UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

EWERTON MONTEIRO SALVADOR

Arquitetura de uma Ferramenta e Técnicas de Visualização para Medições sobre Tráfegos SNMP

Dissertação apresentada como requisito parcial
para a obtenção do grau de
Mestre em Ciência da Computação

Prof. Dr. Lisandro Zambenedetti Granville
Orientador

Porto Alegre, abril de 2008

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Salvador, Ewerton Monteiro

Arquitetura de uma Ferramenta e Técnicas de Visualização para Medições sobre Tráfegos SNMP / Ewerton Monteiro Salvador. – Porto Alegre: PPGC da UFRGS, 2008.

95 f.: il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2008. Orientador: Lisandro Zambenedetti Granville.

1. Gerência de Redes de Computadores. 2. SNMP. 3. Visualização de Informação. I. Granville, Lisandro Zambenedetti. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
Reitor: Prof. José Carlos Ferraz Hennemann
Vice-Reitor: Prof. Pedro Cezar Dutra Fonseca
Pró-Reitora de Pós-Graduação: Profª. Valquíria Linck Bassani
Diretor do Instituto de Informática: Prof. Flávio Rech Wagner
Coordenadora do PPGC: Profa. Luciana Porcher Nedel
Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

À Geórgia, por ter tornado os meus sonhos perfeitos.
À minha família, por ser sempre um porto seguro para mim.

AGRADECIMENTOS

À Deus, por todas as bênçãos e aprendizados que pude receber ao longo do mestrado, tanto para a minha vida acadêmica quanto para a minha vida pessoal;

À minha família, por todo o apoio aos meus estudos e pela paciência para suportar os anos de saudades;

À minha noiva, Geórgia, por todo o amor, carinho e força que ela tem me dado, e também pela paciência nas horas em que não pude ser tão presente quanto gostaria;

Ao professor Lisandro Zambenedetti Granville, que me orientou ao longo do mestrado, por todas as valiosas lições que me fizeram crescer imensamente como pesquisador e como pessoa;

Ao Seu Euclides e à Dona Sílvia, por terem sido verdadeiros pais adotivos tanto para mim quanto para minha noiva durante a nossa estadia na cidade de Porto Alegre;

À todos os meus colegas de mestrado, tanto do Grupo de Pesquisa em Redes de Computadores quanto de outros grupos de pesquisa da UFRGS, por todo o companheirismo, apoio, alegrias, consolos, enfim, por todas as experiências que pudemos partilhar juntos durante a nossa caminhada no mestrado.

“Existem verdades que a gente só pode dizer depois de ter conquistado o direito de dizê-las.”
— JEAN COCTEAU, CINEASTA FRANCÊS

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS

API	*Application Programming Interface*
ASCII	*American Standard Code for Information Interchange*
CG/CR	*Command Generator / Command Responder*
CPU	*Central Processing Unit*
CSS	*Cascading Style Sheets*
CSV	*Comma Separated Values*
DOD	*Department of Defense*
EGP	*External Gateway Protocol*
FTP	*File Transfer Protocol*
HTML	*HyperText Markup Language*
ICMP	*Internet Control Message Protocol*
IP	*Internet Protocol*
IRTF	*Internet Research Task Force*
ISO	*International Standards Organization*
MIB	*Management Information Base*
MRTG	*Multi Router Traffic Grapher*
NMRG	*Network Management Research Group*
NO/NR	*Notification Originator / Notification Receiver*
OID	*Object Identifier*
OSI	*Open Systems Interconnection*
PCAP	*Packet Capture*
POP	*Point of Presence*
RFC	*Request for Comments*
RNP	Rede Nacional de Ensino e Pesquisa
SNMP	*Simple Network Management Protocol*
TCP	*Transmission Control Protocol*

UDP *User Datagram Protocol*

VLAN *Virtual Local Area Network*

WWW *World Wide Web*

XML *eXtended Markup Language*

LISTA DE FIGURAS

RESUMO

Em março de 2006 o *Internet Research Task Force* (IRTF) propôs uma metodologia para medições sobre tráfegos SNMP, a fim de identificar os padrões de uso desse protocolo. Contudo, essa metodologia apresenta algumas limitações, tais como: ausência de técnicas para visualização de dados, não especificação de formas de comparação dos resultados obtidos a partir de vários arquivos de tráfego e falta de integração entre as ferramentas necessárias para a execução da metodologia. Esta dissertação de mestrado apresenta uma proposta de arquitetura para uma ferramenta Web que automatiza, de forma integrada, a execução das etapas da metodologia do IRTF, buscando solucionar os problemas identificados nessa metodologia. Essa arquitura foi implementada em um software denominado *Management Traffic Analyzer*, e técnicas de visualização específicas para os resultados gerados a partir de análises de tráfegos SNMP foram desenvolvidas especialmente para essa ferramenta. Por fim, utilizou-se o *Management Traffic Analyzer* para o estudo de amostras de tráfego SNMP provenientes de duas grandes redes brasileiras: a do ponto de presença da Rede Nacional de Ensino e Pesquisa (RNP) no Rio Grande do Sul, e a da própria RNP.

Palavras-chave: Gerência de Redes de Computadores, SNMP, Visualização de Informação.

Architecture of a Tool and Visualization Techniques for SNMP Traffic Measurements

ABSTRACT

In March 2006 the Internet Research Task Force (IRTF) proposed a methodology for measuring SNMP traffic traces. However, this methodology presents some limitations, such as: absence of data visualization techniques, non specification of forms for comparing results obtained from various traffic traces and lack of integration among the needed tools for the execution of the methodology. This masters dissertation proposes an architecture of a web tool which automatizes, in an integrated fashion, the execution of the IRTF methodology's steps. This architecture was implemented in a software which was called *Managament Traffic Analyzer*, and visualization techniques specific for the results generated by the SNMP traffic traces analyses were developed specifically for this tool. Finally, the *Management Traffic Analyzer* was used for the study of SNMP traffic traces originated from two large brazilian networks: the one of the Brazilian National Education and Research Network (RNP) point of presence in the state of Rio Grande do Sul, and the one of the RNP itself.

Keywords: Network Management, SNMP, Information Visualization.

1 INTRODUÇÃO

O *Simple Network Management Protocol* (CASE; FEDOR; SCHOFFSTAL, 1990) (SNMP) foi proposto há mais de 15 anos, e atualmente é tido como o protocolo padrão *de facto* para o gerenciamento de redes TCP/IP. Apesar de ser amplamente empregado, muito pouco se sabe sobre os padrões de uso desse protocolo nas redes em produção.

Em março de 2006, o *Network Management Research Group* (NMRG), pertencente ao *Internet Research Task Force* (IRTF), publicou a primeira versão do *internet draft* intitulado "*SNMP Traffic Measurements*" (SCHOENWAELDER, 2006), o qual propunha uma metodologia sistemática para medições e geração de estatísticas sobre o uso do SNMP. Uma vez que um *internet draft* possui validade de apenas 6 meses, esse documento foi renovado novamente em janeiro de 2007 (SCHOENWAELDER, 2007a), em dezembro de 2007 (SCHOENWAELDER, 2007b) e por fim em fevereiro de 2008 (SCHOENWAELDER, 2008), estando ele válido até 15 de agosto de 2008. O objetivo dessa metodologia é identificar padrões de utilização do SNMP através da realização de uma série de medições sobre tráfegos de redes em produção, a fim de poder se descobrir características do SNMP que atualmente ainda não são efetivamente conhecidas. Algumas das questões que estão sendo investigadas neste contexto são: quais recursos do protocolo (versões, operações, MIBs, etc.) estão sendo utilizados, como o uso do SNMP difere nos vários tipos existentes de redes de computadores e organizações, quais informações são mais freqüentemente requisitadas e quais são as interações mais típicas que estão sendo empregadas utilizando este protocolo. Devido ao fato desta metodologia ser relativamente recente, existem poucos trabalhos publicados contendo resultados obtidos a partir do estudo de tráfegos SNMP em redes de produção (SCHÖNWÄLDER et al., 2007) (SALVADOR; GRANVILLE, 2008a) (SALVADOR; GRANVILLE, 2008b) (SALVADOR; GRANVILLE, 2008c). Contudo, esses trabalhos apresentam apenas resultados preliminares de análises feitas sobre um determinado conjunto de tráfegos, sendo necessários ainda estudos mais aprofundados sobre uma quantidade maior de tráfegos de gerenciamento, a fim de se obter informações mais sólidas sobre a utilização do SNMP.

Apesar da metodologia proposta pelo IRTF ser de grande relevância para a área de gerenciamento de redes, a mesma ainda possui algumas limitações. Os estudos sobre o SNMP baseados nessa metodologia certamente irão gerar uma grande quantidade de novos dados sobre esse protocolo. Essa nova massa de dados também precisará ser interpretada pelas pessoas que estiverem conduzindo esse tipo de estudo, de forma a responder às questões que ainda se encontram em aberto sobre o SNMP. Uma forma de aumentar a eficiência desse processo de interpretação dos dados é a utilização de técnicas de visualização de informação, as quais permitem que pessoas obtenham *insights* sobre os dados que estão sendo analisados (e.g., detecção de padrões, descoberta de características interessantes, etc.) de uma forma mais rápida e natural, graças às capacidades únicas do

sistema visual humano. Contudo, nenhuma técnica para visualização de informação é descrita pela metodologia original apresentada pelo IRTF.

Ao se buscar por técnicas de visualização contextualizadas no estudo de tráfegos SNMP, percebe-se que ainda são escassas as pesquisas nessa área. De acordo com o conhecimento do autor desta dissertação de mestrado, o trabalho que mais se aproxima dos objetivos da mesma é o de Seong Jin Ahn *et al.* (AHN; YOO; CHUNG, 1999), o qual busca definir uma ferramenta Web destinada a analisar o desempenho de redes TCP/IP a partir da análise dos dados gerados pela monitoração da rede via SNMP. Outras pesquisas na área de visualização de informação no contexto de gerenciamento de redes também podem ser destacadas. Oberheide *et al.* (OBERHEIDE; GOFF; KARIR, 2006) descreveu uma ferramenta para auxiliar a tarefa de gerenciar uma rede através de visualizações sobre *Netflow feeds*. Há ainda o trabalho de Keim *et al.* (KEIM et al., 2006), o qual apresentou um conjunto de ferramentas de visualização para se compreender atividades típicas de comunicação, assim como prever gargalos ou problemas na rede.

Conforme se pode observar, todos os trabalhos sobre visualização de dados apresentados no parágrafo anterior lidam com tráfego de rede de uma forma geral, sem levar em consideração as particularidades do tráfego de gerenciamento e, mais especificamente, do SNMP. O grande problema disso é que as técnicas de visualização que não sejam específicas para o tráfego de gerenciamento de redes não levam em consideração características intrínsecas do protocolo SNMP, como versão utilizada, elementos na *varbind list* e o comportamento de operações básicas como `get-next-request` e `set-request`. Dessa forma, se faz necessária uma adaptação das técnicas de visualização contextualizadas na área genérica de redes de computadores para o contexto específico de gerenciamento de redes.

Outra característica da metodologia original do IRTF é a necessidade de se empregar um conjunto de ferramentas específicas para certas fases do estudo sobre o tráfego SNMP. Esse fato dificulta a utilização da metodologia em si, pois aumenta a complexidade da execução dos diversos passos que compõem esse tipo de estudo, devido ao aumento da quantidade de *softwares* a serem instalados e à necessidade de os usuários aprenderem a lidar com vários tipos de programas. Além disso, as ferramentas indicadas pela metodologia do IRTF disponibilizam para seus usuários apenas interfaces em modo texto, fazendo com que as pessoas tenham mais dificuldade em aprender como utilizá-las. Dessa forma, se torna necessário o desenvolvimento de um *framework* que possa integrar todos esses *softwares* numa única ferramenta, a qual deverá possuir uma interface gráfica dedicada a tornar o seu uso o mais simples possível, a fim de facilitar a execução de análises sobre tráfegos SNMP de redes em produção.

Por fim, também não é fornecida pela metodologia do IRTF nenhuma forma de se comparar os resultados de análises realizadas sobre dois ou mais tráfegos distintos. Essas comparações teriam por objetivo apresentar, para a pessoa que está aplicando a metodologia do IRTF para medições sobre SNMP, as variações de uma amostra de tráfego com relação à outra amostra, permitindo se identificar as diferenças entre as possíveis estratégias de gerenciamento que podem ser empregadas na administração de uma rede. Esse tipo de comparação auxiliaria os pesquisadores a ter uma visão global do uso do SNMP, uma vez que as principais semelhanças e diferenças entre os diversos tráfegos estudados ficariam fortemente evidenciadas após esse tipo de análise. Por essa razão é importante a disponibilização desta funcionalidade no conjunto das ferramentas necessárias para a execução da metodologia do IRTF.

Esta dissertação de mestrado tem por objetivo pesquisar e desenvolver soluções para as

limitações da metodologia para medições sobre tráfegos SNMP que foram destacadas anteriormente, a fim de proporcionar melhorias na eficiência e eficácia desse tipo de estudo. As principais contribuições deste trabalho são: especificação de uma nova arquitetura, baseada em tecnologias Web, voltada à investigação sobre tráfegos SNMP segundo as definições do IRTF; desenvolvimento de técnicas de visualização de informação específicas para o contexto onde está inserido o SNMP; e apresentação de alguns resultados acerca de investigações sobre tráfegos SNMP originados a partir de duas redes de computadores brasileiras. O restante deste trabalho está organizado da seguinte maneira. O capítulo 2 apresenta uma revisão bibliográfica dos principais conceitos relacionados à área em que se encontra essa dissertação de mestrado. Já o capítulo 3 introduz um conjunto de técnicas de visualização de informação desenvolvidas especialmente para serem empregadas no contexto de medições sobre tráfego SNMP, enquanto que o capítulo 4 apresenta a arquitetura de uma ferramenta Web para automatizar a execução da metodologia do IRTF. Os resultados das análises de tráfegos SNMP originados a partir de várias redes em produção, segundo a metodologia do IRTF, são discutidos no capítulo 5. Por fim, o capítulo 6 apresenta as conclusões e os trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Ao longo do desenvolvimento deste trabalho foi empregada uma série de conceitos pertinentes à área de gerenciamento de redes. Devido a isso, será apresentada neste capítulo uma revisão da bibliografia utilizada para a elaboração desta dissertação, a fim de facilitar a compreensão dos conceitos aqui empregados.

2.1 *Simple Network Management Protocol*

Proposto há mais de 15 anos, o *Simple Network Management Protocol* (SNMP) é uma coleção de especificações de gerenciamento de redes que incluem o protocolo em si, a definição das estruturas de dados e os conceitos associados (STALLINGS, 1999). Dentre essas especificações, pode-se destacar a que descreve a primeira versão do protocolo SNMP (CASE; FEDOR; SCHOFFSTAL, 1990) e do *Managament Information Base* (ROSE; MCCLOGHRIE, 1991).

Objetivando explicitar os conceitos básicos referentes ao protocolo SNMP, será apresentada nesta seção a arquitetura de gerenciamento de redes utilizada nas redes TCP/IP, os métodos básicos para solicitação e configuração de valores de objetos SNMP, os conceitos de *polling* e *traps*, a estrutura desse protocolo de gerenciamento em relação ao modelo OSI/ISO e, por fim, as definições de comunidade e *strings* de comunidade.

2.1.1 Arquitetura de Gerenciamento de Redes

Segundo William Stallings (STALLINGS, 1999), o modelo de gerenciamento utilizado pelas redes TCP/IP é composto pelos seguintes elementos fundamentais:

- Estação de gerenciamento;
- Agente de gerenciamento;
- Base de informação de gerenciamento (*Management Information Base* - MIB);
- Protocolo de gerenciamento de rede;

A **estação de gerenciamento**, ou simplesmente **gerente**, é a interface para a monitoração e controle de redes dentro dos sistemas de gerenciamento, a fim de possibilitar a execução dessas operações por seres humanos. No mínimo, essas estações devem possuir:

- Um conjunto de aplicações de gerenciamento para análise de dados, recuperação de falhas, entre outras;
- Uma interface pela qual um operador possa gerenciar a rede;

- A capacidade de fazer com que as definições ditadas pelo gerente de rede se concretizem nos elementos que compõem a rede;

- Um banco de dados de informações extraídas das MIBs de todas as entidades da rede que estão sendo gerenciadas.

Um outro elemento ativo nesse sistema é o **agente de gerenciamento**, ou simplesmente **agente**. Este elemento da arquitetura se encontra instalado em equipamentos como servidores, *bridges*, roteadores, *hubs*, etc. A presença dos agentes de gerenciamento nesses dispositivos os tornam gerenciáveis. Isso acontece porque são esses agentes que fornecem respostas para as requisições e ações originadas a partir de uma estação de gerenciamento.

Os recursos da rede precisam ser gerenciados através de representações próprias, em forma de objetos. Cada objeto é, essencialmente, uma variável de dados que representa um aspecto de um agente gerenciado. A coleção desses objetos é chamada de **Base de Informação de Gerenciamento** (*Management Information Base*, ou MIB). O conjunto de objetos que diz respeito à uma determinada categoria de dispositivos ou serviços que podem ser gerenciados forma uma classe. As MIBs consistem nas diversas classes desses objetos de gerenciamento que foram organizadas de forma padronizada. Por exemplo, o conjunto de roteadores de rede, por possuírem características comuns, forma uma única classe de dispositivos gerenciáveis, possuindo um conjunto padronizado de objetos de gerenciamento. Isso permite que o administrador da rede possa monitorar e configurar cada aspecto gerenciável que seja característico de um roteador, independentemente de modelo ou fabricante.

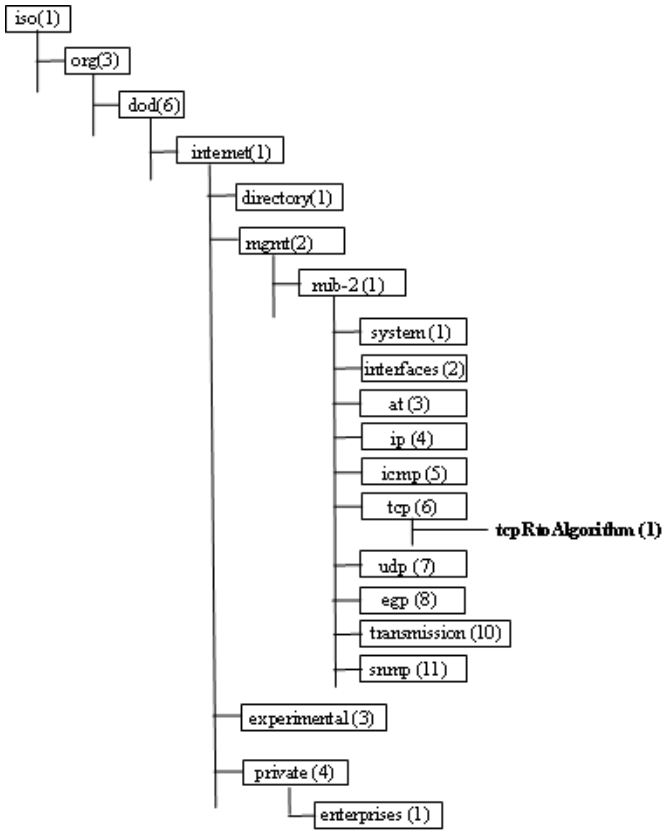


Figura 2.1: Grupos de objetos da MIB-II

A Figura 2.1 (STALLINGS, 1999) apresenta a árvore (MIB *tree*) da MIB-II, com seus objetivos organizados de forma hierárquica. As folhas dessa árvore, como o `tcpRtoAlgorithm`, representam os objetos gerenciados. A estrutura da árvore permite, assim, a organização desses objetos em agrupamentos lógicos relacionados entre si. Os números que estão entre os parêntesis após o nome de cada classe ou objeto possibilitam a representação do caminho a ser percorrido na árvore da MIB até se chegar a uma determinada posição dessa hierarquia. Por exemplo, o caminho para se chegar à classe "tcp" da MIB-II é "1.3.6.1.2.1.6". Dessa forma, sempre que uma mensagem SNMP se referir ao objeto `tcpRtoAlgorithm`, ela irá transmitir na rede o identificador "1.3.6.1.2.1.6.1".

Por fim, existe o elemento que faz a ligação entre a estação de gerenciamento e os agentes gerenciados, que é o **protocolo de gerenciamento de rede**. O protocolo usado para o gerenciamento de redes TCP/IP é o SNMP, o qual inclui as seguintes funcionalidades básicas:

- **Método *get***: possibilita que a estação de gerenciamento requisite valores de objetos dos agentes gerenciados;
- **Método *set***: define valores para os objetos dos agentes gerenciados;
- ***Trap***: permite que o agente notifique a estação de gerenciamento sobre a ocorrência de eventos relevantes.

A Figura 2.2 (STALLINGS, 1999) apresenta um ambiente típico de gerenciamento em redes TCP/IP via SNMP, contendo gerente, agentes, MIBs e o protocolo em si.

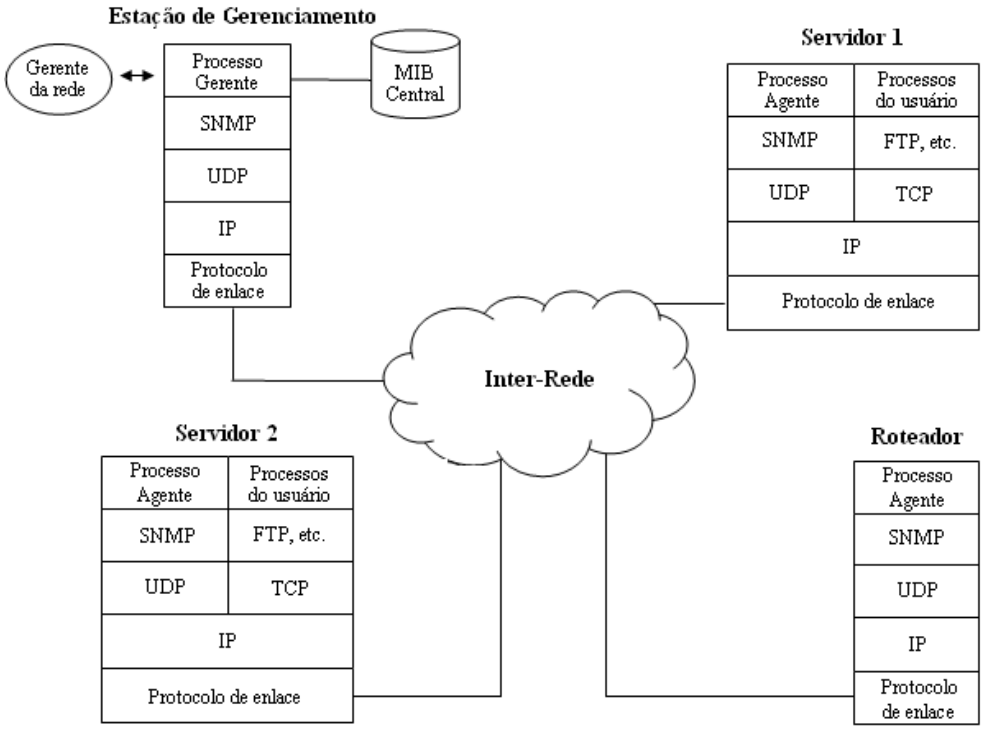


Figura 2.2: Ambiente de rede gerenciado via SNMP

2.1.2 Métodos básicos para solicitação e configuração de valores de objetos SNMP

Dentre as operações do protocolo SNMP, aquela que é mais elementar para recuperação de valores de objetos SNMP é a *get-request*. A funcionalidade dessa operação é simplesmente ordenar a um agente que obtenha o valor de um objeto indicada pelo gerente. Caso o objeto indicado pelo gerente seja válido, o agente responderá à solicitação com uma mensagem do tipo *get-response*, contendo o valor e a identificação do objeto informado pelo gerente durante a requisição. Existe ainda uma variante dessa operação, chamada *get-next-request*, cuja funcionalidade é obter o valor do objeto sucessor daquele indicado pelo gerente. Essa operação é utilizada principalmente na leitura de tabelas de tamanho desconhecido, e a mesma também recebe como resposta uma mensagem do tipo *get-response*.

A operação básica para configuração de valores de objetos é a *set-request*. Seu funcionamento consiste no envio do identificador de um objeto e um determinado valor para um agente da rede. Dessa forma, o agente irá atribuir o valor informado para o objeto indicado pelo gerente.

2.1.3 *Polling* e *Traps*

Polling é o processo utilizado pela estação de gerenciamento para inquirir um agente gerenciado sobre a ocorrência de algum evento importante (falhas ocorridas, necessidade de maior alocação de recursos, etc.).

Devido ao fato de que a estação de gerenciamento geralmente é responsável por uma grande quantidade de agentes, e que cada agente geralmente possui uma grande quantidade de objetos, a realização do processo de *polling* em cada um desses agentes e seus respectivos objetos se torna impraticável. Devido a isso, o protocolo SNMP e suas MIBs associadas foram projetadas para encorajar gerentes de rede a utilizarem o recurso de notificação. Esse recurso permite atribuir ao agente a responsabilidade de notificar a estação de gerenciamento sobre o acontecimento de algum evento não esperado. Por exemplo, pode-se notificar condição de sobrecarga, falha de um dos servidores, etc. As mensagens que comunicam esses eventos aos gerentes SNMP são conhecidas como *traps*.

Uma vez que a estação de gerenciamento tenha sido alertada sobre a ocorrência dessas condições não esperadas (exceções), ela pode executar ações de modo a reagir ao evento ocorrido. Por exemplo, o gerente pode requisitar novas informações do agente que o notificou, ou de agentes próximos, para melhor diagnosticar o problema. Logo em seguida, a estação de gerenciamento poderá enviar requisições para alteração de valores de objetos nos agentes problemáticos, a fim de manter a rede operacional. Dessa forma, o uso de *traps* pode economizar substancialmente capacidade de rede e tempo de processamento nos agentes.

2.1.4 Estrutura do Protocolo SNMP em relação ao modelo OSI/ISO

O *Simple Network Management Protocol* (SNMP) foi desenvolvido para atuar na camada de aplicação, prevista pelo modelo OSI/ISO. Dessa forma, os agentes de gerenciamento das redes TCP/IP implementam o SNMP como protocolo de aplicação, o *User Datagram Protocol* (UDP) como protocolo de transporte, o *Internet Protocol* (IP) como protocolo de rede, e um protocolo adequado para a camada de enlace da rede que está sendo gerenciada.

Devido ao fato do *Simple Network Management Protocol* utilizar o UDP como protocolo da camada de transporte, o SNMP em si não é orientado à conexões. Isso significa

que nenhuma conexão é mantida entre a estação de gerenciamento e seus agentes, ou seja, cada troca de informação é uma transação distinta.

A Figura 2.3 (STALLINGS, 1999) apresenta um esquema que mostra a estrutura do protocolo SNMP em relação ao modelo OSI/ISO, assim como o processo básico de troca de mensagens de gerenciamento.

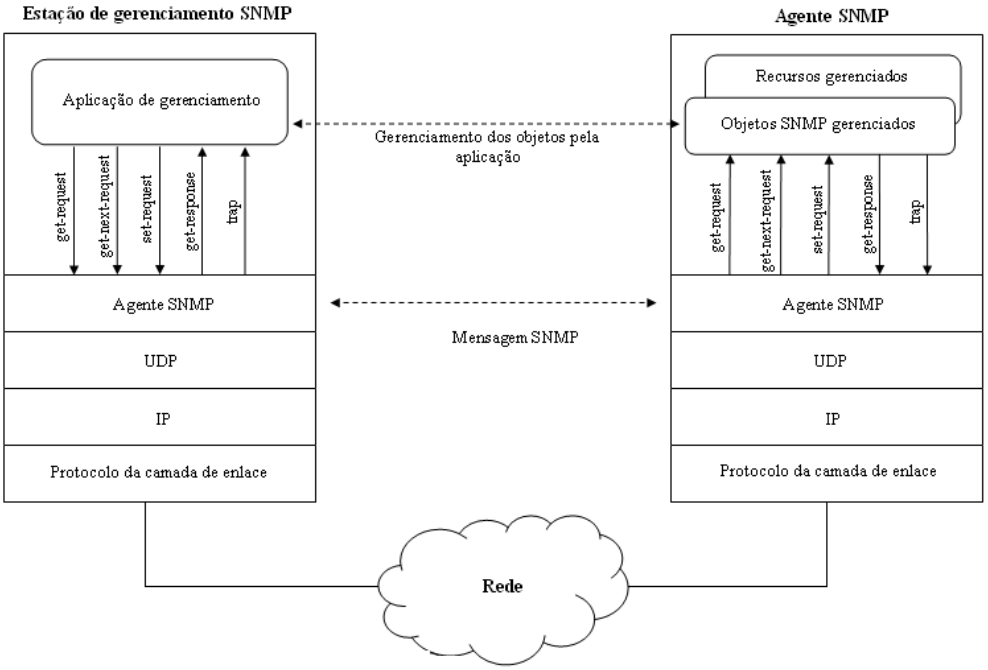


Figura 2.3: Estrutura do protocolo e das mensagens SNMP

2.1.5 Comunidades e *strings* de comunidade

A fim de implementar um sistema de segurança básico, o SNMP faz uso do conceito de comunidades, que basicamente consiste no relacionamento entre um agente e um conjunto de gerentes SNMP. Nesse relacionamento são definidas as regras para autenticação e controle de acesso. Desta forma, cada conjunto de parâmetros para autenticar e controlar o acesso de gerentes SNMP é descrito como uma comunidade. Para ilustrar esse fato, suponha-se um grupo de gerentes da rede que tenha acesso à leitura dos valores dos objetos SNMP em um roteador, e um outro grupo de gerentes que tenha permissão para leitura e gravação de valores desses mesmos objetos. Cada um desses grupos integra uma comunidade distinta das duas ou mais comunidades configuradas no agente SNMP.

O sistema de autenticação dessas comunidades de gerenciamento é bastante trivial. Cada uma dessas comunidades é identificada por uma *string*, denominada ***string* de comunidade**. Dessa forma, quando um gerente SNMP pertencente a uma determinada comunidade deseja se autenticar no agente SNMP, ele apenas informa a *string* de comunidade, que nesses casos funcionam como uma espécie de senha. A partir daí, o agente SNMP assume a comunicação como sendo autêntica e executa as requisições feitas pelo gerente SNMP autenticado.

O grande problema desse esquema de autenticação é que, até a versão 3 do SNMP, essas *strings* de comunidade trafegam na rede sem passar por nenhum processo de criptografia. Sendo assim, essas mensagens SNMP podem ser facilmente capturadas por um monitor de rede (*sniffer*), e as *strings* de comunidade poderão ser descobertas, permitindo

a um atacante obter todos os privilégios de acesso garantidos aos membros daquela comunidade de gerenciamento. Devido a isso, acredita-se que o protocolo SNMP é utilizado quase que exclusivamente para monitoramento da rede, o que significaria que as comunidades costumam ser configuradas de forma a permitir apenas a leitura dos valores dos objetos SNMP, e nunca a escrita. Isso acontece porque, devido à extrema simplicidade do mecanismo de autenticação do protocolo SNMP, a permissão de escrita de valores nesses objetos pode acarretar em uma séria vulnerabilidade na segurança dessas redes.

2.2 Medições sobre Tráfego SNMP

Em março de 2006, foi publicada a primeira versão do *internet draft* intitulado "SNMP Traffic Measurements" (SCHOENWAELDER, 2006), o qual propunha uma metodologia sistemática para medições e geração de estatísticas sobre o uso do SNMP. Esse *internet draft* foi renovado outras 3 vezes (SCHOENWAELDER, 2007a) (SCHOENWAELDER, 2007b) (SCHOENWAELDER, 2008), estando ele válido até 15 de agosto de 2008. Uma das motivações que levaram à publicação dessa metodologia foi o fato de que muitos trabalhos publicados trataram do desempenho apresentado pelo protocolo SNMP, o impacto da segurança do SNMPv3, ou o desempenho relativo apresentado pelo SNMP em relação ao uso de Web Services aplicados na tarefa de gerenciamento de redes de computadores (SCHOENWAELDER, 2006). Entretanto, apesar dessas publicações serem de grande importância para a compreensão do impacto causado pelas decisões tomadas no desenvolvimento das tecnologias tratadas nesses trabalhos, parte da fundamentação dos mesmos pode ser questionada. Os autores desses trabalhos muitas vezes assumem certos padrões de interação na utilização do protocolo SNMP sem possuírem evidências experimentais de que essas interações de fato ocorrem no dia-a-dia das redes em produção. Esse fato pode ser bastante problemático em qualquer pesquisa que se refira ao SNMP, tendo em vista que fatos especulados estão sendo apresentados como fundamentação para esses trabalhos. Até então, nenhuma metodologia sistemática para a detecção desses padrões de interação na utilização do protocolo SNMP foi desenvolvida ou utilizada.

Para exemplificar esse problema, o *internet draft* apresenta como exemplo o fato de que muitos autores usam a `ifTable` da IF-MIB (MCCLOGHRIE; KASTENHOLZ, 1994) ou a `tcpConnTable` da TCP-MIB (RAGHUNARAYAN, 2005) como ponto inicial da sua análise ou comparação. Quando isso ocorre, a argumentação desses autores perde força, porque não existem quaisquer evidências de utilização significativa dessas tabelas nos tráfegos SNMP das redes que estão atualmente em funcionamento. Além disso, ainda é mais incerta a maneira com que as aplicações que fazem uso do SNMP lêem essas tabelas, e quais otimizações elas empregam nesse processo. Outro aspecto do protocolo ainda desconhecido é a verdadeira relação existente entre o processo de *polling* periódico e aperiódico ocorrido nas redes em produção. Também não se conhece, com relação à utilização das MIBs, se o tráfego SNMP real é mais voltado à utilização de objetos padronizados ou de objetos proprietários. Como se pode observar, é importante que as respostas a essas questões sejam obtidas através do emprego de uma metodologia adequada, a fim de se conseguir resultados que não se baseiem mais em especulações, mas sim em fatos cientificamente comprovados.

Nas próximas subseções serão apresentadas as etapas que compõem a metodologia do IRTF, assim como os aspectos do protocolo SNMP que devem ser esclarecidos a partir das análises geradas pelo uso dessa metodologia.

2.2.1 Primeira Fase: Captura do Tráfego SNMP

A primeira etapa descrita pela metodologia do IRTF para medições sobre tráfegos SNMP é a captura do tráfego em si. Essa captura pode ser realizada através do uso de *sniffers* de rede convencionais, tais como o Ethereal (ETHEREAL, 2007), o Wireshark (SHARPE; WARNICKE; LAMPING, 2007) ou o TCPDUMP (JACOBSON; LERES; MCCANE, 2007). O formato utilizado por esses *sniffers* para armazenar o tráfego monitorado é o PCAP (*Packet Capture*).

Os *sniffers* podem ser facilmente configurados através do uso de filtros, a fim de que todos os pacotes SNMP que trafegarem na rede no momento da monitoração sejam capturados. Sugere-se que sejam capturados todos os pacotes UDP que tenham como origem ou destino as portas 161 e 162. Como exemplo, no caso do TCPDUMP esse filtro seria '*udp and (port 161 or port 162)*'. Também é necessária a garantia de que o pacote SNMP seja capturado por completo, ou seja, que o pacote não seja "truncado" (no caso do TCPDUMP, isso é realizado através da opção -s 0).

Com relação ao posicionamento desse *sniffer* na rede, deve-se ter o cuidado de colocá-lo em um ponto estratégico para a captura da maior quantidade possível de pacotes SNMP. Especialmente em redes locais baseadas em *bridges*, é importante garantir que a estação que fará o monitoramento tenha acesso a todas as VLANs (*Virtual LANs*) por onde passe tráfego de gerenciamento. Na maioria dos casos, o *sniffer* deve ser posicionado muito próximo do sistema de gerenciamento, assim como deve se configurar portas de monitoramento específicas nas redes locais baseadas em *bridges*.

Recomenda-se uma semana completa como tempo mínimo para monitoração de um tráfego SNMP. Entretanto, períodos de monitoração ainda maiores são encorajados pela metodologia do IRTF. Caso o tamanho dos arquivos contendo o tráfego SNMP seja muito grande, pode-se dividir o mesmo em vários pedaços através do uso de ferramentas como o TCPSLICE e o PCAPMERGE (JACOBSON; LERES; MCCANE, 2007).

Por fim, para cada seção de monitoração, é necessário que um conjunto de metadados seja armazenado juntamente com os arquivos de tráfego. Esse conjunto de metadados é apresentado na Tabela 2.1.

Tabela 2.1: Metadados relacionadas a cada seção de monitoração

Metadado	Descrição
Name	Nome do arquivo PCAP
Network	Nome da rede monitorada
Organization	Organização que opera a rede monitorada
Contact	Nome e e-mail de um contato
Start-Date	Data de início da monitoração, no formato ISO
End-Date	Data de término da monitoração, no formato ISO
Size	Tamanho do arquivo PCAP, em *bytes*
Description	Breve descrição sobre a rede monitorada

O nome de cada campo apresentado na Tabela 2.1 está em inglês com o objetivo de manter uma padronização com os metadados gerados por outros grupos de pesquisa em diferentes países.

2.2.2 Segunda Fase: Conversão dos Arquivos PCAP

Os arquivos que contêm tráfego SNMP no formato PCAP devem ser convertidos para um formato que atenda aos seguintes requisitos:

- **Legível para humanos**: fazer com que um arquivo de tráfego seja facilmente legível para seres humanos permite que um operador verifique se dados confidenciais não estão presentes nesse novo formato, evitando assim a divulgação dos mesmos;

- **Legível para máquinas**: o novo formato do arquivo também deve ser definido de forma a ser facilmente processado pelo computador, para que a análise dos dados disponíveis nesses arquivos ocorra de forma eficiente.

Uma escolha natural para que esses requisitos sejam atendidos é o formato XML, já que ele é facilmente legível por humanos e apresenta suporte para a maioria das linguagens de programação de alto nível, facilitando assim o desenvolvimento de programas ou *scripts* que realizem as tarefas de análise desses dados. Entretanto, como XML é uma linguagem que emprega o uso de muitas *tags* (causando sobrecarga no processamento), arquivos nesse formato podem ser difíceis de serem processados, se o tráfego neles representado for muito extenso. Por isso, recomenda-se o uso de APIs (*Application Programming Interfaces*) que façam o *streaming* do arquivo XML, a fim de evitar que uma representação completa desse documento na memória se faça necessária, o que inviabilizaria o processamento desses arquivos em diversos tipos de computadores.

Uma alternativa mais leve para o uso do formato XML é o formato CSV (*Comma Separated Values*, ou Valores Separados por Vírgula). O formato CSV consiste basicamente em se armazenar num arquivo de texto puro (ASCII) as informações que compõem uma mensagem SNMP seqüencialmente, numa única linha, com seus valores separados por vírgula. No caso de se utilizar esse tipo de representação, recomenda-se então que sejam armazenadas apenas as informações mais essenciais referentes ao tráfego SNMP, de maneira a facilitar ainda mais o processamento desses arquivos.

2.2.3 Terceira Fase: Filtragem dos Arquivos XML/CSV

Esse é um passo considerado fundamental nessa metodologia, pois quando lida-se com tráfegos SNMP reais (i.e., tráfegos originados por redes em produção), é necessário que se tenha o cuidado de proteger as fontes dessas informações, pois o protocolo SNMP carrega em si muitos dados que, por questões de segurança, não podem ser divulgados. Dessa forma, um processo de remoção/anonimização desses dados sensíveis deve ser especificado, a fim de dar segurança aos operadores que se dispuserem a oferecer amostras de tráfego SNMP pertencentes a suas respectivas redes.

A filtragem desses dados pode ser feita através da análise e alteração da representação do tráfego SNMP no formato XML ou CSV. No caso do formato XML, processadores XSLT padrão, como o xsltproc (XMLSOFT, 2007), podem ser utilizados para esse propósito. Também poderão ser utilizadas bibliotecas de linguagens de programação de alto nível, específicas para o tratamento de documentos XML, para a manipulação desses arquivos. Por exemplo, pessoas familiarizadas com a linguagem Perl poderão utilizar a biblioteca XML::LibXML (XMLSOFT, 2007) para realizar a filtragem dos dados obtidos.

2.2.4 Quarta Fase: Armazenamento do Arquivo PCAP e de sua Representação XML/CSV

Tanto o arquivo PCAP contendo a amostra do tráfego SNMP a ser estudada como a sua representação filtrada em formato XML ou CSV necessitam ser armazenadas em um repositório estável, como por exemplo um CD-ROM ou um DVD-ROM. Esse repositório deve estar sob o controle do grupo que está conduzindo a pesquisa ou do operador da rede que serviu de fonte para os dados coletados.

Essa etapa é necessária porque ao longo da pesquisa, ou até mesmo depois que a mesma tiver ocorrida, poderão ser descobertos problemas em algum processo realizado durante a obtenção, conversão ou análise dos dados, de forma que seja necessário repetir um ou mais passos da metodologia. Por causa disso, os dados originais (arquivo PCAP bruto e arquivos XML/CSV filtrados) devem ser armazenados e preservados de forma a possibilitar a recuperação dessas informações numa eventual necessidade futura. Dessa forma, é preciso se ter muito cuidado para que os dados armazenados permaneçam inalterados até que se tenha absoluta certeza de que eles não serão mais necessários.

Nesse processo de armazenamento, poderá ser utilizado algum algoritmo de compressão sem perdas (comumente encontrado em programas como gzip ou bzip2), ou algum algoritmo de criptografia, a fim de aumentar a segurança dos dados armazenados.

2.2.5 Quinta Fase: Análise dos Arquivos Filtrados

O último passo da metodologia consiste na análise dos arquivos filtrados, a fim de se agregar os dados dos diversos pacotes de gerenciamento contido nesses arquivos e, a partir deles, extrair informações de forma a responder às perguntas sobre a real utilização do protocolo SNMP.

A análise desses dados deverá se dar através da criação e execução de *scripts* que procurem agregar os dados dos pacotes SNMP de forma a permitir a identificação de predominâncias e tendências dentro desse tráfego (por exemplo, qual versão do protocolo é mais utilizada, qual a relação entre o tráfego periódico e o aperiódico, etc). Esses *scripts* deverão ser bem avaliados, a fim de se verificar a corretude dos mesmos. Preferencialmente, esses *scripts* deverão ser publicamente acessíveis, para que terceiros possam auxiliar nesse processo de verificação. Além disso, o compartilhamento desses arquivos também poderá ajudar outros grupos de pesquisa a repetir as análises feitas pelos pesquisadores que criaram esses *scripts*, podendo até estendê-los quando necessário.

Basicamente, os *scripts* de análise dos dados filtrados poderão ser escritos em qualquer linguagem de programação, devido à abundância de bibliotecas para tratamento de arquivos XML existentes entre as linguagens de alto nível. Contudo, recomenda-se que esses *scripts* sejam implementados utilizando-se a linguagem de programação Perl juntamente com a biblioteca XML::LibXML, a fim de se criar um ”vocabulário” comum entre pesquisadores e operadores de rede, e também entre os diversos grupos que estão realizando essa pesquisa pelo mundo. Além disso, Perl possui uma vantagem natural com relação a linguagens de programação como C/C++, porque o mesmo apresenta um menor tempo necessário para o desenvolvimento dos *scripts* de análise.

Por último, relembra-se que arquivos XML muito grandes podem ser problemáticos para serem processados quando se utilizam bibliotecas de tratamento XML que necessitam da representação completa do arquivo na memória. Devido a isso, é recomendável nesses casos a utilização de uma API que faça *streaming* do arquivo XML, para que apenas partes desses arquivos estejam representadas na memória ao longo do processamento.

2.2.6 Aspectos do Tráfego SNMP a serem Analisados

Nesta subseção serão apresentadas algumas questões iniciais a serem respondidas através da análise do tráfego SNMP coletado. Estas são as questões encontradas no *internet draft* que apresenta a metodologia descrita nesse trabalho (SCHOENWAELDER, 2008). Entretanto, o autor desse *draft* deixa bastante claro que não houve nenhum empenho em se fornecer uma lista completa de questões, de forma que várias outras perguntas podem ser acrescentadas a essa lista, o que só viria a contribuir com a metodologia do IRTF.

- **Estatísticas Básicas**: As estatísticas básicas, no âmbito dessa metodologia, compreende o conjunto de informações genéricas a respeito do tráfego SNMP que foi analisado. Dentre os elementos que compõem essa análise, pode-se citar: versões mais utilizadas do protocolo, operações que aparecem com maior freqüência, número de agentes e gerentes na rede, etc. Esse conjunto de informações propicia uma visão geral sobre a amostra do tráfego que está sendo estudada.

- **Relação existente entre tráfego periódico e aperiódico**: Para permitir o gerenciamento de dispositivos de rede, o protocolo SNMP permite a realização de *polling* periódico, a fim de se obter informações do dispositivo desejado. Por outro lado, o protocolo SNMP também permite consultas sob demanda, ou seja, a obtenção de informações dos dispositivos realizada através de solicitações explícitas por parte do administrador da rede. Esse tipo de operação gera tráfego caracterizado como aperiódico. Dessa forma, espera-se encontrar na amostra dos pacotes SNMP esses dois grupos de tráfegos: periódico e aperiódico. É importante se descobrir qual a relação existente entre as quantidades de tráfego encontradas nesses dois grupos.

- **Tamanho da mensagem e distribuição da latência**: As mensagens SNMP possuem seu tamanho restrito pelos mapeamentos da camada de transporte e pelos *buffers* utilizados pelos mecanismos SNMP. Devido a isso, e objetivando subsidiar aprimoramentos nas futuras versões do protocolo SNMP, é interessante se investigar qual a distribuição dos tamanhos das mensagens encontradas na amostra de tráfego SNMP. Além disso, é importante que se compreenda a distribuição da latência, especialmente a distribuição do tempo de processamento pelos dispositivos que estão processando as requisições do protocolo. Algumas implementações do SNMP inferem os atrasos da rede através da medição do tempo de requisição-resposta, abordagem essa que poderá ser validada ou não a partir do estudo desse tempo de processamento das requisições.

- **Níveis de concorrência**: Uma outra característica interessante do protocolo SNMP é a de que ele permite que as estações de gerenciamento requisitem informações de múltiplos agentes concorrentemente. Entretanto, apesar de se saber que esse recurso se encontra disponível, não se conhece se o mesmo é extensamente utilizado, nem os níveis de concorrência que são utilizados nessa operação. Devido a isso, é interessante se identificar nas amostras de tráfego SNMP estudadas os níveis de concorrência tipicamente encontrados, ou seja, descobrir se as requisições de gerenciamento estão sendo feitas em alto nível de concorrência, ou se as aplicações estão preferindo a solicitação de informações de maneira mais seqüencial.

- **Abordagens para leitura de tabelas**: No protocolo SNMP, as tabelas podem ser lidas de diversas maneiras. Dentre elas, a maneira mais ineficiente é a leitura célula

por célula, percorrendo-se as colunas da tabela sequencialmente. Uma opção que apresenta maior eficiência é a leitura de tabelas linha a linha, ou mesmo lendo-se múltiplas linhas de cada vez. Além disso, a leitura dos elementos de índice pode ser suprimida na maioria dos casos. A partir do estudo das amostras de tráfegos obtidas, é possível se descobrir quais abordagens estão de fato sendo colocadas em uso nas redes em produção que utilizam o protocolo SNMP para realizarem seu gerenciamento.

- **Notificações utilizando *traps***: Um importante conceito empregado no protocolo SNMP é o de *traps*. As aplicações de gerenciamento são responsáveis por realizar o *polling* periodicamente entre os dispositivos da rede, a fim de determinar o seu *status*. Através do uso do conceito de *traps*, os dispositivos gerenciáveis podem ser programados para notificarem os gerentes SNMP sobre eventos que tenham ocorrido, para que esses gerentes adotem as medidas aplicáveis àquele evento de maneira mais rápida do que o processo convencional de *polling*. Análises de tráfego SNMP podem identificar exatamente o quanto de *polling* convencional e o quanto de *traps* estão sendo realmente empregados. Uma questão ainda mais particular que deve ser levantada é a identificação de quando as notificações geradas por *traps* levam a uma reação por parte do gerente SNMP. Essa é uma questão importante, porque a rápida reação a eventos na rede é um dos fundamentos que justificam a existência do conceito de *trap* no protocolo SNMP.

- **Módulos MIB populares**: Através da análise dos prefixos dos objetos SNMP encontrados na amostra do tráfego, é possível se descobrir quais são os módulos MIBs mais utilizados, ou mesmo os tipos de notificações definidos por esses módulos. Essa também é uma informação importante a ser descoberta, pois através desse estudo pode-se propiciar um desenvolvimento dessas MIBs (e de outras relacionadas) mais direcionado para a utilização real desse protocolo, ao mesmo tempo em que evita-se investir trabalho em um módulo que tenha pouca ou nenhuma utilização real.

- **Uso de objetos obsoletos**: Atualmente, muitos objetos do protocolo SNMP são considerados obsoletos, pelo fato dos mesmos não conseguirem representar a realidade dos dispositivos de redes atuais. Como exemplo desse fato, podemos citar o objeto `ipRouteTable`, lançado na Internet em 1993 e que posteriormente foi considerado obsoleto por não ser apto a representar roteamento sem uso de classes. Entretanto, apesar de considerados obsoletos, alguns desses objetos continuam sendo citados em publicações populares e até mesmo em trabalhos acadêmicos. Devido a isso, seria interessante verificar se as aplicações SNMP de fato ainda utilizam esses objetos considerados obsoletos, ou se elas foram atualizadas para substituírem esses objetos por suas novas versões.

- **Distribuição do tamanho da codificação**: A fim de se estimar o tamanho das mensagens SNMP, algumas vezes assume-se valores sobre a distribuição do tamanho da codificação dos vários tipos de dados dessas mensagens. Essa estimativa é utilizada para se determinar as restrições de transporte e de *buffers* das implementações do protocolo. Devido a esse fato, também seria importante se descobrir nesse estudo de uso do protocolo SNMP a distribuição do tamanho da codificação dos vários tipos de dados encontrados nas mensagens presentes na amostra do tráfego estudada.

- **Contadores de decontinuidades**: No protocolo SNMP, os contadores podem sofrer descontinuidades (MCCLOGHRIE; PERKINS; SCHOENWAELDER, 1999). O indicador de descontinuidade padrão é o objeto escalar `sysUpTime`, da MIB SNMPv2 (MCCLOGHRIE; KASTENHOLZ, 2002), que também pode ser utilizado para detecção do problema de *counter roll-over*. Alguns módulos MIB apresentam indicadores de descontinuidades mais específicos, como o objeto *ifCounterDiscontinuityTime*, da IF-MIB (MCCLOGHRIE; KASTENHOLZ, 2000). Dessa forma, é interessante o estudo dos propósitos com os quais os objetos apresentados no parágrafo anterior são utilizados pelas aplicações de gerenciamento para o tratamento dos eventos de descontinuidade.

- ***Spin Locks***: Geradores de comandos cooperativos podem fazer uso de travas (*locks*) de alerta para coordenar o uso do protocolo SNMP durante as operações de escrita de informações. O objeto escalar `snmpSetSerialNo` da MIB SNMPv2 (MCCLOGHRIE; KASTENHOLZ, 2002) é o objeto padrão para coordenação desse tipo de operação. Por isso, o estudo da existência de geradores de comandos que se coordenam através do uso desses *spin locks* é relevante dentro do contexto de medições sobre tráfegos SNMP.

- **Criação de linhas em tabelas**: A criação de linhas em tabelas é uma operação que não possui suporte nativo no protocolo SNMP. Entretanto, tabelas conceituais que possuem suporte para a criação de linhas tipicamente fornecem uma coluna de controle que utiliza a convenção textual `RowStatus`, definida na MIB SNMPv2-TC (CASE et al., 1996). O objeto `RowStatus` em si suporta diferentes modos de criação de linhas, tais como os modos *createAndWait* e *createAndGo*. Além disso, diferentes abordagens podem ser utilizadas para inferir se o identificador de instância não possui uma semântica especial associada. Devido a isso, é interessante se estudar quais das diferentes abordagens existentes para criação de linhas em tabelas estão sendo realmente utilizadas nas aplicações de gerenciamento das redes em produção.

Uma vez que foram revisados os conceitos básicos relacionados ao contexto da área de gerenciamento de redes e da metodologia do IRTF para medições sobre tráfegos SNMP, serão apresentados nos próximos capítulos os demais aspectos referentes à pesquisa contida nesta dissertação de mestrado. O primeiro desses aspectos a ser discutido é o conjunto de técnicas de visualização de informação desenvolvido especialmente para os resultados obtidos com as medições sobre tráfegos SNMP, o qual será apresentado no próximo capítulo.

3 TÉCNICAS DE VISUALIZAÇÃO DE INFORMAÇÃO PARA MEDIÇÕES SOBRE TRÁFEGO SNMP

Visualização de informação (do inglês *Information visualization*, ou simplesmente Infovis) é a área do conhecimento que objetiva auxiliar no processo de descoberta e análise de dados através da exploração visual (FEKETE; PLAISANT, 2002). O uso de técnicas de visualização permite que as pessoas façam uso das propriedades do sistema visual humano para explorarem e obterem *insights* sobre um conjunto de dados, reconhecendo padrões, anomalias e demais características interessantes, de forma eficiente e eficaz. Essa área do conhecimento é relativamente nova, mas rapidamente se tornou um campo abrangente e interdisciplinar. Trabalhos na área de visualização de informação são encontrados em grande número na literatura, especialmente na área de recuperação de informação, hypertexto e World Wide Web (WWW), bibliotecas digitais e interface homem-máquina (CHEN, 2004).

Uma técnica de visualização é composta de uma representação visual e, comumente, de um mecanismo de interação associado à essa representação (FREITAS, 2007). A representação visual é baseada numa forma de se mapear os atributos de uma estrutura de dados abstrata para atributos visuais. Geralmente, o nível de abstração de uma representação visual é maior do que o dos dados puros, o que é especialmente útil quando se está explorando grandes conjuntos de dados. Os mecanismos de interação, por sua vez, permitem aos usuários manipularem a representação visual de forma a agilizar e facilitar a exploração do conjunto de dados. A visualização de dados também é um processo que pode ser definido como um conjunto de mapeamentos, onde os dados brutos vão sofrendo transformações até que seja obtida a visualização desses dados, através de primitas geométricas e visuais. O modelo mais conhecido para esse processo é apresentado por Card *et al*. (CARD; MACKINLAY; SHNEIDERMAN, 1999), e está representado na Figura 3.1.

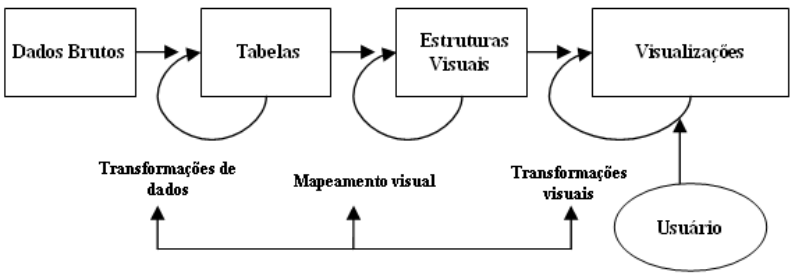


Figura 3.1: Modelo de referência de visualização de Card

Atualmente, tanto a academia quanto a indústria estão cada vez mais interessadas no

desenvolvimento de técnicas de visualização de dados específicas para a área de gerenciamento de redes. Os trabalhos de Oberheide *et al.* (OBERHEIDE; GOFF; KARIR, 2006), Papadopoulos *et al.* (PAPADOPOULOS et al., 2004) e Keim *et al.* (KEIM et al., 2006) são exemplos de estudos que objetivaram o desenvolvimento de técnicas de visualização para tráfegos de redes de computadores. Entretanto, todos esses trabalhos lidam com tráfegos de redes de um modo geral, ou seja, sem levar em consideração as finalidades específicas de cada tipo de tráfego.

Os primeiros trabalhos sobre técnicas de visualizações de tráfego SNMP foram apresentados nos trabalhos de Salvador e Granville (SALVADOR; GRANVILLE, 2008b) (SALVADOR; GRANVILLE, 2008a). Essas técnicas também estão implementadas na ferramenta *Management Traffic Analyzer* e serão descritas nas próximas seções deste capítulo.

3.1 Visualização da Topologia da Rede de Gerenciamento

Técnicas de visualização baseadas em grafos são amplamente utilizadas para a representação de determinados aspectos de uma rede de computadores, como a sua topologia (BECKER; EICK; WILKS, 1995). Para que essa técnica seja adaptada a fim de que possa produzir uma visualização da topologia de uma rede de gerenciamento, é necessário se identificar no tráfego SNMP quais terminais atuam como gerentes, quais atuam como agentes, e quais atuam como ambos. Essa identificação deve ser feita através da análise dos fluxos de mensagens SNMP que compõem o tráfego. Esses fluxos são definidos como o conjunto de mensagens que partiu de uma determinada origem até um determinado destino. Além disso, os fluxos de mensagens SNMP pertencem à duas classes de relacionamento:

- *Command Generator* (CG) / *Command Responder* (CR): relacionamento do tipo requisição/resposta (e.g., *polling*). O nó de origem atua como gerente e o nó de destino atua como agente;

- *Notification Originator* (NO) / *Notification Receiver* (NR): tipo de relacionamento existente numa operação de envio de notificação (e.g., *traps*). O nó de origem atua como agente e o nó de destino como gerente.

Uma vez que gerentes e agentes tenham sido identificados, esses elementos serão representados na visualização através de círculos. O tamanho dos círculos indica o papel que um determinado terminal desempenhou no tráfego estudado. Se um terminal atuou apenas como agente ao longo de todo o tráfego, ele será representado por um círculo menor. Por outro lado, se um terminal atuou como um gerente em qualquer um dos fluxos de mensagens SNMP, ele será representado por um círculo maior, mesmo que este tenha atuado como agente em algum outro fluxo de mensagens. O tamanho do círculo também dependerá do número de mensagens enviadas/recebidas que sejam características de um gerente da rede: quanto maior o número de mensagens desse tipo, maior será o tamanho do círculo na visualização. Isso indica que, quanto maior o círculo, maior é a atuação do nodo como gerente na rede representada.

Uma linha sólida ligando dois nós representa uma conexão entre um nó gerente e um nó estritamente agente (i.e., que atuou como agente em todos os fluxos de mensagens do tráfego). Já uma linha tracejada representa a conexão entre um nó gerente e um nó

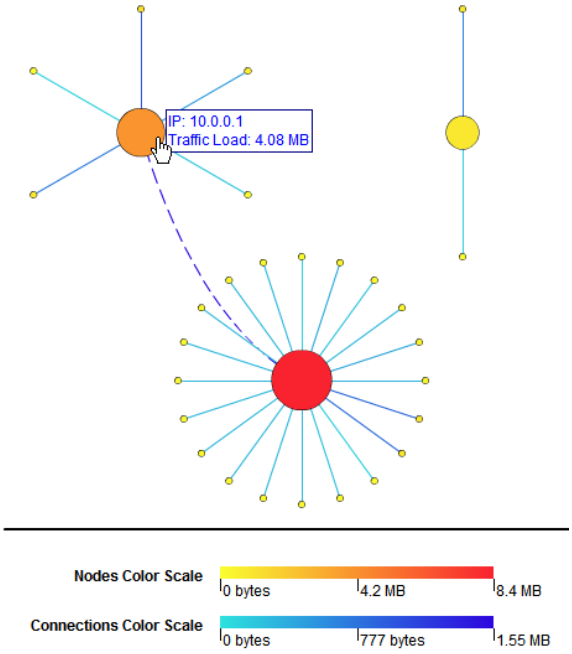


Figura 3.2: Visualização da Topologia da Rede de Gerenciamento

que está atuando naquele relacionamento específico como agente, mas que atuou também como gerente em outros fluxos de mensagens.

Tanto os nós quanto as conexões possuem uma determinada cor que carrega consigo informações sobre carga do tráfego. No caso dos nós, a cor representa a quantidade de tráfego recebida pelo terminal visualizado. Por outro lado, a cor das conexões representa a quantidade de tráfego trocada entre os dois terminais conectados. Na parte inferior da visualização existem 2 barras coloridas, que são as escalas de cores utilizadas para colorir nós e conexões. Foram utilizadas escalas de cores distintas devido à diferença entre as ordens de grandeza dos tráfegos dos nós e das conexões. Todas as cores utilizadas nas visualizações foram definidas no padrão RGB, e elas foram especificadas da seguinte forma (os valores apresentados estão em hexadecimal):

- **Para os nós**: As cores vermelha e azul do padrão RGB possuem seus valores fixos, sendo estes F9 e 2F respectivamente. Para se obter as diversas tonalidades de vermelho encontradas nos nós da visualização, a cor verde do padrão RGB variou entre os valores 23 e FF.

- **Para as conexões**: As cores vermelha e azul do padrão RGB também possuem seus valores fixos, sendo estes 2C e DE, respectivamente. Para se obter as diversas tonalidades de azul encontradas nas conexões da visualização, a cor verde do padrão RGB variou entre os valores 06 e E2.

O posicionamento dos nós na área da visualização é definido segundo o algoritmo abaixo:

- Espalhar os nós que atuam como gerentes no espaço disponível para a visualização;

- Posicionar os nós que atuam exclusivamente como agentes ao redor de seus respectivos gerentes, os quais foram desenhados no passo anterior. A disposição desses nós agentes deverá formar um círculo imaginário;

- Se um determinado nó já foi representado na área da visualização anteriormente, ele não poderá ser desenhado uma segunda vez em outro lugar da visualização. Ele deve ser conectado ao outro nó que forma o seu par da posição onde ele foi desenhado da primeira vez.

Por fim, foram utilizados dois mecanismos de interação nessa visualização: legendas interativas para nós e conexões, e barra de rolagem. As legendas informam para o usuário algumas informações relevantes sobre nós (e.g., endereço IP e carga no nó) e conexões (e.g., tipo de relacionamento do fluxo e carga de tráfego). O usuário irá visualizar essas legendas quando posicionar o ponteiro do mouse sobre um nó ou uma conexão. Por sua vez, a barra de rolagem é o mecanismo que evita que a representação topológica seja truncada quando uma rede de grande porte estiver sendo representada. O resultado dessa análise pode ser observado na Figura 3.2.

Dentre as possíveis aplicações para essa técnica de visualização, destacam-se as seguintes:

- Determinar os nós com maior e menor volume de tráfego, sob o ponto de vista do tráfego de gerenciamento da rede;
- Identificar os gerentes com maior e menor concentração de agentes sob sua responsabilidade, a fim de possibilitar um melhor balanceamento da rede de gerenciamento;
- Identificar nós e conexões que não eram do conhecimento dos administradores da rede, e que possam ser considerados indesejáveis para o correto gerenciamento daquela rede;
- Visualizar a relação existente entre consultas do tipo *polling* (relacionamento do tipo CG/CR) e notificações (relacionamentos do tipo NO/NR).

3.2 Visualização de Objetos SNMP em uma *MIB Tree*

Uma das análises previstas pela metodologia do IRTF é o cálculo do número de vezes em que um determinado objeto SNMP é visto em um tráfego. Através desta análise, é possível se elaborar uma série de estatísticas sobre os objetos SNMP presentes no tráfego, como o conjunto de objetos mais acessados, os menos acessados, as MIBs (*Management Information Bases*) mais importantes, etc. Contudo, uma resposta textual desse tipo de análise possui algumas limitações. Provavelmente a principal delas é a dificuldade que um usuário teria para identificar o relacionamento hierárquico entre dois ou mais objetos, observando apenas seus nomes ou seus OIDs (*Objects Identifiers*). Uma vez que os objetos SNMP são organizados em uma árvore conhecida como *MIB tree*, é desejável que o resultado desse tipo de análise também apresente o conjunto de objetos encontrados no tráfego nesse tipo de estrutura. Devido a isso, foi desenvolvida uma técnica que mistura duas outras técnicas de visualização bastante conhecidas: visualização de árvores e histogramas. Dessa forma, o processo de análise dos resultados desse tipo de estatística se tornará mais eficiente, pois se assume que os administradores de rede estão potencialmente familiarizados com a organização de objetos SNMP em *MIB trees*.

A *MIB tree* desta visualização é desenhada segundo o algoritmo abaixo:

- Classificar os objetos SNMP encontrados no tráfego analisado pelos seus respectivos OIDs;

- Armazenar os objetos SNMP na memória em uma estrutura de dados de árvore, onde os nós folhas contêm as informações relevantes para um determinado objeto, como o nome da MIB a qual pertence este objeto e o número de mensagens SNMP onde esse objeto pôde ser encontrado;
- Realizar um caminhamento infixo pela árvore, desenhando nos nós folhas as informações dos objetos SNMP que são representados por esses nós. Cada nó folha deverá ser identificado por um rótulo formado pelo nome da MIB e o nome do objeto (no formato <nome-da-mib>::<nome-do-objeto>), quando estes forem conhecidos, ou o próprio OID do objeto representado.

Uma vez que a *MIB tree* contendo os objetos SNMP encontrados no tráfego tenha sido desenhada, será necessário representar o número de mensagens relacionadas com cada um dos objetos SNMP em um histograma. As barras do histograma são desenhadas do lado direito dos nós folhas da árvore, os quais representam os objetos SNMP. O tamanho da barra é baseado numa escala criada em tempo real, onde o valor mínimo é 0 e o máximo é o maior número de mensagens contendo um determinado objeto SNMP, dentre todos os objetos listados nos resultados da análise. Por fim, foi empregada uma barra de rolagem como mecanismo de interação para essa visualização, pois muito freqüentemente existe um grande número de objetos distintos em um tráfego SNMP. A Figura 3.3 mostra um exemplo dessa técnica sendo aplicada sobre um tráfego SNMP.

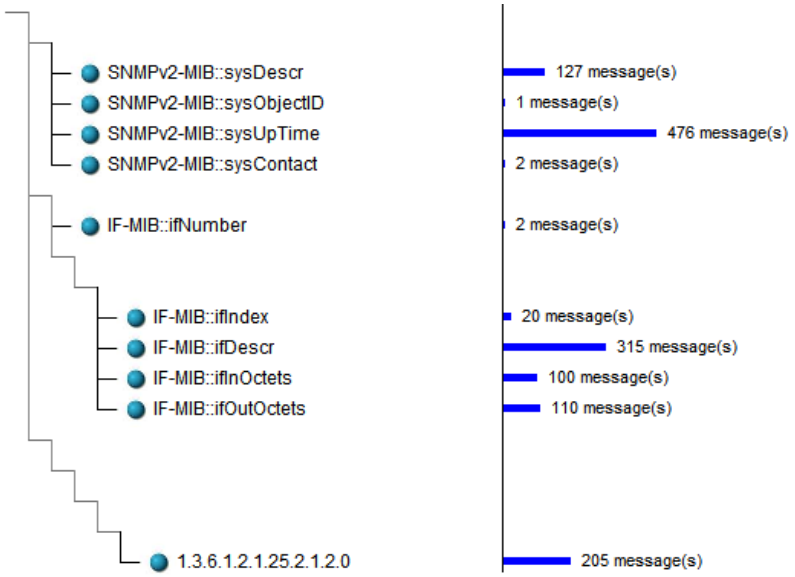


Figura 3.3: Visualização de Objetos do SNMP em uma *MIB Tree*

Dentre as aplicações para este tipo de visualização, destacam-se as seguintes:

- Identificar os objetos com maior e menor número de acessos;
- Identificar as MIBs que são mais utilizadas e as que estão sub-utilizadas;
- Verificar se existem objetos indesejáveis sendo acessados, isto é, objetos que possuem informações consideradas confidenciais e que não deveriam estar sendo acessados;
- Planejar possíveis otimizações nas consultas a objetos SNMP realizadas pelos softwares de gerenciamento.

3.3 Visualização da Quantidade de Mensagens SNMP em Intervalos de 1 Hora

Uma das análises disponibilizadas na ferramenta *Management Traffic Analyzer* é o cálculo da quantidade de mensagens encontradas no tráfego em intervalos de 1 hora. Esse tipo de análise é útil para se identificar o comportamento da quantidade dos diversos tipos de mensagens de gerenciamento trocadas na rede ao longo de um dia.

Os resultados dessa análise são representados em uma visualização que emprega histogramas para apresentar o número de mensagens SNMP transmitidas em cada intervalo de 1 hora do tráfego analisado. Por isso, cada histograma possui 24 barras, onde cada uma dessas barras representa um intervalo de 1 hora. Por exemplo, a primeira barra do histograma representa o intervalo entre 00h00min e 00h59min. O tamanho das barras do histograma é baseado em uma escala que vai de 0 ao número máximo de mensagens encontradas em uma única hora, ao longo de todo o tráfego analisado. Além disso, cada barra é dividida em várias seções, onde cada seção possui uma cor distinta, conforme mostra a Figura 3.4. Existem duas formas de se seccionar as barras do histograma: por versões ou por operações do protocolo SNMP. Dessa forma, o usuário poderá identificar as versões e operações predominantes no seu tráfego, além de ter conhecimento sobre o comportamento da quantidade total de mensagens por hora.

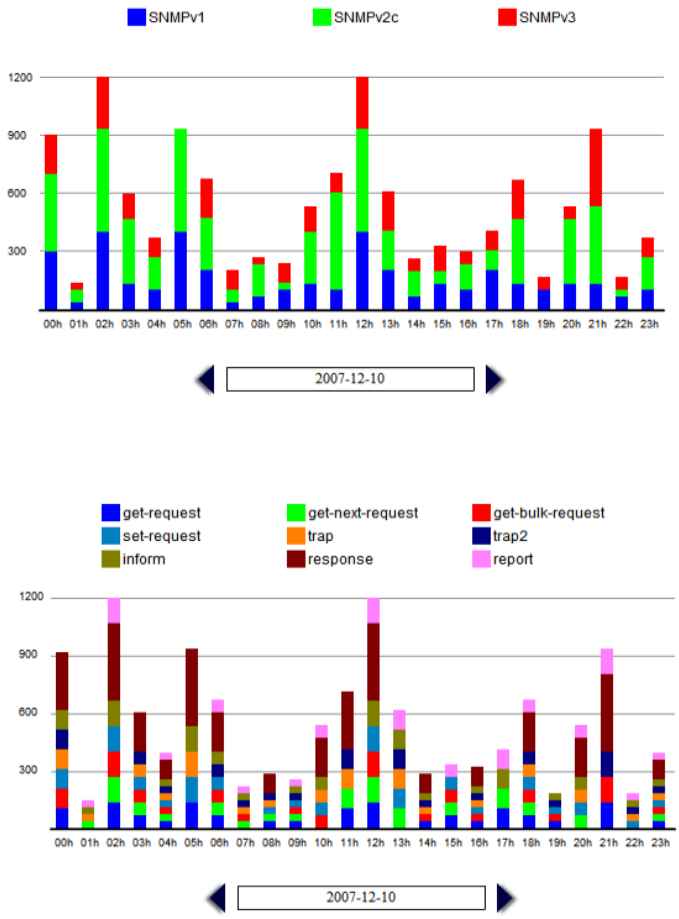


Figura 3.4: Histogramas seccionados por versões e por operações do SNMP

Um mecanismo de interação permite que o usuário saiba a qual dia o histograma que

está sendo apresentado se refere, assim como fornece a possibilidade do usuário navegar por esses dias, de modo a conhecer as quantidades de mensagens em intervalos de 1 hora referentes ao dia selecionado.

Para essa técnica de visualização destacam-se as seguintes aplicações:

- Identificar as operações e as versões do SNMP com maior e menor índice de utilização;
- Determinar o tipo e a versão das mensagens que predominam no tráfego analisado;
- Identificar as horas do dia onde são registrados as maiores e menores quantidades de mensagens SNMP;
- Identificar o comportamento do tráfego ao longo dos dias da semana, a fim de se obter uma visão completa das características da rede de gerenciamento.

A implementação das 3 técnicas de visualização de informação apresentadas neste capítulo estão disponíveis na ferramenta *Management Traffic Analyzer*. A arquitetura dessa ferramenta, assim como uma avaliação preliminar da mesma serão apresentadas no próximo capítulo.

4 FERRAMENTA: *MANAGEMENT TRAFFIC ANALYZER*

Conforme já foi apresentado na introdução desta dissertação de mestrado, a metodologia do IRTF para medições sobre tráfego SNMP possui uma série de limitações que dificultam a sua utilização e diminuem a eficiência da análise dos resultados obtidos. Com o intuito de abordar essas limitações, será apresentada nesta seção uma ferramenta Web para automatizar a execução da metodologia para medições sobre tráfego SNMP. As principais contribuições provindas dessa ferramenta para os estudos de tráfegos SNMP são as seguintes:

- Utilização de um módulo gerador de visualizações, destinado a facilitar o processo de interpretação dos resultados das análises executadas sobre tráfego SNMP;
- Eliminação da necessidade de se utilizar um conjunto de ferramentas para realizar todos os passos da metodologia do IRTF. A arquitetura supre todas as necessidades de *software* fundamentais para a realização das medições sobre tráfegos SNMP;
- Inclusão de um módulo voltado para comparação de resultados de análises de tráfego SNMP. A metodologia original do IRTF não especifica nenhuma forma para se comparar resultados de análises distintas.

Nas próximas seções deste capítulo será apresentada a arquitetura dessa ferramenta, assim como informações sobre a implementação da mesma e uma avaliação preliminar da ferramenta feita em conjunto com um grupo de voluntários que atuam na área de computação.

4.1 Arquitetura

Com o intuito de facilitar os trabalhos de pesquisadores e administradores de rede que desejem utilizar a metodologia do IRTF para medições sobre tráfego SNMP, foi especificada a arquitetura de uma ferramenta que automatize as diversas etapas dessa metodologia. Essa arquitetura é composta de 7 módulos, conforme apresentado na Figura 4.1.

O primeiro elemento que compõe a arquitetura do *software* é a sua **interface Web**. É somente através dela que o usuário poderá interagir com o sistema para executar operações como: criação de conta de usuário, login no sistema, inclusão de novo tráfego, análise e visualização de resultados, etc.

O **gerenciador de tráfego** é a parte da arquitetura responsável por manter registro de todos os tráfegos que o usuário submete à ferramenta Web. A entidade "tráfego", na arquitetura que está sendo desenvolvida, é a junção de um conjunto de metadados

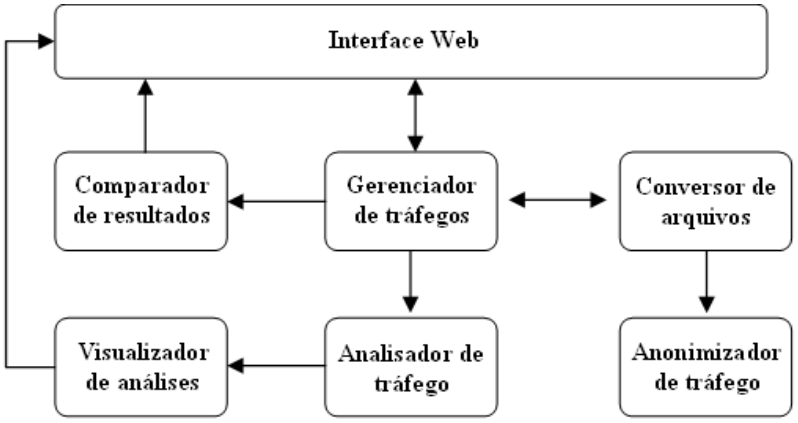


Figura 4.1: Arquitetura de uma ferramenta para medições de tráfego SNMP

que descreve um tráfego em si (esses metadados são os mesmos que são exigidos pela especificação da metodologia do IRTF), e os arquivos que se relacionam àquele tráfego, que podem estar no formato PCAP, XML ou CSV.

Uma vez que já existam arquivos submetidos para um determinado tráfego, o usuário terá à sua disposição uma "sub-ferramenta" denominada **conversor de formato de arquivos**. Com isso será possível, por exemplo, se converter um arquivo gerado a partir de uma seção de monitoramento de uma rede, originalmente no formato PCAP, para os formatos XML ou CSV. Esse tipo de conversão é necessária para permitir que esses arquivos de tráfego possam ser analisados posteriormente, pois a ferramenta não será capaz de analisar arquivos no formato PCAP, conforme já discutido anteriormente.

Durante o processo de conversão de um arquivo, o usuário poderá informar para a ferramenta parâmetros a serem utilizados na anonimização de informações que ele não deseja que sejam divulgadas. A parte da arquitetura responsável por tratar esses parâmetros e coordenar o processo de anonimização é chamada de **anonimizador de tráfego**.

Os parâmetros que poderão ser informados durante o processo de anonimização estão organizados na ferramenta na forma de *perfis de anonimização*. Dessa forma, primeiramente um usuário define um perfil de anonimização, especificando um nome para o mesmo e um conjunto de itens que se deseja anonimizar com a utilização daquele perfil. Dentre as opções que um usuário poderá encontrar para compor o seu perfil de anonimização, pode-se citar as seguintes:

- Remoção do endereço IP da origem da mensagem ou do destino da mesma;
- Remoção da porta da origem da mensagem ou do destino da mesma;
- Remoção da *string* de comunidade;
- Anonimização do endereço IP da origem da mensagem ou do destino da mesma;
- Anonimização da porta da origem da mensagem ou do destino da mesma;

Após os arquivos terem sido devidamente registrados e adequadamente convertidos na ferramenta, os mesmos estarão aptos a serem analisados. A parte da arquitetura responsável por executar essa operação é o **analisador de tráfego**. Para isso, o usuário poderá selecionar um dos tipos de análise disponíveis na ferramenta, e aplicá-la sobre os arquivos disponíveis no formato XML ou CSV. A análise consistirá na execução de um algoritmo

que terá como entrada o arquivo de tráfego, e com base nesse arquivo realizará uma série de buscas e cálculos para chegar ao resultado desejado. Os dados resultantes da análise são armazenados em tabelas específicas da base de dados, para serem utilizados posteriormente em visualizações ou comparações com outras análises. Ao final da operação, o sistema informa ao usuário que os dados resultantes daquela análise estão disponíveis, e apresenta ao mesmo uma lista com todas as visualizações possíveis para aquele resultado.

O módulo analisador de tráfego foi desenvolvido de forma a possuir baixo acoplamento com o restante da ferramenta, permitindo que outras análises possam ser desenvolvidas sem que sejam necessárias modificações no restante do código do sistema. Dessa forma, uma vez que novas análises tenham sido desenvolvidas, é suficiente que sejam feitas algumas poucas modificações na base de dados, como adição de linhas em tabelas ou criação de novas tabelas, para que a nova análise seja reconhecida pelo sistema.

As técnicas de visualização a serem utilizadas na ferramenta Web devem ser desenvolvidas de forma a recuperar os dados resultantes de uma análise diretamente do banco de dados, processar esses dados e exibir a visualização. O módulo da arquitetura que gerencia esse processo é denominado de **visualizador de resultados de análises**. É através desse módulo que são executados os algoritmos que realizarão as transformações e mapeamentos necessários para que os dados da análise sejam representados em estruturas visuais.

Uma análise na ferramenta pode ter uma ou mais técnicas de visualização associadas à mesma para que o usuário possa escolher, dentre as técnicas disponíveis, a que melhor se aplica ao seu caso, conforme mostra a Figura 4.2. De modo análogo ao que ocorre com as análises na ferramenta, também é possível se adicionar novas visualizações. Para isso, também será suficiente a manipulação de determinadas tabelas da base de dados para que o sistema possa reconhecer e disponibilizar uma nova técnica de visualização de dados.

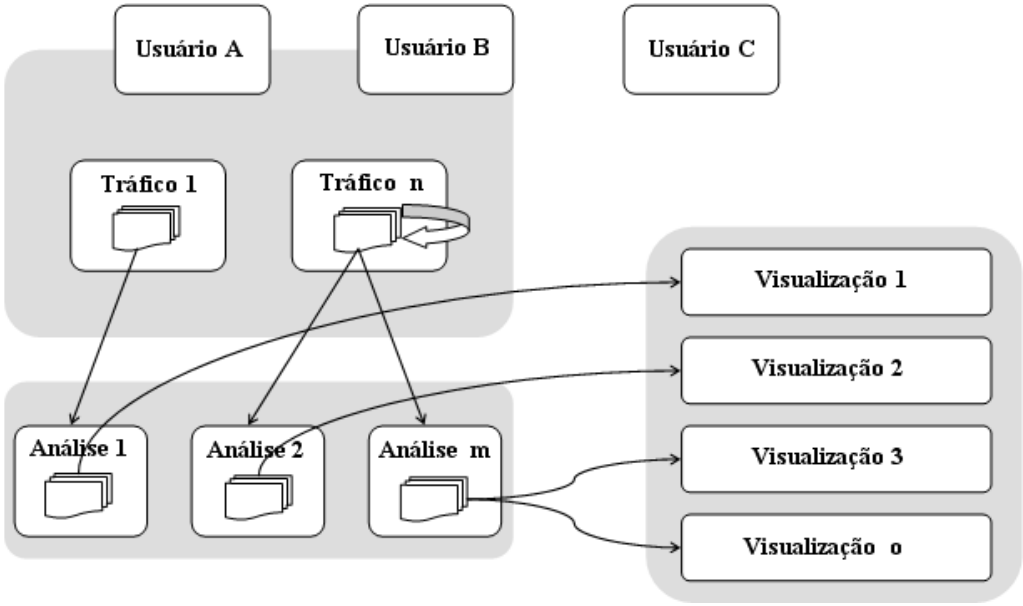


Figura 4.2: Estrutura interna da ferramenta *Management Traffic Analyzer*

Por fim, o usuário terá a opção de realizar comparações entre duas ou mais análises já realizadas. Essa comparação é gerenciada pelo **comparador de resultados de análises**. Uma determinada análise sobre uma amostra de tráfego pode ser comparada com a mesma análise realizada sobre outras amostras de tráfego, pertencentes ao próprio usuário

ou a outras amostras disponíveis no sistema (com o devido consentimento de seus respectivos proprietários). De modo análogo aos casos das análises e visualizações, também é possível se adicionar novas formas de comparação de resultados à ferramenta.

4.1.1 Modelagem da base de dados

Conforme pode ser observado na arquitetura da ferramenta proposta, se faz necessária a utilização de uma base de dados para se armazenar diversas informações necessárias para o funcionamento do *software*. Devido a isso, um modelo de banco de dados também foi desenvolvido, e o mesmo encontra-se representado no diagrama entidade-relacionamento representado na figura 4.3.

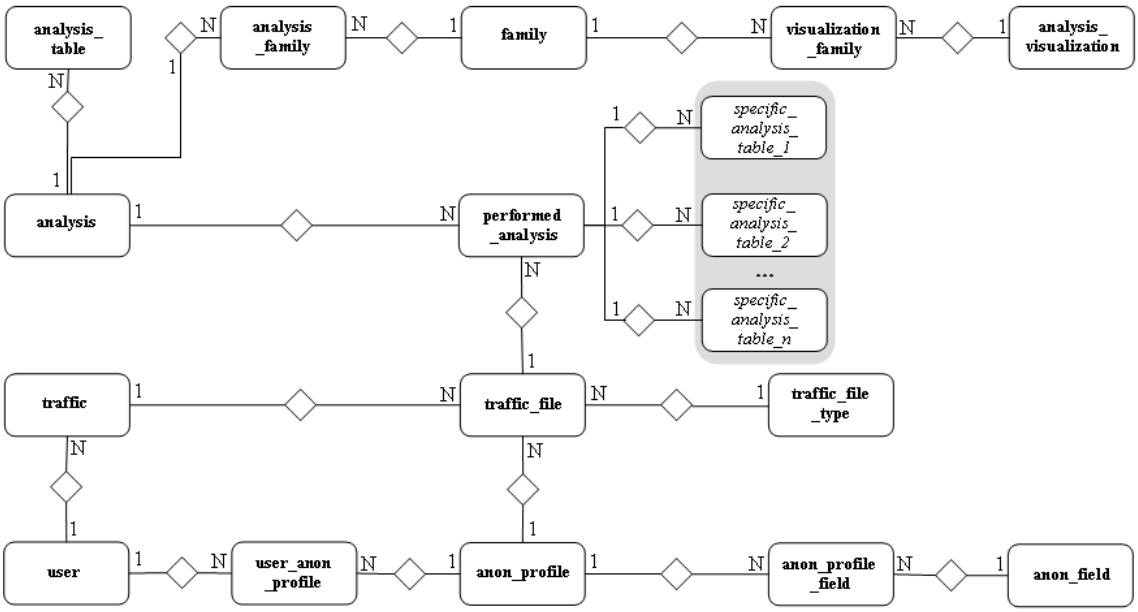


Figura 4.3: Modelo Entidade-Relacionamento para a base de dados da ferramenta proposta

A seguir, apresentaremos de forma resumida a funcionalidade de cada entidade do diagrama da Figura 4.3:

- **user**: armazenar os dados de registro de usuários do sistema;
- **traffic**: armazenar os dados de registro de tráfegos de gerenciamento;
- **traffic_file**: armazenar os arquivos referentes a um determinado tráfego;
- **traffic_file_type**: armazenar informações sobre os formatos de arquivo aceitos pela ferramenta;
- **anon_profile**: armazenar os nomes dos perfis de anonimização disponíveis no sistema;
- **anon_field**: armazenar uma lista de todos os campos que podem ser anonimizados em um tráfego de gerenciamento;
- **user_anon_profile**: relacionar um usuário com os perfis de anonimização que o mesmo pode utilizar;

- **anon_profile_field**: relacionar perfil de anonimização com seus respectivos campos;
- **analysis**: armazenar os dados referentes aos tipos de análises disponíveis no sistema, tais como: nome, descrição e localização no servidor de onde se encontra a parte do código da ferramenta que implementa a referida análise;
- **analysis_table**: armazenar o nome das tabelas que contém os dados resultantes dos tipos de análises disponíveis no sistema;
- **analysis_visualization**: relacionar as análises disponíveis na ferramenta com as visualizações disponíveis para as mesmas;
- **family**: armazenar dados referentes à uma família de análises e visualizações. No contexto desta ferramenta, uma família é definida como um conjunto de análises e visualizações que de alguma forma se relacionam entre si;
- **analysis_family**: relacionar tipos de análises com uma determinada família;
- **visualization_family**: relacionar visualizações disponíveis na ferramenta com uma determinada família;
- **performed_analysis**: armazenar informações sobre todas as análises executadas sobre um determinado tráfego por um usuário;
- ***specific_analysis_table_1, specific_analysis_table_2, ..., specific_analysis_table_n***: essas tabelas estão aqui especificadas de maneira genérica, onde as mesmas representam as tabelas responsáveis por armazenar os dados resultantes de análises específicas. Desta forma, cada análise inserida no sistema deve possuir pelo menos uma tabela relacionada com a tabela **performed_analysis** para receber os dados resultantes do processamento de um arquivo de tráfego.

4.2 Implementação

Como forma de validar a arquitetura proposta neste trabalho, foi implementada uma ferramenta Web de acordo com as especificações apresentadas na seção anterior. A ferramenta, chamada *Management Traffic Analyzer*, automatiza a execução da metodologia de medições sobre tráfego SNMP do IRTF, desde o processo de conversão de arquivos até as análises dos tráfegos submetidos e visualização dos resultados. A Figura 4.4 apresenta uma das telas que compõem a interface gráfica da ferramenta.

Para o desenvolvimento da mesma, foram aproveitadas as funcionalidades da ferramenta SNMPDUMP, que foi desenvolvida em conjunto com a metodologia do IRTF para converter arquivos PCAP para os formatos XML ou CSV e remover ou anonimizar informações sensíveis que possam estar presentes no tráfego. Dessa forma, se faz necessário que o SNMPDUMP esteja instalado no mesmo computador que irá receber a instalação do *Management Traffic Analyzer*. Como o SNMPDUMP é uma ferramenta de interface textual, optou-se por invocar as suas funcionalidades a partir da ferramenta *Management Traffic Analyzer* através de chamadas de sistema. É importante ressaltar que as chamadas ao SNMPDUMP não bloqueiam o funcionamento da ferramenta, permitindo que o usuário faça uso dos recursos dessa ferramenta mesmo durante a utilização do SNMPDUMP.

Figura 4.4: Amostra da interface gráfica da ferramenta *Management Traffic Analyzer*

A ferramenta é capaz de monitorar o estado da execução do SNMPDUMP para manter o usuário ciente da situação do processo de conversão ou anonimização.

A maior parte da ferramenta foi desenvolvida na linguagem PHP. Optou-se pela utilização dessa linguagem devido às suas características de velocidade, robustez, facilidade de conexão com vários bancos de dados, sintaxe similar ao Perl, entre outras. Dentre as partes do software desenvolvidas com PHP, destacam-se:

- Estabelecimento de conexões com a maior parte das tabelas da base de dados, como as de controle de usuários, tráfegos, perfis de anonimização, arquivos, entre outras;
- Invocação da ferramenta SNMPDUMP para conversão e anonimização de arquivos de tráfego;
- Invocação dos *scripts* Perl para análise de arquivos de tráfego;
- Gerenciamento da sessão de utilização da ferramenta por parte do usuário.

A interface gráfica da ferramenta foi implementada utilizando-se a linguagem HTML. Também foram utilizados os recursos da linguagem CSS (*Cascading Style Sheets*), a fim de se obter uma maior separação entre formato e conteúdo, além de possibilitar a criação de detalhes de *design* mais refinados do que se utilizássemos apenas a linguagem HTML.

Seguindo a recomendação da metodologia do IRTF, todos os tipos de análises realizadas pela ferramenta são implementadas em *scripts* escritos na linguagem Perl. Esses *scripts* são invocados pela parte do *software* escrita em PHP, através de chamadas de sistema, e os resultados das análises realizadas são armazenados diretamente na base de dados da ferramenta. De maneira análoga ao que ocorre com o SNMPDUMP, a execução dos *scripts* em Perl não bloqueiam a ferramenta, e esta tem a capacidade de monitorar

o estado da execução do *script*, para manter o usuário informado sobre a situação do processo.

Para a implementação das técnicas de visualização foi utilizada a linguagem Actionscript 2.0 do Macromedia Flash 8. Como se faz necessário acessar o banco de dados para se ter acesso aos resultados das análises, a fim de se criar as visualizações, foi necessária a utilização da extensão PHPObject. Basicamente, essa extensão permite que um script da linguagem Actionscript possa ser integrado com um objeto desenvolvido em PHP. A utilização desse recurso se faz necessária porque o Actionscript 2.0 não possui suporte nativo para se conectar a uma base de dados, e o uso do PHPObject permite que essa conexão seja iniciada por um objeto PHP que é integrado à classe criada no Actionscript.

O módulo comparador de resultados de análise é o único que ainda está em estágio de desenvolvimento, não estando disponível ainda para utilização na ferramenta *Management Traffic Analyzer*. Atualmente estão sendo desenvolvidos os algoritmos para conduzir o *software* na comparação dos resultados das análises que já estão disponibilizadas pela ferramenta. Alguns resultados já foram alcançados, mas os mesmos ainda são muito incipientes para serem citados nesta dissertação de mestrado.

Por fim, foi utilizado o banco de dados MySQL para armazenar todas as informações pertinentes à ferramenta. Foi feita essa escolha para a base de dados devido ao fato de que o MySQL é uma solução *freeware* com ótimo desempenho e que se integra muito bem com a linguagem PHP.

4.2.1 Funcionamento da Ferramenta

Quando um usuário acessa a ferramenta *Management Traffic Analyzer*, a primeira coisa à qual o mesmo terá acesso é a página de boas-vindas e um formulário de autenticação do lado esquerdo da página, conforme mostrado na Figura 4.5. O usuário deverá fornecer um login e uma senha para poder ter acesso às funcionalidades da ferramenta, ou então deverá clicar no link disponibilizado abaixo do formulário de login para criar uma conta. No processo de criação de conta, o usuário deverá informar os seguintes dados: nome, afiliação, país, e-mail, nome de usuário (login) e senha.

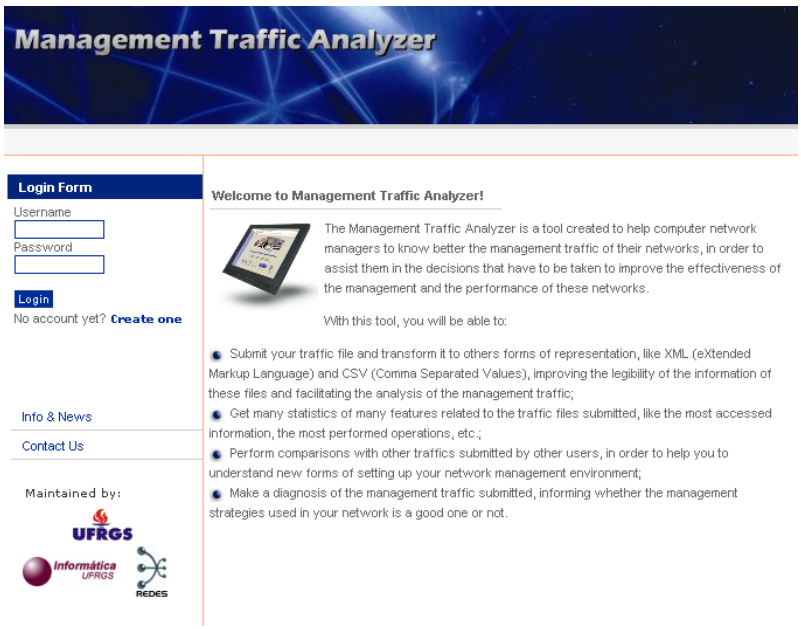


Figura 4.5: Página inicial da ferramenta *Management Traffic Analyzer*

Uma vez autenticado na ferramenta, o usuário irá visualizar uma ”área de trabalho” que é basicamente dividida em duas partes: um menu de acesso às opções do sistema no lado esquerdo, e a visualização da opção que está sendo acessada em um dado momento no lado direito. A primeira opção apresentada a uma pessoa logo após a autenticação da mesma é a ”Your Traffics”, onde apresenta-se a lista de tráfegos já cadastrados na ferramenta disponíveis àquele usuário, assim como botões para se adicionar, editar ou remover tráfegos SNMP. Essa opção também dá acesso a sub-ferramenta para conversão de tráfegos, onde o usuário poderá converter um arquivo no formato PCAP para os formatos XML ou CSV. Além da opção ”Your Traffics”, ainda estão disponibilizados no menu do lado esquerdo da página as opções abaixo, representadas na Figura 4.6:

- **Traffic Statistics**: permite ao usuário realizar, visualizar e remover análises sobre tráfegos SNMP já cadastrados no sistema;
- **Anonymization Profile**: através dessa opção o usuário poderá criar seus perfis de anonimização, ou seja, poderá definir parâmetros para futuras anonimizações de tráfegos SNMP;
- **Info & News**: área reservada para divulgação de informações e notícias relacionadas à ferramenta *Management Traffic Analyzer*;
- **Contact Us**: fornece informações para que o usuário possa entrar em contato com a equipe que mantém o sistema.

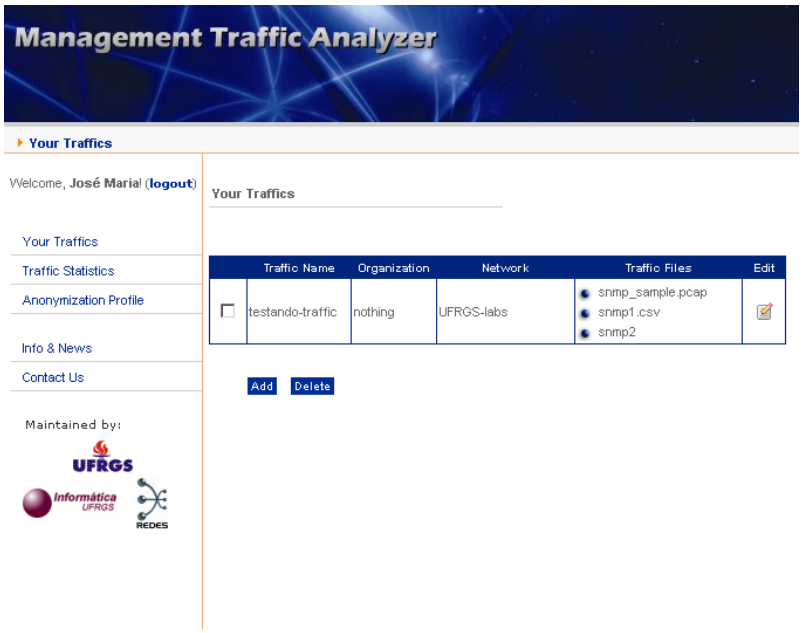


Figura 4.6: Página apresentada após a autenticação do usuário no sistema

Após a realização de uma determinada análise disponibilizada pela ferramenta, o usuário terá opções para a visualização dos dados resultantes do processo. Quando a opção de visualização selecionada faz uso de recursos gráficos, uma determinada região da área de trabalho do usuário fica reservada para a apresentação dessa visualização. Nessa área o usuário poderá explorar todo o gráfico gerado, assim como também poderá interagir com essa visualização de acordo com os recursos de interatividade que a mesma oferecer, conforme representado pela figura 4.7.

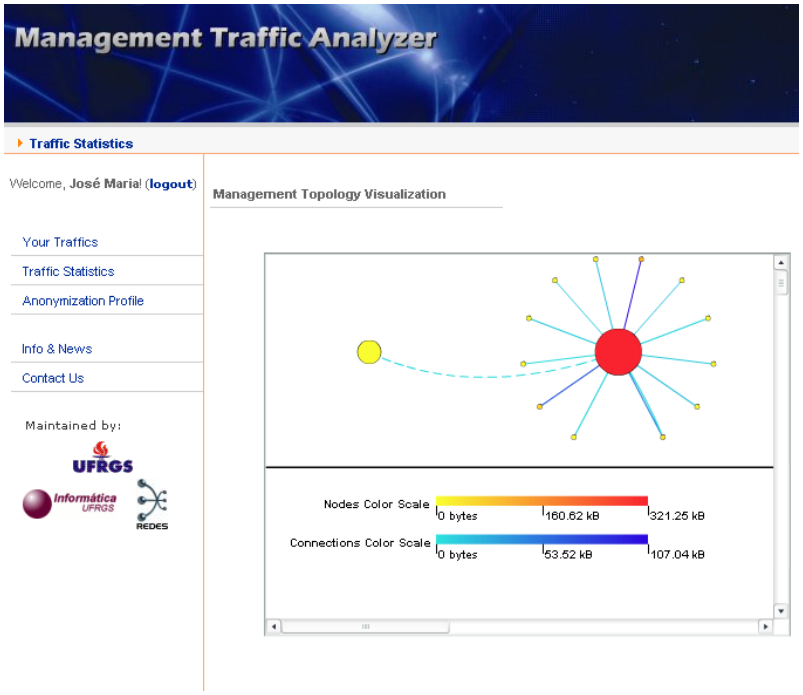


Figura 4.7: Técnica de visualização sendo apresentada pelo sistema

4.3 Avaliação Preliminar da Ferramenta

Com o intuito de se obter uma avaliação da qualidade da ferramenta em seu atual estágio de desenvolvimento, foi conduzida uma pequena pesquisa sobre a usabilidade do *software Management Traffic Analyzer* junto a profissionais e estudantes da área de computação. Essa avaliação, contudo, é de caráter informal e preliminar, tendo em vista que a ferramenta ainda não se encontra em seu estágio final de desenvolvimento e que o emprego de metodologias formais para esse tipo de pesquisa seria inviável devido ao curto espaço de tempo disponível para esse fim. Dessa forma, optou-se pela utilização de questionários para se coletar a opinião de estudantes e profissionais da área de computação de várias cidades do Brasil, e ao mesmo tempo aproveitou-se a oportunidade para se divulgar a ferramenta desenvolvida entre essas pessoas. Após essa coleta, os dados obtidos foram analisados e interpretados, a fim de se inferir o potencial que a ferramenta tem para a área de gerenciamento de redes. É importante também se ressaltar que, como não foram determinados parâmetros para serem comparados aos resultados dessa pesquisa, devido ao seu caráter informal, as conclusões que foram obtidas com as avaliações ainda não constituem evidências irrefutáveis da qualidade da ferramenta. A realização de uma avaliação mais formal, com a utilização de uma metodologia apropriada, se faz necessária no futuro, para que a ferramenta possa ter a sua qualidade atestada de forma mais concreta.

4.3.1 Estrutura do Questionário Aplicado

O questionário utilizado para coletar as opiniões das pessoas consultadas nessa pesquisa foi desenvolvido em forma de páginas web, a fim de permitir que participantes de diversos lugares distintos pudessem ter fácil acesso ao mesmo. Esse questionário foi dividido em 3 partes básicas: dados pessoais, roteiro para utilização da ferramenta e avaliação geral da ferramenta.

A primeira parte do questionário teve por objetivo coletar alguns dados pessoais do

participante, sem que os mesmos fossem identificados, a fim de tornar possível a caracterização da amostra de pessoas pesquisada. Dentre os dados solicitados nessa etapa, pode-se citar idade, sexo, cidade, escolaridade, experiência na área de gerência de redes, entre outros.

Na fase seguinte da pesquisa cada pessoa deveria seguir um roteiro de utilização da ferramenta, que conduziria o voluntário pelas diversas partes do software, desde o cadastro e upload do tráfego a ser pesquisado até a visualização dos resultados. A utilização de um roteiro se fez necessária porque não houve a possibilidade de se treinar o voluntário previamente para o uso da ferramenta. Nessa fase foi fornecido aos usuários uma amostra de tráfego resultante do monitoramento de uma rede experimental gerenciada via SNMP, pertencente ao Instituto de Informática da UFRGS. O monitoramento foi realizado durante um período de 3 horas, resultando na captura de 4373 mensagens SNMP originadas a partir de 14 nós da rede. Além do roteiro, também foi feito um conjunto de 6 perguntas relacionadas aos resultados obtidos com as visualizações geradas pela ferramenta, para avaliar o nível de compreensão do participante e, conseqüentemente, a eficácia da técnica de visualização em si.

Por fim, algumas questões mais gerais sobre a percepção do voluntário quanto à qualidade da ferramenta foram feitas na última fase do questionário. Dentre os itens avaliados nessa fase, estão: facilidade de interpretação das visualizações, facilidade de uso da ferramenta, avaliação da interface gráfica, avaliação do potencial de utilização por pesquisadores e administradores de redes, entre outras. Também foi inserido um campo de comentários, para que os participantes pudessem fazer qualquer observação que eles julgassem pertinentes à pesquisa e que não teria sido contemplada nas demais questões da avaliação. O questionário completo está apresentado no Apêndice A.

4.3.2 Perfil das pessoas consultadas na pesquisa

Para a realização da avaliação da ferramenta *Management Traffic Analyzer* foram enviados e-mails solicitando a contribuição de diversas pessoas que trabalham ou estudam na área de computação. Ao todo 21 pessoas puderam responder ao questionário, onde 18 pessoas são do sexo masculino e 3 do sexo feminino. Já a faixa etária variou dos 23 aos 56 anos de idade. Os voluntários que participaram dessa pesquisa são de 6 locais distintos, conforme apresentado na Figura 4.8.

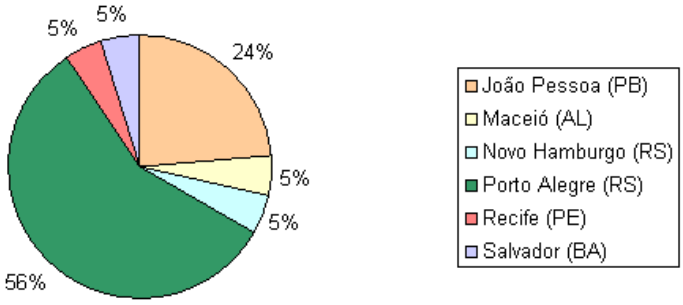


Figura 4.8: Locais de residência dos participantes da pesquisa

A amostra pesquisada contou com representantes de todos os níveis de escolaridade, desde o 2º grau até o doutorado. A distribuição dos participantes da amostra pelos níveis de escolaridade está representada na Figura 4.9. Com relação às funções desempenhadas por essas pessoas, encontram-se na amostra: vários estudantes, vários professores, dois

programadores, um analista de sistemas, um gerente de projetos e um administrador de rede.

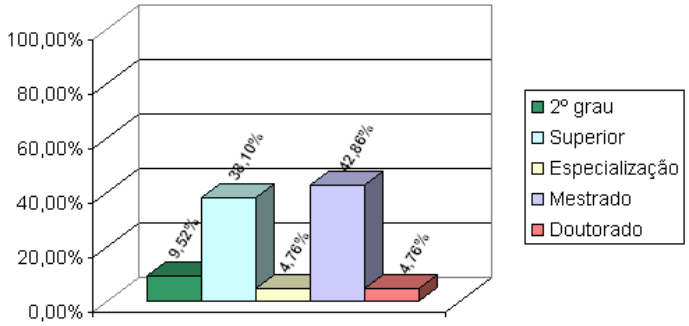


Figura 4.9: Níveis de escolaridade dos participantes da pesquisa

Com relação ao tempo de trabalho na área de computação, a amostra pôde contar com pessoas pertencentes às várias categorias que foram listadas no formulário de pesquisa, conforme mostra a Figura 4.10. Fato semelhante ocorre no que diz respeito à atuação na área de gerenciamento de redes dos voluntários pesquisados, conforme representado na Figura 4.11.

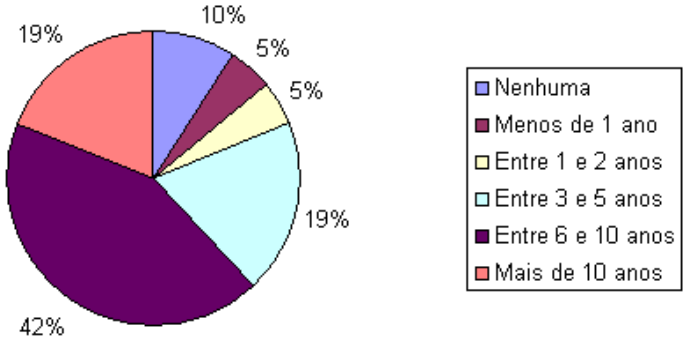


Figura 4.10: Classificação da amostra de acordo com o tempo de trabalho na área de computação

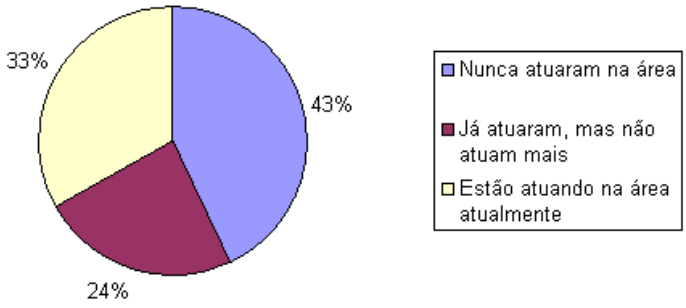


Figura 4.11: Classificação da amostra de acordo com a atuação na área de gerenciamento de redes

Por fim, a classificação da amostra de acordo com o nível de conhecimento sobre o SNMP também foi relativamente equilibrada, contando com representantes em todas as categorias listadas no questionário. Contudo, nota-se que a maioria das pessoas pesquisadas possuem conhecimento básico ou intermediário sobre esse protocolo, conforme pode ser observado na Figura 4.12.

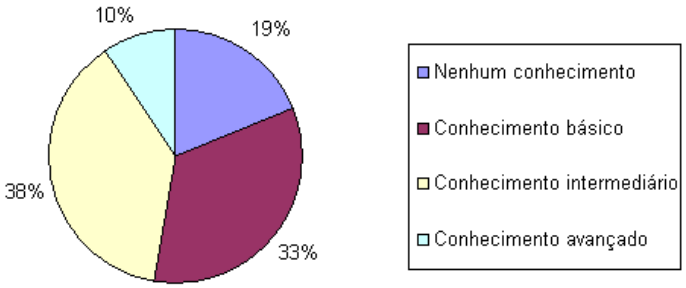


Figura 4.12: Nível de conhecimento da amostra sobre o protocolo SNMP

4.3.3 Resultado da Aplicação dos Questionários

Logo após a conclusão do roteiro de utilização da ferramenta pelos participantes da pesquisa, foi feita uma série de questões de avaliação da ferramenta para os mesmos. A primeira questão que eles responderam perguntava se os participantes haviam compreendido bem a finalidade da ferramenta. Essa pergunta é relevante para a pesquisa porque a ferramenta está inserida num contexto bastante específico da área de gerenciamento de redes, que é a metodologia do IRTF para medições de tráfego SNMP. Por esse motivo, os objetivos da ferramenta podem não ser muito claros para uma pessoa à primeira vista, especialmente se ela não tiver relação com a área de gerenciamento de redes, e a não compreensão da finalidade da ferramenta pode afetar praticamente todos os outros aspectos da avaliação que está sendo realizada.

Os resultados apresentados na Figura 4.13 motraram-se satisfatórios, uma vez que mais de 76% da amostra pesquisada afirmou ter compreendido a finalidade da ferramenta. Outros 24% afirmaram possuir apenas uma vaga idéia sobre os objetivos do *software*, enquanto que nenhum candidato afirmou não ter compreendido nada da finalidade da ferramenta. Como cada pessoa que participou da pesquisa teve acesso à uma breve descrição da ferramenta antes de iniciar os trabalhos de avaliação, essa parcela de 24% pode ser justificada pelos seguintes motivos:

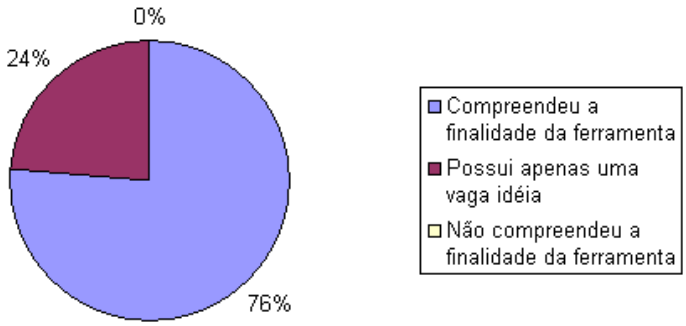


Figura 4.13: Nível de compreensão quanto à finalidade da ferramenta

- Completa falta de afinidade com a área de gerenciamento de redes, a ponto de não ter sido possível se compreender conceitos básicos utilizados na descrição da ferramenta;

- Falta de atenção com o texto inicial da pesquisa, devido à pressa ou à confiança de que se conseguiria compreender o objetivo da ferramenta ao se utilizá-la.

A segunda questão de avaliação disse respeito à facilidade com que os participantes tiveram em utilizar os recursos da ferramenta. Os resultados obtidos podem ser observados na Figura 4.14.

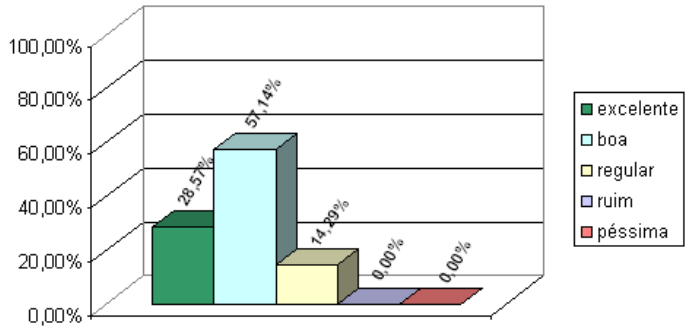


Figura 4.14: Avaliação da facilidade de uso da ferramenta

Essa avaliação pode ser considerada satisfatória, uma vez que mais de 80% da amostra pesquisada avaliou a facilidade de uso da ferramenta como boa ou excelente. Apenas 14,29% das pessoas pesquisadas avaliaram a ferramenta como regular, e nenhuma pessoa avaliou a ferramenta como ruim ou péssima.

O próximo item avaliado foi a qualidade visual da interface gráfica da ferramenta. Conforme mostra a Figura 4.15, 23,81% da amostra considerou a qualidade da interface gráfica excelente, enquanto que os outros 76,19% a consideraram boa. Não houve registro de pessoas que avaliassem a interface como regular, ruim ou péssima.

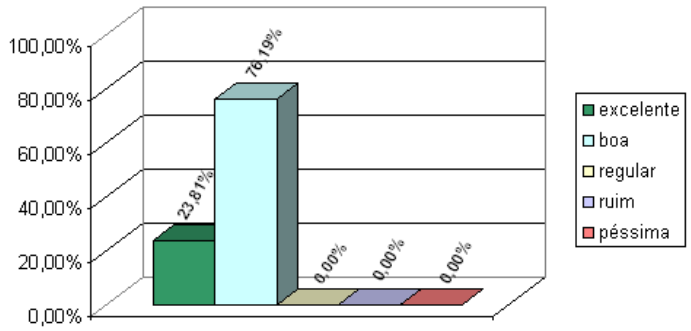


Figura 4.15: Avaliação da interface gráfica da ferramenta

O quarto item avaliado foi a facilidade de compreensão das visualizações geradas pela ferramenta. A finalidade dessa questão foi verificar se os mapeamentos visuais são compreensíveis para as pessoas, uma vez que estas tenham recebido ao menos instruções básicas sobre as técnicas de visualização que seriam utilizadas. Os resultados dessa avaliação são apresentados na Figura 4.16.

Como se pode observar, a maioria dos participantes (43%) declarou ter encontrado muita facilidade na compreensão das visualizações utilizadas pela ferramenta, enquanto que 38% declarou ter encontrado relativa facilidade. Contudo, 19% das pessoas pesquisadas afirmaram ter sentido um pouco de dificuldade em compreender as visualizações. Nenhum participante alegou ter sentido muita dificuldade no processo de compreensão das visualizações.

Um outro item do questionário que buscou avaliar as visualizações da ferramenta foi o conjunto de 6 perguntas contidas no roteiro de utilização do *Management Traffic Analyzer*.

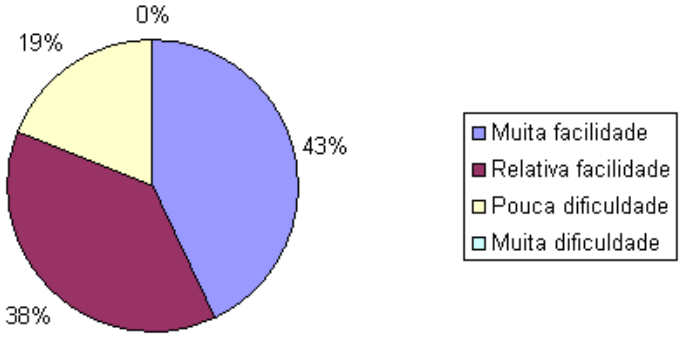


Figura 4.16: Facilidade de compreensão das visualizações da ferramenta

Cada par de perguntas se referia a uma das 3 técnicas de visualização empregadas, ou seja, as questões A e B se referiam à técnica ”Visualização de objetos SNMP em uma *MIB Tree*”, as questões C e D se referiam à técnica ”Visualização da quantidade de mensagens SNMP em intervalos de 1 hora” e as questões E e F se referiam à técnica ”Visualização da Topologia da Rede de Gerenciamento”. O objetivo dessas questões foi avaliar de forma prática, e não baseada apenas no que declaravam os participantes da pesquisa, o nível de compreensão dos dados condificados nas referidas visualizações. As questões respondidas pelas pessoas que participaram da avaliação foram as seguintes:

- **Questão A**: Qual o objeto SNMP mais acessado do tráfego?
- **Questão B**: Qual é a MIB menos utilizada no tráfego?
- **Questão C**: Qual é a hora do dia em que é registrado o maior volume de mensagens SNMPv1?
- **Questão D**: Qual é a operação do protocolo SNMP mais utilizada na rede?
- **Questão E**: Qual o volume de tráfego concentrado pelo principal gerente da rede?
- **Questão F**: Quais os endereços IPs do par de máquinas que mais trocou informações no tráfego?

Apesar do fato de que as respostas para essas questões podem ser identificadas nas visualizações geradas com relativa facilidade, elas possuem um grau de complexidade razoável quando se considera a obtenção dessas informações sem a utilização de técnicas de visualização. A taxa de acertos por parte dos participantes da pesquisa para cada uma dessas questões é apresentada na Figura 4.17.

Observa-se que a maioria das questões tiveram uma alta taxa de acerto (acima de 80%), enquanto que nas questões B e F essa taxa de acerto é menor, apesar de que a maioria dos participantes pôde responder corretamente às mesmas. Levando-se em consideração que os participantes conseguiram essas taxas de acerto sem possuírem nenhum treinamento prévio com a ferramenta, e que na ocasião da pesquisa aquele era o primeiro contato dessas pessoas com a ferramenta avaliada, essas taxas de acerto podem ser consideradas satisfatórias.

Também buscou-se avaliar a eficiência da ferramenta no processo de análise do tráfego SNMP utilizado na pesquisa, através do tempo compreendido entre o momento em que o usuário acessa o roteiro que o guiará na utilização da ferramenta (início da utilização da

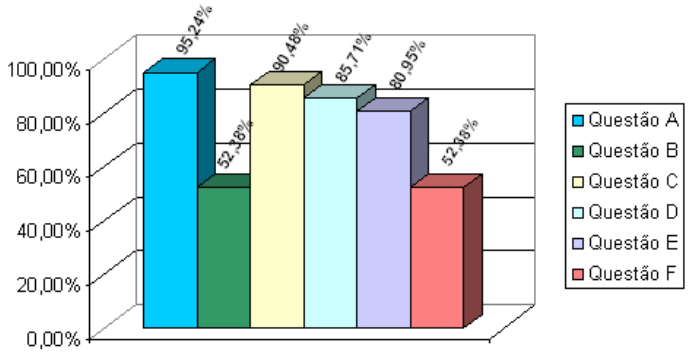


Figura 4.17: Taxa de acertos das questões sobre as visualizações geradas durante a avaliação

ferramenta) e o momento em que o usuário submete as respostas para todas as questões envolvendo as visualizações que foram apresentadas no roteiro (fim da utilização da ferramenta). Dessa forma, a média do tempo gasto se utilizando a ferramenta para efetuar as análises pedidas pelo roteiro , dentre 20 dos 21 participantes, foi de 26,9 minutos. Apenas o tempo de um dos participantes foi excluído dessa média, por ter sido excessivamente pequeno (apenas 1 minuto). Isso provavelmente aconteceu porque o participante deve ter feito a avaliação em dois momentos: primeiramente ele deve ter visto o roteiro para fazer a prática, e num segundo momento ele acessou novamente o roteiro apenas para submeter os resultados. Levando-se em consideração a complexidade do tráfego e das questões, o fato de que esse tempo também envolve a leitura e compreensão do roteiro de utilização da ferramenta, e a inexperiência em gerenciamento de redes de vários dos participantes, a média de tempo registrada pode ser considerada bastante satisfatória para esse tipo de análise.

O item seguinte do questionário procurou saber dos participantes se os mesmos acreditavam que a ferramenta pudesse ser útil para auxiliar pesquisadores a compreenderem melhor a utilização do protocolo SNMP. Todos os participantes afirmaram acreditar que de fato a ferramenta poderá auxiliar nos trabalhos desses pesquisadores, conforme mostrado na Figura 4.18.

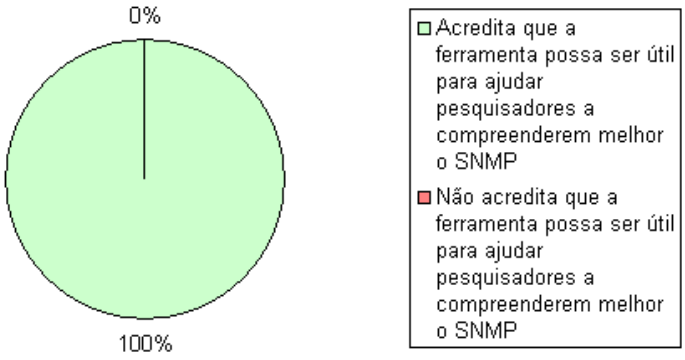


Figura 4.18: Avaliação quanto à utilidade da ferramenta para pesquisadores

Os participantes também foram perguntados pelo questionário se os mesmos acreditavam que a ferramenta pudesse ser útil para facilitar algumas das tarefas desempenhadas por administradores de redes. Novamente, 100% dos participantes afirmaram acreditar no fato de que a ferramenta realmente pode ser útil para o trabalho dos administradores de

redes, conforme mostrado na Figura 4.19.

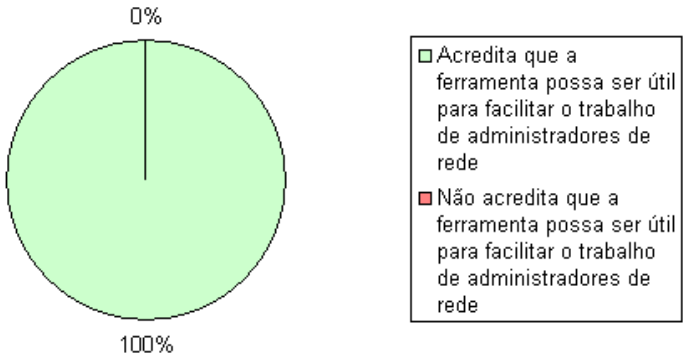


Figura 4.19: Avaliação quanto à utilidade da ferramenta para administradores de redes

Ao final do questionário, foi solicitado à cada participante da pesquisa a sua avaliação geral quanto à ferramenta *Management Traffic Analyzer*. Os resultados dessa avaliação estão representados na Figura 4.20.

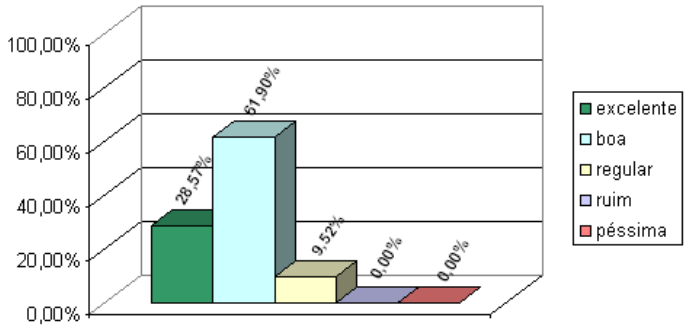


Figura 4.20: Avaliação geral da ferramenta *Management Traffic Analyzer*

Conforme se pode observar, uma boa parte dos participantes (28,57%) consideraram a ferramenta excelente, enquanto que a maioria (61,90%) avaliaram a ferramenta como boa. 9,52% dos participantes consideraram a ferramenta regular, e nenhum participante avaliou a ferramenta como ruim ou péssima. Mais uma vez, esses resultados podem ser considerados bastante satisfatórios para a ferramenta *Management Traffic Analyzer*.

A ferramenta apresentada neste capítulo já foi utilizada na prática para a produção de resultados gerados à partir de medições sobre tráfegos SNMP de redes em produção. O próximo capítulo apresenta os resultados fornecidos pela ferramenta *Management Traffic Analyzer* após a análise de amostras de tráfego SNMP de duas grandes redes brasileiras.

5 RESULTADOS DE MEDIÇÕES SOBRE TRÁFEGOS SNMP

Visando-se validar as visualizações apresentadas nesta dissertação de mestrado, e também procurando contribuir com a comunidade de gerenciamento de redes para a descoberta de padrões de utilização do *Simple Network Management Protocol*, submeteu-se amostras de tráfego SNMP de duas grandes redes brasileiras às análises da ferramenta *Management Traffic Analyzer*. Essas amostras foram obtidas pelos próprios operadores dessas redes, através da realização de seções de monitoramento do tráfego de gerenciamento por pelo menos uma semana, conforme recomendação da metodologia do IRTF.

A primeira rede que foi estudada utilizando-se a metodologia do IRTF através da ferramenta *Management Traffic Analyzer* foi a do ponto de presença da Rede Nacional de Pesquisa (RNP) no estado do Rio Grande do Sul. Posteriormente, o autor desta dissertação obteve amostras de tráfego de uma outra rede ainda maior: a da própria RNP. Os resultados obtidos através da análise desses tráfegos serão descritos nas próximas seções deste capítulo.

5.1 Tráfego SNMP do POP da RNP no Rio Grande do Sul

O tráfego SNMP estudado nessa seção foi fornecido pelo ponto de presença (POP) da RNP no Rio Grande do Sul (POP-RS). Essa rede fornece acesso à Internet para instituições de ensino públicas e privadas, além de centros de pesquisa, no estado do Rio Grande do Sul. Os operadores dessa rede forneceram um arquivo de tráfego já anonimizado e no formato CSV, o qual foi produto do monitoramento do tráfego SNMP realizado entre os dias 10 e 19 de julho de 2006. Os resultados das análises e visualizações sobre o tráfego do POP-RS da RNP serão discutidos nas subseções a seguir.

5.1.1 Topologia da Rede de Gerenciamento

A primeira análise realizada sobre o tráfego do POP-RS foi a da topologia da rede de gerenciamento, conforme mostra a Figura 5.1. Como se pode observar na visualização gerada, existe um gerente que possui grande destaque na rede analisada. Esse gerente está representado por um círculo vermelho, devido ao fato de o mesmo possuir a maior carga de tráfego de toda a rede (145,74MB). É importante se observar que a maioria dos terminais gerenciados na rede (agentes) são monitorados por esse gerente. Outro fato que merece destaque é a existência de 4 outros nós na rede que também assumem papéis de gerente, apesar da pequena quantidade de mensagens relacionadas a esses nós. Dentre esses 4 nós, aquele que mais chama à atenção é o que está representado logo abaixo do gerente principal na topologia visualizada: apesar desse nodo ser monitorado pelo gerente principal da rede, existe uma quantidade considerável de mensagens SNMP partindo dele

para solicitar informações do gerente principal. Ao se investigar mais detalhadamente o arquivo de tráfego no formato CSV, identificou-se que o referido nodo estava realizando *polling* no gerente principal, solicitando informações de dois objetos SNMP: o escalar `SNMPv2-MIB::sysUpTime` e a tabela `UCD-SNMP-MIB::IaLoadInt`. Muito provavelmente esses dois objetos estão sendo acessados para verificar se o gerente principal está em operação, através do objeto `SNMPv2-MIB::sysUpTime`, e qual a carga de tráfego desse gerente, através do objeto `UCD-SNMP-MIB::IaLoadInt`.

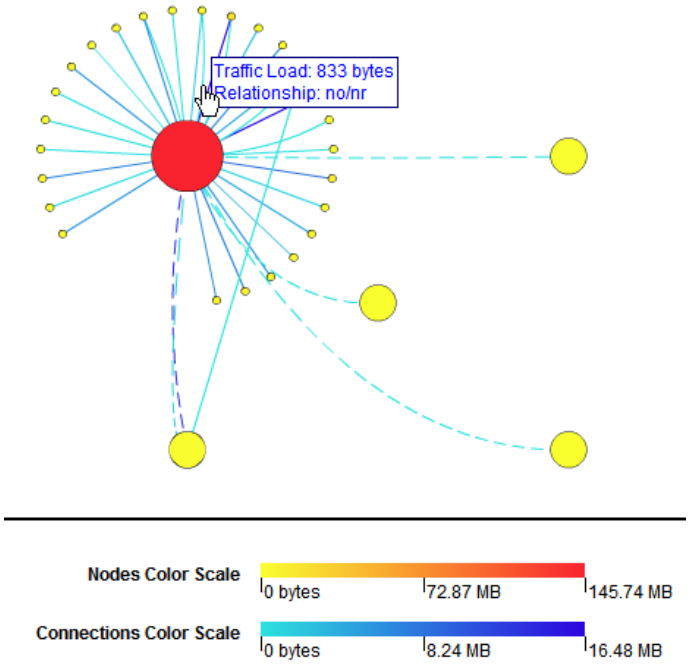


Figura 5.1: Topologia da rede de gerenciamento do POP-RS

A análise das conexões entre os nodos da visualização indica que a maior quantidade de *bytes* transferidos ocorre entre o nodo do gerente principal da rede e um dos agentes gerenciados por ele (16,48 MB). O relacionamento representado por essa conexão, assim como a maioria dos demais relacionamentos, é do tipo CG/CR (Command Generator/Command Responder). Apenas 6 conexões pertencem à relacionamentos do tipo NO/NR, indicando que na rede estudada existe uma predominância do processo de *polling* sobre o processo de notificações (*traps*).

5.1.2 Objetos SNMP Utilizados

Após a execução da análise dos objetos SNMP encontrados no tráfego do POP-RS, obteve-se uma *MIB tree* contendo 128 nós folhas representando os objetos encontrados. Uma vez que essa árvore é muito grande para ser representada nesta dissertação de mestrado, aplicou-se a técnica de visualização para os 25 objetos mais representativos do tráfego. A visualização gerada ao final desse processo está representada na Figura 5.2.

Ao se analizar a visualização da *MIB tree* nota-se que os 3 objetos SNMP mais utilizados são: HOST-RESOURCES-MIB::hrSWRunPerfMem (1.200.353 mensages), IF-MIB::ifOutOctets (424.157 mensagens) e IF-MIB::ifInOctets (422.998 mensagens). Este fato indica que a preocupação mais proeminente dos administradores da rede é a monitoração do uso da memória dos dispositivos, e também o acompanhamento do volume de tráfego da rede. Por outro lado, também existem objetos relacionados ao monitoramento de

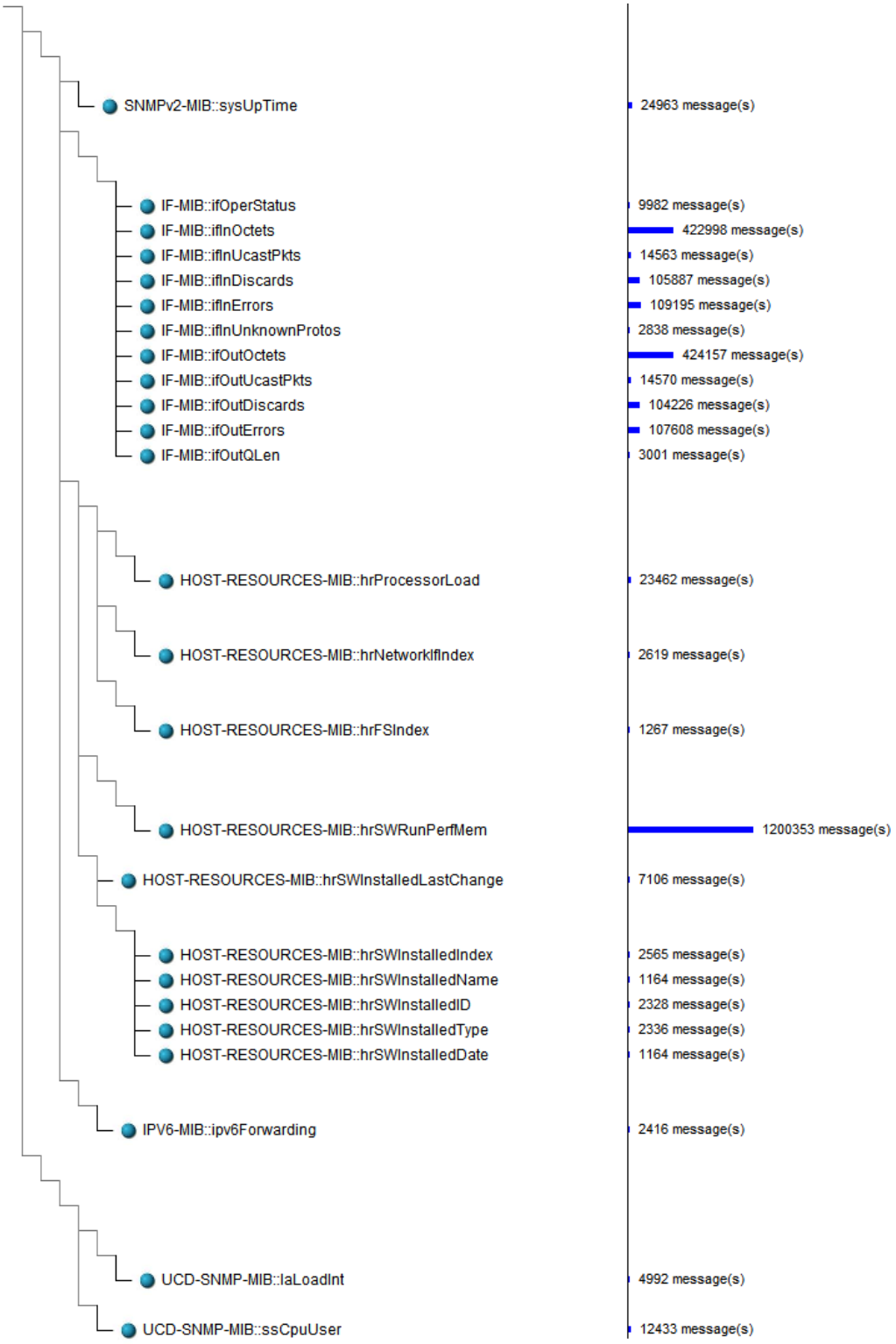


Figura 5.2: *Mib-tree* da rede de gerenciamento do POP-RS

erros na rede que são intensamente utilizados, tais como: IF-MIB::ifOutErrors (107.608 mensagens), IF-MIB::ifInDiscards (105.887 mensagens) e IF-MIB::ifOutDiscards (104.226 mensagens). A visualização também deixa claro que as MIBs mais utilizadas são a IF-MIB e a HOST-RESOURCES-MIB. O uso de outras MIBs é bem menos significativo no tráfego estudado.

5.1.3 Número de Mensagens SNMP por Intervalos de 1 Hora

Como o tráfego do POP-RS é resultante de uma monitoração das mensagens SNMP ao longo de 10 dias, obteve-se com a aplicação da técnica de visualização do número de mensagens SNMP por intervalos de 1 hora o total de 10 pares de histogramas. Cada par de histogramas é formado por um histograma seccionado por versões e um histograma seccionado por operações do protocolo SNMP. Contudo, para esta dissertação de mestrado foram selecionados os 3 pares de histogramas mais representativos da amostra, os quais podem ser vistos na Figura 5.3.

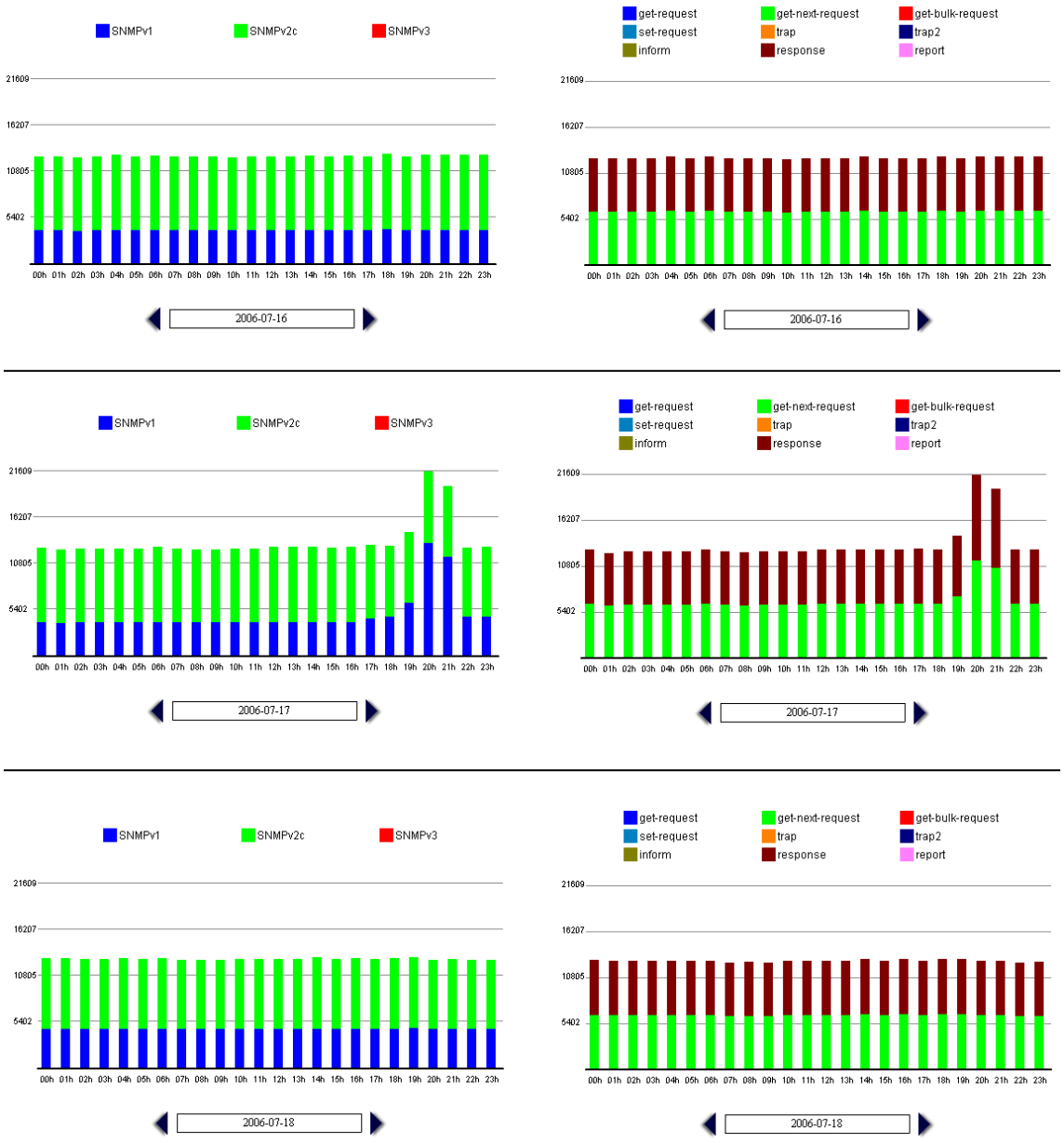


Figura 5.3: Número de mensagens SNMP por hora da rede de gerenciamento do POP-RS

A análise das visualizações geradas indica que a quantidade de mensagens SNMP por intervalos de 1 hora é praticamente constante em toda a amostra de tráfego. Este fato mostra a preferência dos administradores de rede pelo uso de *polling*. Outra característica da visualização que demonstra essa preferência é o fato de que o tráfego é formado quase que em sua totalidade por mensagens `get-next-request` e suas respectivas

respostas (`get-response`). O número de mensagens de `get-requests` e `traps` é muito pequeno, sendo este o motivo dessas operações não terem aparecido nos histogramas gerados pela ferramenta. O fato de não terem sido registradas mensagens do tipo `set-request` demonstra que na rede estudada o SNMP está sendo usado apenas para fins de monitoramento, e não de configuração de dispositivos. Observa-se também que a amostra estudada é composta por mensagens pertencentes às versões SNMPv1 e SNMPv2c, esta última sendo maioria no tráfego. Não foram registradas mensagens da versão SNMPv3.

Uma outra característica das visualizações geradas que chama à atenção é uma espécie de anomalia entre as 19:00h e 21:59h do dia 17 de julho de 2006: a quantidade de mensagens SNMP aumenta significativamente nesse período. Esta anomalia pode ser observada no segundo par de histogramas da Figura 5.3. Investigações realizadas nos arquivos de tráfego no formato CSV acerca desse fenômeno revelaram que um conjunto de operações `get-next-request` sobre objetos da HOST-RESOURCES-MIB foram efetuadas em vários agentes da rede, resultando no crescimento observado na visualização.

Após as considerações sobre as visualizações geradas a partir do tráfego do POP-RS, conclui-se que o tráfego SNMP estudado é formado predominantemente por mensagens de *polling* geradas pelos sistemas de gerenciamento empregados na rede, em detrimento da utilização de sistemas de notificação (e.g., `traps`). Também no caso dessa rede observa-se que o SNMP é utilizado apenas para monitoração dos dispositivos gerenciados, e nunca para configuração dos mesmos.

5.2 Tráfego SNMP da RNP

Nesta seção serão apresentados os resultados das análises das amostras do tráfego de gerenciamento da Rede Nacional de Pesquisa (RNP). Essa rede provê serviço Internet para diversas instituições de ensino e pesquisa em todo o país, com facilidades de trânsito nacional e internacional, em uma infra-estrutura com alta largura de banda e suporte a aplicações avançadas. Essas amostras consistem num conjunto de 13 arquivos anonimizados, no formato CSV, com capturas de tráfegos realizadas entre os dias 22 de junho e 5 de julho de 2007.

5.2.1 Topologia da Rede de Gerenciamento

Ao se observar a topologia da rede de gerenciamento encontrada nos 13 arquivos de tráfego SNMP analisados, percebe-se que a mesma permanece relativamente invariável no intervalo de tempo onde esse tráfego foi monitorado. Tomando como exemplo a topologia observada no dia 22 de junho de 2007, identificou-se que a rede possui um gerente principal que gerencia a maior quantidade de nós, e que também é o terminal que concentra a mais alta carga de tráfego (148,26 MB). Existe também um segundo nó que atua como um gerente menos importante da rede. A quantidade de nós conectados a esse gerente é bem menor do que a quantidade de nós conectados ao gerente principal, assim como a carga de tráfego nesse nó também é inferior (997,46 KB). Destaca-se ainda a existência de dois nós que gerenciam exclusiva e simultaneamente um único nó na rede.

A diferença mais perceptível que pode ser encontrada analisando-se os outros arquivos de tráfego SNMP da RNP é a aparição de um grupo de terminais formado por 1 nó gerente e cerca de 6 nós agentes conectados a esse gerente. Esse grupo aparece pela primeira vez no arquivo de tráfego relativo ao dia 27 de junho de 2007. A Figura 5.4 apresenta a visualização resultante das análises dos tráfegos da RNP relativos aos dias 22 e 27 de

junho de 2007 (da esquerda para a direita, respectivamente).

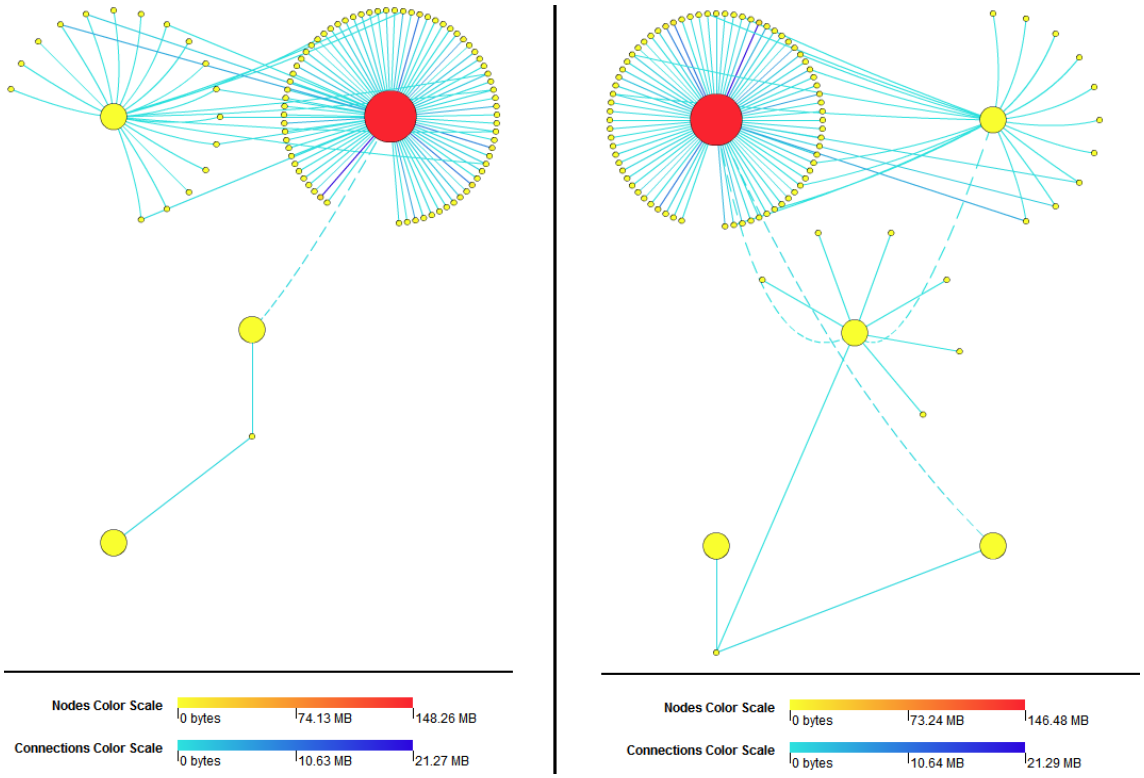


Figura 5.4: Topologias da rede de gerenciamento da RNP

Com relação às conexões representadas na topologia, percebe-se que aquelas onde passa a maior quantidade de tráfego são as que conectam o principal nó gerente da rede com determinados nós agentes. A maior quantidade de tráfego registrada em uma conexão foi de 21,29 MB. No geral, cerca de 80% das conexões são do tipo CG/CR, enquanto cerca de 20% são do tipo NO/NR. Isso mostra uma preferência por parte dos administradores de rede em realizar *polling* nos dispositivos gerenciados, ao invés de utilizar notificações (*traps*).

5.2.2 Objetos SNMP Utilizados

Após a análise para identificação dos objetos SNMP encontrados nos tráfegos ter sido executada, observou-se que os objetos e a quantidade de mensagens associadas a estes são relativamente invariáveis nos arquivos de tráfego estudados. Devido a essa constância, será apresentada nesta dissertação de mestrado a árvore de objetos SNMP (*MIB Tree*) de apenas um dos arquivos de tráfego. Além disso, reduzimos o tamanho dessa árvore de modo a exibir apenas os 25 objetos SNMP mais representativos. O resultado dessa visualização pode ser observado na Figura 5.5.

A observação da árvore da Figura 5.5 mostra que os objetos SNMP mais utilizados são: IF-MIB::ifDescr (250.241 mensagens) e IF-MIB::ifType (245.464 mensagens). Provavelmente esses objetos são acessados para se verificar se o dispositivo consultado está operacional ou não, o que indicaria uma grande preocupação em testar o funcionamento dos dispositivos gerenciados da rede. Outros objetos que são intensivamente utilizados fornecem dados sobre o tráfego da rede. São eles: IF-MIB::ifOutUcastPkts (217.491 mensagens), ifInUcastPkts (217.479 mensagens), IF-MIB::ifHCOutOctets (203.737 mensagens) e IF-MIB::ifHCInOctets (203.730 mensagens). Muito provavelmente algum *soft-*

ware de monitoramento de rede como o MRTG é o responsável pelos *pollings* feitos sobre esses objetos.

Figura 5.5: Os 25 objetos SNMP mais acessados na rede da RNP

Por fim, é possível observar na *MIB tree* que os três últimos objetos listados estão representados pelos seus OIDs, o que indica que a ferramenta não foi capaz de identificar os nomes desses objetos representados. A partir de buscas na Internet constatou-se que os objetos "1.3.6.1.4.1.9.2.1.57.0" e "1.3.6.1.4.1.9.2.1.58.0" pertencem à uma MIB proprietária da empresa *Cisco Systems Inc.*, e os mesmos tem por finalidade monitorar o uso da CPU de dispositivos de rede fabricados pela própria Cisco, como roteadores. Já o objeto "1.3.6.1.4.1.2636.3.1.13.1.8.9.1.0.0" também pertence à uma MIB proprietária chamada JUNIPER-MIB, da empresa *Juniper Networks*. A finalidade desse objeto também é monitorar o uso da CPU de roteadores fabricados pela *Juniper Networks*.

5.2.3 Número de Mensagens SNMP por Intervalos de 1 Hora

De modo análogo à análise dos objetos SNMP presentes no tráfego (discutida na subseção anterior), a análise da distribuição da quantidade de mensagens SNMP em intervalos de 1 hora também se mostrou relativamente constante nos arquivos de tráfego estudados. Portanto, mais uma vez serão representados os histogramas gerados a partir de apenas um dos arquivos fornecidos, correspondente ao monitoramento na RNP realizado no dia 28 de junho de 2007. Os histogramas resultantes dessa visualização podem ser vistos na Figura 5.6, seccionados por versões e por operações do protocolo SNMP (de cima para baixo, respectivamente).

A quantidade de mensagens é praticamente constante ao longo das horas completas em que houve monitoramento de tráfego. A última barra do histograma é menor devido à interrupção do processo de monitoração do tráfego de gerenciamento, que ocorreu em meados das 22h. O maior tráfego SNMP na rede registrado em 1 hora é de 101185 mensagens, observado às 6h do dia 28 de junho de 2007. Essa constância observada na quantidade de mensagens em todas as horas completas onde houve monitoramento mostra uma preferência dos sistemas de gerenciamento por consultas (*polling*) realizadas periodicamente nos dispositivos. Isso é confirmado também pelo grande número de mensagens do tipo *get-request*, *get-next-request* e *get-bulk-request*.

O histograma seccionado de acordo com as versões encontradas do protocolo SNMP mostram que o SNMPv2c é a versão predominante no tráfego. Em menor quantidade, encontra-se mensagens SNMPv1. Não foram identificadas mensagens SNMPv3 nos arquivos analisados.

Já o histograma seccionado de acordo com as operações do SNMP mostra que o tráfego é formado em sua maior parte por mensagens `response`. Em seguida aparecem as mensagens `get-request` e `get-next-request`, respectivamente. Em menor quantidade temos mensagens `get-bulk-request` e `trap2`. Mensagens de `trap` podem ser encontradas no tráfego, mas em quantidade insignificante, sendo esse o motivo de elas não serem representadas no histograma. O fato de mensagens `response` serem maioria no tráfego é justificável, pois esse tipo de mensagem sempre é emitida quando um dispositivo responde à consultas `get-request`, `get-next-request` ou `get-bulk-request`. Não foram observadas mensagens `set-request`, o que indica que o protocolo está sendo utilizado exclusivamente para monitoramento da rede, e não para configuração da mesma.

De um modo geral, podemos concluir que o tráfego SNMP da RNP é formado quase que em sua totalidade por mensagens periódicas de *polling* geradas pelos sistemas de gerenciamento utilizados na rede. O uso de notificações é raro, e praticamente não exerce influência sobre o tráfego como um todo. Por fim, o SNMP é utilizado exclusivamente para a realização do monitoramento sobre os dispositivos e serviços da rede. Muito pro-

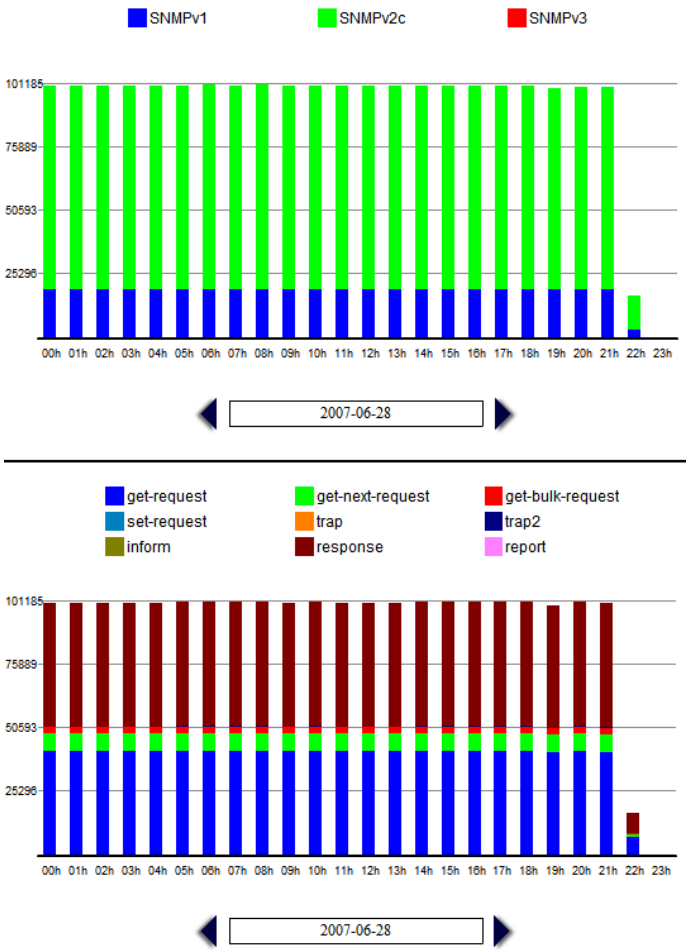


Figura 5.6: Histogramas do tráfego SNMP na RNP

vavelmente esse protocolo não é utilizado para configuração de dispositivos por questões de segurança.

6 CONCLUSÕES E TRABALHOS FUTUROS

Atualmente, várias abordagens de gerenciamento de redes de computadores estão bem consolidadas, tais como as que empregam o protocolo SNMP, e algumas outras já se encontram em um estado avançado de desenvolvimento, como as que fazem uso de *Web Services*. Entretanto, a indústria e a academia ainda não conhecem de maneira satisfatória as características da utilização dessas abordagens no "mundo real", ou seja, nas redes em produção. Nesse contexto, a metodologia do IRTF surge para orientar e incentivar o processo de análises de tráfegos de gerenciamento, a fim de se identificar as características da utilização do protocolo SNMP. O descobrimento desses padrões de uso é importante tanto para administradores de rede, que poderão aproveitar esses dados para fazerem um melhor planejamento de sua infra-estrutura de gerenciamento da rede, quanto para fabricantes e pesquisadores, que poderão desenvolver soluções muito mais voltadas para as reais necessidades das redes atuais.

Contudo, ainda existe espaço para melhorias no que diz respeito à metodologia de medições sobre o protocolo SNMP. A arquitetura de uma ferramenta Web proposta neste trabalho apresenta uma forma de integrar os vários *softwares* que anteriormente eram necessários para a execução dessa metodologia, acrescenta novas funcionalidade e, com isso, aumenta a acessibilidade e a eficiência desse tipo de estudo. Além disso, a inclusão de módulos voltados para a geração de visualizações e comparações de resultados de análises permite a obtenção de dados mais completos com maior facilidade.

O emprego das três técnicas de visualização de informação propostas nesta dissertação de mestrado objetivou tornar o trabalho de interpretação dos resultados das análises mais eficiente e eficaz. A partir dessas técnicas, o usuário da ferramenta poderá analisar os dados obtidos de forma mais natural, pois graças à eficácia do sistema visual humano a descoberta de padrões e características interessantes se dará de forma simplificada.

Através do desenvolvimento da ferramenta *Management Traffic Analyzer*, demonstrou-se a viabilidade da implementação da arquitetura proposta. O estágio atual da implementação dessa ferramenta já permite que a mesma possa ser utilizada por pesquisadores e administradores de rede que tenham interesse em realizar medições sobre tráfegos SNMP, o que consistitui uma importante contribuição para a comunidade de gerenciamento de redes. Quanto mais confiáveis e acessíveis forem os instrumentos que permitam a utilização da metodologia proposta pelo IRTF para o estudo de uso do protocolo SNMP, maior será o número de tráfegos de gerenciamento estudados, e mais consistentes e representativos serão os resultados obtidos sobre as características de utilização do SNMP.

A ferramenta implementada foi utilizada para realizar um estudo sobre arquivos de tráfego SNMP pertencentes à duas redes brasileiras: o ponto de presença da Rede Nacional de Ensino e Pesquisa (RNP) no estado do Rio Grande do Sul, e a rede da própria RNP. Algumas características interessantes e comuns aos dois tráfegos puderam ser confirma-

das através desse estudo, tais como:

- Preferência pela utilização de consultas periódicas aos dispositivos, a fim de identificar os parâmetros de funcionamento dos mesmos (*polling*);
- Baixa utilização dos recursos de notificação do protocolo SNMP (e.g., *traps*);
- Forte tendência para a geração de tráfego periódico do SNMP. O tráfego aperiódico praticamente não influiu nos resultados das análises;
- Não utilização da versão 3 do protocolo SNMP, o que constitui um fato interessante, pois teoricamente as versões SNMPv1 e SNMPv2c são históricas, enquanto que a versão SNMPv3 é considerada como o padrão em vigência;
- Utilização do protocolo SNMP exclusivamente para monitoramento dos recursos da rede, e nunca para configuração de serviços e dispositivos. A mais provável explicação para isso é a falta de segurança nas operações para configuração de elementos da rede disponibilizadas até a versão 2 do protocolo SNMP.

A avaliação preliminar da ferramenta desenvolvida também apresentou resultados interessantes e satisfatórios. A maioria das pessoas consultadas nessa pesquisa avaliou de maneira positiva (conceitos ”bom” ou ”excelente”) itens como: facilidade de uso da ferramenta, qualidade da interface gráfica, facilidade de compreensão das visualizações geradas, entre outros. Os participantes da pesquisa também responderam questões de interpretação das visualizações geradas durante o processo de avaliação da ferramenta, e o índice de acertos de 4 das 6 questões foi superior a 80%, enquanto que as outras 2 questões obtiveram índice de acertos superior a 50%. Levando-se em consideração a completa inexperiência dessas pessoas que responderam aos questionários quanto à utilização da ferramenta e ao uso da metodologia do IRTF, os resultados mostram que o software desenvolvido é capaz de transformar as medições sobre tráfegos SNMP em um processo relativamente fácil, intuitivo e eficiente. Contudo, não se pode esquecer que esses resultados obtidos com a pesquisa preliminar ainda podem ser melhorados, o que leva à necessidade de se procurar um maior refinamento das técnicas de manipulação e análise dos arquivos de tráfegos proporcionados pela ferramenta.

Existe ainda muito espaço para o desenvolvimento da metodologia para estudo de uso do protocolo SNMP, tanto na parte das ferramentas que auxiliam nos processos de análise dos tráfegos, quanto na quantidade de análises de tráfegos SNMP que ainda necessitam ser realizadas, afinal, o sucesso da metodologia do IRTF também depende de que uma quantidade significativa de redes de gerenciamento sejam monitoradas e estudadas. Para trabalhos futuros, pretende-se ampliar o conjunto de técnicas de análise e visualização de tráfegos SNMP disponibilizadas na ferramenta *Management Traffic Analyzer*. Também se fará necessária a conclusão da implementação do módulo comparador de resultados de análises, a fim de se possibilitar a geração de resultados mais refinados sobre as características da utilização do protocolo SNMP nas diversas redes investigadas. Se faz necessária também uma avaliação mais formal e completa da ferramenta desenvolvida e apresentada nesta dissertação de mestrado, através da utilização de uma metodologia apropriada, de um maior número de pessoas para avaliarem a ferramenta, e de um ambiente de avaliação mais controlado, a fim de se obter resultados mais significativos quanto à qualidade do *Management Traffic Analyzer*. Por fim, planeja-se buscar amostras de tráfegos SNMP oriundas de outras redes em produção, a fim de tornar os resultados sobre a utilização real do protocolo SNMP mais significativos para a comunidade de gerenciamento de redes.

REFERÊNCIAS

AHN, S.; YOO, S. K.; CHUNG, J. W. Design and Implementation of a Web-based Internet Performance Management System Using SNMP MIB-II. **International Journal of Network Management**, [S.l.], 1999.

BECKER, R. A.; EICK, S. G.; WILKS, A. R. Visualizing Network Data. IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, [S.l.], v.1, n.1, p. 16-18, Mar. 1995.

CARD, S.; MACKINLAY, J. D.; SHNEIDERMAN, B. **Readings in Information Visualization**: Using Vision to Think. San Francisco, CA: Morgan Kaufmann, 1999.

CASE, J. D.; FEDOR, M. L.; SCHOFFSTAL, J. D. **Simple Network Management Protocol (SNMP)**. RFC 1157. [S.l.]: Internet Engineering Task Force, Network Working Group, 1990.

CASE, J. et al. **Textual Conventions for Version 2 of the Simple Network Management Protocol (SNMPv2)**: RFC 1903. [S.l.]: Internet Engineering Task Force, Network Working Group, 1996.

CHEN, C. **Information Visualization**: beyond the horizon. Londres: Springer, 2004.

ETHEREAL: A network protocol analyzer. Disponível em: <http://www.ethereal.com>. Acesso em: 20 fev. 2008.

FEKETE, J. D.; PLAISANT, C. Interactive Information Visualization of a Million Items. In: IEEE SYMPOSIUM ON INFORMATION VISUALIZATION, INFOVIS, 2002. **Proceedings…** [S.l.: s.n.], 2002. p. 117-124

FREITAS, C. M. Dal S. Visualização de Informações e a Convergência de Técnicas de Computação Gráfica e Interação Humano-Computador. In: BREITMAN, K. K.; KOWALTOWSKI, T. (Ed.). **Atualizações em Informática 2007**. Rio de Janeiro: SBC, 2007. p. 171-220.

JACOBSON, V.; LERES, C.; MCCANE, S. **Tcpdump**. Disponível em: <http://www.tcpdump.org>. Acesso em: 20 fev. 2008.

KEIM, D. A. et al. Monitoring Network Traffic with Radial Traffic Analyzer. In: IEEE SYMPOSIUM ON VISUAL ANALYTICS SCIENCE AND TECHNOLOGY, 2006, Baltimore, Maryland, USA. **Proceedings…** [S.l.]: IEEE Computer Society, 2006. p.123-128

MCGLOGHRIE, K.; KASTENHOLZ, F. **Evolution of the Interfaces Group of MIB-II**: RFC 1573. [S.l.]: Internet Engineering Task Force, Network Working Group, 1994.

MCGLOGHRIE, K.; KASTENHOLZ, F. **The Interfaces Group MIB**: RFC 2863. [S.l.]: Internet Engineering Task Force, Network Working Group, 2000.

MCGLOGHRIE, K.; PERKINS, D.; SCHOENWAELDER, J. **Structure of Management Information Version 2 (SMIv2)**: RFC 2578. [S.l.]: Internet Engineering Task Force, Network Working Group, 1999.

OBERHEIDE, J.; GOFF, M.; KARIR, M. Flamingo: Visualizing Internet Traffic. In: IEEE/IFIP NETWORK OPERATIONS AND MANAGEMENT SYMPOSIUM, NOMS, 10., 2006, Vancouver, Canada. **Proceedings…** [S.l.: s.n.], 2006. p. 150-161.

PAPADOPOULOS, C. et al. CyberSeer: 3D audio-visual immersion for network security and management. In: ACM WORKSHOP ON VISUALIZATION AND DATA MINING FOR COMPUTER SECURITY, VIZSEC/DMSEC, 2004. **Proceedings…** New York: ACM Press, 2004. p. 90-98.

PRESUHN, R. **Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)**: RFC 3418. [S.l.]: Internet Engineering Task Force, Network Working Group, 2002.

RAGHUNARAYAN, R. **Management Information Base for the Transmission Control Protocol (TCP)**: RFC 4022. [S.l.]: Internet Engineering Task Force, Network Working Group, 2005.

ROSE, M.; MCCLOGHRIE, K. **Concise MIB Definitions**: RFC 1212. [S.l.]: Internet Engineering Task Force, Network Working Group, 1991.

SALVADOR, E. M.; GRANVILLE, L. Z. Using Visualization Techniques for SNMP Traffic Analyses. In: IEEE SYMPOSIUM ON COMPUTERS AND COMMUNICATIONS, ISCC, 2008. **Proceedings…** [S.l.: s.n.], 2008a.

SALVADOR, E. M.; GRANVILLE, L. Z. An Investigation of Visualization Techniques for SNMP Traffic Traces. In: IEEE/IFIP NETWORK OPERATIONS AND MANAGEMENT SYMPOSIUM, NOMS, 2008. **Proceedings…** [S.l.: s.n.], 2008b.

SALVADOR, E. M.; GRANVILLE, L. Z. Arquitetura de uma Ferramenta e Técnicas de Visualização para Medições sobre Tráfego SNMP. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, SBRC, 26., 2008. **Anais…** [S.l.: s.n.], 2008c.

SCHOENWAELDER, J. et al. SNMP Traffic Analysis: Approaches, Tools, and First Results. In: IFIP/IEEE INTERNATIONAL SYMPOSIUM ON INTEGRATED NETWORK MANAGEMENT, IM, 10., 2007. **Proceedings…** [S.l.]: IEEE, 2007. p. 323-332.

SCHOENWAELDER, J. **SNMP Traffic Measurements**: Internet Draft. [S.l.: s.n.], 2006.

SCHOENWAELDER, J. **SNMP Traffic Measurements**: Internet Draft. [S.l.: s.n.], 2007a.

SCHOENWAELDER, J. **SNMP Traffic Measurements**: Internet Draft. [S.l.: s.n.], 2007b.

SCHOENWAELDER, J. **SNMP Traffic Measurements**: Internet Draft. [S.l.: s.n.], 2008.

SHARPE, R.; WARNICKE, E.; LAMPING, U. **Wireshark User's Guide**. Disponível em: <http://www.wireshark.org/docs/wsug_html_chunked/>. Acesso em: 20 fev. 2008

STALLINGS, W. **SNMP, SNMPv2, SNMPv3 and RMON 1 and 2.** Reading, Massachussets: Addison-Wesley, 1999.

XMLSOFT. **The XSLT C Library for Gnome**. Disponível em: <http://xmlsoft.org>. Acesso em: 23 fev. 2008.

APÊNDICE A FORMULÁRIO DE AVALIAÇÃO

Antes de tudo, gostaríamos de agradecer pela sua disponibilidade em contribuir com a nossa pesquisa. Muito obrigado!

O objetivo deste questionário é avaliar a ferramenta ***Management Traffic Analyzer***, a qual foi criada com o objetivo de analisar amostras de tráfego voltadas para o gerenciamento de rede, e que por sua vez sejam formadas por mensagens do protocolo SNMP (Simple Network Management Protocol). Essas análises são realizadas pela ferramenta através da conversão do arquivo de tráfego original obtido através de um sniffer (formato PCAP) em formatos que possam ser mais facilmente processados (XML ou CSV), e em seguida uma série de scripts são executados sobre esses arquivos convertidos, de forma a extrair uma série de dados estatísticos. A partir dessas análises, é possível se identificar padrões e características do uso deste protocolo, que até o presente momento ainda não puderam ser evidencialmente determinadas. Para a interpretação desses dados, a ferramenta também oferece um conjunto de visualizações específicas para cada tipo de análise, a fim de tornar o reconhecimento das características do tráfego mais rápido e eficaz.

O tempo estimado para que você responda a esse questionário é entre 15 e 20 minutos, e o mesmo se encontrará dividido em 3 partes:

- **1ª Parte - Dados Pessoais**: nesta parte serão solicitados alguns dados sobre você, como idade, formação, experiência na área de computação, entre outros. Contudo, não será solicitado nenhum dado que possa identificar você, como nome, CPF, etc.;
- **2ª Parte - Roteiro para Utilização da Ferramenta**: aqui você encontrará um roteiro de utilização da ferramenta, assim como algumas perguntas sobre os dados que você irá visualizar, para verificar se os mesmos foram corretamente compreendidos ou não. O roteiro se faz necessário como uma maneira de simplificar a utilização da ferramenta, uma vez que o fornecimento de um manual completo explicando cada uma das funcionalidades seria algo que levaria muito tempo, o que não é interessante para essa fase de avaliação do software;
- **3ª Parte - Avaliação Geral da Ferramenta**: uma vez conhecido o software, serão feitas algumas perguntas sobre o mesmo, de forma a avaliar determinados aspectos que puderam ser experimentados durante a etapa anterior;

Para que a presente avaliação possa transcorrer da melhor forma possível, é necessário tornar explícitas algumas observações:

- Na segunda parte do questionário, é fundamental ler cada um dos passos a serem seguidos por completo e com bastante atenção, pois a execução correta dos mesmos,

na ordem correta, é o que irá garantir o sucesso da etapa. Além disso, é natural que você sinta dificuldade com alguns dos conceitos encontrados na ferramenta se você não é da área de redes de computadores. Contudo, pedimos que mesmo assim você tente responder às questões, utilizando a sua intuição para tentar identificar o que está sendo pedido. A idéia do emprego de técnicas de visualização de informação também é fazer com que as informações possam ser obtidas intuitivamente, e isso também está sendo avaliado com este questionário;

- É necessário também se ter consciência de que a ferramenta Management Traffic Analyzer ainda está em fase de desenvolvimento. Por ser apenas um protótipo, uma série de erros ainda pode ser encontrado no software. Isso torna ainda mais importante a observância dos passos da utilização da ferramenta na 2ª etapa, a fim de se evitar que você se depare com um desses possíveis bugs.

Lembramos ainda que em caso de dúvidas, sugestões ou críticas você poderá escrever um e-mail para o autor dessa pesquisa (**emsalvador@inf.ufrgs.br**).

A.1 1ª Parte - Dados Pessoais

Sexo: ( )Masculino ( )Feminino

Idade:

Cidade:

Estado:

Escolaridade: ( )1º grau ( )2º grau ( )Superior ( )Especialização ( )Mestrado ( )Doutorado ( )Pós-doutorado

Função que exerce atualmente:

Tempo em que trabalha na área de computação: ( )Nunca trabalhei na área ( )Menos de 1 ano ( )Entre 1 e 2 anos ( )Entre 3 e 5 anos ( )Entre 6 e 10 anos ( )Mais de 10 anos

Qual a sua relação com a área de gerência de redes? ( )Já atuei, mas não atuo mais ( )Estou atuando nessa área atualmente ( )Nunca atuei na área

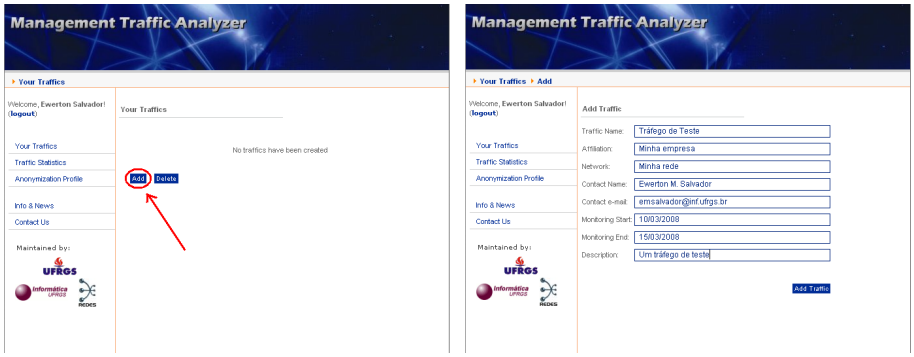
Como você classifica o seu conhecimento sobre o protocolo SNMP (*Simple Network Management Protocol*)? ( )Não sei nada sobre ele ( )Nível de conhecimento básico ( )Nível de conhecimento intermediário ( )Nível de conhecimento avançado

A.2 2ª Parte - Roteiro para Utilização da Ferramenta

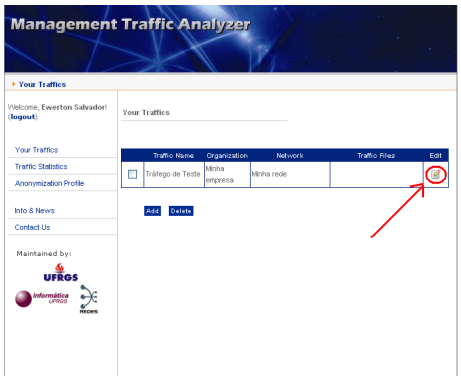
1. Acesse a ferramenta através do link **http://noc.inf.ufrgs.br/mtAnalyzer/** e se conecte com o login e a senha que você recebeu por e-mail;

2. Assim que você efetuar o login no sistema, você acessará automaticamente a opção "Your Traffics". Contudo, nenhum tráfego terá sido cadastrado ainda, pois sua

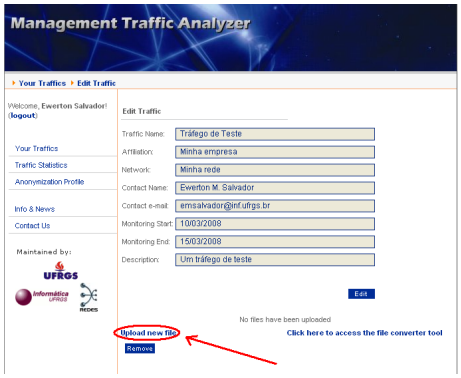
conta foi recém-criada. Para criar um tráfego, pressione o botão “Add”. Como esse roteiro tem por finalidade apenas avaliar a usabilidade da ferramenta, você pode ficar à vontade para preencher os campos do formulário que irá surgir com os dados que desejar;



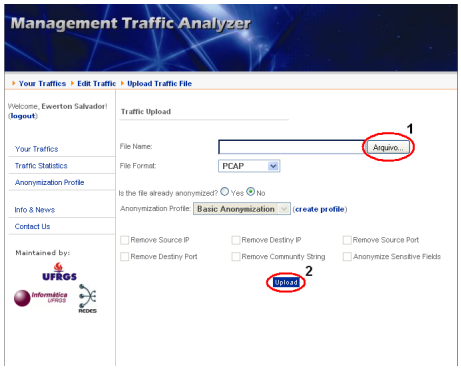
3. Após o novo tráfego ter sido cadastrado, você retornará para a opção “our Traffics”, onde você verá em uma tabela o tráfego que você acabou de registrar. Clique no ícone da última coluna da tabela, denominada “Edit”;



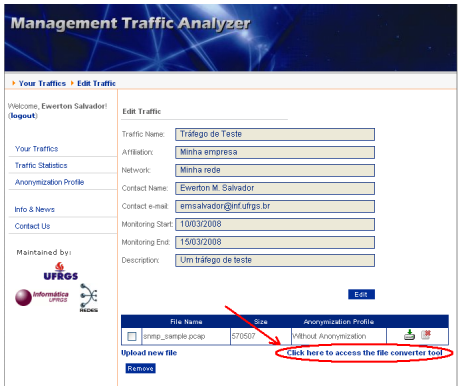
4. Em seguida será apresentado o formulário de registro de tráfego com os dados preenchidos anteriormente, e abaixo desse formulário é exibida uma seção para upload de arquivos contendo tráfego SNMP. Clique no link “Upload new file” e submeta o arquivo “snmp_sample.pcap”, que deve ser obtido através do link **http://noc.inf.ufrgs.br/mtAnalyzer/snmp_sample.pcap**;



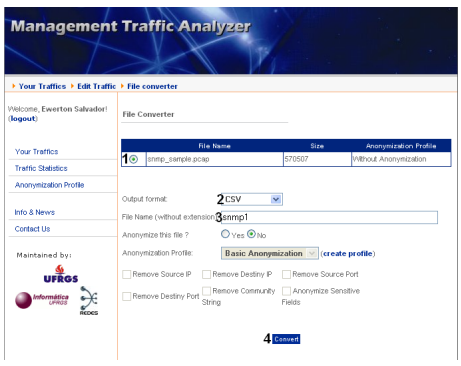
5. Selecione o arquivo snmp_sample.pcap pressionando o botão “Arquivo...”. Após ter selecionado o arquivo, pressione o botão “Upload” deixando as outras opções com seus valores padrão. Aguarde o término do upload do arquivo;



6. Após o upload ter sido concluído, você irá retornar para a tela de edição do tráfego cadastrado. Logo abaixo da tabela que exibe os arquivos submetidos ao tráfego, você encontrará a opção “Click here to access the file convertion tool”, que serve para converter formatos de arquivo de tráfego. Clique nessa opção;

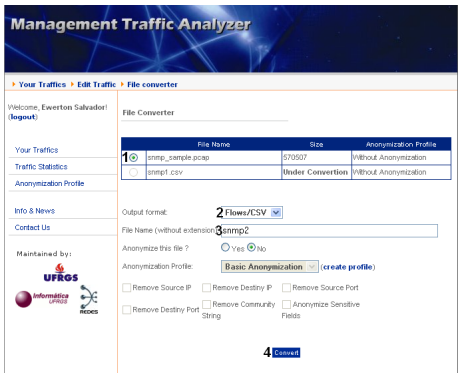


7. Em seguida, selecione o arquivo snmp_sample.pcap na tabela da parte superior da página. Depois, selecione o formato “CSV” (comma separated values) no campo “Output format”. Após, digite o nome de arquivo “snmp1” na caixa “File Name (without extension)”. Por fim, clique no botão “Convert”. Observe que após você pressionar o botão, ao invés de aparecer o tamanho do arquivo convertido na tabela, você obterá a mensagem “Under Convertion”. Ignore essa mensagem, pois nos próximos passos ela irá deixar de aparecer;

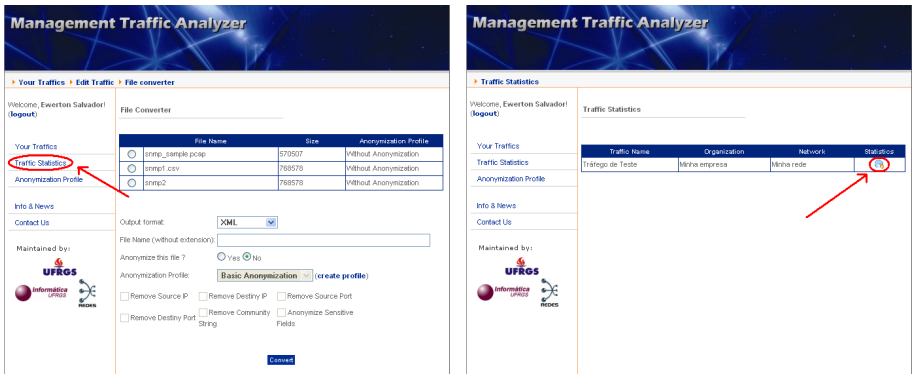


8. Agora você irá basicamente repetir o passo anterior, porém selecionando um formato diferente. Selecione na tabela da parte superior da página o arquivo snmp_sample.pcap, e em seguida selecione a opção “Flows/CSV” na linha “Output format”. Por fim,

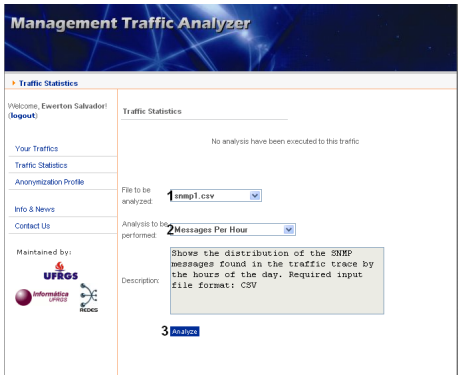
digite mais uma vez “snmp2” na caixa “File Name (without extension)” e pressione “Convert”;

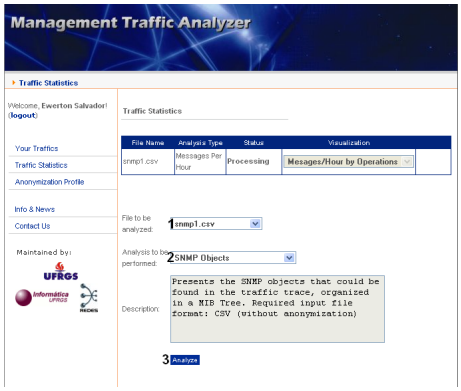


9. Uma vez terminada a conversão dos arquivos, você deverá acessar a opção “Traffic Statistics” no menu localizado no lado esquerdo da página. Ao serem listados seus tráfegos na ferramenta, você deverá clicar no ícone em formato de um gráfico em pizza, na última coluna da tabela contendo os tráfegos;



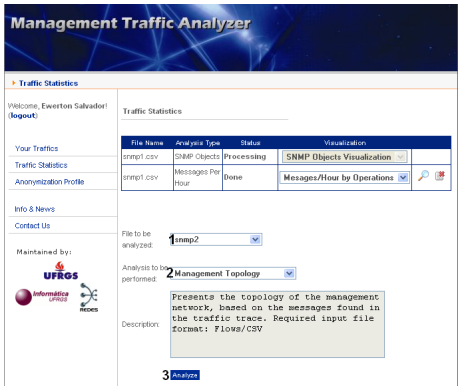
10. No campo “File to be analyzed” selecione o arquivo “snmp1.csv”, e no campo “Analysis to be performed” selecione a opção “Messages Per Hour”. Por fim, pressione o botão “Analyze”. Logo após você ter pressionado o botão, na tabela de análises realizadas você verá que o status daquela que você acabou de criar existirá a mensagem “Processing”. Por enquanto, ignore essa mensagem, pois nos próximos passos você verá que a mesma terá deixado de aparecer;



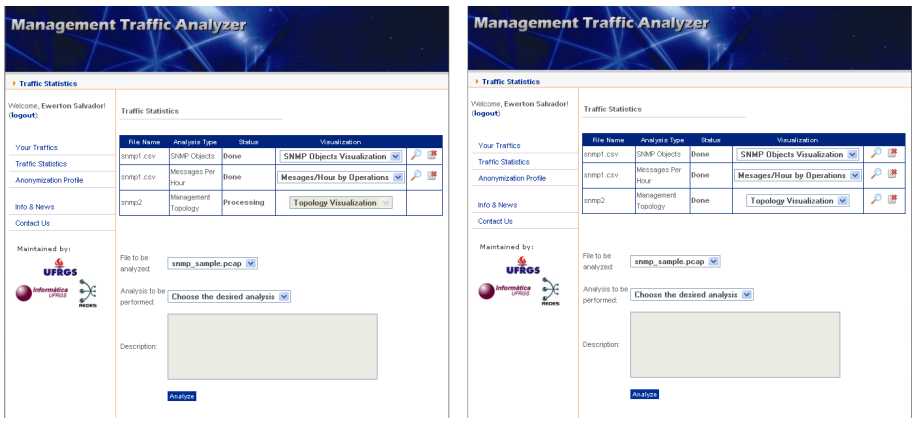


11. Novamente, selecione o arquivo “snmp1.csv” no campo “File to be analyzed”, e no campo “Analysis to be performed” selecione a opção “SNMP Objects”. Pressione o botão “Analyze”;

12. Na última análise a ser feita neste roteiro, selecione o arquivo “snmp2” no campo “File to be analyzed”, e em seguida selecione a opção “Management Topology” no campo “Analysis to be performed”. Pressione o botão “Analyze”;



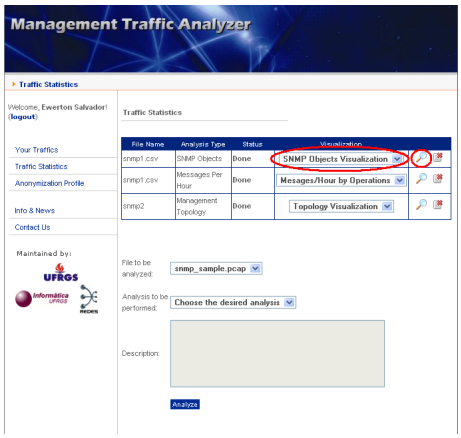
13. Uma vez concluído o passo anterior, você deverá atualizar a página da ferramenta para atualizar o status das análises que estão sendo realizadas, até que todas fiquem com o status “Done”. A mudança na página após de atualizar seu conteúdo (pressione a tecla F5 para isso) deverá ser como mostrado abaixo;



14. A partir deste ponto, você deverá responder a algumas perguntas simples sobre o tráfego SNMP que está sendo estudado, a partir das visualizações geradas pela

ferramenta. Essa visualizações são acessadas a partir da seleção de uma das visualizações disponíveis para a análise realizada, e posteriormente clicando-se na imagem de uma lupa à direita da caixa de seleção, conforme mostrado na figura abaixo. Caso você não entenda a pergunta, ou não encontre uma resposta para a mesma, indique isso na caixa de resposta (digite algo como ”Não sei“ ou “Não consegui encontrar a resposta”). Contudo, é importante que você tente da melhor forma possível encontrar respostas para essas questões, utilizando a própria intuição, caso você não seja da área de gerenciamento de redes.

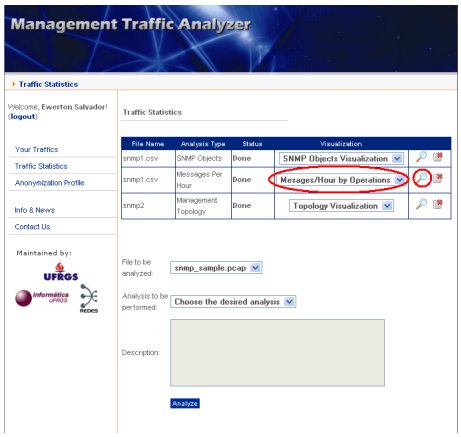
Com base nas visualizações geradas a partir da análise dos objetos SNMP (SNMP Objects, a primeira da tabela de análises), responda:



a) Qual o objeto SNMP mais acessado do tráfego?

b) Qual é a MIB menos utilizada no tráfego? (na visualização, observe que cada folha da árvore contém o nome da MIB e o nome do objeto SNMP, no formato “nome-da-MIB::nome-do-objeto-SNMP”)

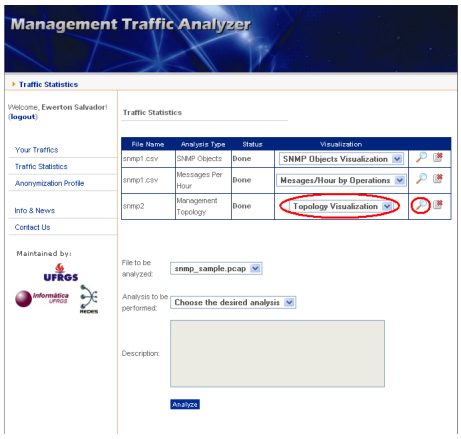
15. Baseado(a) nas visualizações geradas a partir da análise das mensagens/hora (Messages per Hour, na segunda linha da tabela de análises), responda às questões abaixo. **Observe que para essa análise existem DUAS visualizações disponíveis na caixa de seleção**. Você deverá selecionar cada uma delas, e após cada seleção você deverá clicar na lupa.



c) Qual é a hora do dia em que é registrado o maior volume de mensagens SNMPv1?

d) Qual é a operação do protocolo SNMP mais utilizada na rede?

16. Por fim, responda às questões abaixo baseado na análise da topologia da rede de gerenciamento (Management Topology, na última linha da tabela de análises):



e) Qual o volume de tráfego concentrado pelo principal gerente da rede?

f) Quais os endereços IPs do par de máquinas que mais trocou informações no tráfego?

A.3 3ª Parte - Avaliação Geral da Ferramenta

1. Você compreendeu a finalidade da ferramenta? ( )Não ( )Apenas possuo uma vaga idéia ( )Sim

2. Como você avalia a facilidade de uso da ferramenta? ( )Péssima ( )Ruim ( )Regular ( )Boa ( )Excelente

3. Como você avalia a interface gráfica da ferramenta? ( )Péssima ( )Ruim ( )Regular ( )Boa ( )Excelente

4. Com que facilidade você conseguiu interpretar as visualizações geradas após as análises dos arquivos? ( )Foi muito difícil ( )Foi um pouco difícil ( )Foi relativamente fácil ( )Foi muito fácil

5. Você acredita que essa ferramenta pode ser útil para auxiliar pesquisadores a compreenderem melhor o uso do protocolo SNMP? ( )Sim ( )Não

6. Você acredita que essa ferramenta pode ser útil para facilitar o trabalho de administradores de rede? ( )Sim ( )Não

7. Qual é a sua avaliação geral da ferramenta? ( )Péssima ( )Ruim ( )Regular ( )Boa ( )Excelente

8. Caso você queira fazer comentários adicionais sobre a ferramenta avaliada, utilize o espaço ao lado (opcional).

Muito obrigado pela sua contribuição!

Caso você deseje escrever algum comentário para os autores do projeto, envie e-mail para **emsalvador@inf.ufrgs.br**.

APÊNDICE B ARTIGOS PUBLICADOS

Este apêndice é dedicado à apresentação dos artigos desenvolvidos a partir dos resultados obtidos ao longo do curso do mestrado. O primeiro artigo apresenta as técnicas de visualização de informação desenvolvidas especificamente para o uso em conjunto com a metodologia do IRTF para medições sobre tráfegos SNMP. Além disso, também foi apresentada a arquitetura da ferramenta desenvolvida ao longo do mestrado, assim como resultados obtidos à partir do uso dessa ferramenta para análise dos arquivos de tráfego SNMP cedidos pela Rede Nacional de Ensino e Pesquisa (RNP)

- **Título**: *Arquitetura de uma Ferramenta e Técnicas de Visualização para Medições sobre Tráfego SNMP*
- **Evento**: *26º Simpósio Brasileiro de Redes de Computadores, 2008. SBRC 2008*
- **URL**: http://www.sbrc2008.ufrj.br/
- **Data**: De 26 a 30 de maio de 2008
- **Local**: Rio de Janeiro, Brasil

No segundo artigo apresentado neste apêndice foram discutidas as técnicas de visualização criadas especificamente para o uso com a metodologia do IRTF para medições sobre tráfego SNMP. Adicionalmente, o artigo também apresenta e discute os resultados obtidos à partir da análise dos arquivos de tráfego SNMP cedidos pela Rede Nacional de Ensino e Pesquisa (RNP).

- **Título**: *Using Visualization Techniques for SNMP Traffic Analyses*
- **Evento**: *IEEE Symposium on Computers and Communications, 2008. ISCC'08*
- **URL**: http://www.comsoc.org/iscc/2008/
- **Data**: De 6 a 9 de julho de 2008
- **Local**: Marraqueche, Marrocos

Arquitetura de uma Ferramenta e Técnicas de Visualização para Medições sobre Tráfego SNMP

Ewerton Monteiro Salvador, Lisandro Zambenedetti Granville

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre, RS – Brasil

{emsalvador, granville}@inf.ufrgs.br

Abstract. *In march 2006 the IRTF proposed an approach for the measurement of SNMP traffic. However, this approach has some limitations, such as: absence of data visualization techniques and lack of integration among the necessary tools for supporting the approach. This paper proposes an architecture for a Web-based tool that automates, in an integrated fashion, the execution of the IRTF approach's steps. Visualization techniques have been also developed in order to properly present the results of the analyses of SNMP traffics. An implementation of this architecture, named Management Traffic Analyzer, has been used for studying SNMP traffic samples from the Brazilian National Education and Research Network (RNP).*

Resumo. *Em março de 2006 o IRTF propôs uma metodologia para medição de tráfego SNMP. Contudo, essa metodologia possui algumas limitações, tais como: ausência de técnicas para visualização de dados e falta de integração entre as ferramentas necessárias para a execução da metodologia. Este artigo apresenta uma proposta de arquitetura para uma ferramenta Web que automatiza, de forma integrada, a execução das etapas da metodologia do IRTF. Também foram desenvolvidas técnicas de visualização para os resultados gerados a partir de análises de tráfegos SNMP. Uma implementação dessa arquitetura, denominada Management Traffic Analyzer, foi utilizada para o estudo de amostras de tráfego SNMP da Rede Nacional de Ensino e Pesquisa (RNP).*

1. Introdução

O *Simple Network Management Protocol* [Case et al. 1990] (SNMP) foi proposto há mais de 15 anos, e atualmente é tido como o protocolo padrão *de facto* para o gerenciamento de redes TCP/IP. Apesar de ser amplamente utilizado, muito pouco se sabe de fato sobre os padrões de uso desse protocolo nas redes em produção.

Em março de 2006, o *Network Management Research Group* (NMRG), pertencente ao *Internet Research Task Force* (IRTF), publicou um documento no formato *internet draft* intitulado "*SNMP Traffic Measurements*" [Schoenwaelder 2006], o qual propunha uma metodologia sistemática para medições e geração de estatísticas sobre o uso do SNMP. O objetivo dessa metodologia é identificar padrões de utilização do SNMP a fim de poder se descobrir características deste protocolo que atualmente ainda não são efetivamente conhecidas. Algumas das questões que estão sendo investigadas neste contexto são, por exemplo: quais recursos do protocolo (versões, operações, MIBs, etc.) estão sendo utilizados, como o uso do SNMP difere nos vários tipos existentes de redes de

computadores e organizações, quais informações são mais freqüentemente requisitadas e quais são as interações mais típicas que estão sendo empregadas utilizando o protocolo [Schoenwaelder et al. 2007].

Apesar da metodologia proposta pelo IRTF ser de grande relevância para a área de gerenciamento de redes, a mesma ainda possui algumas limitações. Os estudos sobre o SNMP baseados nessa metodologia certamente irão gerar uma grande quantidade de novos dados sobre esse protocolo. Essa nova massa de dados também precisará ser interpretada pelas pessoas que estiverem conduzindo o estudo, de forma a responder às questões que ainda se encontram em aberto sobre o SNMP. Uma forma de aumentar a eficiência desse processo de interpretação dos dados é a utilização de técnicas de visualização de informação, as quais permitem que pessoas obtenham *insights* sobre os dados que estão sendo analisados (e.g., detecção de padrões, descoberta de características interessantes, etc.) de uma forma mais rápida e natural, graças às capacidades únicas do sistema visual humano. Contudo, nenhuma técnica de visualização de informação é descrita pela metodologia original apresentada pelo NMRG.

Outra característica da metodologia original do IRTF é a necessidade de se empregar um conjunto de ferramentas específicas para certas fases do estudo sobre o tráfego SNMP, o que dificulta a utilização da metodologia em si. O desenvolvimento de um framework que possa integrar todos esses softwares numa única ferramenta é algo bastante desejável, uma vez que simplificaria consideravelmente a execução desse tipo de estudo. Por fim, também não é fornecida pela metodologia do IRTF nenhuma forma de se comparar os resultados de análises realizadas sobre dois ou mais tráfegos distintos. Essas comparações teriam por objetivo apresentar, para a pessoa que está realizando o estudo, as variações de uma amostra de tráfego SNMP com relação à outra amostra, permitindo se identificar as diferenças entre as possíveis estratégias de gerenciamento que podem ser empregadas na administração de uma rede.

Neste artigo propomos algumas soluções para essas deficiências da metodologia do IRTF através do desenvolvimento da arquitetura de uma ferramenta Web para automatizar, de forma integrada, a execução das etapas da abordagem. As principais contribuições deste trabalho são: especificação de uma nova arquitetura, baseada em tecnologias Web, voltada à investigação sobre tráfego SNMP segundo as definições do IRTF; desenvolvimento de técnicas de visualização de informação específicas para o contexto onde está inserido o SNMP; e apresentação dos primeiros resultados acerca das investigações sobre o tráfego SNMP da Rede Nacional de Ensino e Pesquisa (RNP). O restante deste artigo está organizado da seguinte maneira. A seção 2 apresenta os trabalhos que estão relacionados com este artigo. A seção 3 apresenta os principais conceitos relacionados à área de investigação de tráfego de gerenciamento, enquanto que a seção 4 apresenta a arquitetura de uma ferramenta Web para automatizar a execução da metodologia do IRTF. A seção 5 descreve as técnicas de visualização desenvolvidas para serem utilizadas em conjunto com a metodologia proposta pelo IRTF. Os primeiros resultados das análises de um tráfego SNMP da RNP são apresentados na seção 6. Por fim, a seção 7 apresenta conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Ainda são poucos os trabalhos que tratam da investigação de tráfego de gerenciamento nas redes em produção. Um desses trabalhos é o de Schönwälder *et al.* [Schoenwaelder et al. 2007], que apresenta os primeiros resultados sobre o uso do SNMP nas redes em produção. Nele são publicadas informações obtidas através da aplicação direta da metodologia do IRTF sobre diversas amostras de tráfego SNMP. Contudo, o estudo descrito pelo trabalho de Schönwälder ainda está em andamento, e os resultados apresentados são apenas preliminares.

Os trabalhos sobre técnicas de visualização de dados aplicados à área de gerenciamento de redes também são escassos. De acordo com o conhecimento dos autores deste artigo, o trabalho que mais se aproxima dos objetivos desta pesquisa é o artigo de Seong Jin Ahn *et al.* [Ahn et al. 1999]. O objetivo do trabalho de Seong Jin Ahn *et al.* é definir uma ferramenta Web destinada a analisar o desempenho de redes TCP/IP, através da análise dos dados gerados pela monitoração da rede, utilizando o SNMP.

Oberheide *et al.* [Oberheide et al. 2006] descreveu uma ferramenta para auxiliar a tarefa de gerenciar uma rede através de visualizações sobre *Netflow feeds*. Um dos aspectos dessa ferramenta que se assemelha à proposta dos autores deste trabalho é a preferência pelo uso de um conjunto de técnicas de visualização e ferramentas de manipulação associadas, ao invés da utilização de uma técnica de visualização única.

Conforme se pode observar, todos os trabalhos sobre visualização de dados descritos nesta seção lidam com tráfego de rede de uma forma geral, sem levar em consideração as particularidades do tráfego de gerenciamento e, mais especificamente, do SNMP. O grande problema disso é que as técnicas de visualização para tráfego de rede não específicos não levam em consideração características intrínsecas do protocolo SNMP, como versão utilizada, elementos na *varbind list* e o comportamento de operações básicas como `get-next-request` e `set-request`. Dessa forma, se faz necessária uma adaptação das técnicas de visualização de informações sobre redes de computadores para o contexto onde está inserido o protocolo SNMP.

3. Metodologia para Análise de Tráfego de Gerenciamento

Conforme já afirmado anteriormente, a primeira metodologia sistemática para análise de tráfego de gerenciamento foi proposta pelo IRTF em março de 2006 [Schoenwaelder 2006]. Essa metodologia é composta de etapas bem definidas e sua relativa simplicidade facilita a sua implementação. Entretanto, conforme também já foi afirmado anteriormente, a metodologia do IRTF e suas ferramentas auxiliares apresentam uma série de restrições, tais como: não especificar técnicas para a visualização dos dados obtidos, exigir o uso de um determinado conjunto de softwares e não especificar formas de comparação entre resultados de análises. A seguir, serão descritas brevemente as etapas que compõem a metodologia do IRTF para medições de tráfego SNMP.

1. **Captura do tráfego de gerenciamento** – A captura do tráfego de gerenciamento a ser analisado pode ser realizada através do uso de *sniffers* de rede convencionais, como o Wireshark;
2. **Conversão dos arquivos `pcap`** – A necessidade de se converter o arquivo de tráfego no formato `pcap` para um outro formato (XML ou CSV) vem da

constatação de que o formato `pcap` não é suficientemente legível, nem para humanos nem para máquinas. Fazer com que o arquivo de tráfego seja "legível por seres humanos" permite que um operador verifique se dados confidenciais não estão presentes nos dados analisados. Já a expressão "legível por máquinas" refere-se à facilidade com que um arquivo em um determinado formato tem de ser processado pelo computador. Um arquivo de fácil processamento permite que a análise de seus dados seja realizada de forma eficiente. Dois possíveis formatos para armazenar os tráfegos SNMP que atendem a esses pré-requisitos são XML (*eXtended Markup Language*) e CSV (*Comma Separated Values*);

3. **Filtragem dos arquivos XML/CSV** – Um dos problemas existentes em se publicar tráfegos de gerenciamento para pesquisa é o fato de que comumente são encontradas nos mesmos informações sensíveis (e.g., senhas, nomes de usuários, informações sigilosas) que não podem ser divulgadas. Devido a isso, é necessária a utilização de um meio para proteger as fontes que estejam fornecendo amostras de tráfego de gerenciamento através de uma filtragem nos dados desses tráfegos, a fim de se remover e/ou anonimizar essas informações sensíveis;
4. **Armazenamento do arquivo `pcap` e de sua representação XML/CSV** – Durante o período em que a metodologia aqui proposta estiver sendo aplicada, ou até mesmo após esse período, poderão ser descobertos problemas em algum processo realizado durante a obtenção, conversão ou análise dos dados, de forma que será necessário verificar e/ou repetir um ou mais passos da metodologia. Por causa disso, os dados originais (arquivo `pcap` "bruto" e arquivos XML/CSV filtrados) devem ser armazenados e preservados de forma a possibilitar a recuperação dessas informações numa eventual necessidade futura;
5. **Análise dos arquivos filtrados** – O último passo da metodologia consiste na análise dos arquivos filtrados, a fim de se agregar os dados do tráfego de gerenciamento contidos nesses arquivos e, a partir deles, extrair informações que ajudem a responder às questões em aberto sobre a utilização prática do SNMP. A análise dos dados se dá através da execução de programas ou *scripts* que procuram agregar os dados do tráfego de gerenciamento de forma a fornecer informações úteis, como predominâncias e/ou tendências dentro desses tráfegos (e.g., qual versão do protocolo é mais utilizada, qual a relação entre o tráfego periódico e o aperiódico e quais os objetos mais acessados).

4. Arquitetura de uma Ferramenta para Medições sobre Tráfegos SNMP

Conforme já foi apresentado na introdução deste artigo, a metodologia do IRTF para medições sobre tráfego SNMP possui uma série de limitações que dificultam a sua utilização e diminuem a eficiência da análise dos resultados. Com o intuito de abordar essas limitações, será apresentada nesta seção a arquitetura de uma ferramenta Web para automatizar a execução da metodologia para medições sobre tráfego SNMP. As principais contribuições provindas dessa arquitetura para os estudos sobre tráfego SNMP são as seguintes:

- Utilização de um módulo gerador de visualizações, destinado a facilitar o processo de interpretação dos resultados das análises executadas sobre tráfego SNMP;
- Eliminação da necessidade de se utilizar um conjunto de ferramentas para realizar todos os passos da metodologia do IRTF. A arquitetura supre todas as necessidades de software fundamentais para a realização das medições sobre o tráfego SNMP;

- Inclusão de um módulo voltado para comparação de resultados de análises de tráfego SNMP. A metodologia original do IRTF não especifica nenhuma forma para se comparar resultados de análises distintas.

A Figura 1 apresenta os principais componentes que formam a arquitetura.

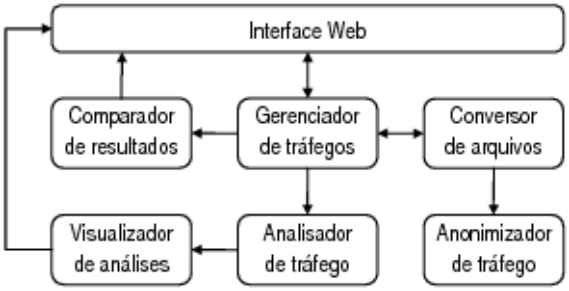


Figura 1. Arquitetura de uma ferramenta para medições de tráfego SNMP

O primeiro elemento que compõe a arquitetura do software é a sua **interface Web**. É somente através dela que o usuário poderá interagir com o sistema para executar operações como: criação de conta de usuário, login no sistema, inclusão de novo tráfego, análise e visualização de resultados, etc.

O **gerenciador de tráfego** é a parte da arquitetura responsável por manter registro de todos os tráfegos que o usuário submete à ferramenta Web. A entidade "tráfego", na arquitetura que está sendo desenvolvida, é a junção de um conjunto de metadados que descreve um tráfego em si (i.e., a rede onde o monitoramento foi realizado, o período em que se deu esse monitoramento, o nome de uma pessoa responsável pela administração da rede monitorada, etc.), e os arquivos que se relacionam àquele tráfego, que podem estar no formato `pcap`, XML ou CSV.

Uma vez que já existam arquivos submetidos para um determinado tráfego, o usuário terá à sua disposição uma "sub-ferramenta" – o **conversor de formato de arquivos**. Com isso o usuário poderá, por exemplo, converter um arquivo gerado a partir de uma seção de monitoramento de uma rede, originalmente no formato `pcap`, para os formatos XML ou CSV. Esse tipo de conversão é necessária para permitir que esses arquivos de tráfego possam ser analisados posteriormente, pois a ferramenta não será capaz de analisar arquivos no formato `pcap`, conforme já discutido anteriormente.

Durante o processo de conversão de um arquivo, o usuário poderá informar para a ferramenta parâmetros a serem utilizados na anonimização de informações que ele não deseja que sejam divulgadas. A parte da arquitetura responsável por tratar esses parâmetros e coordenar o processo de anonimização é chamada de **anonimizador de tráfego**.

Os parâmetros que poderão ser informados durante o processo de anonimização estão organizados na ferramenta na forma de *perfis de anonimização*. Dessa forma, primeiramente um usuário define um perfil de anonimização, especificando um nome para o mesmo e um conjunto de itens que se deseja ocultar com a utilização daquele perfil.

Após os arquivos terem sido devidamente registrados e adequadamente convertidos na ferramenta, os mesmos estarão aptos a serem analisados. A parte da arquitetura responsável por executar essa operação é o **analisador de tráfego**. Para isso, o usuário

poderá selecionar um dos tipos de análise disponíveis na ferramenta, e aplicá-la sobre os arquivos disponíveis no formato XML ou CSV. Os dados resultantes da análise são armazenados em tabelas específicas da base de dados, para serem utilizados posteriormente em visualizações ou comparações com outras análises. Ao final do processo, o sistema informa ao usuário que os dados resultantes daquela análise estão disponíveis, e apresenta ao mesmo uma lista com todas as visualizações possíveis para aquele resultado.

O módulo analisador de tráfego foi desenvolvido de forma a possuir baixo acoplamento com o restante da ferramenta, permitindo que outras análises possam ser desenvolvidas sem que sejam necessárias modificações no restante do código do sistema. Dessa forma, uma vez que novas análises tenham sido desenvolvidas, é suficiente que sejam feitas algumas poucas modificações na base de dados, como adição de linhas em tabelas ou criação de novas tabelas, para que a nova análise seja reconhecida pelo sistema.

As técnicas de visualização a serem utilizadas na ferramenta Web devem ser desenvolvidas de forma a recuperar os dados resultantes de uma análise diretamente do banco de dados, processar esses dados e exibir a visualização. O módulo da arquitetura que gerencia esse processo é denominado de **visualizador de resultados de análises**.

Uma análise na ferramenta pode ter uma ou mais técnicas de visualização associadas à mesma para que o usuário possa escolher, dentre as técnicas disponíveis, a que melhor se aplica ao seu caso. De modo análogo ao que ocorre com as análises na ferramenta, também é possível se adicionar novas visualizações. Para isso, também será suficiente a manipulação de determinadas tabelas da base de dados para que o sistema possa reconhecer e disponibilizar uma nova técnica de visualização de dados.

Por fim, o usuário terá a opção de realizar comparações entre duas ou mais análises já realizadas. Essa comparação é gerenciada pelo **comparador de resultados de análises**. Uma determinada análise sobre uma amostra de tráfego pode ser comparada com a mesma análise realizada sobre outras amostras de tráfego, pertencentes ao próprio usuário ou a outras amostras disponíveis no sistema (com o devido consentimento de seus respectivos proprietários). Também de modo análogo aos casos das análises e visualizações, também é possível se adicionar novas formas de comparação de resultados à ferramenta.

5. Técnicas de Visualização de Informação para os Resultados das Análises

Atualmente, tanto a academia quanto a indústria estão cada vez mais interessadas no desenvolvimento de técnicas de visualização de dados específicas para a área de gerenciamento de redes. Os trabalhos de Oberheide *et al.* [Oberheide et al. 2006] e Papadopoulos *et al.* [Papadopoulos et al. 2004] são exemplos de estudos que objetivaram o desenvolvimento de técnicas de visualização para tráfegos de redes de computadores. Entretanto, todos esses trabalhos lidam com tráfegos de redes de um modo geral, ou seja, sem levar em consideração as finalidades específicas de cada tipo de tráfego.

Os primeiros trabalhos sobre técnicas de visualizações de tráfego SNMP foram apresentados no trabalho de Salvador e Granville [Salvador and Granville 2008]. Essas técnicas também estão implementadas na ferramenta *Management Traffic Analyzer* e serão descritas brevemente nas próximas subseções.

5.1. Visualização da Topologia da Rede de Gerenciamento

Técnicas de visualização baseadas em grafos são amplamente utilizadas para representar determinados aspectos de uma rede de computadores, como a sua topologia [Becker et al. 1995]. Para adaptarmos essa técnica a fim de que ela possa produzir uma visualização da topologia de uma rede de gerenciamento, é necessário identificarmos no tráfego SNMP quais terminais atuam como gerentes, quais atuam como agentes, e quais atuam como ambos. Essa identificação deve ser feita através da análise dos fluxos de mensagens SNMP que compõem o tráfego. Esses fluxos são definidos como o conjunto de mensagens que partiu de uma determinada origem até um determinado destino. Além disso, os fluxos de mensagens SNMP pertencem à duas classes de relacionamento: *Command Generator* (CG) / *Command Responder* (CR) e *Notification Originator* (NO) / *Notification Receiver* (NR). O nó de origem em um relacionamento do tipo CG/CR atua como gerente e o destino atua como agente. Por outro lado, em um relacionamento do tipo NO/NR, a fonte atua como agente e o destino como gerente.

Uma vez que gerentes e agentes tenham sido identificados, esses elementos serão representados na visualização através de círculos. O tamanho dos círculos indica o papel que um determinado terminal desempenhou no tráfego estudado. Se um terminal atuou apenas como agente ao longo de todo o tráfego, ele será representado por um círculo menor. Por outro lado, se um terminal atuou como um gerente em qualquer um dos fluxos de mensagens SNMP, ele será representado por um círculo maior, mesmo que este tenha atuado como agente em algum outro fluxo de mensagens. O tamanho do círculo também dependerá do número de mensagens enviadas/recebidas que sejam características de um gerente da rede: quanto maior o número de mensagens desse tipo, maior será o tamanho do círculo na visualização. Isso indica que, quanto maior o círculo, maior é a atuação do nodo como gerente na rede representada.

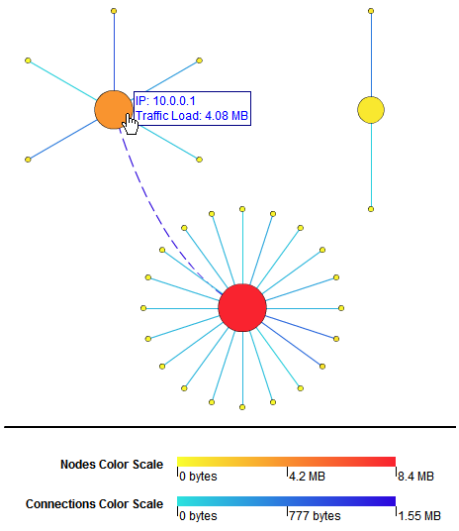


Figura 2. Visualização da Topologia da Rede de Gerenciamento

Uma linha sólida ligando dois nós representa uma conexão entre um nó gerente e um nó estritamente agente (i.e., que atuou como agente em todos os fluxos de mensagens do tráfego). Já uma linha tracejada representa a conexão entre um nó gerente e um nó

que está atuando naquele relacionamento específico como agente, mas que atuou também como gerente em outros fluxos de mensagens.

Tanto os nós quando as conexões possuem uma determinada cor que carrega consigo informações sobre carga do tráfego. No caso dos nós, a cor representa a quantidade de tráfego recebido por aquele determinado terminal. Por outro lado, a cor das conexões representa a quantidade de tráfego trocado entre os dois terminais conectados. Na parte inferior da visualização existem 2 barras coloridas, que são as escalas de cores utilizadas para coloris nós e conexões. Foram utilizadas escalas de cores distintas devido à diferença entre as ordens de grandeza dos tráfegos dos nós e das conexões. O resultado dessa análise pode ser observado na Figura 2.

Por fim, foram utilizados dois mecanismos de interação nessa visualização: legendas interativas para nós e conexões, e barra de rolagem. As legendas informam para o usuário algumas informações relevantes sobre nós (e.g., endereço IP e carga no nó) e conexões (e.g., tipo de relacionamento do fluxo e carga de tráfego). O usuário irá visualizar essas legendas quando posicionar o ponteiro do mouse sobre um nó ou uma conexão. Por sua vez, a barra de rolagem é o mecanismo que evita que a representação topológica seja truncada quando uma rede de grande porte estiver sendo representada.

5.2. Visualização de Objetos SNMP em uma *MIB Tree*

Uma das análises sobre tráfegos SNMP previstas pela metodologia do IRTF é o cálculo do número de vezes em que um determinado objeto SNMP é visto em um tráfego. Através desta análise, é possível se elaborar uma série de estatísticas sobre os objetos SNMP presentes no tráfego, como o conjunto de objetos mais acessados, os menos acessados, as MIBs (*Management Information Bases*) mais importantes, etc. Contudo, uma resposta textual desse tipo de análise possui algumas limitações. Provavelmente a principal delas é a dificuldade que um usuário teria de identificar o relacionamento hierárquico entre dois ou mais objetos, observando apenas seus nomes ou seus OIDs (Objects Identifiers). Uma vez que os objetos SNMP são organizados em uma árvore conhecida como *MIB tree*, é desejável que o resultado desse tipo de análise também apresente o conjunto de objetos encontrados no tráfego nesse tipo de estrutura. Devido a isso, foi desenvolvida uma técnica que mistura duas técnicas de visualização bastante conhecidas: visualização de árvores e histogramas. Dessa forma, o processo de análise dos resultados desse tipo de estatística se tornará mais eficiente, pois se assume que os administradores de rede estão potencialmente familiarizados com a organização de objetos SNMP em *MIB trees*.

Uma vez que a *MIB tree* contendo os objetos SNMP encontrados no tráfego tenha sido desenhada, será necessário representar o número de mensagens relacionadas com cada um dos objetos SNMP em um histograma. As barras do histograma são dese-nhadas do lado direito dos nós folhas da árvore, os quais representam os objetos SNMP. O tamanho da barra é baseado numa escala criada em tempo real, onde o valor mínimo é 0 e o máximo é o maior número de mensagens contendo um determinado objeto SNMP, dentre todos os objetos listados nos resultados da análise. Por fim, foi empregada uma barra de rolagem como mecanismo de interação para essa visualização, pois muito freqüentemente existe um grande número de objetos distintos em um tráfego SNMP. A Figura 3 mostra um exemplo dessa técnica sendo aplicada sobre um tráfego SNMP.

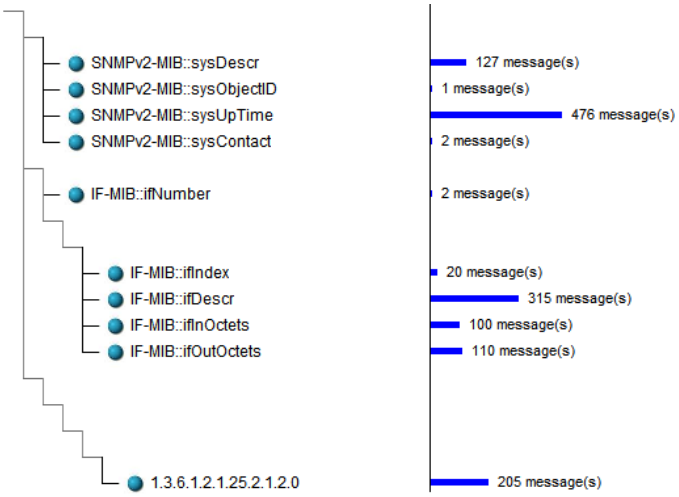


Figura 3. Visualização de Objetos do SNMP em uma *MIB Tree*

5.3. Visualização da Quantidade de Mensagens SNMP em Intervalos de 1 Hora

Uma das análises disponibilizadas na ferramenta *Management Traffic Analyzer* é o cálculo da quantidade de mensagens encontradas no tráfego em intervalos de 1 hora. Esse tipo de análise é útil para se identificar o comportamento da quantidade dos diversos tipos de mensagens de gerenciamento trocadas na rede ao longo de um dia.

Os resultados desse tipo de análise são representados em uma visualização que emprega histogramas para apresentar o número de mensagens SNMP transmitidas em cada intervalo de 1 hora do tráfego analisado. Por isso, cada histograma possui 24 barras, onde cada uma dessas barras representa um intervalo de 1 hora. Por exemplo, a primeira barra do histograma representa o intervalo entre 00h00min e 00h59min. O tamanho das barras do histograma é baseado em uma escala que vai de 0 ao número máximo de mensagens encontradas em uma única hora, ao longo de todo o tráfego analisado. Além disso, cada barra é dividida em várias seções, onde cada seção possui uma cor distinta, conforme mostra a Figura 4. Existem duas formas de se seccionar as barras do histograma: por versões ou por operações do protocolo SNMP. Dessa forma, o usuário poderá identificar as versões e operações predominantes no seu tráfego, além de ter conhecimento sobre o comportamento da quantidade total de mensagens por hora.

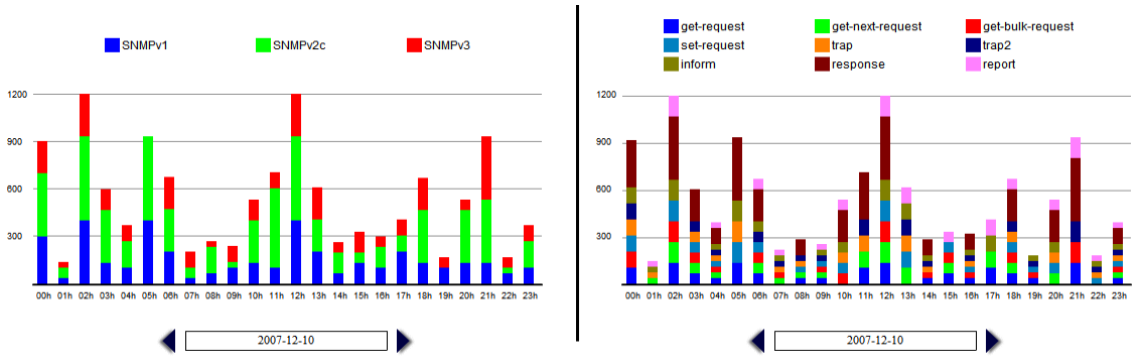


Figura 4. Histogramas seccionados por versões e por operações do SNMP

Um mecanismo de interação permite que o usuário saiba à qual dia o histograma

que está sendo apresentado se refere, assim como fornece a possibilidade do usuário navegar por esses dias, de modo a conhecer as quantidades de mensagens em intervalos de 1 hora referentes ao dia selecionado.

6. Resultados das Análises de Tráfegos SNMP

Como forma de validar as propostas deste artigo, foi implementada uma ferramenta Web de acordo com as especificações da arquitetura apresentada anteriormente. A ferramenta, chamada *Management Traffic Analyzer*, automatiza a execução da metodologia de medições sobre tráfego SNMP do IRTF, desde o processo de conversão de arquivos até as análises dos tráfegos submetidos e visualização dos resultados.

Para o desenvolvimento da mesma, foram aproveitadas as funcionalidades da ferramenta SNMPDUMP, desenvolvida em conjunto com a metodologia do IRTF para converter arquivos `pcap` para os formatos XML ou CSV e remover ou anonimizar informações sensíveis que possam estar presentes no tráfego. Dessa forma, se faz necessário que o SNMPDUMP esteja instalado no mesmo computador que irá receber a instalação do *Management Traffic Analyzer*. O núcleo da ferramenta foi desenvolvido na linguagem PHP, enquanto os scripts de análise de tráfego foram desenvolvidos em Perl, e as visualizações implementadas como aplicações Macromedia Flash através da linguagem ActionScript. Para armazenamento dos dados das análises foi utilizado o MySQL.

Já para validar as análises e visualizações implementadas no *Management Traffic Analyzer*, submetemos amostras de tráfego SNMP coletadas da Rede Nacional de Ensino e Pesquisa (RNP) à ferramenta, a fim de analisarmos os resultados gerados pela mesma. Essas amostras consistem num conjunto de 13 arquivos anonimizados, no formato CSV, com capturas de tráfegos realizadas entre os dias 22 de junho e 5 de julho de 2007.

Nas próximas subseções apresentaremos alguns dos resultados das análises sobre os tráfegos cedidos pela RNP. Foram escolhidos os resultados mais relevantes, uma vez que não seria possível a inclusão de todos os resultados devido à restrições de espaço.

6.1. Topologia da Rede de Gerenciamento

Ao se observar a topologia da rede de gerenciamento encontrada nos 13 arquivos de tráfego SNMP analisados, percebe-se que a mesma permanece relativamente invariável no intervalo de tempo onde esse tráfego foi monitorado. Tomando como exemplo a topologia observada no dia 22 de junho de 2007, identificamos que a rede possui um gerente principal que gerencia a maior quantidade de nós, e que também é o terminal que concentra a mais alta carga de tráfego (148,26 MB). Existe também um segundo nó que atua como um gerente menos importante da rede. A quantidade de nós conectados a esse gerente é bem menor do que a quantidade de nós conectados ao gerente principal, assim como a carga de tráfego nesse nó também é inferior (997,46 KB). Destaca-se ainda a existência de dois nós que gerenciam exclusiva e simultaneamente um único nó na rede.

A diferença mais perceptível que pode ser encontrada analisando-se os outros arquivos de tráfego SNMP da RNP é a aparição de um grupo de terminais formado por 1 nó gerente e cerca de 6 nós agentes conectados a esse gerente. Esse grupo aparece pela primeira vez no arquivo de tráfego relativo ao dia 27 de junho de 2007. A Figura 5 apresenta a visualização resultante das análises dos tráfegos da RNP relativos aos dias 22 e 27 de junho de 2007 (da esquerda para a direita, respectivamente).

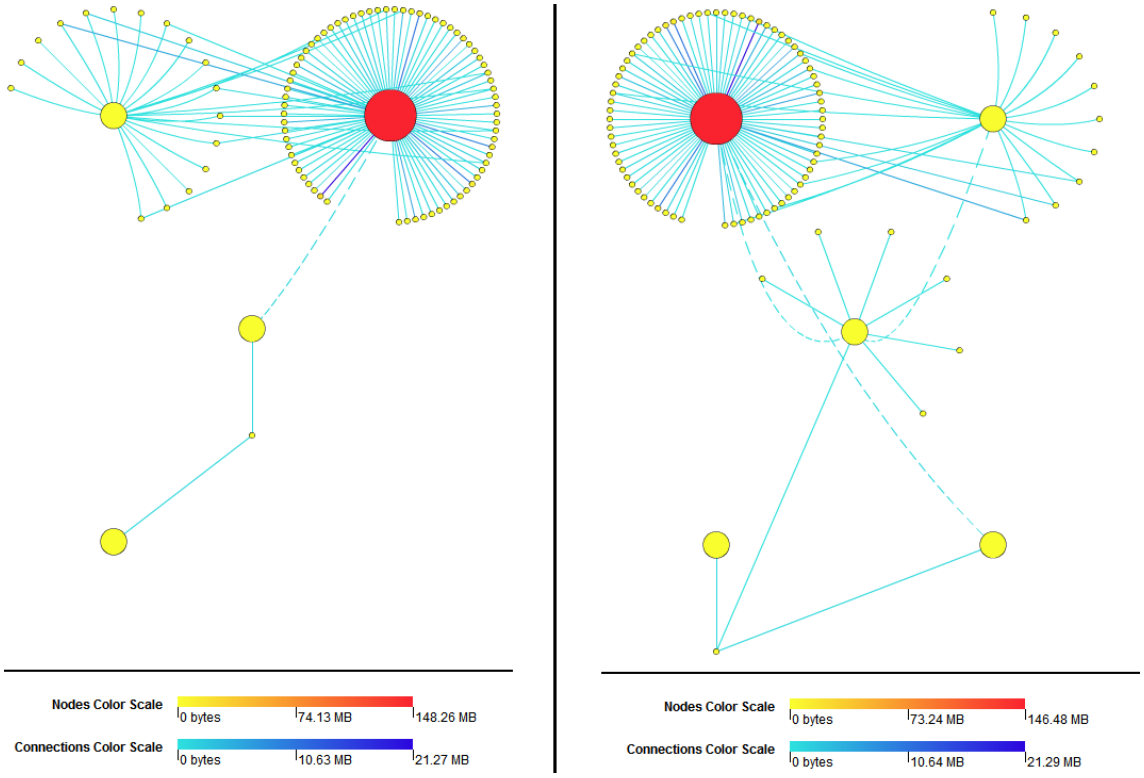


Figura 5. Topologias da rede de gerenciamento da RNP

Com relação às conexões representadas na topologia, percebemos que aquelas onde passa a maior quantidade de tráfego são as que conectam o principal nó gerente da rede com determinados nós agentes. A maior quantidade de tráfego registrada em uma conexão foi de 21,29 MB. No geral, cerca de 80% das conexões são do tipo CG/CR, enquanto cerca de 20% são do tipo NO/NR. Isso mostra uma preferência por parte dos administradores de rede em realizar *polling* nos dispositivos gerenciados, ao invés de utilizar notificações (*traps*).

6.2. Objetos SNMP Utilizados

Após a análise para identificação dos objetos SNMP encontrados nos tráfegos ter sido executada, observou-se que os objetos e a quantidade de mensagens associadas a estes são relativamente invariáveis nos arquivos de tráfego estudados. Devido a essa constância, será apresentada nesse artigo a árvore de objetos SNMP (*MIB Tree*) de apenas um dos arquivos de tráfego. Além disso, reduzimos o tamanho dessa árvore de modo a exibir apenas os 15 objetos SNMP mais representativos. O resultado dessa visualização pode ser observado na Figura 6.

A observação da árvore da Figura 6 mostra que os objetos SNMP mais utilizados são: IF-MIB::ifDescr (250.241 mensagens) e IF-MIB::ifType (245.464 mensagens). Provavelmente esses objetos são acessados para se verificar se o dispositivo consultado está operacional ou não, o que indicaria uma grande preocupação em testar o funcionamento dos dispositivos gerenciados da rede. Outros objetos que são intensivamente utilizados fornecem dados sobre o tráfego da rede. São eles: IF-MIB::ifOutUcastPkts (217.491 mensagens), ifInUcastPkts (217.479 mensagens), IF-MIB::ifHCOutOctets (203.737 men-

sagens) e IF-MIB::ifHCInOctets (203.730 mensagens). Muito provavelmente algum software de monitoramento de rede como o MRTG é o responsável pelos *pollings* feitos sobre esses objetos.

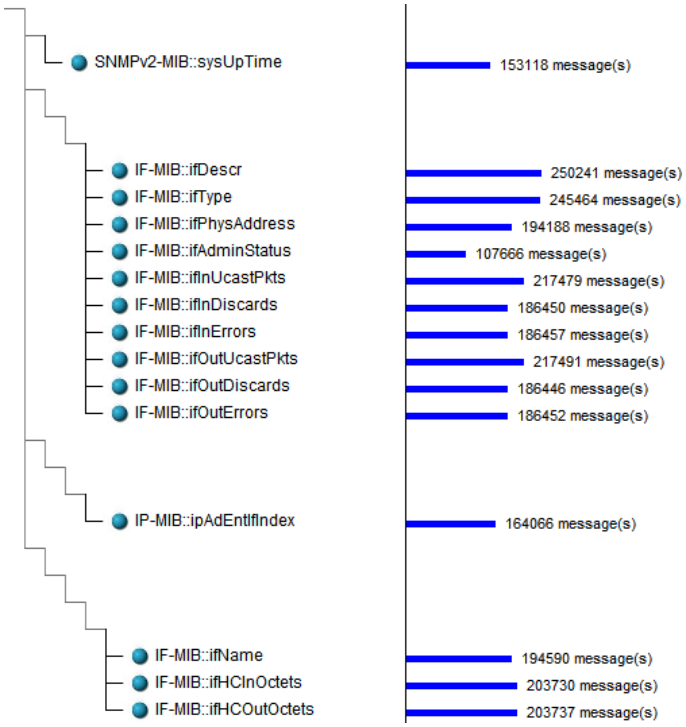


Figura 6. Os 15 objetos SNMP mais acessados na rede da RNP

6.3. Número de Mensagens SNMP por Intervalos de 1 Hora

De modo análogo à análise dos objetos SNMP presentes no tráfego (discutida na subseção anterior), a análise da distribuição da quantidade de mensagens SNMP em intervalos de 1 hora também se mostrou relativamente constante nos arquivos de tráfego estudados. Portanto, mais uma vez serão representados apenas os histogramas gerados a partir de apenas um dos arquivos fornecidos, correspondente ao monitoramento na RNP realizado no dia 28 de junho de 2007. Os histogramas resultantes dessa visualização podem ser vistos na Figura 7, seccionados por versões e por operações do protocolo SNMP (da esquerda para a direita, respectivamente).

A quantidade de mensagens é praticamente constante ao longo das horas completas em que houve monitoramento de tráfego. A última barra do histograma é menor de-vido à interrupção do processo de monitoração do tráfego de gerenciamento, que ocorreu em meados das 22h. O maior tráfego SNMP na rede registrado em 1 hora é de 101185 mensagens, observado às 6h do dia 28 de junho de 2007. Essa constância observada na quantidade de mensagens em todas as horas completas onde houve monitoramento mostra uma preferência dos sistemas de gerenciamento por consultas (*polling*) realizadas periodicamente nos dispositivos. Isso é confirmado também pelo grande número de mensagens do tipo *get-request*, *get-next-request* e *get-bulk-request*.

O histograma seccionado de acordo com as versões encontradas do protocolo SNMP mostram que o SNMPv2c é a versão predominante no tráfego. Em menor quantidade, encontra-se mensagens SNMPv1. Não foram encontradas mensagens SNMPv3, o

que constitui um fato interessante, pois teoricamente as versões SNMPv1 e SNMPv2c são históricas, enquanto que a versão SNMPv3 é considerada como o padrão em vigência.

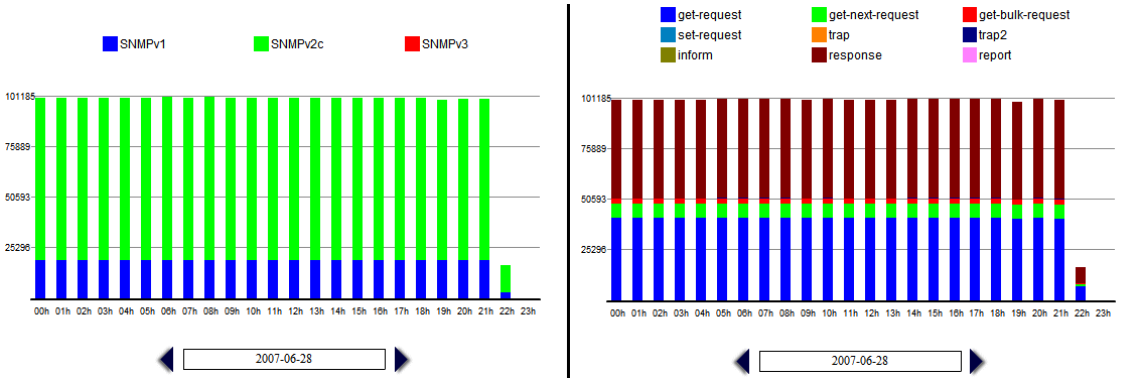


Figura 7. Histogramas do tráfego SNMP na RNP

Já o histograma seccionado de acordo com as operações do SNMP mostra que o tráfego é formado em sua maior parte por mensagens `response`. Em seguida aparecem as mensagens `get-request` e `get-next-request`, respectivamente. Em menor quantidade temos mensagens `get-bulk-request` e `trap2`. Mensagens de `trap` podem ser encontradas no tráfego, mas em quantidade insignificante, sendo esse o motivo de elas não serem representadas no histograma. O fato de mensagens `response` serem maioria no tráfego é justificável, pois esse tipo de mensagem sempre é emitida quando um dispositivo responde à consultas `get-request`, `get-next-request` ou `get-bulk-request`. Não foram observadas mensagens `set-request`, o que indica que o protocolo está sendo utilizado exclusivamente para monitoramento da rede, e não para configuração da mesma.

De um modo geral, podemos concluir que o tráfego SNMP da RNP é formado quase que em sua totalidade por mensagens periódicas de *polling* geradas pelos sistemas de gerenciamento utilizados na rede. O uso de notificações é raro, e praticamente não exerce influência sobre o tráfego como um todo. Por fim, o SNMP é utilizado exclusivamente para a realização do monitoramento sobre os dispositivos e serviços da rede. Muito provavelmente esse protocolo não é utilizado para configuração de dispositivos por questões de segurança.

7. Conclusões e Trabalhos Futuros

Atualmente, várias abordagens de gerenciamento de redes de computadores estão bem consolidadas, tais como as que empregam o protocolo SNMP, e algumas outras já se encontram em um estado avançado de desenvolvimento, como as que fazem uso de Web Services. Entretanto, a indústria e a academia ainda não conhecem de maneira satisfatória as características da utilização dessas abordagens no “mundo real”, ou seja, nas redes em produção.

Contudo, certamente existe espaço para melhorias no que diz respeito à metodologia de medições sobre o protocolo SNMP. A arquitetura de uma ferramenta Web proposta neste trabalho apresenta uma forma de integrar os vários softwares que anteriormente eram necessários para a execução dessa metodologia, acrescentar novas funcionalidade e, com isso, aumentar a acessibilidade e a eficiência desse tipo de estudo.

O emprego das três técnicas de visualização de informação propostas neste artigo tornou o trabalho de interpretação dos resultados das análises mais eficiente e eficaz. A partir dessas técnicas, o usuário da ferramenta poderá analisar os dados obtidos de forma mais natural, pois graças às características do sistema visual humano a descoberta de padrões e características interessantes se dará de forma simplificada.

Através da implementação da ferramenta *Management Traffic Analyzer*, demonstrou-se a viabilidade da implementação da arquitetura proposta. Além disso, a ferramenta implementada foi utilizada para realizar um estudo sobre arquivos de tráfego SNMP pertencentes à Rede Nacional de Ensino e Pesquisa (RNP), segundo a metodologia proposta pelo IRTF. Algumas características interessantes puderam ser confirmadas através desse estudo, tais como: preferência pela utilização de *polling* ao invés de notificações (*traps*), não utilização da versão 3 do protocolo SNMP e utilização de operações de forma predominantemente periódica.

Como trabalhos futuros, pretende-se ampliar o conjunto de técnicas de análise, visualização e comparação de tráfegos SNMP disponibilizadas na ferramenta *Management Traffic Analyzer*. Além disso, planeja-se buscar amostras de tráfegos SNMP oriundas de outras redes em produção, a fim de tornar os resultados obtidos mais representativos quanto à utilização atual do protocolo SNMP em outros contextos.

Referências

Ahn, S. J., Yoo, S. K., and Chung, J. W. (1999). Design and Implementation of a Web-based Internet Performance Management System Using SNMP MIB-II. *International Journal of Network Management*.

Becker, R. A., Eick, S. G., and Wilks, A. R. (1995). Visualizing Network Data. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):16–28.

Case, J. D., Fedor, M. L., and Schoffstal, J. D. (1990). Simple Network Management Protocol (SNMP). RFC 1157. [S.l.]: Internet Engineering Task Force, Network Working Group.

Oberheide, J., Goff, M., and Karir, M. (2006). Flamingo: Visualizing Internet Traffic. In *Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP*, pages 150–161, Vancouver, Canada.

Papadopoulos, C., Kyriakakis, C., Sawchuk, A., and He, X. (2004). CyberSeer: 3D audio-visual immersion for network security and management. In *VizSEC/DMSEC '04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, pages 90–98, New York, NY, USA. ACM Press.

Salvador, E. M. and Granville, L. Z. (2008). An Investigation of Visualization Techniques for SNMP Traffic Traces. In *IEEE/IFIP Network Operations and Management Symposium, 2008. NOMS '08. (aceito como short-paper)*.

Schoenwaelder, J. (2006). SNMP Traffic Measurements. *Internet Draft*.

Schoenwaelder, J., Pras, A., Harvan, M., Schippers, J., and van de Meent, R. (2007). SNMP Traffic Analysis: Approaches, Tools, and First Results. *Proceedings of the 10th IFIP/IEEE International Symposium on Integrated Network Management*.

Using Visualization Techniques for SNMP Traffic Analyses

Ewerton Monteiro Salvador, Lisandro Zambenedetti Granville
Federal University of Rio Grande do Sul
Institute of Informatics – II
Av. Bento Gonçalves, 9500 – Porto Alegre, RS – Brazil
Postal Code: 91501-970
Contact Phone: +55 51 8461-6794
{emsalvador, granville}@inf.ufrgs.br

***Abstract*—The network management area is currently dealing with huge amounts of information, which are produced, for example, by large scale and high-speed networks, heterogeneous devices, and monitoring and notification systems. Researchers and network administrators are frequently supported by information visualization techniques in the task of analyzing these large data sets. The Simple Network Management Protocol (SNMP) is the *de facto* standard for TCP/IP networks management. Despite its importance, there are no specific visualizations defined for SNMP traffic traces. In this paper we present a study on techniques for visualizing SNMP trace files, motivated by the fact that general purpose network traffic visualizations available today are not suitable for SNMP observation. Our proposed techniques have been prototyped in a software tool called Management Traffic Analyzer, which has been used to analyze and visualize SNMP traces.**

I. Introduction

Nowadays, network administrators and operators rely on complex management systems to support daily management processes. These systems can compute large sets of information about the network operations, re-organize and analyze them, and presenting such information to the network operator, in order to help him/her in managing the entire network. Due to that, the network management area is increasingly dealing with huge amounts of information.

Given the current data explosion that the various areas of the knowledge are dealing, information visualization techniques have been developed. Through these techniques, one can detect interesting features and patterns in an efficient and effective way. Once the network management area also faces an explosion of the data to be handled, the use of visualization techniques is of great importance to this area as well.

There are some investigations that present visualization techniques being used in the computer networks context. For instance, Oberheide *et al.* [1] presented a technique for helping the task of operating and managing a network by visualizing Internet traffic received in the form of Netflow feeds. Keim *et al.* [2] developed a visualization toolkit for understanding typical network communication activities in order to predict potential performance bottlenecks. Papadopoulos *et al.* [3] introduced 3D interactive representations and immersive spatial audio for abstracting huge amounts of network security and management information. Despite the fact that these works have addressed network traffic visualizations, all of them deal with general purpose network traffic.

There are situations where visualizations for a particular kind of traffic (e.g., management, VoIP and peer-to-peer traffics) are required. A good example of the need for specific management traffic visualizations can be observed in the approach for SNMP Traffic Measurements [4] developed by the Network Management Research Group (NMRG) of the Internet Research Task Force (IRTF). This approach specifies a methodology for discovering usage patterns for the Simple Network Management Protocol (SNMP) [5] by analyzing large scale SNMP traffic traces. Which features are being used, how SNMP usage differs in the various types of networks and organizations, which information is frequently queried, and what typical SNMP interactions are being performed are some of the questions that can be investigated. The amount of data generated by these SNMP traffic measurements potentially enables some human insights for identifying usage patterns, whereas these insights could be more easily achieved through the use of visualization techniques.

General purpose network traffic visualization techniques are not applicable for this kind of SNMP study, because they do not consider many elements that are intrinsic characteristics of SNMP. As examples of these characteristics, we can mention the SNMP message format (protocol version, varbind list, error-status, etc.) and the behavior of the basic SNMP operations (walks, get-bulk requests, get-next requests, set requests, etc.). As far as we are aware of, the only study approaching specific SNMP traffic visualization was carried out by *Salvador and Granville* [6], and only some preliminary results were published so far.

In this paper we take one step further the SNMP traffic visualization subject by presenting more detailed information regarding techniques for visualizing the results of analyses performed accordingly to the NMRG's approach for SNMP traffic measurements [4]. In addition, new SNMP traffic traces are studied using a tool developed for our investigation called Management Traffic Analyzer, and the data generated from

this study are shown through the visualization techniques introduced here.

The rest of this paper is organized as follows. Section II presents the background that supports our proposals. The visualizations techniques for the data obtained from the SNMP measurement approach are presented in Section III. Section IV introduces the results obtained from the analyses of some SNMP traces, while related work is discussed in Section V. Finally, conclusions and future work are provided in Section VI.

II. Background

In this section we first introduce the main concepts of the visualization area. Then, the IRTF's SNMP traffic measurements approach is described. Finally, a web tool for automating the IRTF's approach execution, which contains the implementations of the visualization techniques described in this work, is presented.

A. Information Visualization

Information visualization, or infovis, is a field of study that aims at supporting discovery and analysis of data through visual exploration [7]. The use of visualization techniques allows people to use the properties of the human visual system to explore and have insights into a data set, recognizing patterns, anomalies, and gaps in a faster and more effective way.

A visualization technique is composed of a visual representation and, commonly, of an interaction mechanism associated with it. The visual representation is based on a form of mapping the attributes of an abstract data structure to visual attributes. Usually, the level of abstraction of a visual representation is higher then the raw data, which is especially useful when someone is exploring large data sets. Interaction mechanisms, in turn, allow users to manipulate the visual representation in order to experience a faster and easier exploration of data sets.

B. SNMP Traffic Measurements Approach

In March 2006 the Network Management Research Group (NMRG) of the Internet Research Task Force (IRTF) released the internet draft called "SNMP Traffic Measurements" [4] [8] [9]. This draft presents the first systematic approach for researching the use of SNMP in production networks, based on traffic measurements. The objective of that research is to identify usage patterns of SNMP, like frequently queried information, volume of usually transmitted SNMP messages, approaches for reading object values used by network management tools, the relationship between periodic and aperiodic traffic in SNMP, among others.

That study is important in several scenarios, out of which we highlight two examples. The information obtained from the analysis is valuable for protocol and MIB (Management Information Base) designers because they would understand which features of SNMP are actually being used. Also, developers and researchers of network management tools would have the required information to identify potential improvements on their tools/researches [10].

C. Analysis Tool: Management Traffic Analyzer

The visualization techniques presented in this paper have been implemented in a prototype software tool called **Management Traffic Analyzer**. This tool automates the execution process of the SNMP traffic measurement approach, from file conversions to trace analyses.

An important feature of the Management Traffic Analyzer is the integration, in a single Web application, of the various steps defined in the NMRG's approach. Previously to our tool, one would need to use various pieces of software in order to convert and perform the analyses that he/she would like to do, such as using the SNMPDUMP [4] tool and a collection of Perl scripts.

The SNMPDUMP software can convert a raw PCAP file to the XML or CSV formats, as well as remove or anonymize sensitive data that may be present in traffic traces. Due to that, the Management Traffic Analyzer tool was implemented over the SNMPDUMP, invoking its functions whenever necessary. The analyses are performed by scripts written in the Perl language, and the results of these analyses are stored in a MySQL database. Finally, the visualization techniques have been implemented as Adobe Flash applications using the ActionScript language.

III. Visualization Techniques

Although the techniques presented in this section are not in fact new to the information visualization area, to the best of our knowledge there are no previous studies indicating whether they are appropriated for SNMP traffic visualization. The rest of this section presents three visualizations based on techniques that have been adapted in order to meet the specific needs of our SNMP traffic measurement study.

A. Visualization of Management Network Topologies

Computer networks, in general, belong to a class of phenomena whose underlying feature is that of connection. Visualization technics based on graphs are widely used in this class of phenomena, being very suitable for representing specific aspects of a computer network such as its topology [11] [12].

In order to produce a visualization of a management network topology, the adapted technique must identify, among the messages in the SNMP traffic trace, which terminals act as managers, agents, and both. We can identify these terminals by splitting the entire traffic trace into various SNMP message flows. Each SNMP message flow is composed of the set of messages between a source and destination addresses that belongs to a Command Generator (CG) / Command Responder (CR) relationship, or a Notification Originator (NO) / Notification Receiver (NR) relationship. The source of a CG/CR relationship acts as a manager, and the destination acts as an agent. On the other hand, in a NO/NR relationship, the source acts as an agent, and the destination acts as a manager.

The nodes' circles have two visual parameters with encoded information: size and color. The circle's size indicates the role that a terminal performed in the set of all message flows that it has participated. If it only acted as an agent, it will be represented by a small circle. However, if the terminal acted as a manager in any of the message flows, it will be represented by a larger circle, even if it acted as an agent in any other message flow. The circle's size also depends on the manager intrinsic messages that the node has originated/received: when the number of these messages increases, the circle's size increases too. This indicates that the larger the circle, the stronger the "manager role" of a terminal.

The circle's color represents the traffic load. If a terminal has a low traffic load, then it will have a lighter color. However, if a node has a high traffic load, then it will have a stronger color. The color varies accordingly to a traffic load scale, which goes from 0 to the maximum amount of bytes received and transmitted by a node on the traffic trace. This scale is then mapped into a color scale that varies from a light color (yellow) to a strong color (red), depending on the value of the traffic load scale.

The connection line is solid when connecting a manager to a node that acted as an agent during the whole traffic trace. A dotted line represents the connection between a manager to a node that also acted as a manager in another message flow of the traffic trace. An example of the nodes and connections representation is presented in Figure 1.

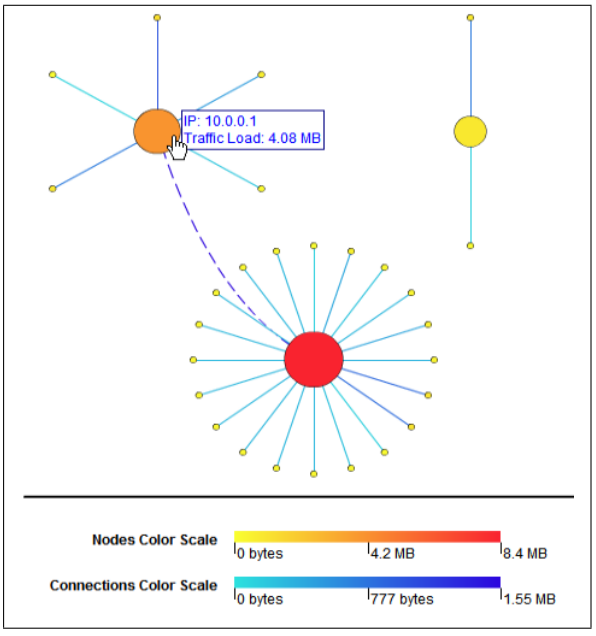


Fig. 1. Nodes and connections in the visualization technique

The drawn connections also carry traffic load information in their color. One more time, the lighter the connection's traffic load, the lighter its color (e.g., light blue). The heavier the connection's traffic density, the stronger its color (e.g., dark blue). Another color scale was used for the connections, because the traffic load of these connections may be of other order of magnitude if compared to the traffic load of the nodes.

Two interaction mechanisms were used in this visualization: nodes and connections caption, and scrolling. The captions inform to the user some relevant information about the nodes (e.g., IP address and traffic load) or the connections (e.g., type of the flow and traffic load). A user can visualize a caption by pointing the mouse over a node or a connection. In its turn, the scrolling mechanism prevents the truncation of the network topology representation, allowing the user to explore all the visualization independently from the size of the visualized management network.

B. MIB-tree Visualization

A possible analysis for an SNMP traffic trace is the computation of the number of times that an SNMP object is seen in the messages' varbind lists of this trace. This is a useful analysis because it is usually difficult to notice the hierarchical relationship between the SNMP objects of the traffic trace only by observing the OIDs (identifiers) of these objects. Since SNMP objects are organized in a Management Information Base (MIB) tree, it would be much easier if one could also visualize the result of this analysis in a MIB tree, like the one seen in Figure 2. Because of this, we proposed a visualization technique for presenting the analyses results in a way that the insight potential would be increased. This visualization is a combination of two other widely used visualization techniques: trees visualizations [13] and histograms [14]. Currently, many management applications use trees to present MIB modules and its objects. We take advantage of this already common structure and visualize the SNMP traffic in a tree as well, assuming that network administrators may be already familiar with SNMP objects organized in this way.

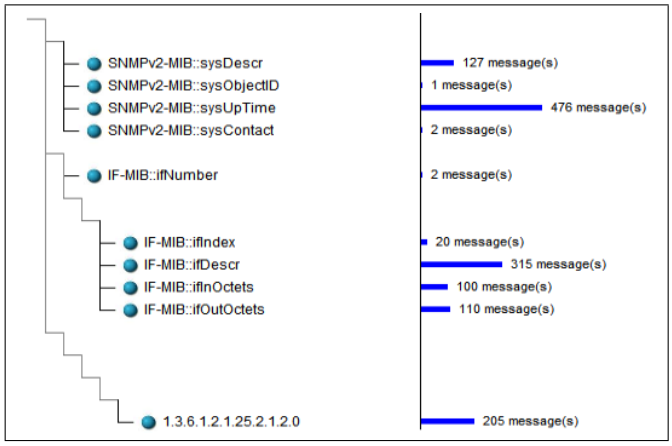


Fig. 2. MIB-tree visualization technique

The execution of the script that implements the SNMP objects analysis creates a list of objects, each one composed of an OID, a MIB object name (when this name is known by the script), and the number of messages that contains this object. Once the MIB tree is drawn, it is necessary to represent the number of messages related to each SNMP object in a histogram. This histogram's bars, which will represent the

number of messages, must be drawn at the right side of the leaf node that represents the SNMP object. The size of the bar will be based on a scale created in real time, where the minimum value is 0 and the maximum value is the biggest message number related to an SNMP object among the ones listed in the analysis result. Finally, the scrolling interaction mechanism must be used in this visualization again because very often there are a great number of SNMP objects in a SNMP Traffic Trace, which leads to big trees representations.

C. Visualization of SNMP Messages per 1-hour Intervals

In this last visualization, we used histogram graphics to present the results of analyses that provides the number of SNMP messages transmitted in 1-hour intervals. Each histogram has 24 bars, representing the 24 1-hour intervals of a day. For example, the first bar represents the time interval between 00:00h and 00:59h. Moreover, each bar is divided in various sections, where each section has a unique color, as shown in Figure 3. There are two ways of sectioning the bars of the histogram: by protocol versions or operations. In this way, the user will be able to identify the predominant versions and operations in his/her traffic, as well as to understand the behavior of the total amount of SNMP messages per hour.

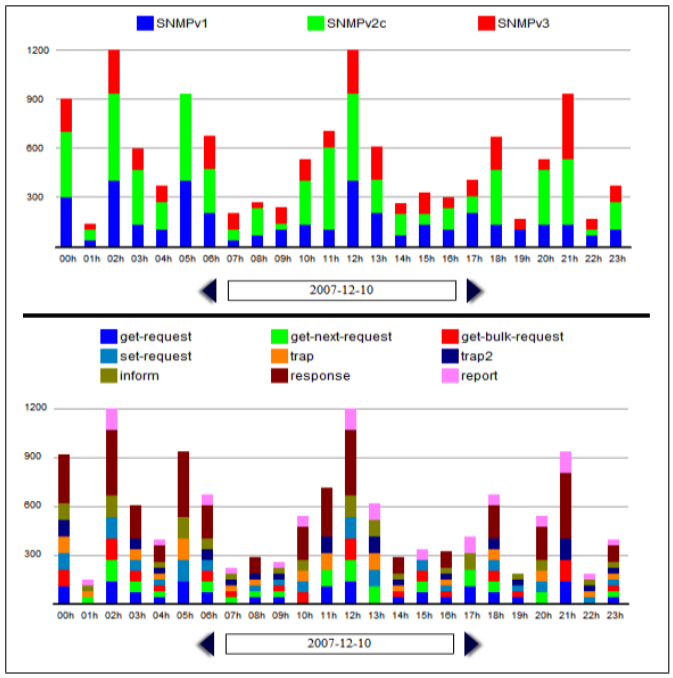


Fig. 3. SNMP Messages Per Hour Visualization

An interaction mechanism allows the user to see the day that is being visualized in the histogram, as well as to browse through the set of days available in the SNMP traffic trace, in order to visualize the number of SNMP messages per hour in the selected day. At the left and right sides of the text box that shows the day being represented, one can found buttons allowing him/her to go forward and backward in the traffic trace days.

IV. Analyses and Visualizations Results

In order to provide an example of the potential of use of the visualizations proposed in this work, we have employed them over the results of some measurements of a real production network. The results provided by the use of the visualization techniques are discussed in the rest of this section.

The SNMP traffic trace used in our study case was provided by the Brazilian Education and Research Network (RNP) backbone. This organization provides Internet access to public or private education institutions and research centers in Brazil. The RNP's operators provided a set of 13 anonymized SNMP traffic traces in the CSV format, each file containing 1 day traffic information. The traffic monitoring happened between June 22th and July 5th 2007. The processes of analyzing this trace and visualizing the results have been performed using the Management Traffic Analyzer tool presented before.

A. Management Network Topology

The topology analyses of the 13 traffic traces show that the RNP's management network topology is relatively constant in the whole time interval of the monitored traffic. The topology found in Figure 4 was obtained from June 27th traffic trace, and it shows the topology most commonly found in RNP's traffic traces set. One can identify that RNP's network has a main manager, which can be easily identified in the figure by the biggest circle. This manager monitors the biggest number of nodes, and it concentrates the biggest amount of traffic (146.48 MB). There is also two less important managers, with smaller number of managed nodes and small amount of traffic (less than 1.5 MB). We can also highlight the existence of two managers that monitors exclusively a single node of the network.

The connections which concentrate the biggest amounts of traffic are the ones connecting the main manager with its managed nodes. The biggest amount of traffic found in a single connection was 21.29MB. In general, approximately 80% of the connections belong to CG/CR type, while 20% belong to NO/NR type. This evidences the preference of the network administrators in using polling in the managed devices, instead of using notifications (traps).

B. Used SNMP Objects

After analyzing all the 13 traffic files of RNP's network, one could observe that the sets of SNMP objects found in each file were almost the same. Due to this, we will present in this subsection a MIB tree of only one traffic file. Also, the size of the tree was reduced in order to present the 15 most representative SNMP objects, due to space constraints. The result of this visualization can be seen in Figure 5.

By observing the resulting MIB tree, one would discover that the most used objects are: IF-MIB::ifDescr (250,241 messages) and IF-MIB::ifType (245,464 messages). These objects are probably accessed for verifying the status of the network devices. Other objects that are intensively used are: IF-MIB::ifOutUcastPkts (217,491 messages), ifInUcastPkts

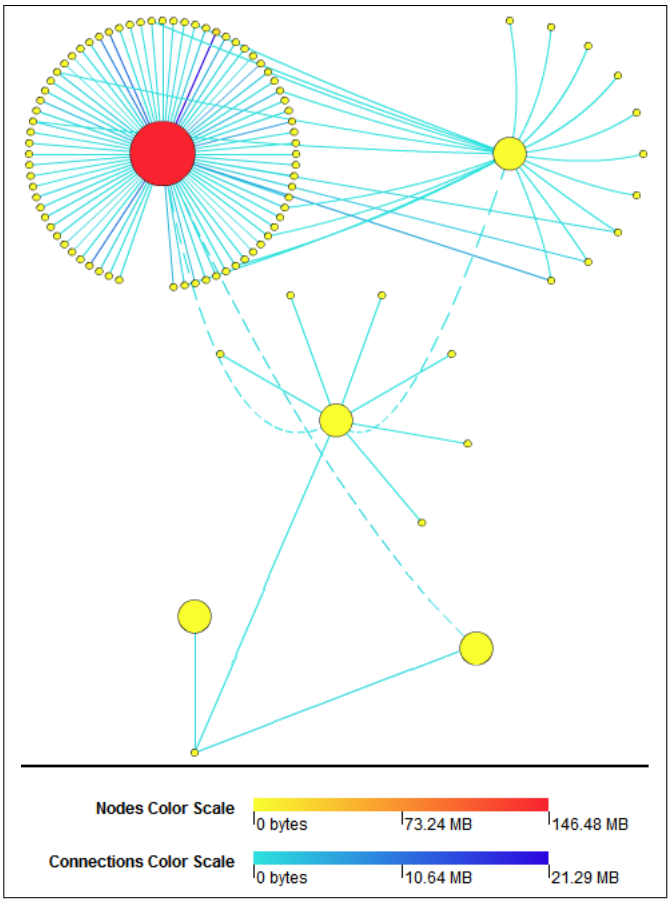


Fig. 4. Management Topology of RNP's network

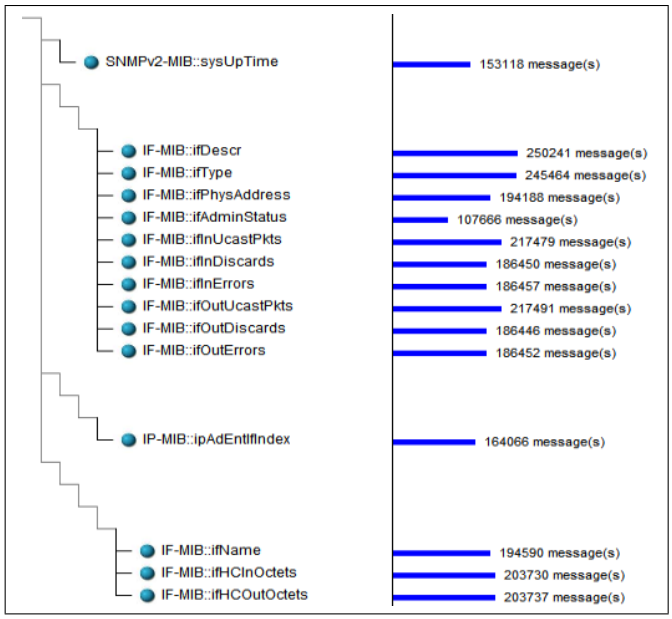


Fig. 5. The most accessed objects of RNP's network

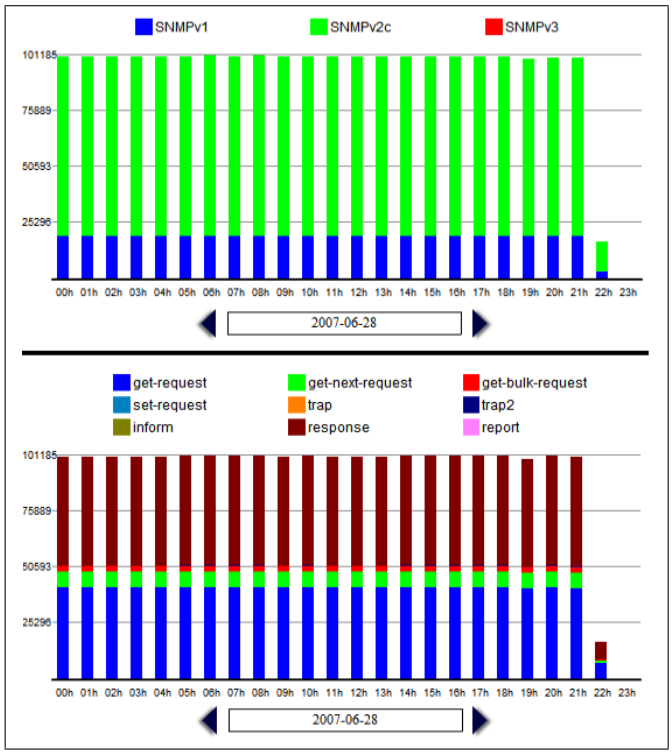


Fig. 6. SNMP Traffic Histogram of RNP's network

(217,479 messages), IF-MIB::ifHCOutOctets (203,737 messages) and IF-MIB::ifHCInOctets (203,730 messages). All these objects provide data of the network traffic.

C. Number of SNMP Messages per 1-Hour Intervals

The analyses of the distribution of SNMP messages per 1-hour intervals in the RNP's traffic traces demonstrated that each one of the 13 studied trace files are very similar to each other. Due to that we will again present only the histograms related to just one of the traffic traces, which was monitored in June 28th, 2007. The resulting histograms of this analysis are shown in Figure 6.

The amount of messages is pratically constant through the hours of the studied traffic file. The last bar of the histogram is smaller due to the interruption of the monitoring process, which occured approximately at 22h. The biggest amount of SNMP traffic is of 101,185 messages, which can be observed at 6h of June 28th, 2007. The constant number of messages in pratically every hour of the histogram shows a preference of the management systems for periodically polling the managed devices. This fact is also confirmed by the big number of `get-request`, `get-next-request` and `get-bulk-request` messages.

The histogram sectioned by SNMP versions shows that SNMPv2c is predominant in the traffic trace. SNMPv1 is also used, but with a smaller number of messages. There were no SNMPv3 messages in RNP's SNMP traffic, which is an interesting fact, since SNMPv1 and SNMPv2c are considered historic and SNMPv3 is considered the current official standard.

In its turn, the histogram sectioned by the operations of SNMP shows that RNP's traffic is composed mostly by `response` messages. `Get-request`

and `get-next-request` come in second and third places, respectively. In a smaller number, one can identify `get-bulk-request` and `trap2` messages. `Trap` messages can be found in the traffic trace, but its number is insignificant, which is the reason for this type of message not appearing in the histogram. The fact that `response` messages are majority in the traffic is reasonable, since this type of messages is sent in response of every `get-request`, `get-next-request` and `get-bulk-request` message. There were no `set-request` in the traffic trace, which indicates that the protocol is being used exclusively for monitoring the network, and not for configuring it.

V. Related Work

Nowadays, there are some published works dealing with visualization techniques for the network management area [1] [2] [3]. However, to the best of our knowledge, there are no work investigating specific visualization techniques for SNMP traffic traces.

Oberheide *et al.* describe a tool for helping the task of operating and managing a network by visualizing Netflow feeds. One of this work's aspects that are very similar to ours is the use of a combination of visualization techniques with navigation and data-filtering controls in a single tool.

Keim *et al.* [2] presented a visualization toolkit for understanding typical network communication activities and in anticipating potential performance bottlenecks. The use of visualization techniques for better understanding sniffed networks makes his work very similar to ours. The basic difference between these works is the fact of Keim's one being suitable for general network traffic, while our objective is to deal with specific SNMP management traffic.

The first results of the usage of the IRTF's approach for SNMP traffic measurements were published in the work of Schoenwaelder *et al.* [10]. The main difference between his work and ours is the chosen focus: the authors presented the results obtained from the use of IRTF's approach in a set of traffic traces collected from various networks, while our focus was to present visualization techniques for visualizing results originated from the application of this approach.

VI. Conclusions and Future Work

The use of visualization techniques has great potential in the network management area. The huge amount of information generated by the various devices and systems that compose the modern networks demand more and more efforts to make all management data friendlier for network operators, and management system vendors and researches. The IRTF's approach for SNMP traffic measurements is an important step and exposes even boldly the need for the development of specific visualization techniques. The new information about SNMP usage collected in real world production networks can be more easily observed with the employment of visualization techniques over the results of SNMP traffic analyses.

This work proposed an initial set of three techniques for visualizing specific information related to the IRTF's SNMP traffic measurements approach. These visualizations are able to graphically encode many of the information generated as result of the SNMP traffic traces analyses. We also demonstrated how the visualizations proposed in this paper can make the analysis of traffic trace measurements easier and more efficient.

The Management Traffic Analyzer tool is still under development. For future work, we plan to increase the set of visualization techniques available in the tool, in order to improve the insight potential over SNMP traffic traces. We also intend to add in the Management Traffic Analyzer tool support for measuring other types of management traffic (e.g., Web services traffic, ICMP traffic, etc.), in order to increase the knowledge of the usage patterns of network management traffic that can be found in the current and future production networks.

References

[1] J. Oberheide, M. Goff, and M. Karir, "Flamingo: Visualizing Internet Traffic," in *Network Operations and Management Symposium, 2006. NOMS 2006. 10th IEEE/IFIP*, Vancouver, Canada, Apr. 2006, pp. 150–161.

[2] D. A. Keim, F. Mansmann, J. Schneidewind, and T. Schreck, "Monitoring Network Traffic with Radial Traffic Analyzer," in *Visual Analytics Science And Technology, 2006 IEEE Symposium On*, Baltimore, Maryland, USA, Nov. 2006, pp. 123–128.

[3] C. Papadopoulos, C. Kyriakakis, A. Sawchuk, and X. He, "Cyberseer: 3D audio-visual Immersion for Network Security and Management," in *VizSEC/DMSEC '04: Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*. New York, NY, USA: ACM Press, 2004, pp. 90–98. [Online]. Available: http://dx.doi.org/10.1145/1029208.1029223

[4] J. Schoenwaelder. (2006) SNMP Traffic Measurements. Internet draft. [Online]. Available: http://www3.tools.ietf.org/html/draft-irtf-nmrg-snmp-measure-00

[5] J. D. Case, M. L. Fedor, and J. D. Schoffstal, "Simple Network Management Procotol (SNMP)."

[6] E. M. Salvador and L. Z. Granville, "An Investigation of Visualization Techniques for SNMP Traffic Traces," in *Network Operations and Management Symposium, 2008. NOMS 2008. 11th IEEE/IFIP*, Salvador, Brazil, Apr. 2008.

[7] J. D. Fekete and C. Plaisant, "Interactive Information Visualization of a Million Items," in *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, 2002, pp. 117–124.

[8] J. Schoenwaelder. (2007) SNMP Traffic Measurements. Internet draft. [Online]. Available: http://www3.tools.ietf.org/html/draft-irtf-nmrg-snmp-measure-01

[9] ——. (2007) SNMP Traffic Measurements. Internet draft. [Online]. Available: http://www3.tools.ietf.org/html/draft-irtf-nmrg-snmp-measure-02

[10] J. Schoenwaelder, A. Pras, M. Harvan, J. Schippers, and R. van de Meent, "SNMP Traffic Analysis: Approaches, Tools and First Results," in *Proc. IEEE 10th IFIP/IEEE International Symposium on Integrated Network Management*, Munich, Germany, May 2007, pp. 323–332.

[11] R. A. Becker, S. G. Eick, and A. R. Wilks, "Visualizing Network Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 1, no. 1, pp. 16–28, Mar. 1995.

[12] R. Spence and A. Press, *Information Visualization*. Addison Wesley, Dec. 2000.

[13] T. Hill, J. Noble, and J. Potter, "Visualising the Structure of Object-Oriented Systems," in *Visual Languages, 2000. Proceedings. 2000 IEEE International Symposium on*, 2000, pp. 191–198.

[14] Z. Konyha, K. Matkovic, D. Gracanin, M. Jelovic, and H. Hauser, "Interactive Visual Analysis of Families of Function Graphs," *Transactions on Visualization and Computer Graphics*, vol. 12, no. 6, pp. 1373–1385, 2006.