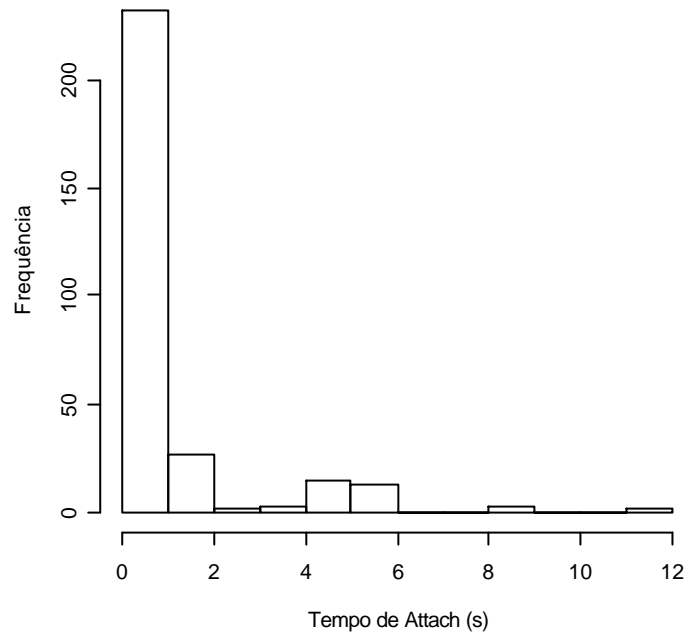
Composição Agent/PDN - 25/09/07



Composição Agent/PDN - 26/09/07



Composição Agent/PDN - 29/09/07



Composição Agent/PDN - 30/09/07

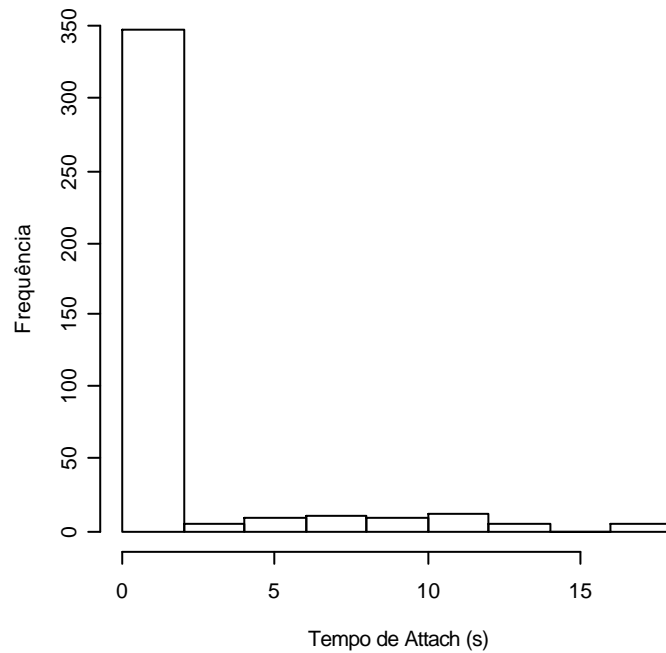


Figura 6.5. Tempos de composição Agente/PDN.

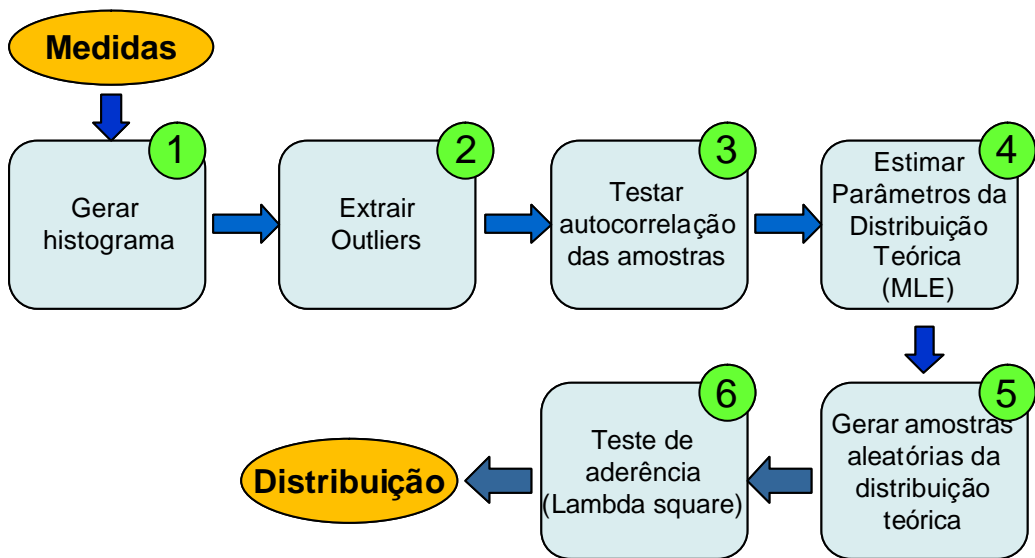
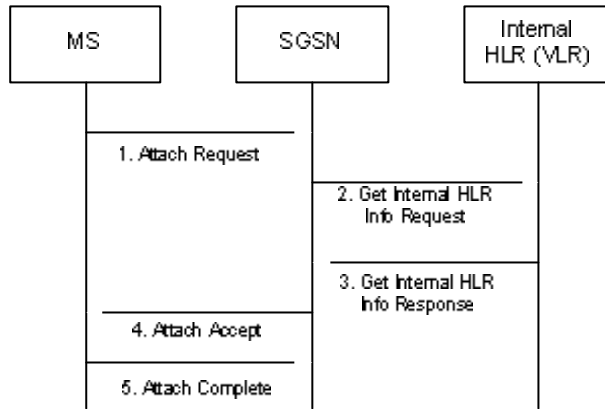
A Tabela 6.1 mostra os valores em segundos da média e desvio padrão para os tempos de attach obtidos, representando a duração da composição Agente/PDN.

Medida	Média (s)	Desvio Padrão (s)
Dia 25/09/07	0.60511	1.024647069
Dia 26/09/07	1.877488	3.23084624
Dia 29/09/07	0.99366	1.666248793
Dia 30/09/07	1.681943	3.228085789

Tabela 6.1. Média e Desvio Padrão para composição Agente/PDN.

Pode-se perceber nos histogramas da Figura 6.5 a presença de Outliers. Um Outlier é uma observação que está a uma distância anormal dos outros valores em uma amostra aleatória de uma população. Em outras palavras, um outlier é uma observação que está numericamente distante do resto dos dados. A presença de outliers dificulta a análise estatística.

Neste caso em particular, os outliers correspondem a tempos de attach na casa das dezenas de segundos. Estes tempos elevados de attach ocorrem quando as informações de usuário ainda não se encontram no banco de dados interno do SGSN, chamado de VLR (Visitor Location Register), forçando o SGSN a trocar mensagens com outros elementos da rede GPRS (antigo SGSN, HLR, EIR) para obter estas informações. O tempo gasto para a troca destas mensagens contribui para aumentar o tempo total para a conclusão do procedimento de attach. Para a maioria dos casos, porém, as informações dos usuários já se encontram no VLR interno ao SGSN (obtidas em attaches anteriores), não havendo necessidade de consulta a outros elementos (conforme ilustra a Figura 6.6), o que reduz bastante o tempo de attach.

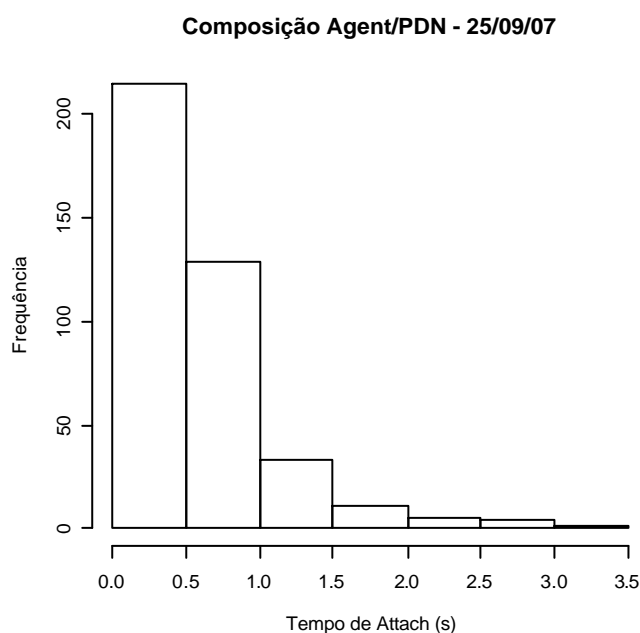


A extração dos outliers do conjunto de observações foi realizada utilizando-se o método conhecido como “box plot” [17]. O procedimento consiste em desenhar uma caixa entre o maior e o menor quartil com uma linha sólida através da caixa representando a média. Os seguintes quartis, chamados de cercas (*fences*, em inglês), são necessários para identificar valores extremos na distribuição:

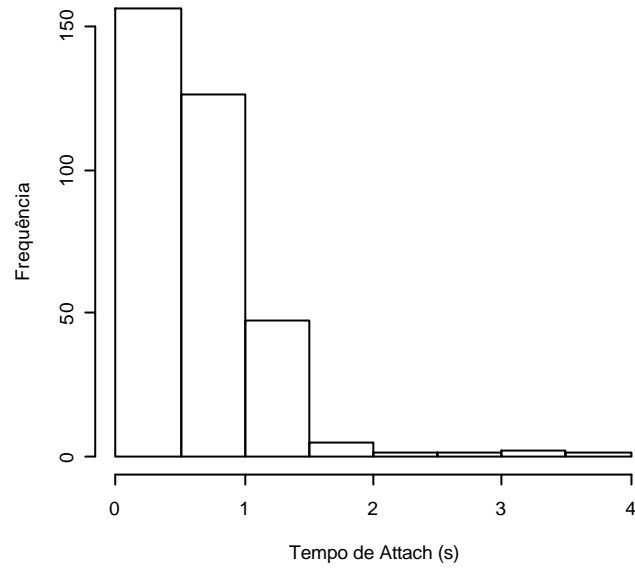
- Lower inner fence = $Q1 - 1.5*IQ$.
- Upper inner fence = $Q3 + 1.5*IQ$.
- Lower outer fence = $Q1 - 3*IQ$.
- Upper outer fence = $Q3 + 3*IQ$.

Onde Q1 e Q3 são o primeiro e terceiro quartil, respectivamente. IQ é o range interquartil ($Q3 - Q1$). Um ponto além de uma “cerca mais interna” (inner fence) em qualquer dos lados é considerado um outlier mediano (*mild outlier*). Um ponto além de uma “cerca mais externa” (outer fence) é considerado um outlier extremo (*extreme outlier*). Neste trabalho, excluimos apenas os outliers extremos ($Q3 + 3*IQ$).

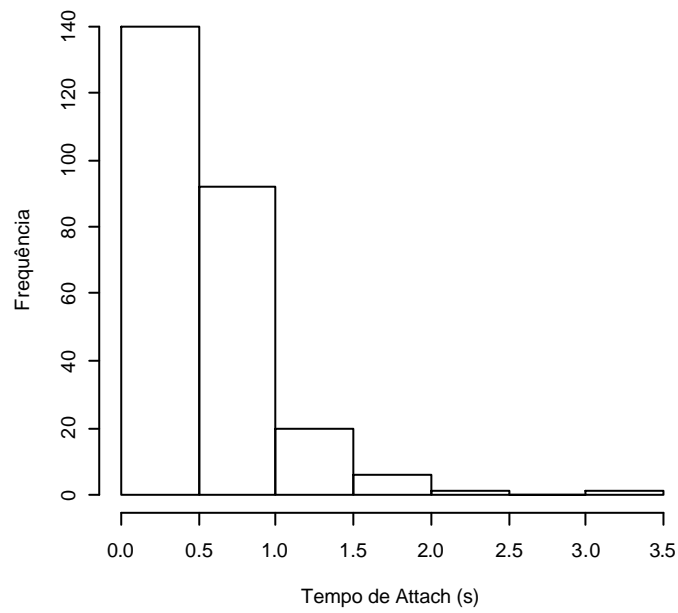
A Figura 6.8 mostra os novos histogramas com os outliers extremos excluídos.



Composição Agent/PDN - 26/09/07



Composição Agent/PDN - 29/09/07



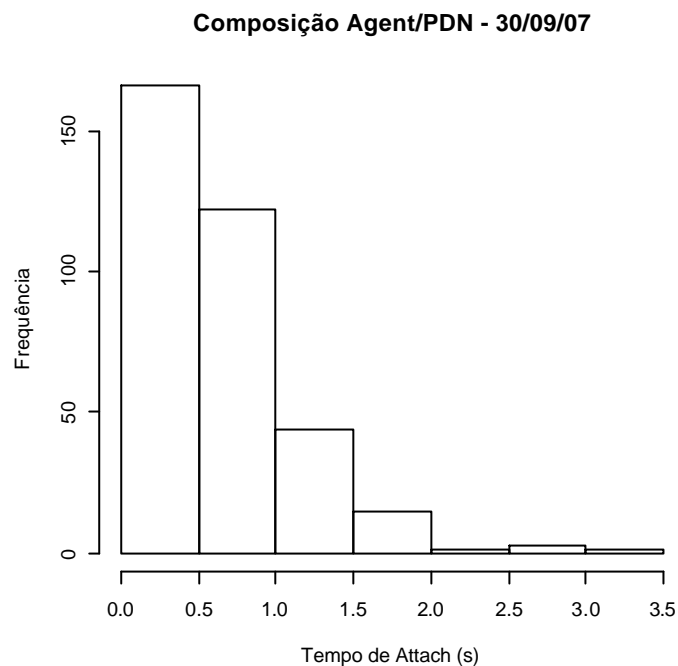


Figura 6.8. Tempos de composição Agente/PDN com extração de outliers.

A Tabela 6.2 mostra os novos valores de média e desvio padrão para os tempos de composição Agente/PDN obtidos após extração dos outliers.

Medida	Média (s)	Desvio Padrão (s)
Dia 25/09/07	0.535894	0.593122021
Dia 26/09/07	0.602139	0.564445039
Dia 29/09/07	0.472285	0.511741432
Dia 30/09/07	0.585582	0.573248671

Tabela 6.2. Média e Desvio Padrão para composição Agente/PDN (sem outliers).

6.4.2. Teste de aderência (Goodness-of-fit)

Para poder utilizar estes valores medidos de duração de composição no modelo do simulador, os dados foram testados usando-se o teste de aderência I^2 (lambda square) para descobrir se eles seguem alguma distribuição estatística conhecida (Distribuição Normal, Lognormal, Exponencial, etc.). O teste I^2 é o teste de aderência mais adequado para distribuições contínuas, com tamanho da amostra grande e valores de amostra auto-correlacionados. A auto-correlação das amostras será demonstrada na seção 7.2.2.1 através de teste de auto-correlação.

Utilizamos também técnicas visuais para analisar as amostras. Foram utilizados o histograma, o gráfico Q-Q (QQ Plot) e a Função de Densidade de Probabilidade (PDF).

6.4.2.1. Teste de Auto-Correlação

Utilizamos inicialmente o teste de aderência de Kolmogorov-Smirnov, porém os resultados não foram confiáveis. Foram realizados testes de aderência para as distribuições Normal, Lognormal, Exponencial, Poisson, Gamma e Chi-squared. Em todos os casos, a hipótese *Null* foi rejeitada. Os parâmetros da distribuição teórica foram estimados a partir do conjunto de medidas através do método Analógico que consiste em estimar os parâmetros da distribuição teórica aplicando-se a mesma função aos dados empíricos. Por exemplo, estimamos a média desconhecida de uma população normal usando a média das medidas.

Com os resultados obtidos com o teste de Kolmogorov-Smirnov, suspeitamos que os valores das medidas fossem autocorrelacionados, o que explicaria os resultados obtidos. Sabe-se que o teste de Kolmogorov-Smirnov apresenta resultados não confiáveis quando os valores da amostra são autocorrelacionados. Além disso, para se aplicar o teste de Kolmogorov-Smirnov a distribuição teórica deve ser completamente especificada, isto é, os parâmetros da distribuição não podem ser derivados das amostras.

A auto-correlação das amostras foi comprovada através de testes de auto-correlação disponível no software R. A seqüência de comandos no software R para realização do teste de auto-correlação para o conjunto de medidas de tempo de attach do dia 25/09/07 (vetor *x_attach_250907*) é descrito a seguir.

- `acf(x_attach_250907, type= c("correlation"), plot=TRUE)`
- `autocor= acf(x_attach_250907, type= c("correlation"))`

A Tabela 6.3 mostra os resultados.

Lag	Valor da Função de Auto-Correlação	Lag	Valor da Função de Auto-Correlação
0	1.000	13	0.765
1	0.974	14	0.754
2	0.954	15	0.742
3	0.934	16	0.732
4	0.916	17	0.722
5	0.898	18	0.712
6	0.880	19	0.701
7	0.862	20	0.692
8	0.844	21	0.683
9	0.825	22	0.674
10	0.805	23	0.665
11	0.792	24	0.656
12	0.779	25	0.647

Tabela 6.3. Valores para função de auto-correlação.

O resultado do teste pode ser melhor visualizado em forma gráfica (Figura 6.9) através da construção de um gráfico com os valores da função de auto-correlação (Auto-Correlation Function – ACF).

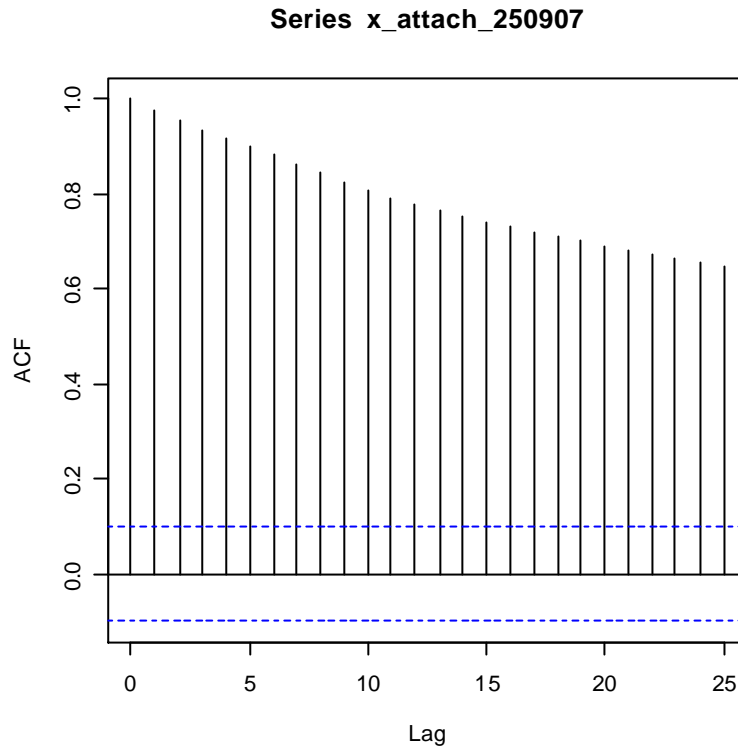


Figura 6.9. Função de auto-correlação.

Pelo gráfico da Figura 6.9, os valores das medidas não seriam auto-correlacionados se os valores de ACF (?) estivessem próximos de zero, o que não ocorre neste caso. Portanto, podemos afirmar que os valores das medidas são auto-correlacionados. Desta forma optou-se pelo teste de aderência I^2 (lambda square).

Podemos também fazer uma avaliação analítica para decidir se as medidas são auto-correlacionadas considerando-se o valor de ACF para Lag=1. Se uma variável aleatória é correlacionada, a função de auto-correlação é normalmente máxima para um lag=1. Se $|g(1)| < 2/\sqrt{n}$ (onde n é o número de observações no conjunto de medidas, neste caso, 397 observações), então as medidas são consideradas não correlacionadas.

Para a série x_attach_250907 em estudo, $|g(1)| = 0.974 > 2/\sqrt{397} = 0.1$. Portanto, podemos afirmar que as medidas são auto-correlacionadas.

6.4.2.2. Estimativa de parâmetros da distribuição teórica

Os parâmetros das distribuições teóricas utilizadas nos testes de aderência foram estimados utilizando-se o método de Máxima Verossimilhança (Maximum Likelihood Estimation – MLE). No método MLE, temos uma variável aleatória com uma função densidade de probabilidade (PDF) $f(x, \theta)$ conhecida, descrevendo a população. Devemos estimar o vetor de parâmetros constantes e desconhecidos θ de acordo com os dados amostrados: x_1, x_2, \dots, x_n . O método MLE inicia com a expressão matemática conhecida como uma função de verossimilhança dos dados da amostra. Basicamente, a verossimilhança de um conjunto de dados é a probabilidade de se obter aquele conjunto de dados em particular dada a distribuição teórica escolhida. Esta expressão contém os parâmetros desconhecidos. Aqueles valores dos parâmetros que maximizam a função de verossimilhança da amostra são conhecidos como a estimativa de máxima verossimilhança ou MLE.

Define-se a função de verossimilhança como

$$L(x_1, x_2, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta) \quad (7.1)$$

O método MLE consiste em encontrar θ que maximize $L(x_1, x_2, \dots, x_n, \theta)$.

O método MLE está implementado no software R através da função *fitdistr*. Utilizamos esta função para estimar os parâmetros das distribuições teóricas, conforme exemplos a seguir.

Para distribuição Normal:

```
fitdistr (x_attach_250907, "normal")  
  
mean      sd  
0.53589421 0.59237454  
(0.02973043) (0.02102259)
```

Valores de média (mean) e desvio padrão (Standard Deviation – sd) estimados.

Para distribuição Lognormal:

```
fitdistr (x_attach_250907, "lognormal")  
  
meanlog    sdlog  
-1.77688774  1.83884954  
( 0.09228921) ( 0.06525833)
```

Para distribuição Exponencial:

```
fitdistr (x_attach_250907, "exponential")  
  
rate  
  
1.86603995  
(0.09365386)
```

Para distribuição Chi-squared:

```
fitdistr (x_attach_250907, dchisq, list  
(df=0.05))  
  
df  
  
0.82593750  
(0.03827867)
```

6.4.2.3. Teste de aderência I^2

A medida de discrepância I^2 é uma estatística que depende do intervalo (ou bin) do histograma das amostras. O tamanho escolhido para a largura do intervalo afeta a precisão da estatística. Foram utilizadas técnicas bem conhecidas para calcular com precisão a

largura do intervalo. A largura do intervalo para um conjunto de amostras com uma distribuição enviesada deve ser calculada diferentemente daqueles com uma distribuição simétrica. Utilizamos uma técnica sugerida por Scott [19] para cálculo da largura dos intervalos.

Os testes revelaram que a melhor aderência (com menor discrepância I^2) ocorreu para uma distribuição Chi-squared. Mostramos a seguir o procedimento de teste para este caso, com a seqüência de funções do software R e suas saídas.

```
rfit <- fitdistr(x_attach_250907, "chi-squared", list (df=5))
```

Esta função estima parâmetros para distribuição Chi-squared (df) usando método MLE. O parâmetro df (degree of freedom) é o grau de liberdade, que é dado por $n-r-1$, onde n é o número de intervalos no histograma e r é o número de parâmetros a serem estimados. Para a série `x_attach_250907` com os valores de medidas de tempos de attach do dia 25/09/07, $n=7$ e $r=1$. Logo, $df=7-1-1=5$.

A função “list” mostra o valor estimado usando o método MLE para o parâmetro df .

```
list (rfit)
      df
0.82592773
(0.03827827)
```

A função `rchisq` gera `length(x_attach_250907)` amostras aleatórias seguindo uma distribuição Chi-squared com os parâmetros estimados anteriormente.

```
dist <- rchisq(length(x_attach_250907), df=rfit$estimate[1])
```

A função *lambdasquare* calcula os valores das discrepâncias χ^2 e I^2 e a variância. Estes valores são armazenados na variável *rlambda*.

```
rlambda <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_  
250907","Lambda-square")
```

A função *list* lista estes valores.

```
list (rlambda)  
[[1]]$x2  
[1] 62.53659  
[[1]]$lambda2  
[1] 0.1437074  
[[1]]$dev  
[1] 0.04422132
```

O valor da variável `$lambda2` representa o grau de discrepância entre a distribuição empírica (medidas) e a distribuição analítica (teórica). A aderência da distribuição empírica à distribuição analítica é maior quanto mais próximo de zero for o valor de `$lambda2`. Podemos afirmar que uma distribuição empírica segue a distribuição analítica com os parâmetros estimados pelo MLE se o valor da discrepância estiver entre 0 e 1. No caso em análise, como o valor da discrepância é 0.1437074, pode-se afirmar que as medidas seguem uma distribuição Chi-squared com grau de liberdade=0.82592773.

A Figura 6.10 a seguir ilustra as funções densidade de probabilidade (PDF) para a distribuição empírica e analítica Chi-squared.

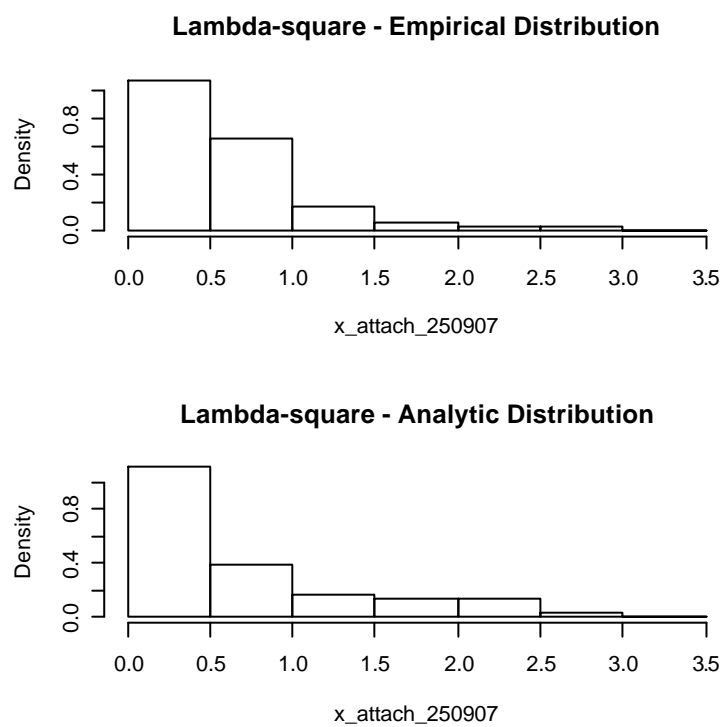


Figura 6.10. Comparação PDFs para distribuição empírica e analítica.

A Figura 6.11 a seguir ilustra o gráfico QQ para distribuição Chi-squared.

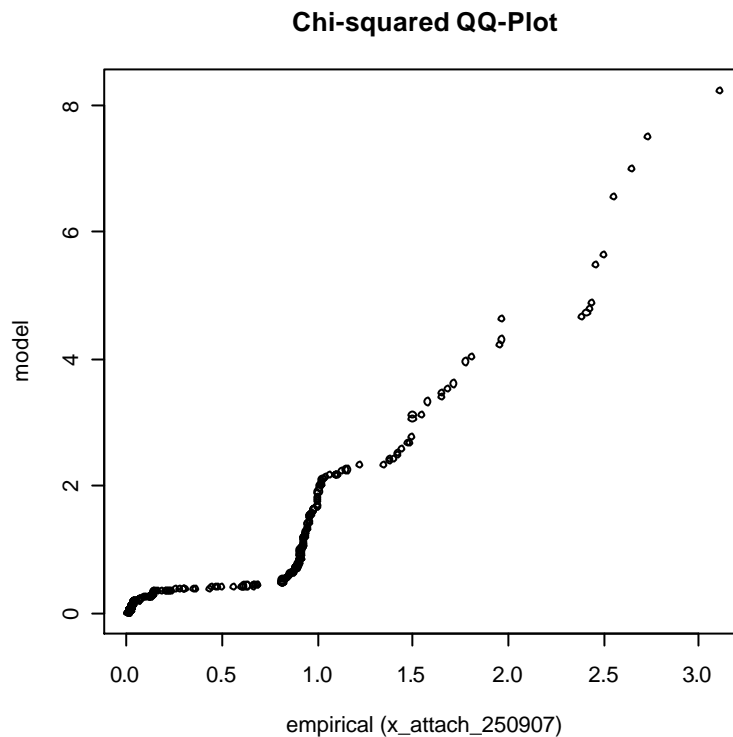


Figura 6.11. QQ-Plot para distribuição Chi-squared.

Os resultados dos demais testes de aderência estão descritos no Apêndice A.

6.5. COMPOSIÇÃO PDN/PDN

Como não existe na rede GPRS um procedimento equivalente a uma composição do tipo PDN/PDN, utilizamos os resultados obtidos por Jennifer Lima [19] em um protótipo em laboratório. O cenário usado por Lima para realizar as medidas dos tempos de composição PDN/PDN está ilustrado na Figura 6.12. Este cenário representa duas PDNs, cada uma composta por um P-Node. O experimento consistiu em compor estas duas PDNs e coletar o tempo decorrido para completar este processo.

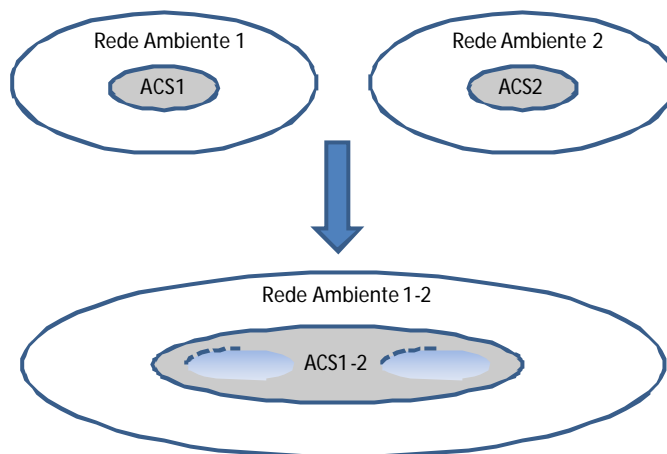


Figura 6.12. Composição PDN/PDN.

Os parâmetros usados nas medições estão representados na Tabela 6.4. Ambas as PDNs são servidores Linux. As medidas foram repetidas 1000 vezes, isto é, foram obtidos 1000 tempos de composição PDN/PDN. Foram definidos seis níveis diferentes de carga que representam diferentes tipos de serviço sendo providos por cada PDN (1, 10, 20, 30, 40 e 50 serviços).

Parâmetro	Valor
Número de Replicações	1000
Número de PDNs	2
Número de P-Nodes	1
Carga	L1, L2, L3, L4, L5 e L6

Tabela 6.4. Parâmetros utilizados nas medições.

A carga é determinada pela quantidade de informação armazenada na PDN e segue os valores mostrados na Tabela 6.5. Os valores são proporcionais ao número de serviços. Em outras palavras, o número de serviços determina o número dos outros parâmetros. Esta proporção é mostrada na Tabela 6.6.

Parâmetros	Valores
-------------------	----------------

	L1	L2	L3	L4	L5	L6
Acessos	1	10	20	30	40	50
Serviços	1	10	20	30	40	50
Grupos	2	20	40	60	80	100
Usuários	100	1000	2000	3000	4000	5000
Políticas	5	50	100	150	200	250

Tabela 6.5. Parâmetros.

Parâmetros	Valores
Acessos	1
Serviços	1
Grupos	2
Usuários	100
Políticas	5

Tabela 6.6. Proporções.

Os resultados das medidas realizadas mostraram que os tempos de composição variaram entre 8 e 30 segundos dependendo da quantidade de informação (definida pelos serviços) a ser publicada na nova PDN. Este resultado confirma que uma composição do tipo PDN/PDN é uma operação custosa e toma um tempo considerável para ser completada. A Tabela 6.7 mostra os valores da média e desvio padrão para os dados coletados, considerando diferentes níveis de carga.

Carga	Média (ms)	Desvio Padrão (ms)
1 serviço	8731.03	181.673
10 serviços	9387.34	322.2046
20 serviços	11818.92	1301.812
30 serviços	17297.37	2228.206
40 serviços	21552.05	3765.524
50 serviços	29740.15	4653.615

Tabela 6.7. Médias e Desvios Padrão para composição PDN/PDN.

Após as medições dos tempos de composição, os dados foram testados estatisticamente usando-se o teste de aderência de Kolmogorov-Smirnov para saber se eles seguem alguma distribuição conhecida. Contudo, todas as hipóteses testadas foram rejeitadas. Assumindo isso, a alternativa foi construir uma distribuição empírica. A distribuição empírica foi construída considerando-se o número de usuários em cada PDN e as medidas realizadas, conforme ilustrado na Tabela 6.8. Esta tabela significa que dependendo do número de usuários envolvidos em uma dada composição (composta de duas PDNs), será considerada uma distribuição empírica diferente nas simulações.

Número total de usuários	Distribuição
<200	Dist1
200-2000	Dist2
2000-4000	Dist3
4000-6000	Dist4
6000-8000	Dist5
>8000	Dist6

Tabela 6.8. Distribuições.

As distribuições empíricas foram construídas considerando os quantiles de 1 a 100 de cada medida específica realizada e foi usado para determinar o tempo de composição das composições PDN/PDN.

6.6. DURAÇÃO DE SERVIÇO

6.6.1. Rede GPRS Brasil Telecom

Foram realizadas medidas na rede GPRS para a duração de Contextos PDP (sessões de dados GPRS). Estes valores foram utilizados para modelar a duração de serviço no simulador.

A duração do Contexto PDP corresponde à diferença temporal entre o recebimento das mensagens *Activate PDP Context Accept* e *Deactivate PDP Context Accept*. O procedimento de Ativação de Contexto PDP está ilustrado na Figura 6.13.

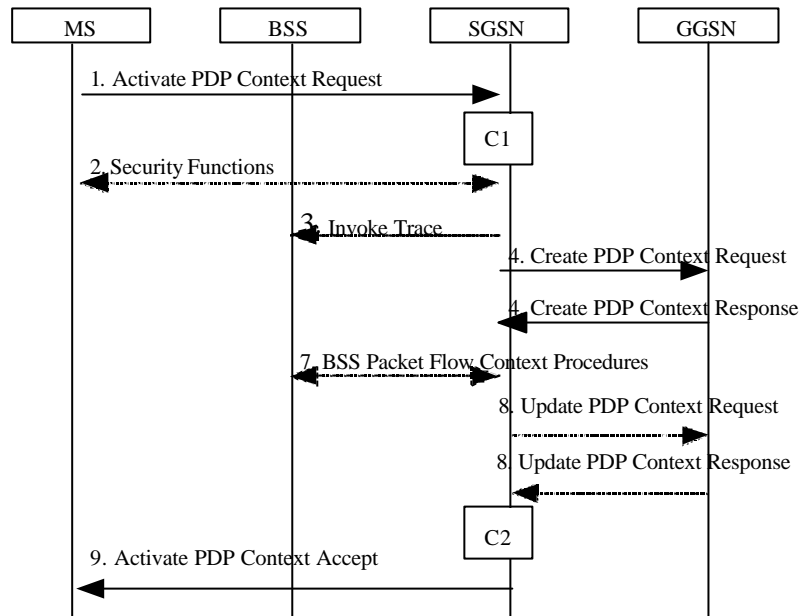


Figura 6.13. Procedimento de Ativação de Contexto PDP.

O procedimento de Desativação de Contexto PDP está ilustrado na Figura 6.14.

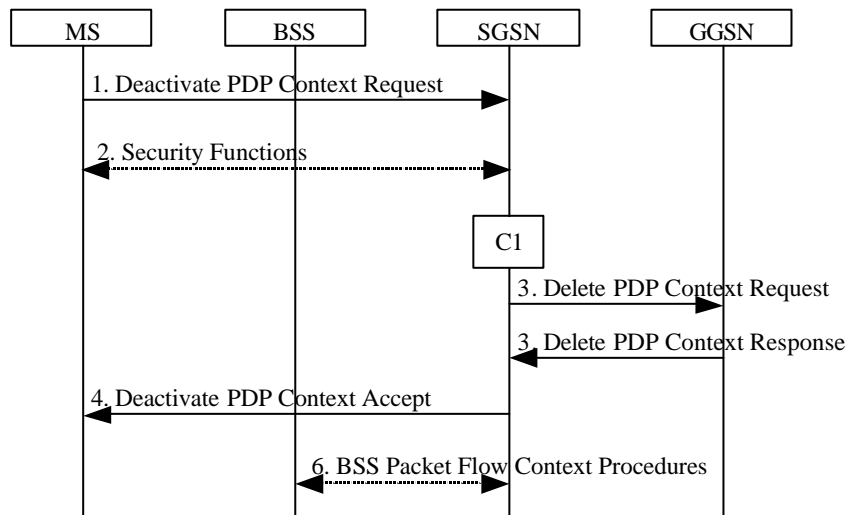


Figura 6.14. Procedimento de Desativação de Contexto PDP.

Foram realizados sete conjuntos de medidas para a duração do Contexto PDP (tempo em que o Contexto PDP permaneceu ativo), em dias e horários distintos.

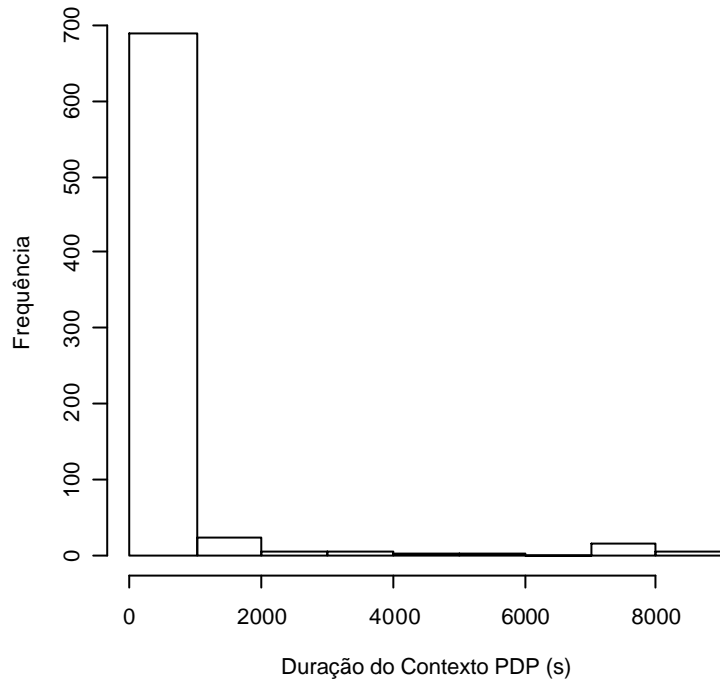
A Tabela 69 mostra os valores em segundo da média e desvio padrão para os dados coletados representando a duração do serviço.

Medida	Média (seg.)	Desvio Padrão (seg.)
Medida 1	203,751	852,0428978
Medida 2	169,1292	742,5270529
Medida 3	530,4069	2229,850219
Medida 4	167,8316	610,3679061
Medida 5	195,2046	689,9869983
Medida 6	138,7143	554,9101996
Medida 7	42,48061	99,25378701

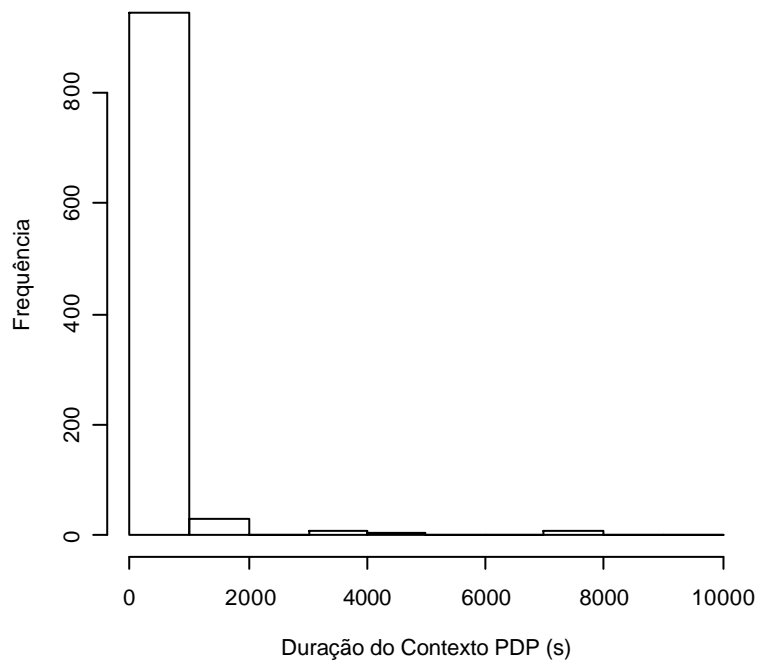
Tabela 6.9. Média e Desvio Padrão para a duração do serviço.

De maneira análoga ao item 7.2.1, foram identificados outliers nas amostras de duração de Contexto PDP, o que pode ser visualizado nos histogramas da Figura 6.15 a seguir:

Duração de Serviço - 30/08/07



Duração de Serviço - 03/09/07



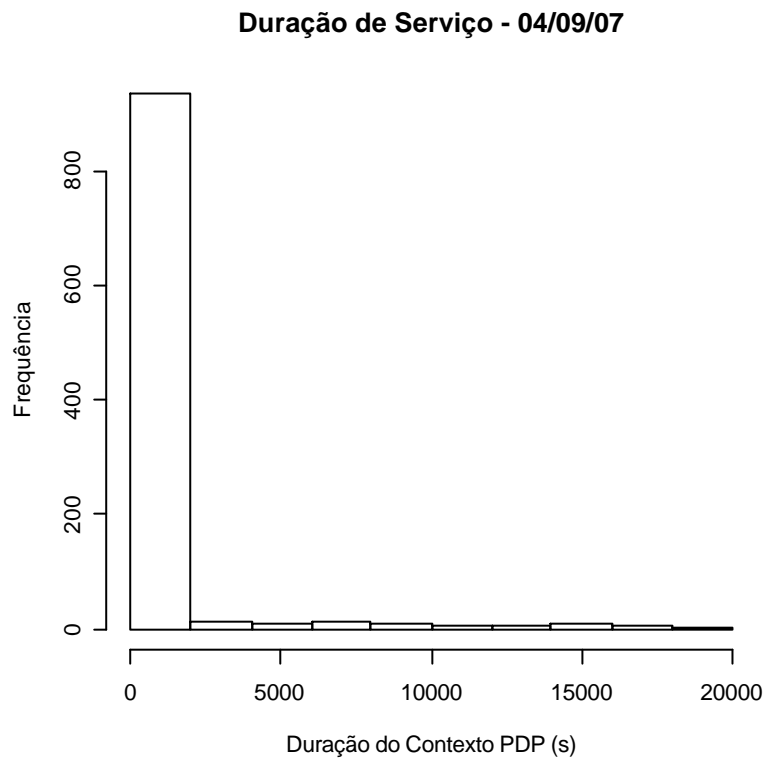
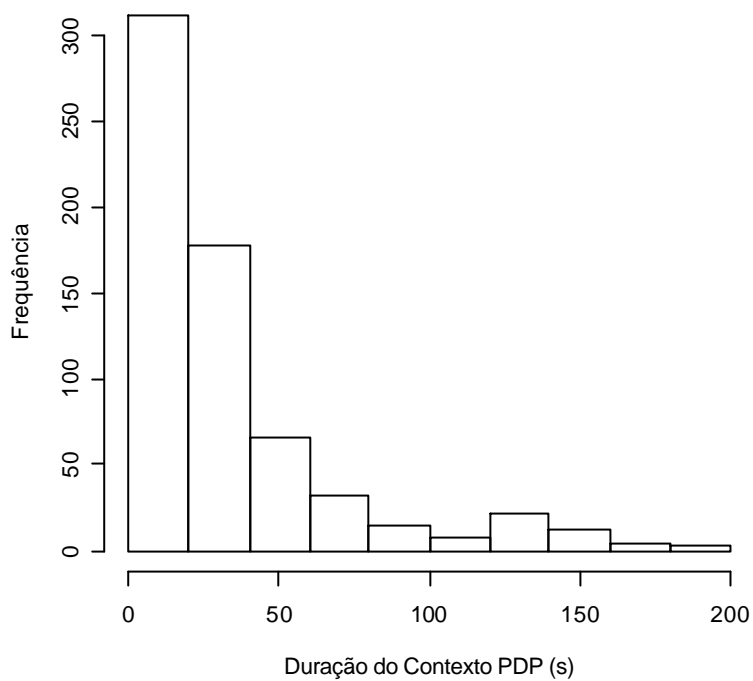


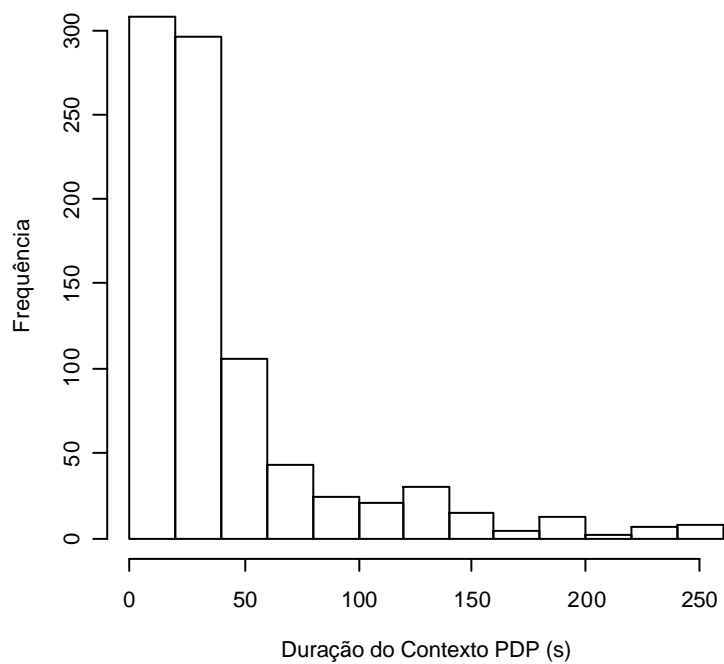
Figura 6.15. Duração de Serviço.

Seguindo o procedimento descrito na seção 7.2.1, foram extraídos os outliers mais extremos das amostras. A Figura 6.16 mostra os novos histogramas com os outliers extremos excluídos.

Duração de Serviço - 30/08/07



Duração de Serviço - 03/09/07



Duração de Serviço - 04/09/07

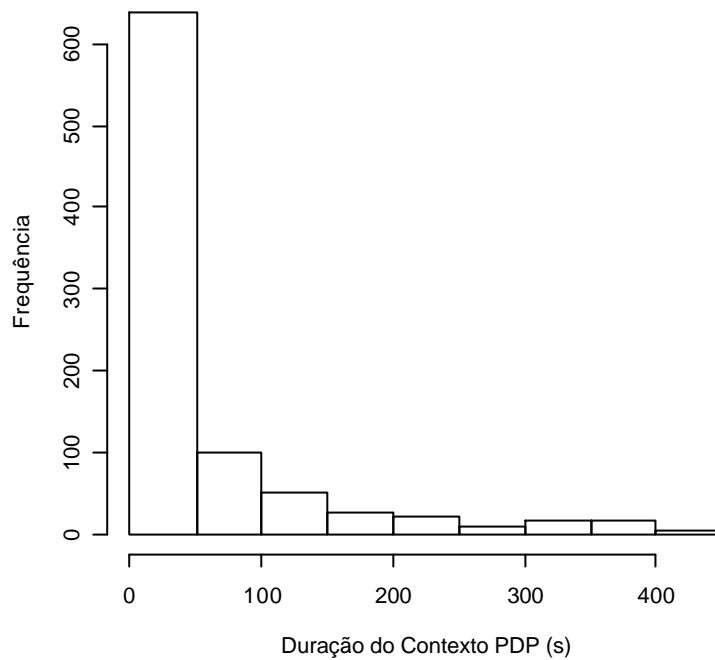


Figura 6.16. Duração de Serviço com extração de outliers.

A Tabela 6.10 mostra os novos valores de média e desvio padrão para duração de serviço obtidos após extração dos outliers.

Medida	Média (seg.)	Desvio Padrão (seg.)
Medida 1	43.632	45.54324399
Medida 2	59.38488	75.98503506
Medida 3	30.17580647	31.90752584
Medida 4	32.23064772	37.99875758
Medida 5	28.13109758	31.18580806
Medida 6	29.061225	30.31577855
Medida 7	32.33590139	36.53150114

Tabela 6.10. Média e Desvio Padrão para a duração do serviço com extração de outliers.

Utilizamos para simulação a média das médias obtidas em cada conjunto de medidas como sendo o valor da duração média de serviço. O valor calculado foi de 36.42 segundos, que arredondaremos para *36 segundos*.

6.7. TAXA DE SOLICITAÇÃO DE SERVIÇO

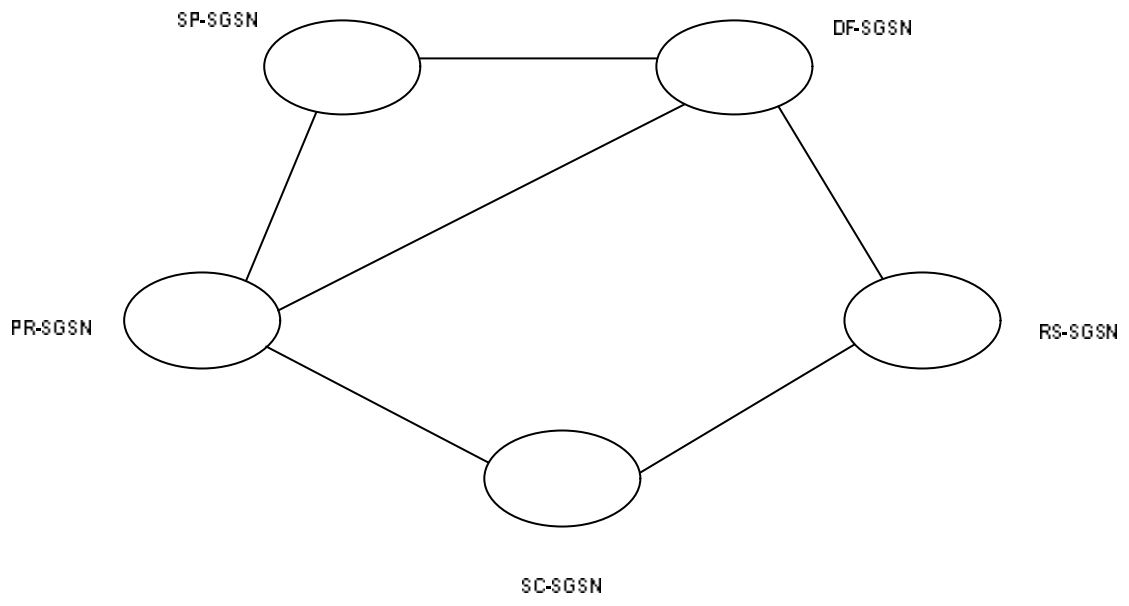
Foram realizados oito conjuntos de medidas para a taxa de solicitação de serviço, que no caso de uma rede GPRS corresponde à taxa de solicitação para estabelecimento de Contextos PDP. A taxa média entre as oito medidas foi de 6.89 solicitações / minuto.

6.8. AMBIENTE DE SIMULAÇÃO E CONFIGURAÇÃO DO EXPERIMENTO

Esta seção apresenta a topologia escolhida para rodar a simulação, bem como o ambiente de simulação, parâmetros, fatores e níveis que compreendem as configurações do experimento, e as métricas obtidas.

6.8.1. Topologia de Rede

Consideramos um cenário futuro para a rede GPRS da Brasil Telecom, onde os elementos SGSN (considerados aqui como PDNs) estão interligados entre si de acordo com uma topologia, conforme mostrado na Figura 6.17.



A Tabela 6.12 mostra os parâmetros utilizados na simulação. Os valores foram escolhidos com base em simulações preliminares.

Parâmetro	Descrição	Valor
Tempo de simulação	Tempo total de simulação em horas	1h
Número de replicações	Número de replicações de uma série de experimentos	30 replicações
Serviços	Serviços oferecidos pelas redes que compõem a topologia	VoIP e QoS
Duração	Duração média de serviço em cada rede Agente	36 segundos
Peso (Weight)	Importância do serviço em cada rede Agente	1 para todos os serviços

Tabela 6.12. Parâmetros de simulação.

Uma composição acontece toda vez que um serviço é solicitado por um usuário. Este trabalho considera um serviço de VoIP e um serviço de QoS para gerar uma carga e solicitações de composição no simulador. A duração do serviço segue uma distribuição exponencial com média igual à média obtida para a duração do Contexto PDP na rede GPRS da Brasil Telecom. O peso é o mesmo para todos os serviços, isto é, cada rede agente pode ser escolhida como alvo com a mesma probabilidade. O tempo de simulação é de 1 hora e o experimento foi replicado 30 vezes.

A Tabela 6.13 representa os fatores e os níveis usados na simulação. Os fatores são variáveis que afetam o resultado do experimento. Cada fator tem um conjunto de valores

alternativos (níveis). Os níveis são os valores que cada fator pode assumir, isto é, cada nível é uma alternativa para o fator correspondente [17].

Fator	Descrição	Níveis
Topologia	Topologia a ser simulada	Topologia da Rede GPRS evoluída
Tempo de overtime de composição	Tempo que a rede espera para estabelecer a decomposição	0, 10, 50 e 100%
Composições Agente/PDN por PDN (controle de admissão)	Porcentagem de composições do tipo Agente/PDN que cada PDN pode aceitar	10, 50 e 100%
Número de Redes Agentes	Número total de redes Agente na topologia	2000, 5000 e 10000 agentes
Carga	Frequência com que cada serviço é solicitado	0.25, 1, 3 e 7 solicitações / minuto

Tabela 6.13. Fatores e Níveis de Simulação.

O fator Tempo de Overtime de Composição é o tempo em que uma composição é mantida ativa quando nenhum serviço está usando ela. É uma maneira de evitar decomposições prematuras e é calculado sobre o tempo que a composição deve ser terminada por inatividade (0.5 de overtime significa que as redes permanecerão compostas por um tempo adicional de 50% do tempo de uso do serviço).

Não há diferença entre Agente e redes Agente, isto é, as redes Agente consistem de um usuário solicitando um serviço. Por simplificação, ambos são chamados Agente.

É realizado um controle de admissão através da restrição do número de composições Agente/PDN por PDN: 10% de composições Agente/PDN por PDN significa que apenas 10% das redes Agente conectadas com a rede PDN poderão compor.

O número de redes Agentes na topologia é dividido entre as redes PDNs de acordo com o tamanho da população da localidade. Por exemplo, o nó SP-SGSN na topologia atende a uma população maior (Cidade de São Paulo) do que o nó localizado no Estado de Santa Catarina (SC-SGSN). Portanto, o nó de São Paulo terá proporcionalmente mais agentes do que o nó de Santa Catarina porque receberá mais solicitações de serviço. Utilizamos dados de população por Estados fornecidos pelo IBGE (Instituto Brasileiro de Geografia e Estatística), conforme Tabela 6.14.

Estado	População (habitantes)	Proporção do total da população	Nó (PDN)
São Paulo	39 827 570	0.57	SP-SGSN
Distrito Federal	2 455 903	0.04	DF-SGSN
Paraná	10 284 503	0.15	PR-SGSN
Santa Catarina	5 866 252	0.09	SC-SGSN
Rio Grande do Sul	10 582 840	0.15	RS-SGSN

Tabela 6.14. População por Estado brasileiro e proporção de redes Agentes.

O último fator é a carga que representa a frequência com que cada serviço é solicitado em cada rede. Variamos a carga desde 0.25 solicitações por minuto até um valor médio de carga próximo ao valor medido na rede GPRS da Brasil Telecom (6.89 solicitações/minuto).

6.8.3. Rede UMTS

Foram realizadas também simulações com dados¹ de duração média do serviço de Vídeo Stream em redes móveis de 3ª Geração (3G) baseada no sistema UMTS (Universal Telecommunications System). A duração média deste serviço é de 177 segundos.

Procurou-se avaliar a influência da duração do serviço sobre o desempenho das composições de rede, com o objetivo de se verificar a viabilidade das Redes de Ambiente em um cenário constituído por sistemas UMTS.

6.8.4. Métricas consideradas

As métricas são saídas observáveis de um experimento. O seguinte conjunto de métricas foi primeiramente definido no começo deste trabalho e foi atualizado à medida em que o trabalho progredia, até que o conjunto final foi definido.

Solicitação de Serviço: Representa o número de solicitações para um serviço específico. Estas solicitações são feitas pelas redes Agente.

Solicitação de Serviço Bloqueada: Representa o número de solicitações de serviço que falharam.

Tempos de Resposta: Tempo que um usuário espera para que o serviço se torne disponível. Este tempo inclui a duração das composições solicitadas, isto é, todas as composições que são necessárias para fornecer este serviço.

Solicitação de Composição: Representa o número de composições solicitadas entre duas redes adjacentes.

Número de Composições: Esta métrica representa o número de composições bem sucedidas. A composição pode ser do tipo Agente/PDN ou PDN/PDN.

¹ Dados fornecidos pela empresa Alcatel-Lucent, a partir de medidas realizadas em redes UMTS na Europa.

Composições Desfeitas: Número de composições que precisam ser desfeitas por causa de uma falha de serviço.

Recomposição: Número de composições que precisam ser refeitas após serem decompostas.

Duração da Composição: Duração da composição.

Probabilidade de Composição: Probabilidade de uma solicitação de serviço gerar uma composição.

Composições Ociosas: Número de composições que não estão sendo usadas por nenhuma rede agente num momento específico.

Número de Decomposições: Número total de decomposições.

6.9. RESULTADOS DA SIMULAÇÃO

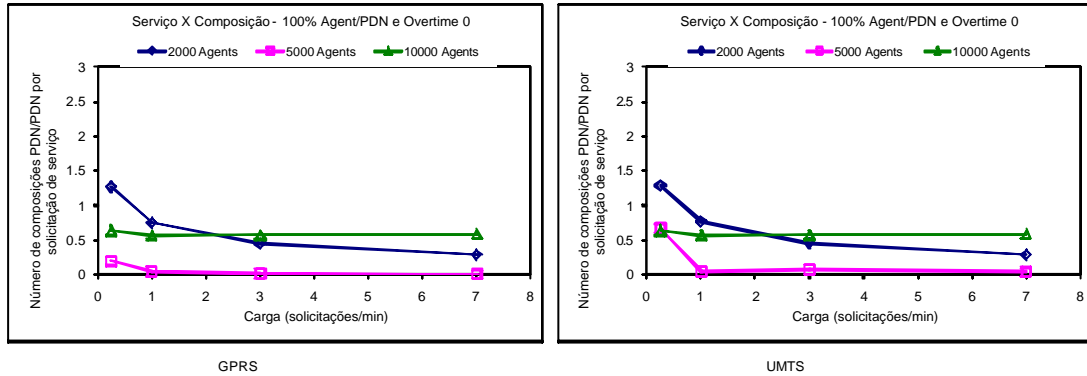
Esta seção apresenta a avaliação do desempenho de composição e decomposição em Redes de Ambiente. Os resultados consideram as métricas apresentadas na seção 7.6.3. Os valores médios das métricas são apresentados nos gráficos com 99% de intervalo de confiança, os quais foram calculados seguindo [20].

As próximas seções, divididas em Escalabilidade, Tempo de Resposta e Estabilidade da Composição, explicam como o processo de composição influencia o serviço percebido pelos usuários, considerando a topologia utilizada.

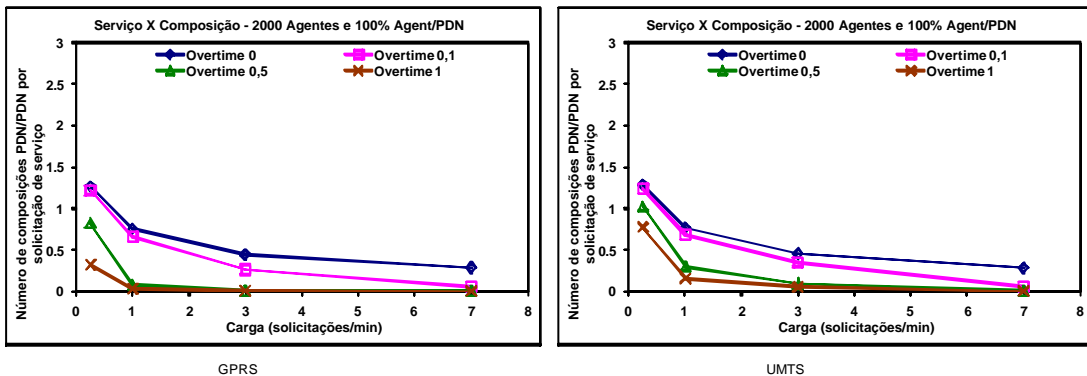
6.9.1. Escalabilidade

Esta seção avalia a escalabilidade da composição de redes em Redes de Ambiente. O objetivo desta avaliação é analisar se o sistema permanece escalável, não degradando o tempo de resposta para os usuários.

A escalabilidade será analisada com base nos gráficos da Figura 6.18 (a) e (b). Estes gráficos consideram que não há controle de admissão.



(a)



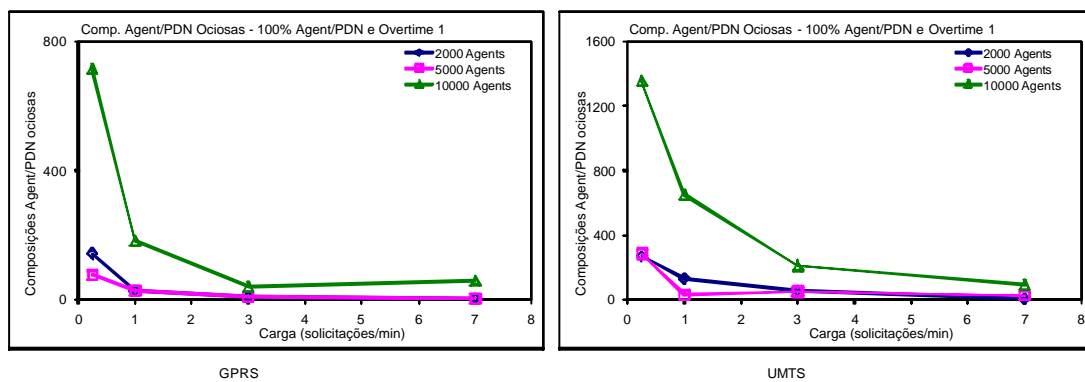
(b)

Figure 6.18. Número de composições por solicitação de serviço por tamanho de rede (a) e por overtime (b).

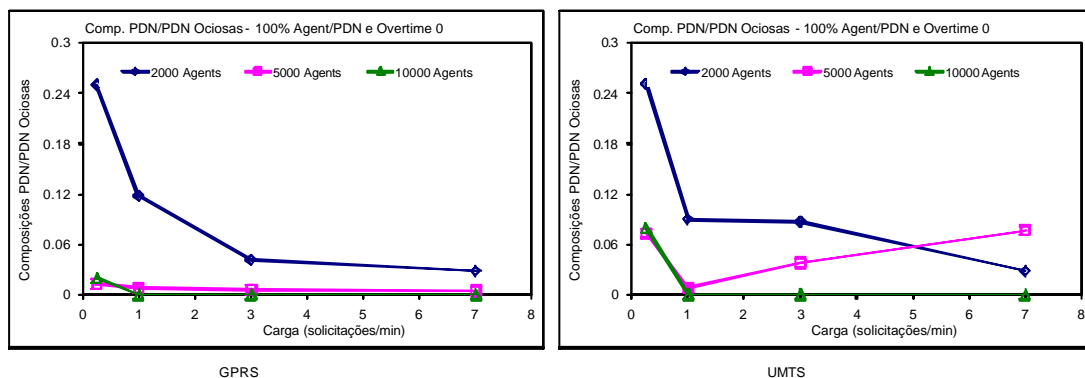
A Figura 6.18 (a) mostra que a relação entre o número de composições estabelecidas e as solicitações de serviço tende a diminuir quando a carga aumenta para todos os tamanhos de rede. A razão é que tendo um número maior de agentes produz o efeito de manter as composições em uso todo tempo, evitando decomposições desnecessárias. Estes resultados mostram que mesmo que o sistema tenha mais usuários solicitando serviços, existem menos solicitações de composições, o que tem um efeito positivo para a escalabilidade do sistema. Pode-se melhorar os resultados usando-se a função de overtime, como mostra a Figura 6.18 (b). O comportamento esperado é diminuir o número de composições quando a

carga aumenta. Este comportamento é melhorado dando-se um tempo de vida extra às composições. Estes resultados mostram que é uma boa estratégia dar um tempo extra às composições para se reduzir o número de composições e os custos associados a elas.

Quando as redes permanecem compostas umas às outras por longos períodos de tempo, as composições podem ficar ociosas, isto é, as composições são estabelecidas, mas nenhum usuário as está usando. Como resultado, o recurso é alocado, mas não é utilizado, o que pode ser prejudicial quando é usado controle de admissão, uma vez que os usuários serão bloqueados porque não é permitida composições adicionais, embora haja recursos ociosos. A Figura 6.19 (a) mostra o gráfico com o número médio de composições do tipo Agente/PDN ociosas com overtime 1, o pior caso.



(a)



(b)

Figure 6.19. Composições Agente/PDN ociosas com overtime 1 (a) e Composições PDN/PDN ociosas sem overtime (b).

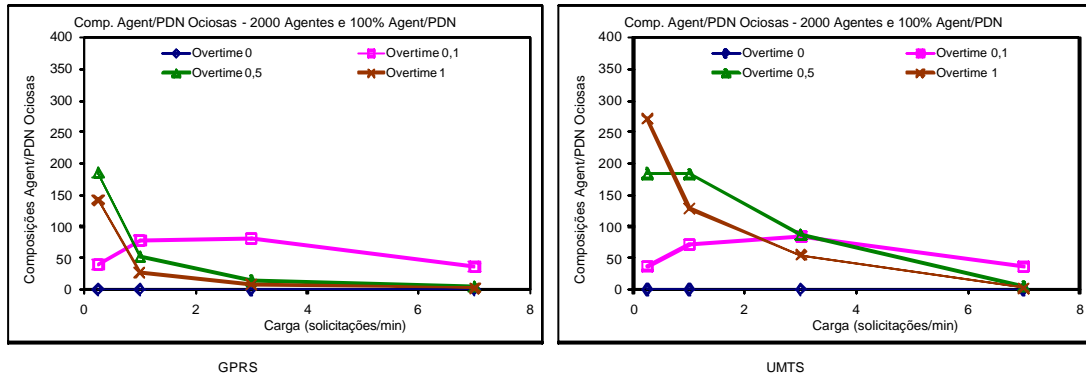
À medida que a carga aumenta, o número de composições Agente/PDN ociosas diminui. Isso ocorre porque quando a carga aumenta, aumenta o número de solicitações de serviço e, conseqüentemente, mais composições estão sendo usadas pelos usuários que solicitam os serviços, fazendo com que o número de composições ociosas diminua. Nota-se para o caso da rede UMTS um aumento significativo no número de composições ociosas para uma rede grande (10000 agentes) e numa situação de baixa carga (0.25 solicitações/minuto). Isso se deve ao fato de que para a rede UMTS, a duração média do serviço (serviço de vídeo streaming) é de 177 segundos, bem maior do que os 36 segundos do serviço GPRS. Como o overtime é um percentual sobre a duração do serviço, em redes UMTS as composições ficarão por mais tempo ativas, contribuindo para a ociosidade.

A Figura 6.19 (b) mostra o número médio de composições ociosas do tipo PDN/PDN sem overtime. As composições ociosas acontecem porque as composições PDN/PDN duram um tempo significativo e algumas vezes mais de uma composição precisa ser estabelecida ao longo de um caminho antes que os serviços possam ser providos aos usuários. O número de composições ociosas diminui com o aumento da carga e da rede, chegando a zero para cargas elevadas.

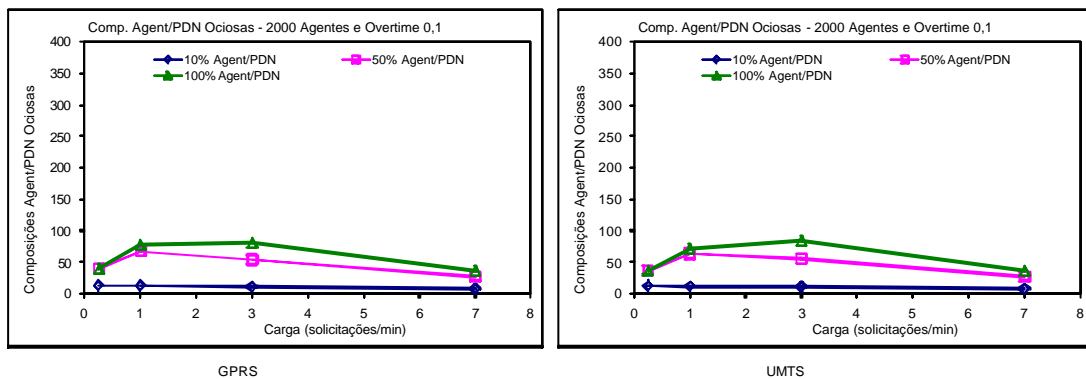
As composições ociosas prejudicam a escalabilidade uma vez que os recursos estão ocupados e quando um usuário o solicita, ocorre uma falha (isto é, o serviço é bloqueado). Contudo, os resultados mostram que o número de composições ociosas diminui quando a carga e a rede aumentam, o que favorece a escalabilidade.

Vamos analisar a escalabilidade de uma Rede de Ambiente considerando o overtime. Nesta avaliação consideraremos apenas as composições Agente/PDN já que os resultados mostraram que as composições PDN/PDN ociosas tendem a diminuir até zero em todos os casos observados. O gráfico da Figura 6.20 (a) mostra o número de composições ociosas para quatro valores distintos de overtime e sem controle de admissão, para uma rede com 2000 agentes. Quando não existe overtime configurado não há composições ociosas. Nota-se para o caso da rede UMTS que o número de composições ociosas aumentou significativamente para uma situação de baixa carga (0.25 solicitações/minuto) e overtime 1. Isso se deve ao fato de que para a rede UMTS, a duração média do serviço (serviço de vídeo streaming) é de 177 segundos, bem maior do que os 36 segundos do serviço GPRS.

Como o overtime é um percentual sobre a duração do serviço, em redes UMTS as composições ficarão por mais tempo ativas, contribuindo para a ociosidade.



(a)



(b)

Figure 6.20. Composições ociosas sem controle de admissão (a) e Composições ociosas com overtime 0.1 (b).

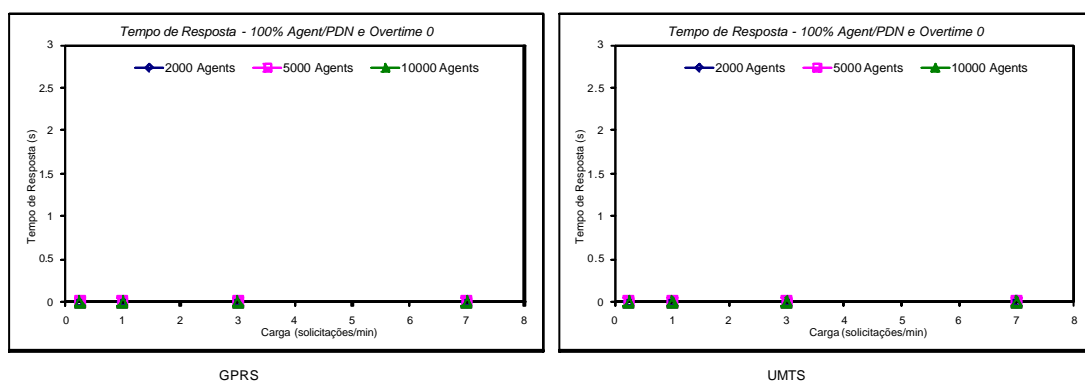
A Figura 6.20 (b) mostra a influência do controle de admissão no número de composições ociosas para um overtime 0.1, o pior caso. Quando a rede é bastante restritiva (apenas 10% de composições Agente/PDN são permitidas), o número de composições ociosas tende a diminuir com o aumento da carga. O número de composições ociosas é mantido abaixo de 50. Porém, quando a rede admite um número maior de composições Agente/PDN, temos um número maior de composições ociosas, que diminui com o aumento da carga. Os

resultados obtidos para o serviço de vídeo streaming na rede UMTS são similares aos obtidos na rede GPRS.

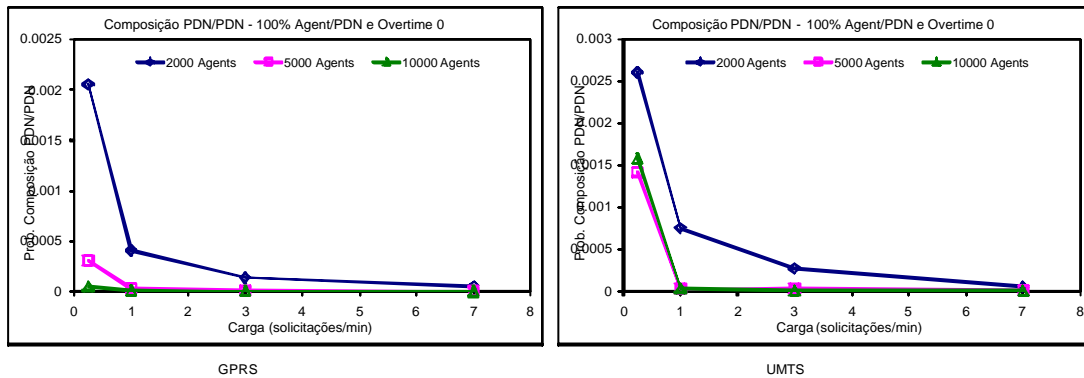
Com base nos resultados apresentados, é necessário que o administrador da rede escolha o que será privilegiado: o tempo de resposta, garantindo que o usuário possa usar o serviço o mais rápido possível; ou a taxa de bloqueio de serviço, garantindo que quase todos os usuários que solicitam um serviço possam utilizá-lo. Esta escolha está relacionada com a configuração adequada do overtime e do nível do controle de admissão.

6.9.2. Tempo de Resposta

O tempo de resposta consiste na quantidade de tempo que um usuário aguarda para que o serviço solicitado esteja disponível. O objetivo desta avaliação é analisar a influência da composição no tempo de resposta e determinar se a composição pode degradar ou não a experiência do usuário. Foram gerados gráficos, Figura 6.21, com os resultados para as redes GPRS e UMTS variando-se a carga sobre a rede e o tamanho da rede (número de agentes conectados a ela), sem controle de admissão e sem overtime (overtime=0).



(a)



(b)

Figure 6.21. Tempo de resposta para diferentes cargas e tamanhos de rede (a) e a probabilidade de composição dada uma solicitação de serviço (b).

A Figura 6.21 (a) mostra o tempo de resposta médio sob diferentes condições de carga para redes de três tamanhos. Como mostra o gráfico, o tempo de resposta foi computado como zero pelo simulador, o que significaria que as composições foram estabelecidas instantaneamente e o serviço estaria disponível imediatamente, sem nenhuma espera por parte do usuário. Numa rede real, porém, existe sempre um atraso (associado ao tempo de propagação do sinal no meio de transmissão, tempos de filas em comutadores de pacotes ao longo do caminho, processamento das solicitações de composição nas redes, etc.). Este resultado de tempo de resposta zero é devido às características da topologia escolhida para simulação. A topologia usada tem poucos nós (apenas 5) e os caminhos entre nós são curtos. Jennifer Lima[18] realizou simulações para a topologia da rede RNP (Rede Nacional de Pesquisas) com 27 nós (27 PDNs), conforme Figura 6.22. A Figura 6.23 mostra o tempo de resposta para esta topologia, onde se necessita um número maior de composições antes que o usuário possa iniciar o serviço, o que contribui para aumentar o tempo de resposta. Nota-se, contudo, que o tempo de resposta tem valores elevados (na ordem de 2 segundos) apenas para redes pequenas (2000 agentes) e em situações de baixa carga (em torno de 0.01 solicitações/minuto). Para redes com grande quantidade de usuários e com cargas elevadas, caso da rede GPRS estudada, o tempo de resposta tende a ser desprezível.

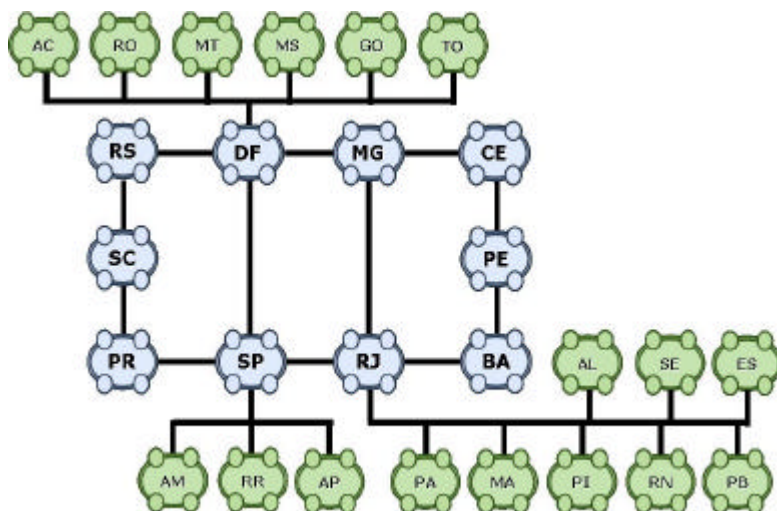


Figure 6.22. Topologia da RNP com 27 PDNs.

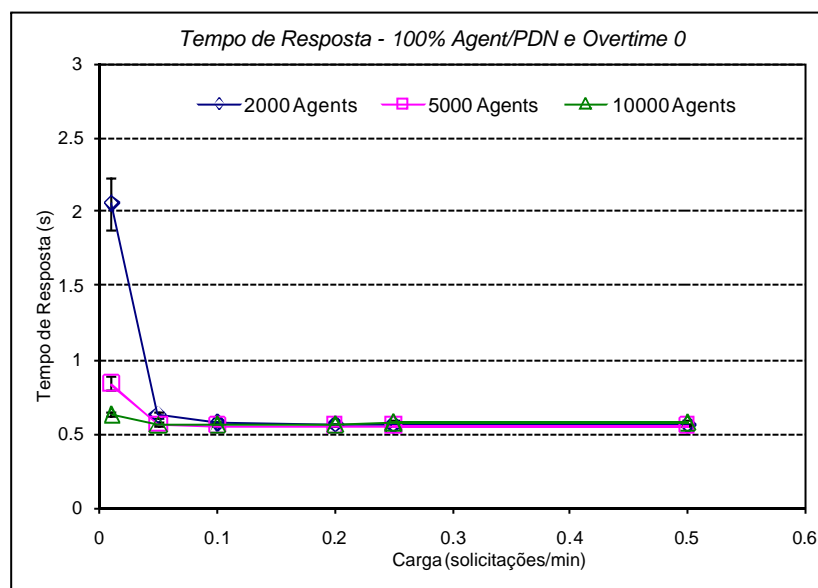


Figure 6.23. Tempo de resposta para uma topologia com 12 PDNs [18].

Quando a carga é baixa e o tamanho da rede é pequeno, poucas solicitações de serviço são geradas e haverá poucos usuários usando as composições. Uma composição é terminada quando seus usuários deixam de usá-la, a menos que seja configurado um tempo de overtime. Se uma nova solicitação de usuário precisa que a mesma composição esteja operacional, a composição será refeita e o tempo de resposta para este usuário aumentará. Isso não acontecerá quando a carga for alta porque existirão usuários precisando da composição na maior parte do tempo, o que impede que aconteçam decomposições

prematuras. O mesmo ocorre quando o tamanho da rede aumenta. Nestes casos, provavelmente uma nova solicitação de serviço não necessitará que novas composições sejam feitas, o que pode ser visto no gráfico da Figura 6.21 (b), que mostra a probabilidade de composição dada uma solicitação de serviço. Aumentando a carga e o tamanho da rede a probabilidade diminui, menos composições serão necessárias por solicitação e menor será o tempo de resposta.

Um overtime bem configurado pode contribuir para o desempenho de rede. Será importante, num cenário real de Redes de Ambiente, entender qual o melhor valor de overtime.

O principal desafio é encontrar o melhor valor de overtime considerando dois aspectos principais: tempo de resposta para os usuários e o número de solicitações de serviço bloqueadas. A Figura 6.24 mostra que quando a rede é restritiva, ou seja, permite-se que apenas 10% de agentes possam compor, a taxa de bloqueio de serviço é maior do que 80%, mesmo para uma carga baixa de 0.25 solicitações/minuto. Esta taxa diminui quando a rede é menos restritiva. Por exemplo, quando se permite que 50% de agentes possam compor, a taxa de bloqueio de serviço é menor do que 50% para a mesma carga de 0.25 solicitações/minuto. Isso ocorre porque os agentes já estão usando a composição e quando outros agentes a solicita ocorre uma falha. Os resultados obtidos para a rede UTMS são similares aos obtidos na rede GPRS.

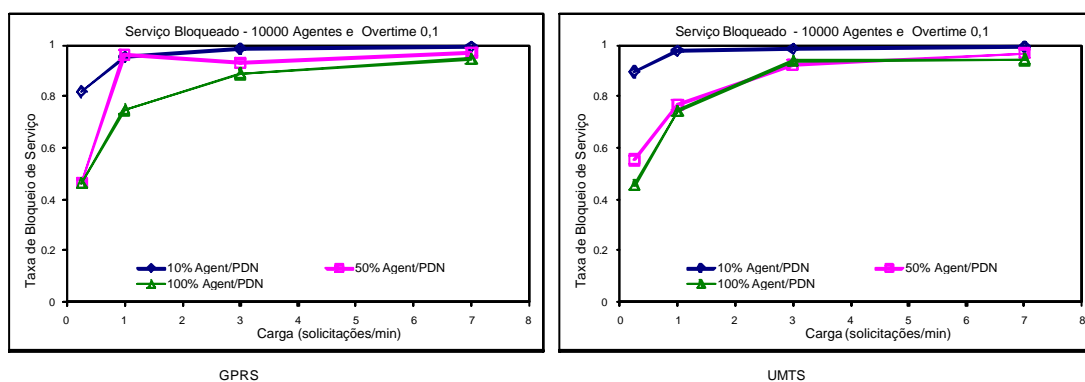
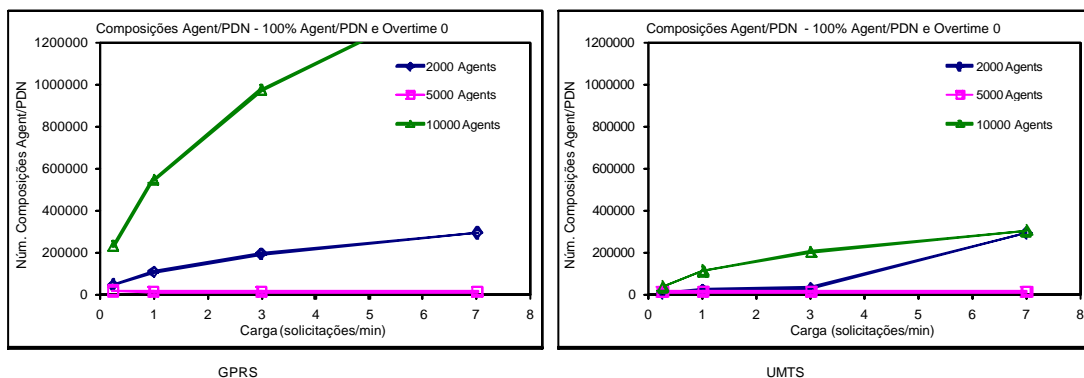


Figure 6.24. Taxa de bloqueio de serviço.

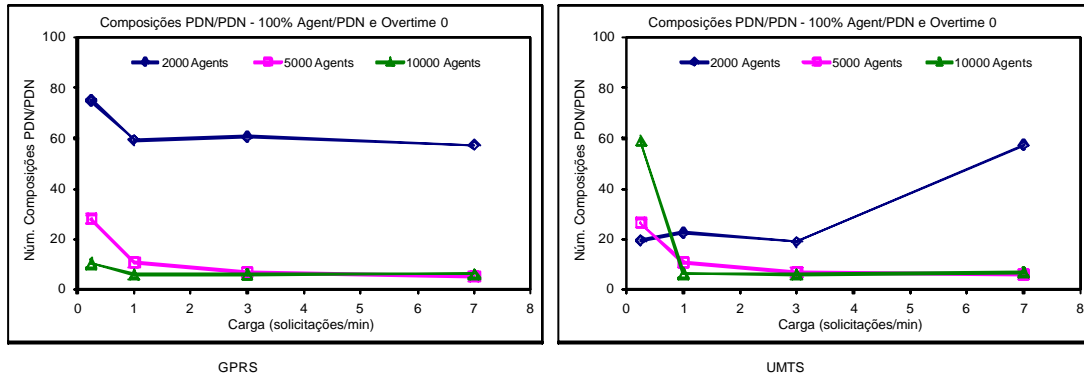
6.9.3. Estabilidade da Composição

Esta seção avalia a estabilidade das composições e sua influência no comportamento da rede. A estabilidade está relacionada à duração de uma composição. Quanto mais usuários acessam um serviço, por mais tempo as composições permanecerão ativas e conseqüentemente mais estáveis elas serão. A estabilidade das composições pode, por exemplo, afetar a quantidade de recursos de software e hardware necessários para estabelecer e manter as composições necessárias. Composições instáveis significa mais negociações entre redes e mais tempo gasto processando-as, o que aumenta o overhead da rede e o tempo de indisponibilidade do serviço. O objetivo é descobrir se as composições do tipo PDN/PDN são estáveis, ocorrendo menos freqüentemente, o que poderia diminuir o tempo de resposta e o custo de composição.

O gráfico da Figura 6.25 (a) mostra o número de composições Agente/PDN para diferentes cargas para três tamanhos de rede. Quando a carga e o tamanho da rede aumentam, aumenta também o número de composições Agente/PDN porque existem mais solicitações de serviço, o que gera mais solicitações de composição. Porém, este tipo de composição toma pouco tempo para ser estabelecida. Para o caso da rede UMTS, o número de composições Agente/PDN cresce numa taxa bem menor.



(a)



(b)

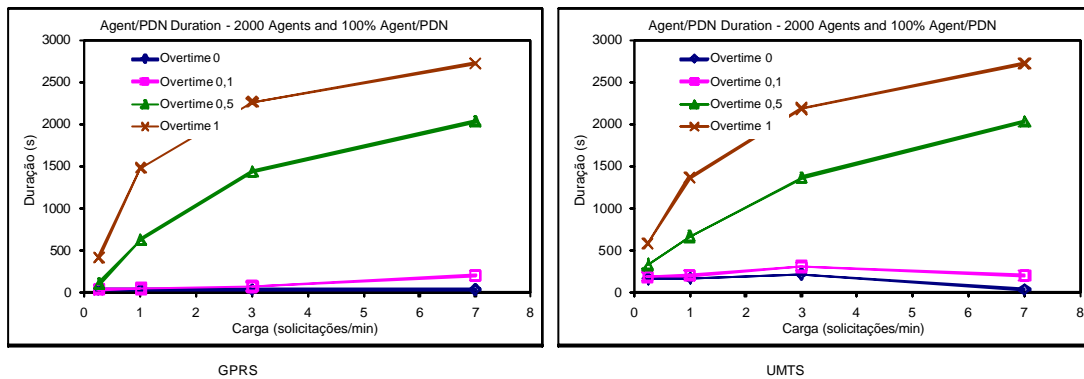
Figure 6.25. Número de composições Agente/PDN (a) e composições PDN/PDN

(b).

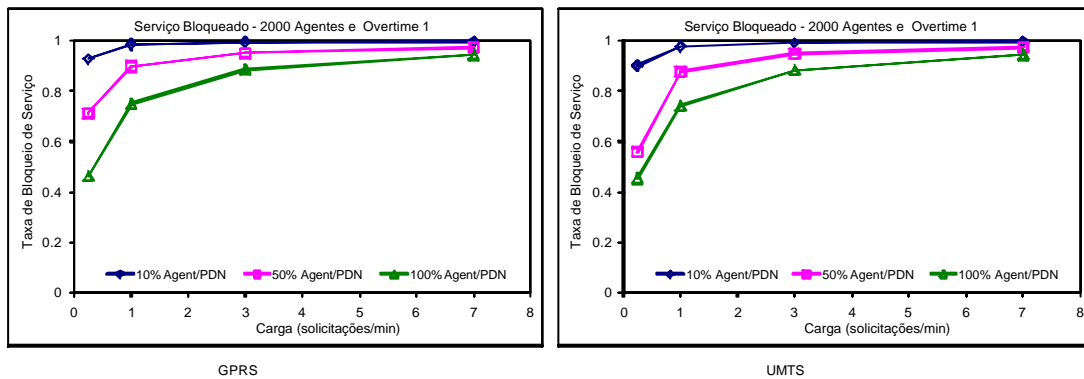
A figura 6.25 (b) mostra o número de composições PDN/PDN variando-se a carga e o tamanho da rede. Como mostra o gráfico, quando a carga e o tamanho da rede aumentam, diminui o número de composições PDN/PDN. Isso acontece porque quando um usuário precisa usar a composição PDN/PDN, a mesma já está estabelecida porque outros usuários a solicitaram anteriormente. Em outras palavras, estas composições acontecem basicamente uma única vez. Os resultados mostram que o número de composições PDN/PDN é baixo, o que é bom para o sistema uma vez que este tipo de composição toma um tempo considerável para ser realizada.

É importante reduzir o número de composições a serem realizadas. Uma forma de fazer isso é aumentar a duração de uma composição, o que pode ser obtido usando-se a função de overtime, como visto em resultados prévios. Outra possibilidade seria utilizar um sistema de cache das informações de contexto de usuários que demandarão composições do tipo PDN/PDN, ou seja, usuários em roaming (visitantes). Seria necessário construir uma função Custo que considere o custo de armazenamento das informações de contexto dos usuários, o custo de troca de sinalização entre redes para realizar as composições e o custo de Qualidade de Serviço (QoS) associado ao aumento do tempo de resposta experimentado pelos usuários enquanto aguardam o estabelecimento das composições para iniciarem o uso dos serviços.

A Figura 6.26 (a) mostra a duração média de composições Agente/PDN usando overtime e sem controle de admissão. Este gráfico considera uma rede de 2000 agentes.



(a)



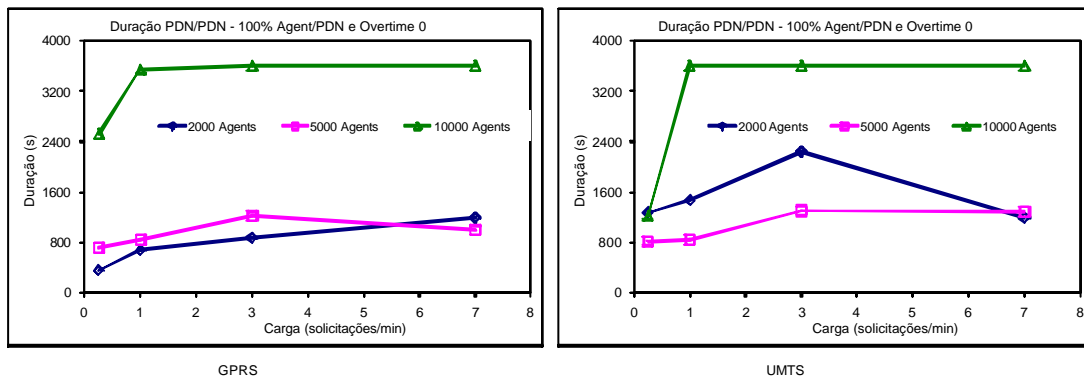
(b)

Figure 6.26. Duração das composições Agente/PDN sob diferentes overtimes (a) e Taxa de bloqueio de serviço (b).

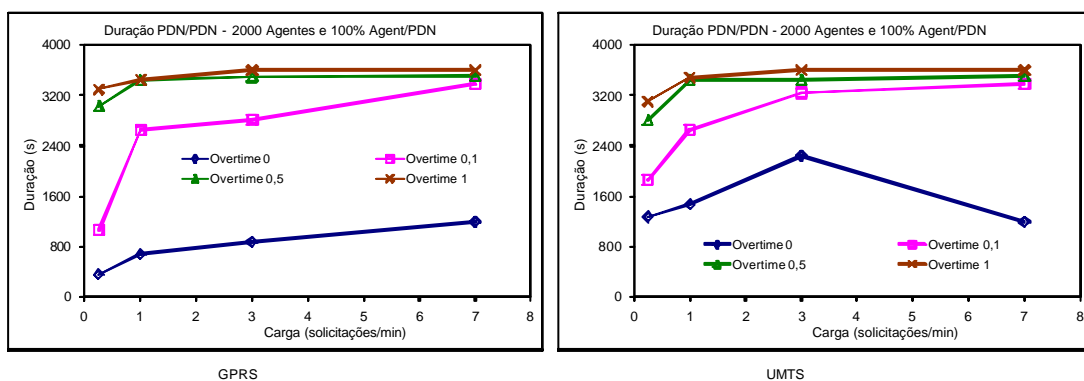
A Figura 6.26 (a) mostra o que acontece com a duração de uma composição Agente/PDN quando o overtime é considerado. Pode-se observar que a duração não muda significativamente para valores pequenos de overtime (por exemplo, 0.1), mas quando o overtime aumenta para 0.5 ou 1, a duração das composições aumenta como esperado. Este é um resultado interessante porque o overtime não afeta a taxa de bloqueio de serviço quando não existe controle de admissão, e se a composição permanece estabelecida por um período maior de tempo, isso ajudará a reduzir o tempo de resposta para os usuários.

Contudo, quando o controle de admissão é configurado, a taxa de bloqueio de serviço é afetada como mostrado na Figura 6.26 (b). Quando a rede é restritiva, com 10% de composições Agente/PDN permitidas, a taxa de bloqueio de serviço é da ordem de 90% para uma carga de 0.25 solicitações/minuto e praticamente 100% para de 7 solicitações/minuto (carga média medida na rede GPRS considerada). Estas taxas diminuem para 45% e 90%, respectivamente, quando não há restrições. Portanto o administrador da rede tem a responsabilidade de escolher o que é melhor para a sua rede: priorizar o tempo de resposta para os usuários usando a configuração de overtime ou não prover todos os recursos da rede, mas bloquear algumas solicitações de serviço.

Com relação à duração das composições PDN/PDN, a Figura 6.27 (a) mostra que quando a carga e o tamanho da rede aumentam a duração das composições também aumenta. A Figura 6.27 (b) mostra que, para cargas mais elevadas com overtime configurado, as durações das composições estão próximas ao tempo de simulação (3600 segundos), significando que as composições são realizadas na inicialização e permanecem ativas durante quase todo o período de operação da rede. Em geral a duração das composições aumenta com o aumento do overtime, como esperado.



(a)



(b)

Figure 6.27. Duração das composições PDN/PDN para cargas diferentes (a) e para diferentes overtimes (b).

7. CONCLUSÕES E TRABALHOS FUTUROS

Esta dissertação apresentou uma nova proposta de evolução de rede denominada Redes de Ambiente que introduz o conceito de composição de redes, o qual busca uma melhor cooperação entre as redes para oferta de serviços e compartilhamento de recursos entre redes pertencentes a domínios tecnológicos e administrativos distintos. Também apresentou o PBMAN, um modelo de gerência baseada em políticas para Redes de Ambiente que utiliza a tecnologia Peer-to-Peer.

É importante a análise de desempenho de composição de redes para se determinar a viabilidade das Redes de Ambiente, já que se espera que a composição de redes seja extensivamente utilizada durante uma interação típica entre um usuário e a rede. Os trabalhos de pesquisa sobre este conceito são recentes e, como ainda não existem implementações de Redes de Ambiente, optou-se por utilizar simulações para este estudo. Utilizou-se um simulador de uso específico, chamado ANCSim, desenvolvido pelo GPRT (Grupo de Pesquisa em Redes e Telecomunicações) da Universidade Federal de Pernambuco (UFPE). Utilizou-se dados coletados em redes móveis reais, redes GPRS e UMTS, para se obter os resultados e avaliar o desempenho da composição de redes de acordo com algumas métricas escolhidas.

Os resultados obtidos mostraram que o conceito de composição de redes não representa um obstáculo à implementação de Redes de Ambiente uma vez que as composições do tipo PDN/PDN, que tendem tomar um tempo considerável para serem realizadas, acontecem tipicamente apenas quando o primeiro usuário tenta acessar os serviços de uma rede remota. Este fato mostra a estabilidade das composições, que diminui o tempo de resposta experimentado pela maior parte dos usuários. Adicionalmente, uma vez que composições PDN/PDN ocorrem apenas uma única vez, o custo associado ao processamento e finalização deste tipo de composição é reduzido. Em resumo, o conceito de composição de redes é uma solução interessante para lidar com a complexidade de configuração quando redes diferentes estão cooperando entre si.

A principal contribuição deste trabalho foi a avaliação do desempenho de composição de redes para valores de duração de serviço, taxa de solicitação de serviço e duração de

composição Agente/PDN (login na rede) obtidos numa rede real (rede GPRS da Brasil Telecom).

Podemos destacar também como contribuição:

- Caracterização da carga média na rede GPRS da Brasil Telecom, que poderá ser utilizada no dimensionamento de um sistema de Autenticação, Autorização e Accounting (AAA) a ser implantado na rede.
- Caracterização da duração média de serviço na rede GPRS da Brasil Telecom, que poderá ser utilizada no planejamento da rede.

Um tópico de pesquisa futuro seria um estudo do uso de cache de informações de contexto de usuários (políticas) obtidas em composições PDN/PDN anteriores, com o objetivo de melhorar a experiência do usuário através da redução do tempo de resposta e redução dos custos, para os administradores das redes, associados à realização deste tipo de composição, por exemplo, o uso da banda do canal de comunicação para troca de sinalização entre as redes durante o estabelecimento da composição.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Wireless World Initiative – WWI, <http://www.wireless-world-initiative.org>, accessed at 07/02/2006.
- [2] Abramowicz H., Niebert N., Mohr W., “The Wireless World Initiative: A Framework for Research on Systems Beyond 3G”, IST Mobile and Wireless Summit 2004, June 2004.
- [3] Niebert N., “AN Framework Architecture”, Deliverable D1.5, WWI Ambient Network Project, December 2005.
- [4] Kampmann M., Vorwerk M., Kleis M., Schmid S., Herborn S., Agüero R. e Choque J., “Multimedia delivery in Ambient Networks”, ACM Multimedia, 2005.
- [5] Niebert N., “Ambient Networks Project Overview and Dissemination Plan”, Deliverable D1.1, WWI Ambient Network Project, June 2004.
- [6] Strassner, J., How Policy Empowers Business-Driven Device Management. In Third International Workshop on Policies for Distributed Systems and Networks. IEEE Computer Society Press, 2002.
- [7] Verma, D. C., “Simplifying Network Administration using Policy-Based Management”, IEEE Network, March 2002.
- [8] Yavatkar, R., Pendarakis, D. & Guerin, R., “A Framework for Policy Based Admission Control,” RFC 2753, January 2000.
- [9] Law, K. L. E. & Saxena, K., “Scalable Design of a policy-Based Management System and its Performance”, IEEE Communications Magazine, 2003.
- [10] Kamienski, C., Fidalgo, J., Sadok, D., Lima, J., Pereira, L. & Ohlman, B., “PBMAN: A Policy-based Management Framework for Ambient Networks”, IEEE Workshop on Policies for Distributed Systems and Networks (Policy 2006), June 2006.
- [11] Kamienski, C., Sadok, D., Fidalgo, J., Lima, J. & Ohlman, B., “On the Use of Peer-to-Peer Architectures for the Management of Highly Dynamic Environments”, 3rd IEEE International Workshop on Mobile Peer-to-Peer Computing (MP2P’06), March 2006.
- [12] Gummadi, K., Gribble, S., Ratnasamy, S., “The Impact of DHT Routing Geometry on Resilience and Proximity”, ACM SIGCOMM 2003, August/September 2003.
- [13] Rocha Jr., J., Fidalgo, J., Dantas, R., Oliveira, L., Kamienski, C. & Sadok, D., “X-Peer: A Middleware for Peer-to-Peer Applications”, 1st Brazilian Workshop on Peer-to-Peer (WP2P), May 2005, (in Portuguese).
- [14] FreePastry, Rice University, <http://freepastry.rice.edu>, accessed at 09/12/2005.

- [15] Rowstron, A. & Druschel, P., "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems", 18th IFIP/ACM Intl. Conference on Distributed Systems Platforms, October 2001.
- [16] Schmid, S., "Smart – Draft Architecture and Multimedia Routing", Deliverable D5.1, WWI Ambient Network Project, February 2005.
- [17] Montgomery, Douglas C., "Applied Statistics and Probability for Engineers", John Wiley & Sons, pp. 207-208, 2004.
- [18] Lima, Jennifer Silva do Monte, "Performance Analysis of Network Composition in Ambient Networks", UFPE, MSc Thesis, 2007.
- [19] Scott, D. W., "Multivariate Density Estimation", John Wiley and Sons, 1992.
- [20] Jain, R., "The Art of Computer Systems Performance Analysis", John Wiley & Sons, pp. 204-207, 1991.
- [21] Soong, T.T., "Fundamentals of Probability and Statistics for Engineers", John Wiley & Sons, pp. 327-330, 1991.

APÊNDICE

A. RESULTADOS DOS TESTES DE ADERÊNCIA

Este apêndice apresenta os resultados completos dos testes de aderência realizados com o software R. Os testes de aderência foram baseados no teste lambda square que mede a discrepância I^2 entre a distribuição empírica (medidas) e a distribuição teórica.

Composição Agente/PDN

Distribuição exponencial

```
rfit <- fitdistr(x_attach_250907, "exponential") # estima parâmetros para distribuição  
Exponencial (rate) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
rate
```

```
1.86603995
```

```
(0.09365386)
```

```
> dist <- rexp(length(x_attach_250907), rate=rfit$estimate[1]) # Gera  
length(x_attach_250907) amostras aleatórias seguindo uma distribuição Exponencial com  
os parâmetros estimados pelo método MLE.
```

```
> rlambda <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2  
50907","Lambda-square") # calcula os valores das discrepâncias  $c^2$  e  $I^2$  e a variância.  
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 269.0766
```

```
[[1]]$lambda2
```

```
[1] 0.6615032
```

```
[[1]]$dev
```

```
[1] 0.07266903
```

Como o valor da discrepância é 0.6615032, pode-se afirmar que as medidas seguem uma distribuição Exponencial com $\text{rate}=1.86603995$.

A Figura A.1 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

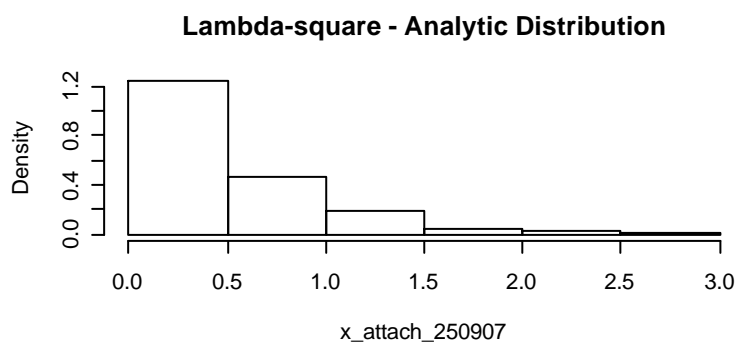
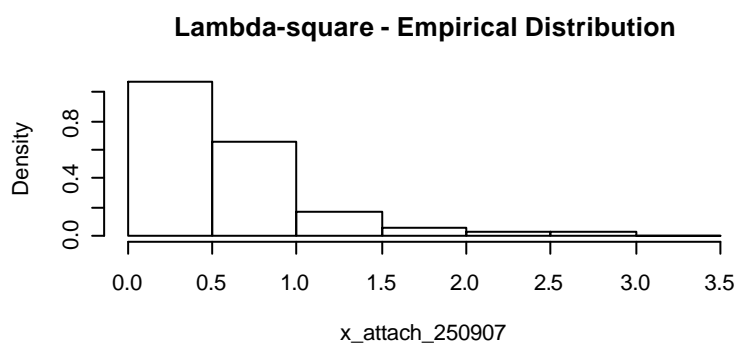


Figura A1. Comparação PDFs para distribuição empírica e analítica.

A Figura A.2 a seguir ilustra o gráfico QQ para distribuição Exponencial.

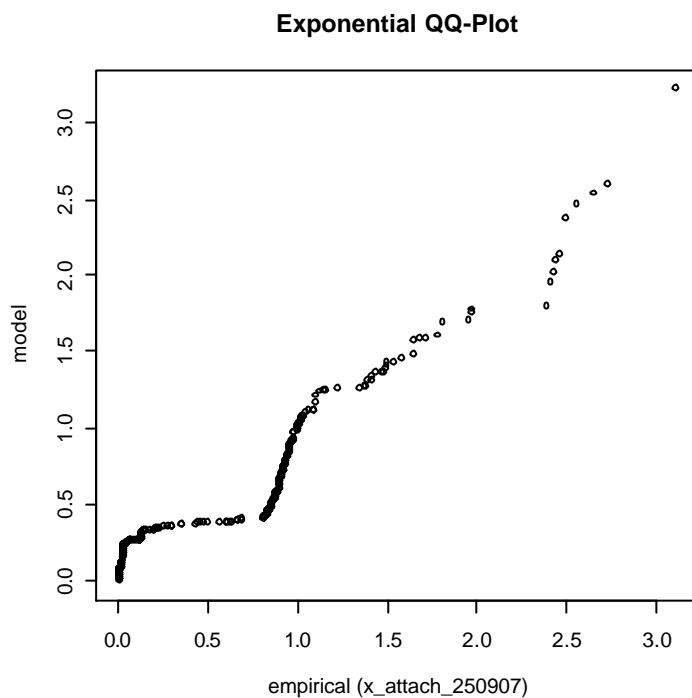


Figura A.2. QQ-Plot para distribuição Exponencial.

Distribuição Gamma

```
> rfit <- fitdistr(x_attach_250907, "gamma" ) # estima parâmetros para distribuição  
Gamma (shape e rate) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
shape rate
```

```
0.54380507 1.01476232
```

```
(0.03211372) (0.09143734)
```

```
> dist <- rgamma(length(x_attach_250907), shape=rfit$estimate[1], rate=rfit$estimate[2])  
# Gera length(x_attach_250907) amostras aleatórias seguindo uma distribuição Gamma  
com os parâmetros estimados pelo método MLE.
```

```
>                               rlambda                               <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2  
50907","Lambda-square") # calcula os valores das discrepâncias  $c^2$  e  $I^2$  e a variância.  
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 292.3251
```

```
[[1]]$lambda2
```

```
[1] 0.7258113
```

```
[[1]]$dev
```

```
[1] 0.0863494
```

Como o valor da discrepância é 0.7258113, pode-se afirmar que as medidas seguem uma distribuição Gama com shape=0.54380507 e rate=1.01476232.

A Figura A.3 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

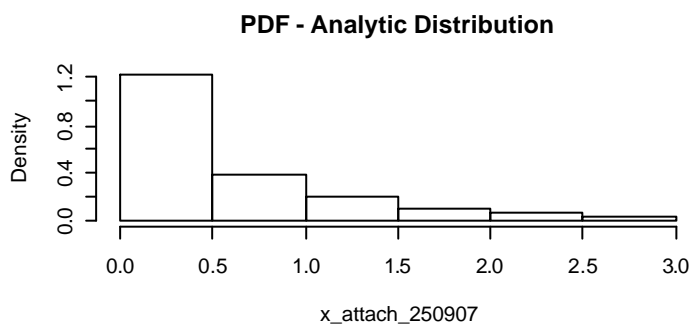
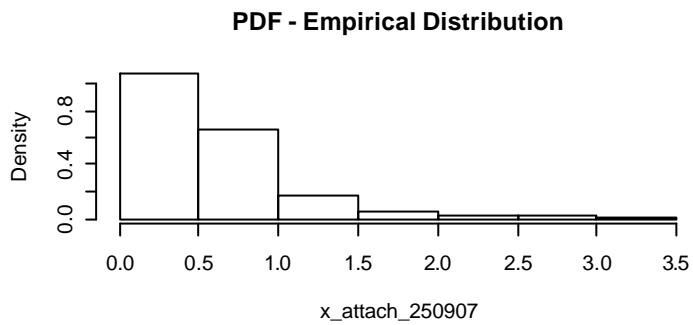


Figura A.3. Comparação PDFs para distribuição empírica e analítica.

A Figura A.4 a seguir ilustra o gráfico QQ para distribuição Gamma.

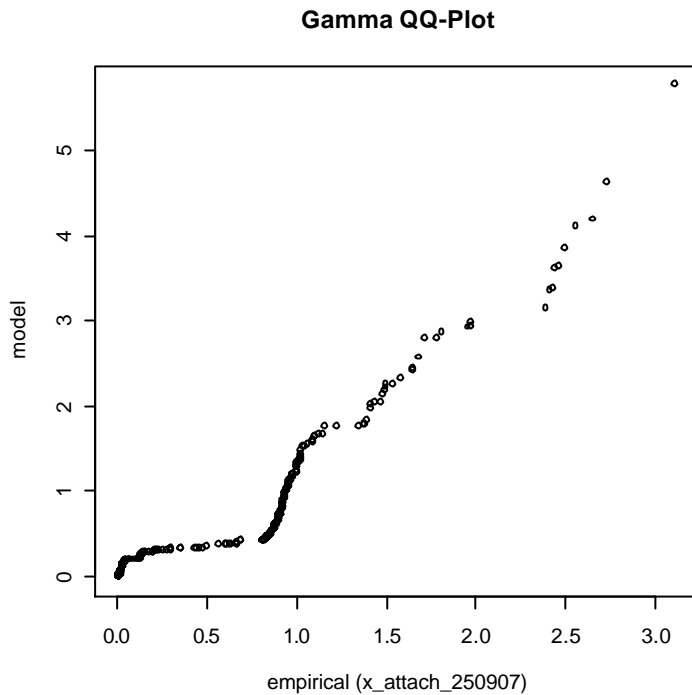


Figura A.4. QQ-Plot para distribuição Gamma.

Distribuição Weibull

```
> rfit <- fitdistr(x_attach_250907, "weibull") # estima parâmetros para distribuição
Weibull (shape e scale) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
  shape    scale
```

```
0.66113996 0.41333155
```

```
(0.02723381) (0.03309997)
```

```
> dist <- rweibull(length(x_attach_250907), shape=rfit$estimate[1], scale=rfit$estimate[2])
# Gera length(x_attach_250907) amostras aleatórias seguindo uma distribuição Weibull
com os parâmetros estimados pelo método MLE.
```

```
>                               rlambda                               <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2  
50907","Lambda-square") # calcula os valores das discrepâncias  $\mathbf{c}^2$  e  $I^2$  e a variância.  
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 92.11733
```

```
[[1]]$lambda2
```

```
[1] 0.2087424
```

```
[[1]]$dev
```

```
[1] 0.0728552
```

Como o valor da discrepância é 0.2087424, pode-se afirmar que as medidas seguem uma distribuição Weibull com shape=0.66113996 e scale=0.41333155.

A Figura A.5 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

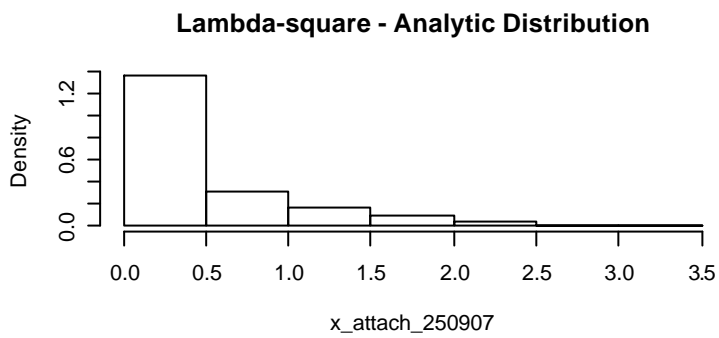
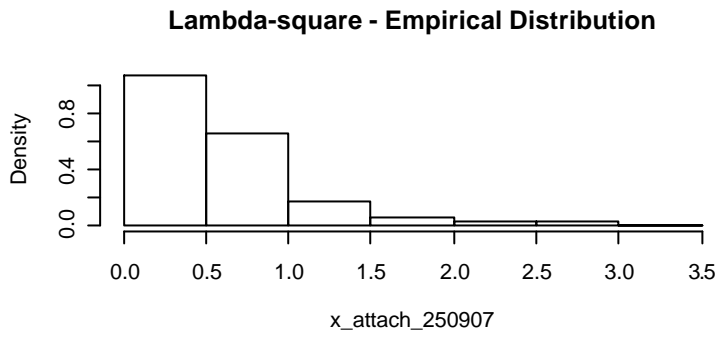


Figura A.5. Comparação PDFs para distribuição empírica e analítica.

A Figura A.6 a seguir ilustra o gráfico QQ para distribuição Weibull.

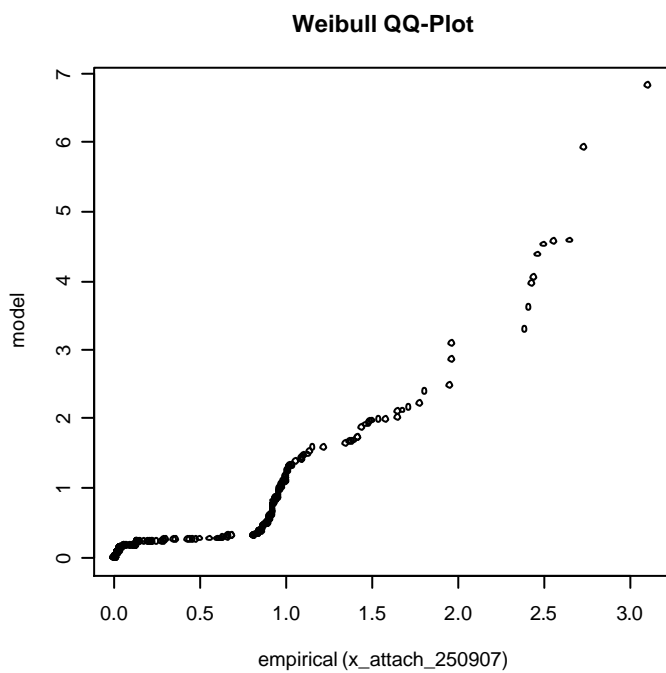


Figura A.6. QQ-Plot para distribuição Weibull.

Distribuição Normal

```
> rfit <- fitdistr(x_attach_250907, "normal") # estima parâmetros para distribuição Normal
(mean e sd) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
mean    sd
```

```
0.53589421 0.59237454
```

```
(0.02973043) (0.02102259)
```

```
> dist <- rnorm(length(x_attach_250907), mean=rfit$estimate[1], sd=rfit$estimate[2]) #
Gera length(x_attach_250907) amostras aleatórias seguindo uma distribuição Normal com
os parâmetros estimados pelo método MLE.
```

```
> rlambda <- lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2
50907","Lambda-square") # calcula os valores das discrepâncias  $c^2$  e  $I^2$  e a variância.
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 5360.147
```

```
[[1]]$lambda2
```

```
[1] 13.41999
```

[[1]]\$dev

[1] 1.886940

Como o valor da discrepância é 13.41999, pode-se afirmar que as medidas não seguem uma distribuição Normal com média=0.53589421 e desvio padrão=0.59237454, pois a discrepância é maior que 1.

A Figura A.7 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

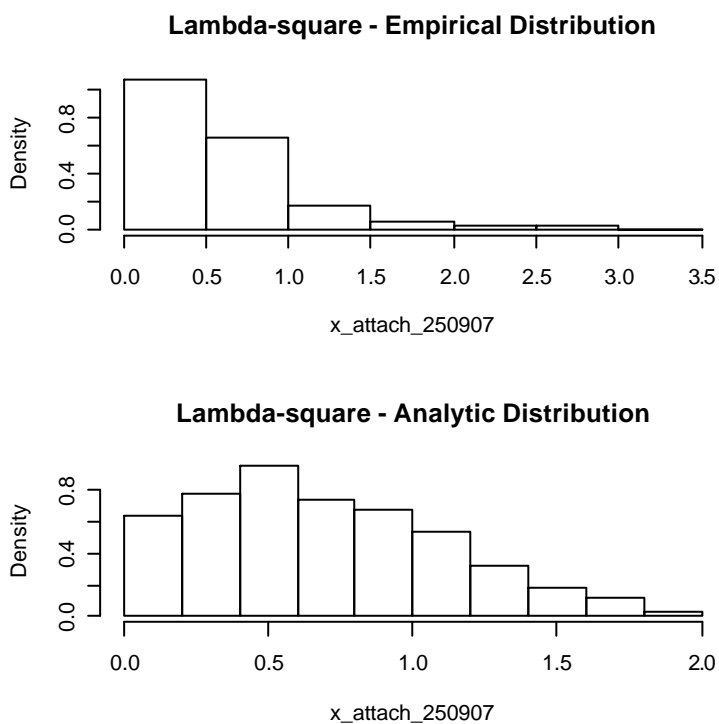


Figura A.7. Comparação PDFs para distribuição empírica e analítica.

A Figura A.8 a seguir ilustra o gráfico QQ para distribuição Normal.

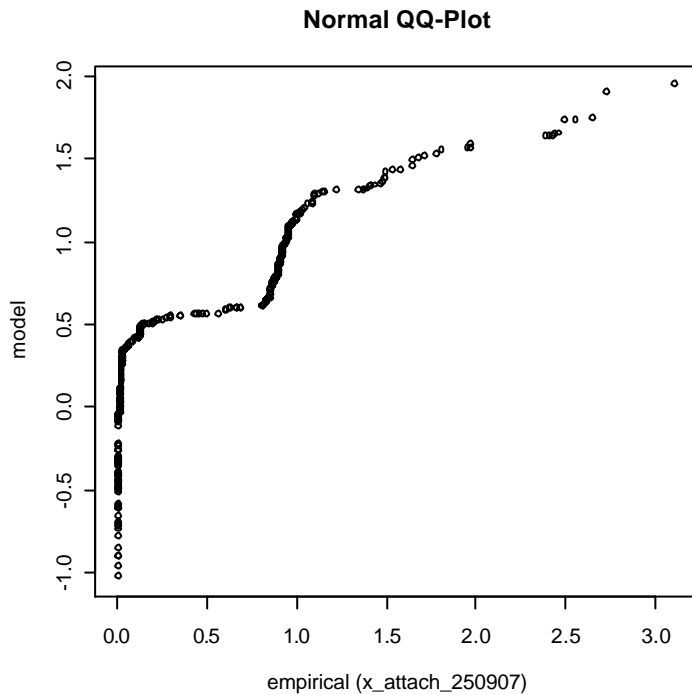


Figura A.8. QQ-Plot para distribuição Normal.

Distribuição Lognormal

```
> rfit <- fitdistr(x_attach_250907, "lognormal") # estima parâmetros para distribuição
Lognormal (meanlog e sdlog) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
meanlog    sdlog
```

```
-1.77688774  1.83884954
```

```
( 0.09228921) ( 0.06525833)
```

```
> dist <- rlnorm(length(x_attach_250907), meanlog=rfit$estimate[1],
sdlog=rfit$estimate[2]) # Gera length(x_attach_250907) amostras aleatórias seguindo uma
distribuição lognormal com os parâmetros estimados pelo método MLE.
```

```
>                               rlambda                               <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2  
50907","Lambda-square") # calcula os valores das discrepâncias  $\mathbf{c}^2$  e  $\mathbf{I}^2$  e a variância.  
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 180.0586
```

```
[[1]]$lambda2
```

```
[1] 0.4281264
```

```
[[1]]$dev
```

```
[1] 0.1120177
```

Como o valor da discrepância é 0.4281264, pode-se afirmar que as medidas seguem uma distribuição Lognormal com média=-1.77688774 e desvio padrão=1.83884954.

A Figura A.9 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

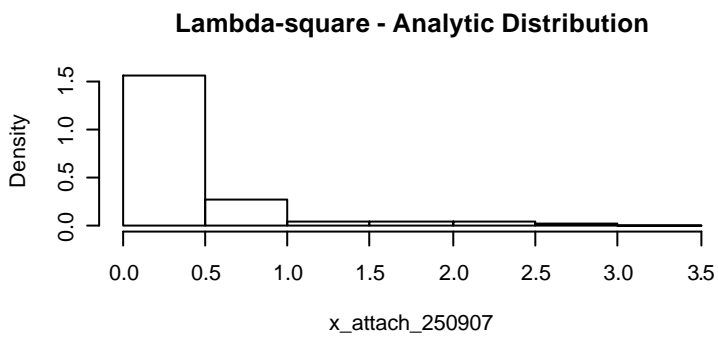
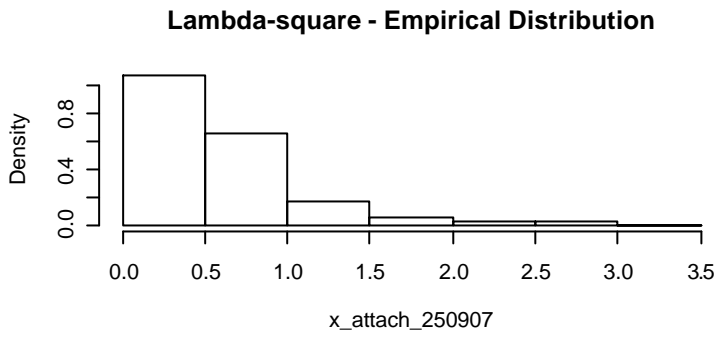


Figura A.9. Comparação PDFs para distribuição empírica e analítica.

A Figura A.10 a seguir ilustra o gráfico QQ para distribuição Lognormal.

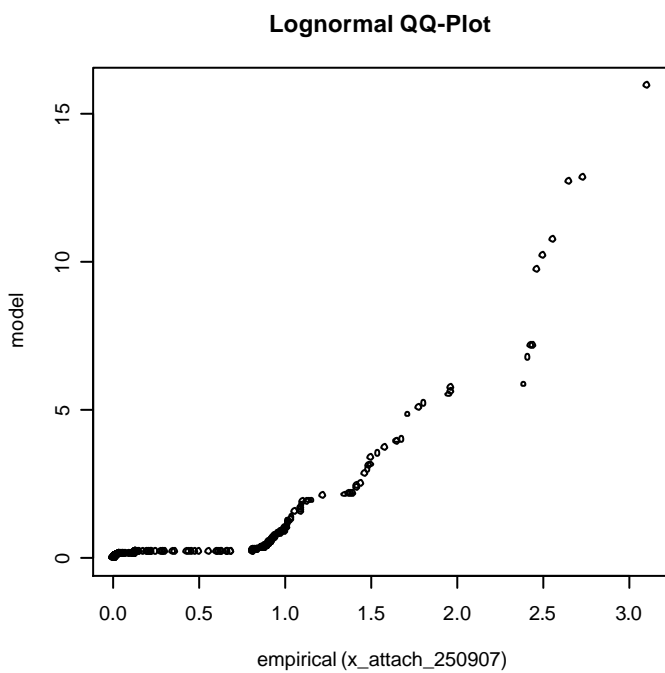


Figura A.10. QQ-Plot para distribuição Lognormal.

Distribuição Chi-Squared

> rfit <- fitdistr(x_attach_250907, "chi-squared", list (df=5)) # estima parâmetros para distribuição Chi-squared (*df*) usando método MLE. O parâmetro *df* (degree of liberty) é o grau de liberdade, que é dado por $n-r-1$, onde n é o número de intervalos no histograma e r é o número de parâmetros a serem estimados. Neste caso, $n=7$ e $r=1$. Logo, $df=7-1-1=5$.

```
> list (rfit)
```

```
[[1]]
```

```
df
```

```
0.82592773
```

```
(0.03827827)
```

```
> dist <- rchisq(length(x_attach_250907), df=rfit$estimate[1]) # Gera  
length(x_attach_250907) amostras aleatórias seguindo uma distribuição Chi-squared com  
os parâmetros estimados pelo método MLE.
```

```
> rlambda <-  
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2  
50907","Lambda-square") # calcula os valores das discrepâncias  $c^2$  e  $I^2$  e a variância.  
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] 62.53659
```

λ^2

0.1437074

σ

0.04422132

Como o valor da discrepância é 0.1437074, pode-se afirmar que as medidas seguem uma distribuição Chi-squared com grau de liberdade=0.82592773.

A Figura A.11 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

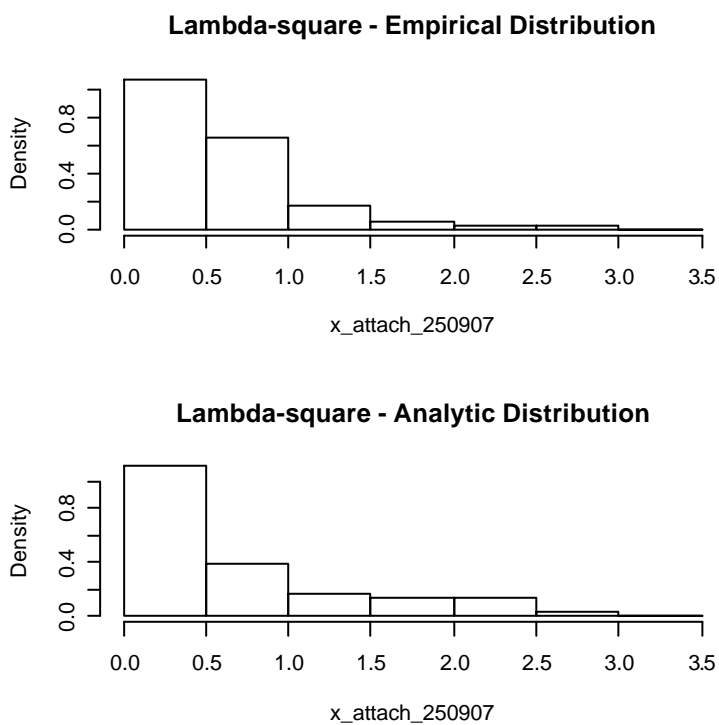


Figura A.11. Comparação PDFs para distribuição empírica e analítica.

A Figura A.12 a seguir ilustra o gráfico QQ para distribuição Chi-squared.

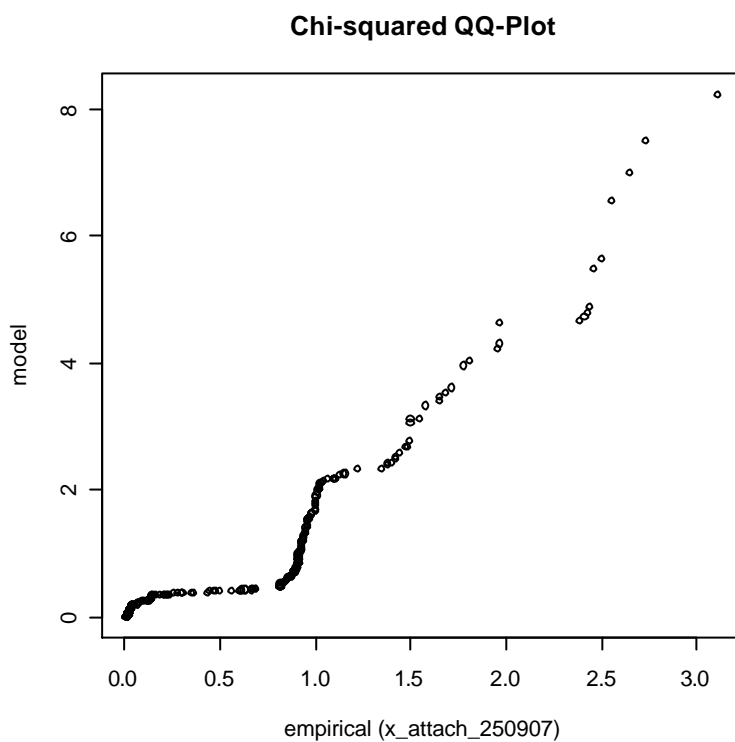


Figura A.12. QQ-Plot para distribuição Chi-squared.

Distribuição Poisson

```
> rfit <- fitdistr(x_attach_250907, "poisson") # estima parâmetros para distribuição Poisson  
(lambda) usando método MLE.
```

```
> list (rfit)
```

```
[[1]]
```

```
lambda
```

```
0.53589421
```

```
(0.03674043)
```

```
> dist <- rpois(length(x_attach_250907), lambda=rfit$estimate[1]) # Gera
length(x_attach_250907) amostras aleatórias seguindo uma distribuição Poisson com os
parâmetros estimados pelo método MLE.
```

```
> rlambda <-
lambdasquare(x_attach_250907,dist,0,hist.scott(x_attach_250907)$breaks,1,1,"x_attach_2
50907","Lambda-square") # calcula os valores das discrepâncias  $c^2$  e  $I^2$  e a variância.
Estes valores são armazenados na variável rlambda, conforme a seguir.
```

```
> list (rlambda)
```

```
[[1]]
```

```
[[1]]$x2
```

```
[1] NA
```

```
[[1]]$lambda2
```

```
[1] NA
```

```
[[1]]$dev
```

```
[1] NA
```

A Figura A.13 a seguir ilustra as funções densidade de probabilidade (PDF) para as distribuições empírica e analítica.

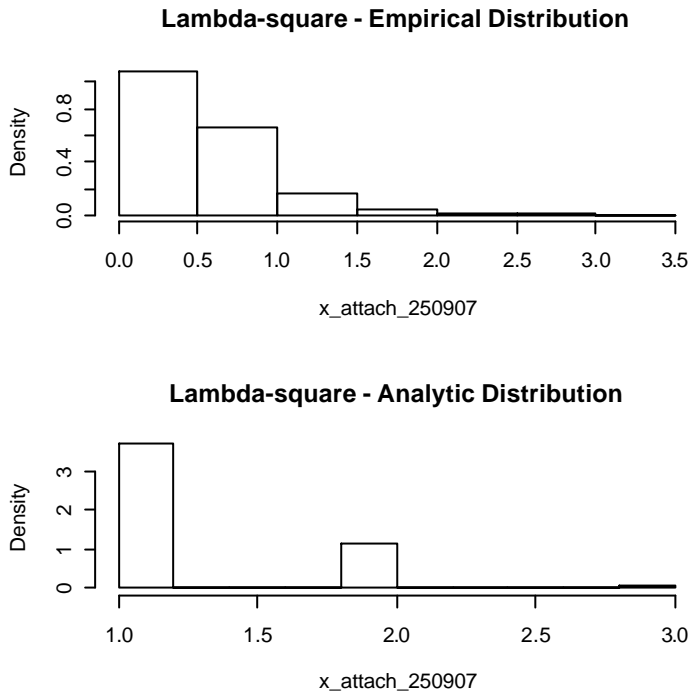


Figura A.13. Comparação PDFs para distribuição empírica e analítica.

A Figura A.14 a seguir ilustra o gráfico QQ para distribuição Poisson.

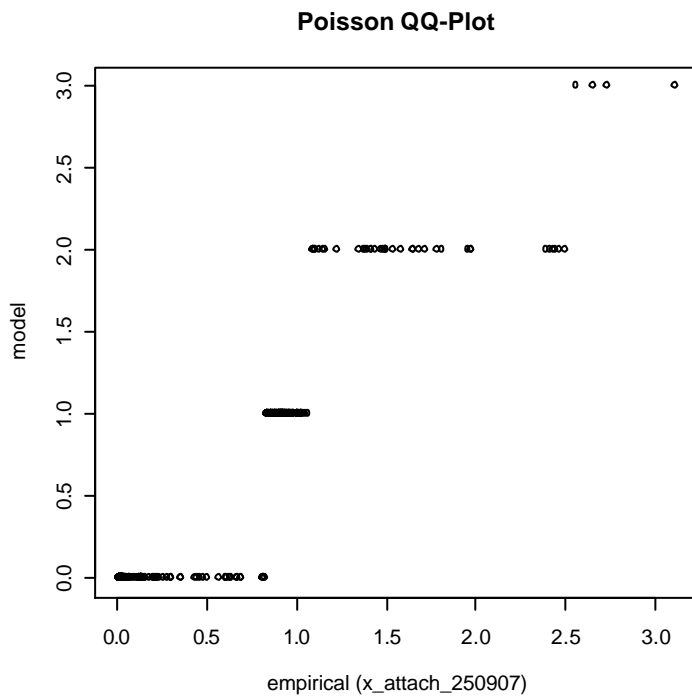


Figura A.14. QQ-Plot para distribuição Poisson.