UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Jader Wallauer

DETECÇÃO DE COMPORTAMENTO NO SISTEMA CATARINENSE DE TELEMEDICINA

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação

Prof. Dr. rer.nat. Aldo von Wangenheim

Florianópolis, outubro de 2008.

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Jader Wallauer

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciências da Computação, área de concentração Sistemas de Computação, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

_____

Prof. Dr. Frank Augusto Siqueira

Coordenador do Curso

Banca examinadora

_____

Prof. Dr. rer.nat. Aldo von Wangenheim

Orientador

_____

Prof. Dr. Ricardo Felipe Custódio

_____

Profa. Dra. Christiane Anneliese G. von Wangenheim

_____

Prof. Dr. Luiz Felipe de Souza Nobre

Ao Professor Aldo, pela oportunidade de fazer parte do projeto Cyclops e desenvolver um grande trabalho.

A todos os integrantes do Cyclops pelo apoio, conversas, e trocas de experiência.

Sou imensamente grato ao meu pai Jordan Paulo Wallauer por acreditar em mim e estar sempre ao meu lado ter tido o trabalho de corrigir o meu português bem como a forma deste trabalho, dando um pouco de sua experiência.

Sou grato a minha mãe Zuleika Wallauer (in memoria), que infelizmente não pode me acompanhar nos passos da vida e que sei estar orgulhosa por tudo que já passei.

A minha gratidão eterna a minha mãe Martha T. B. Wallauer que assumiu corajosamente um papel que somente ela poderia cumprir melhor que qualquer outra

pessoa que pude conhecer e que conhecerei em minha vida, sendo a mãe dos sonhos, que sempre com um carinho, palavra amiga, atenção e amor incondicional, cuidando de mim e do meu irmão. Sem ela seria impossível sequer pensar em estar neste ponto em minha vida.

A meu grande amigo Daniel “Caju” D. Abdala, que se incluem na classe de irmão, e que estive sempre ao meu lado desde o início.

A meu grande irmão Jaury Wallauer, um exemplo em minha vida.

Agradeço de forma especial a Daniela Dencker Leal Wallauer, amiga, companheira, colaboradora e incentivadora. Sem ela seria muito difícil dar mais este passo.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS

ADT	*Anomaly Detection Techniques*
FN	Falsos Negativos
FP	Falsos Positivos
HP	Hewlett-Packard
HTTP	*Hypertext Transfer Protocol*
IBL	*Instance Based Learn*
IDS	*Intrusion Detection System*
IP	*Internet Protocol*
Mbps	Mega bits por segundo
PHP	*Personal Home Page*
RAID	*Redundant Array of Independent Disks*
RAM	*Random Access Memory*
RNP	Rede Nacional de Pesquisa
SATA	*Serial Advanced Technology Attachment*
SCSI	*Small Computer System Interface*
SQL	*Structured Query Language*

SRI	*Stanford Research Institute*
SSL	*Secure Sockets Layer*

RESUMO

Este trabalho apresenta uma pesquisa de métodos para detecção de comportamento, assim como, descreve o desenvolvimento de um *framework* para a detecção de comportamento de usuários em sistemas web e seqüência de ações suspeitas que possam indicar um uso incorreto do mesmo ou uma possível falha de segurança. Deste modo buscou-se incrementar a segurança de sistemas web e avaliar a aplicação da técnica probabilística de Mahalanobis nesse contexto.

O *framework* proposto foi desenvolvido utilizando como métrica de similaridade, a distância de Mahalanobis, de forma a validar suplementarmente, possíveis benefícios advindos de métricas estatísticas na comparação de assinaturas de comportamento.

Foi executada uma extensa validação tendo como base protótipo do sistema aqui descrito, para funcionamento em conjunto com o portal de Telemedicina do Estado de Santa Catarina, que também serviu de base de coleta de dados para validação.

Como resultado obteve-se um *framework* modular que pode ser facilmente portado para diferentes sistemas web que possuam características similares ao Portal de Telemedicina. Obteve-se, ainda, na utilização da técnica proposta como forma de detecção de comportamento, um baixo índice de detecção de comportamento válido marcado como inválido (Falso Negativo) e também, baixos índices de detecção de comportamento inválido marcadas como válido (Falsos Positivos).

Palavras-chave: Detecção de Comportamento, Sistemas Online, Assinaturas de Comportamento, Segurança de Sistemas, Métrica Estatística.

ABSTRACT

This work presents a research about detection of behavior methods, as well as it describes the development of a framework for user behavior detecting in web systems and sequence of actions that may indicate a suspected misuse or a possible security breach.

The proposed framework was developed using dissimilarity metric, the Mahalanobis distance, to supplementary validate, potential benefits derived from statistical metric comparison for behavior signatures.

An extensive validation was performed based on the prototype system described here, for operation in conjunction with the Telemedicine portal of the State of Santa Catarina, which also served as the basis for collection of validation data.

As results, we have obtained a modular framework that can be easily adapted to different web based systems presenting a given set of features similar to the Telemedicine Portal. The proposed behavior detection technique was able to achieve a low detection rate of valid behavior marked as invalid (False Negatives) and also a low detection rate of invalid behavior marked as valid (False Positives) .

Key-words: Behavior Detection, Online System, Behavior Signature, System Security, Statistical Metrics.

Capítulo 1

INTRODUÇÃO

Através dos anos, sistemas web evoluíram a partir de meras paginas pessoais para sistemas computacionais complexos seguindo a evolução das redes de computadores e sua onipresença. A internet e os sistemas web se tornaram ferramentas importantes a serviço de empresas, bancos, governos e milhares de usuários que os utilizam diariamente. Sistemas IDS (do acrônimo – *Intrusion Detection Systems*), ou seja, sistemas de detecção de intrusão, que são ferramentas responsáveis pelo monitoramento e vigilância de redes de computadores. Estes sistemas analisam as atividades de rede, computadores e usuários, procurando por elementos que indiquem um possível comportamento malicioso ou funcionamento malicioso de algum sistema. Basicamente o objetivo de um IDS é analisar eventos ocorridos em sistemas computacionais, ocasionado por outros sistemas e/ou usuários, e detectar ataques (Micarelli & Santonetti, 2007).

Os primeiros sistemas que apresentaram maior nível de segurança, ou seja, utilizavam algo além da tupla de autenticação, usuário e senha, foram concebidos sob tecnologias como *Secure Socket Layer* (SSL) (Weaver, 2006) e/ou em conjunto com a Protocolação Digital, onde, a primeira, busca garantir confiabilidade nas transações entre o sistema e seus usuários, através de um canal criptográfico que faz com que a mensagem enviada do servidor para o cliente e do cliente para o servidor torne-se praticamente ilegível para qualquer receptador externo, e, a segunda, trata da integridade dos dados armazenados ou enviados para o sistema, ou seja, o conjunto de dados utilizado pelo sistema deve ser mantido sem alteração em seu conteúdo, podendo ser aferidos a qualquer instante com o uso do *ticket* (certificado enviado por uma protocoladora digital) recebido pela protocoladora (equipamento responsável pela certificação digital) devidamente certificada (Brocardo et all, 2008).

Ataques para roubo de informações são usualmente praticados contra o usuário (ITL, 1999), parte mais fraca do sistema, pois este, geralmente, faz uso de senhas de acesso que podem ser facilmente quebradas. Isso acontece, em virtude da utilização de

programas maliciosos elaborados para obter informações dos computadores pessoais como, por exemplo, Cavalo de Tróia e *Spywares*.

Programas do tipo Cavalo de Tróia (*Trojans*) são disfarçados de programas legítimos e instalados diretamente no computador do usuário, fazendo com que a pessoa que o criou ou enviou possa controlar ou apenas observar as ações daquele. Alguns exemplos famosos de *trojans* são o Keylogger que armazena e envia para o "dono" do *trojan* todos os dados digitados no teclado; e o Backdoor que visa o total controle da máquina do usuário (Franz, 2007).

Já os programas espiões (*Spywares*), são responsáveis por "espionar" enviando informações pessoais sem autorização, como senhas, números de contas bancárias, entre outros dados que podem comprometer a segurança do usuário (Ames, 2004).

Existem, ainda, os *Cookies,* ferramentas disponíveis nos navegadores de internet mais modernos. São utilizados para manter informações de autenticação de requisições HTTP e informações do usuário, quando este acessa sites dos mais variados tipos, não necessitando repassar todas as informações para uma nova requisição (Liu et al., 2005).

Sistemas de pesquisa como o Google e sistemas de comércio eletrônico entre outros, utilizam *cookies,* para guardar informações dos usuários. Tais informações podem atuar como meio de rastreamento de atividades, já que podem manter o histórico de navegação e utilização do navegador, comprometendo tanto a privacidade como a segurança.

Nos sistemas online, a detecção de comportamento apresenta-se mais e mais como uma necessidade para garantir a segurança dos dados, tal como será definido a seguir.

1.1 DEFINIÇÃO DO PROBLEMA

À medida que os sistemas web tornam-se mais práticos, rápidos e acessíveis, disponibilizando uma gama de funções para usuários, eles também impõe uma nova linha de dificuldades no que tange a quesitos de segurança, ou seja, fazer com que as

informações trocadas ou armazenadas sejam acessadas apenas pelas pessoas com a devida autorização.

Uma das formas de se induzir maior grau de segurança nesses sistemas, é controlar as diversas ações que os usuários podem tomar no decorrer da sua utilização, definindo desta forma, a área de pesquisa: detecção de comportamento (Kemmerer & Vigna, 2002).

A detecção de comportamento permite o controle de todos os atos executados pelos usuários, sinalizando, desta maneira, comportamentos suspeitos ou inesperados para um dado usuário, baseado em comportamentos prévios ou definidos como padrão.

Um sistema de telemedicina deve possuir maneiras rápidas e eficientes de verificar o usuário que o está utilizando, como usuário e senha, porém, por se tratar de informações sigilosas e cruciais para a decisão médica, existe a necessidade de um controle maior de sua utilização e acesso (Wallauer et al, 2008). Algumas alternativas seriam a utilização de leitura de íris (Perez et al., 2007) ou ainda leitura de digital (Younhee et al., 2003), mas, ambos necessitam de *hardware* específico e com custo adicional para as estações de trabalho e tornando o acesso restrito apenas à maquinas que contenham tal aparelho.

Este trabalho visa ao estudo e desenvolvimento dessa tecnologia com o foco de aplicação nos sistemas de saúde, conforme será discutido adiante.

1.2 MOTIVAÇÃO

O comportamento de um indivíduo pode ser medido e analisado com base no seu histórico de navegação. Permitindo que seja detectada uma possível invasão do sistema ainda nos estágios iniciais de sua ocorrência, o que possibilita uma intervenção automática ou assistida, onde o sistema responde a invasão, sem consultar, ou com consulta ao administrador do mesmo, para determinar a resposta necessária ou suficiente para cada caso.

Este trabalho foi motivado pela necessidade de se auferir maior segurança ao sistema on-line de gerenciamento de consultas e exames do Estado de Santa Catarina, O modelo aqui proposto foi desenvolvido, no entanto, de modo a torná-lo adaptável para virtualmente qualquer sistema online.

Pretende-se com ele, investigar de maneira sistemática a área de detecção de comportamento e, potencialmente, propor inovações nos métodos comumente utilizados.

1.3 OBJETIVOS DO TRABALHO

1.3.1 Objetivo Geral

Este trabalho visa a proposição de um *framework* para detecção de comportamento anômalo de usuários em sistemas de web.

1.3.2 Objetivos Específicos

1. Incrementar a segurança em sistemas web;
2. Desenvolver um modelo de arquivamento de ações alternativo ao modelo de arquivo de auditoria (log);
3. Acompanhar a evolução das atividades dos usuários de sistemas web;
4. Propor um modelo estatístico de te tratamento de informações para detecção de comportamento;
5. Avaliar o número de casos necessários e suficientes para utilização de métricas estatísticas.

1.4 HIPÓTESE

Verificar a possibilidade de estudo de comportamento e identificação de detecção de intrusão ou anomalias de uso do sistema, com base em comparações feitas entre as ações de um usuário no sistema e a sua assinatura, gerada a partir de um histórico de comportamento.

1.5 MATERIAL E MÉTODOS

Para a elaboração do *framework* de detecção de comportamento de usuários, objetivo deste trabalho, utilizou-se como base de estudo o Portal de Telemedicina desenvolvido pelo Cyclops Group na Universidade Federal de Santa Catarina em parceria com a Secretaria de Estado da Saúde de Santa Catarina (Maia et al., 2006).

Esse Portal é constituído, basicamente, por um sistema desenvolvido em ambiente web que possibilita acesso aos dados de todos os pacientes, instituições de saúde, e profissionais de saúde que interagem diretamente com a Rede Catarinense de Telemedicina, que será mais bem descrito no Anexo 2.

Para tanto, um conjunto de 969 profissionais de saúde (total de profissionais que utilizam o Portal de Telemedicina entre o período de janeiro de 2008 e dezembro de 2008) tiveram todo o histórico de utilização do Portal, armazenado em um banco de dados, utilizado posteriormente por esse trabalho.

Um total mínimo de 9.000 ações foram analisadas para cada profissional que utiliza o Portal de Telemedicina, limitando o conjunto de usuários do sistema à 15 indivíduos que possuíam maior volume de ações, o motivo deste limite está descrito no Capítulo 4.

Do conjunto de dados analisados, metade foi usado para gerar o padrão (assinatura) de cada indivíduo, e o restante, para testar o grau de erro e/ou acerto dessa assinatura, durante o processo de validação da metodologia proposta.

No contexto desse projeto foi utilizado um servidor com Processador AM2 3500Ghz, 2GB de memória RAM, 6 discos SATA 320GB em RAID 5 e 2 placas de rede Gigabit (1000 Mbps), com apache versão 2.0 e PHP 5.2.0 e um servidor HP modelo DL380 G4, 2x Processador Intel 3400GHz com um núcleo, 2GB memória RAM. 6 discos SCSI 300GB em RAID 5, 2 placas de redes Gigabit com banco de dados PotgreSQL versão 8.2, ambos ligados através de um switch Baseline 2825 – 3com 10/100/1000 e com link de saída para a internet de 100 mbps direto com a RNP.

Tratando-se de ambiente web, os testes foram realizados nos dois navegadores de internet mais utilizados nos dia de hoje, o Internet Explorer 7.0.6001.1800 (Internet Explorer, 2008) e o Mozilla Firefox 2.0.0.14 (Mozilla Firefox, 2008).

Foi utilizado a ferramenta pgAdmin III versão 1.8.2 (pgAdmin, 2008) para visualização e criação das tabelas de banco de dados, verificação de funcionamento de consultas para obtenção de informação contida no histórico de utilização do Portal de Telemedicina.

Também foi criado uma média geral de erros/acertos, e foi mensurada o grau de confiança geral da técnica proposta.

Os dados utilizados neste trabalho provém do Portal de Telemedicina do Estado de Santa Catarina. Um total de 598.296 ações realizadas por 969 profissionais que operam diariamente este sistema, no período de janeiro de 2008 à dezembro de 2008, serviram como base para este estudo.

A partir do conjunto de ações armazenados, foi gerada uma assinatura, ou, padrão de utilização do sistema utilizando métrica estatística de Mahalanobis (Mahalanobis, 1936).

Após a obtenção das assinaturas dos usuários, um conjunto de dados sintéticos foi gerado a partir de todo o conjunto original, reproduzindo ações do usuário a fim de testar a eficiência da detecção de comportamento, como descrito na sessão **Erro! Fonte de referência não encontrada.**.

Foi utilizado um artefato de software conectado a cada parte mapeada do sistema, coletando ações, horário, endereço IP do cliente, usuário em questão, sendo estas as informações necessárias para a análise do comportamento do usuário. Já que para uma análise de comportamento é necessário armazenar todas as ações, com a componente temporal, a fim de ordená-las e distanciá-las entre si, e a componente IP, como possível forma de identificar o computador de origem dos eventos.

1.6 CONTRIBUIÇÕES

Durante a realização deste trabalho, contribuições relevantes à área de aplicação da pesquisa foram desenvolvidas. Dentre elas:

- abordagem de uso de Banco de Dados para arquivamento de ações de usuários em contraste com o modelo de mineração dos *logs* gerados pelo servidor web;
- acoplamento de módulo de software para a captura dos eventos de usuários em contraste com os modelos *code embedded* (código embutido, onde o código faz parte do aplicativo, misturado internamente com o código original), e análise de arquivos de *log* (Arquivos de *log*, são arquivos que armazenam informações provenientes da utilização dos sistemas, conforme configurado pelo administrador do sistema alvo) comumente utilizados;
- uma prospecção para modelos estatísticos de tratamento de informações no contexto de segurança em redes, em específico para detecção de comportamento;
- disponibilização para o meio acadêmico de um *framework* para análise de comportamento baseado em histórico de ações de usuários.

1.7 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em 4 partes como segue:

Parte I – Introdução

Definição do problema de detecção de comportamento, a motivação deste trabalho, os objetivos, tanto o geral, quanto os específicos, a hipótese que delineou toda a pesquisa, o material e os métodos utilizados, e as contribuições relevantes. Compreende apenas o capítulo 1.

Parte II – Fundamentação Teórica

Nesta parte, encontra-se toda a fundamentação teórica não trivial necessária ao entendimento do trabalho, onde as técnicas comumente utilizadas para detecção de comportamento são apresentadas de forma geral, assim como é apresentada uma opção de abordagem estatística para o problema. Em todo o capítulo, são apresentadas

referências, em caso de necessidade de aprofundamento de assuntos de interesse do escopo apresentado. Compreende o capítulo 2.

Parte III – Metodologia de Desenvolvimento

Apresentação de todos os passos necessários para a geração da assinatura comportamental de cada usuário de um sistema web, descrição de cada etapa de desenvolvimento da abordagem proposta, estratégias de mapeamento e registro das ações do usuário, estratégia de adaptação e normalização dos valores armazenados para a técnica proposta, método de criação da assinatura (padrão) individual dos usuários do sistema e, por fim, descreve-se como utilizar a assinatura para verificar se as ações executadas por um usuário pertencem ao seu padrão de comportamento. Compreende o capítulo 3.

Parte IV – Validação, Resultados e Conclusões

Nesta parte do trabalho são apresentados os testes que foram realizados para verificar o grau de acerto da técnica proposta, assim como os resultados obtidos até o fechamento desta pesquisa e as conclusões que foram possíveis a respeito do assunto. Compreende os capítulos 4 e 5.

Capítulo 2
FUNDAMENTAÇÃO TEÓRICA

Para o desenvolvimento deste trabalho, foi executada uma revisão da literatura de modo a fundamentar a idéia inicial de detecção de comportamento com modelo modular para acoplamento do *framework* de maneira relativamente fácil à, potencialmente, qualquer sistema web.

Nesta seção apresentamos os principais tópicos teóricos relevantes para tal trabalho, iniciando com uma fundamentação detalhada da detecção de comportamento, sua relevância e dificuldades inerentes, comumente encontradas. Apresentamos a seguir a detecção de comportamento no momento atual e finalizamos com a métrica estatística como uma opção de medida de desvio comportamental em sistemas web.

2.1 DETECÇÃO DE COMPORTAMENTO

Visando melhorar, não somente, mas inclusive, a segurança de sistemas web, foram criadas técnicas para detecção de comportamento, também conhecidas como ADT – *Anomaly Detection Techniques* (Patcha, 2007).

Nos primórdios da detecção de intrusão, era dos administradores a difícil e repetitiva tarefa de monitorar, através de um console, as atividades dos usuários. Apesar de suficiente para a época, essa forma de detecção de intrusão era possível apenas para pequenos sistemas com poucos usuários, ou seja, não escalável. Segundo kemmerer e Vigna que fizeram uma revisão sobre detecção de comportamento, no final dos anos 70, início dos anos 80, os administradores literalmente imprimiam arquivos de auditoria para analise por inspeção visual e, desta forma, encontrar evidencias de invasão ou comportamento malicioso. Em pouco tempo formava-se uma pilha de papel com um metro ou mais, tornando o trabalho de auditoria uma tarefa obviamente demorado, repetitivo e altamente suscetível a falhas. Por conta do conjunto consideravelmente grande de informações contido em arquivos de auditoria, eles passaram a ser utilizados apenas como método forense para determinar o que causou um incidente particular de

segurança, o que inviabilizava detecção em tempo real, ou seja, a detecção da invasão ou anomalia no momento em que ele ocorre.

Kemmerer e Vigna ainda relatam que, nos anos 90 pesquisadores desenvolveram sistemas de detecção que faziam a revisão dos arquivos de auditoria automaticamente e em tempo real. Possibilitando a detecção e resposta a ataques no momento em que eles estavam ocorrendo, ou ainda prevenindo ataques antes mesmo que eles ocorressem, utilizando duas formas difundidas de realizar detecção de comportamento conhecidas como detecção de anomalia e detecção de mau uso (Kemmerer & Vigna 2002).

Detecção de mau uso – corresponde a detecção de ataques previamente conhecidos, onde os dados coletados após um ataque são utilizados como base de conhecimento (assinatura) para detecção do mesmo conjunto de ataques no futuro. Esta metodologia, no entanto, apresenta uma falha grave, uma vez que não se conhece todos os possíveis ataques e que novos ataques são criados diariamente. Já a abordagem de detecção de anomalia tem como objetivo a prevenção do ataque. Ao invés de modelar um comportamento anormal (ataque), modelamos o que seria um comportamento normal, assim, em comparação do modelo de cada usuário, podemos determinar quais ações futuras pertencem a este usuário e quais são anomalias de comportamento (Micarelli & Sansonetti, 2007).

Um sistema de detecção de comportamento é um artefato de software que visa identificar comportamentos recorrentes de um dado usuário. Tal sistema pode ser parte de um sistema maior, ou simplesmente um módulo ou componente (Balajinath & Raghavan, 2001). Normalmente aplicando à segurança de sistemas web pode, ainda, possuir uma funcionalidade adicional no que se refere à identificar como anômala uma dada sessão de utilização do sistema. Com base em assinaturas de comportamento previamente registradas para um dado usuário, pode disparar um gatilho para a sinalização de uma possível intrusão ou utilização errônea do sistema que pode tanto alertar o administrador do sistema, como responder automaticamente a ameaça de acordo com diretrizes anteriormente programadas.

2.1.1 Estado da Arte em Detecção de Comportamento

Na literatura existem diferentes abordagem para o tema detecção de anomalia. As mais comuns dizem respeito a métodos estatísticos, mineração de dados, redes neurais e sistema especialistas (Noel et al., 2002).

Métodos estatísticos estão entre os mais difundidos e mais utilizados. Eles se amparam em um conjunto de variáveis que alteram o valor com o passar do tempo e que são capturadas a partir da interação entre o usuário e o computador ou entre o usuário e a rede. Com base nessas variáveis é montado um perfil de utilização do sistema durante sua operação padrão (normal). Desvios de comportamento associados a este perfil serão considerado uma anomalia e marcado como tal. Parâmetros como entrada, saída, tempo de utilização, recursos com os determinados períodos de tempo, são utilizados nessa abordagem (Micarelli & Sansonetti, 2007).

Alguns sistemas associados aos métodos estatísticos são EMERALD (Porras, Neumann, 1997) e IDES (Javitz & Valdes, 1991). EMERALD foi desenvolvido no Stanford Research Institute (SRI), com o objetivo de detecção de intrusão em grandes redes de computadores focado em escalabilidade de sistema. As técnicas utilizadas por ele configuram um sistema hibrido de detecção de comportamento, utilizando simultaneamente detecção de anomalia e detecção de mau uso através de um sistema baseado em regras e um detector de anomalia com base em medida estatística.

O IDES, também do SRI, assim como o EMERALD, é um sistema hibrido que utiliza arquivo de auditoria de sistema como forma de capturar e caracterizar as atividades individuais de cada usuário e detectar desvios do comportamento esperado. Ele utiliza os dados referentes a entrada e saída do sistema relacionado ao horário e ainda o conjunto recursos do sistema que foram utilizados.

Outras propostas são mostradas em Carmo e Costa (2007), onde é apresentada uma o uso do conceito de confiança para controle de acesso e identificação de usuários em sistemas web, com base em um experimento realizado em ambiente de teste. Além de propor o uso de algumas técnicas lineares e estatísticas para calcular as distâncias de

similaridade entre os padrões, o autor também as utiliza em conjunto para o cálculo do índice de confiança.

Para tanto, o autor utiliza três técnicas bem definidas: Cadeias de Markov, Distância de Levenshtein e Distância de Frobenius.

Cadeias de Markov, é um modelo estatístico que depende somente do estado presente para determinar o próximo estado, sem levar em consideração os passos que levaram o processo para o estado atual, podemos descrever como um grafo onde os o conjunto de vértices alcançáveis possuem a mesma probabilidade de ser o próximo estado, sem levar em conta nenhuma das transições envolvidas no processo. Para saber mais (Mohamed & Gader, 2000).

Distância de Levenshtein, por sua vez, é uma medida linear que calcula a distância de transformação entre duas cadeias de caracteres, onde através de sucessivas transformações no primeiro conjunto de caracteres para formar o segundo, computa cada transformação e resultando em um valor, quão mais similares forem as cadeias de caracteres, menor será o número de transformações, resultando em um valor de distância baixo, para saber mais (Gilleland, 2008). No caso da proposta do Carmo e do Costa, através da análise das informações salvas pelo servidor de páginas, calculando a distância de Levenshtein entre cada linha do arquivo de auditoria do servidor web, segundo eles, seria possível determinar o fluxo de navegação.

Distância de Frobenius, consiste em um problema com norma esférica, que estende a noção de norma vetorial para matrizes (Kock, 2003). Ainda segundo Carmo e Costa, esta norma é muitas vezes mais fácil de ser calculada do que normas induzidas.

Já no trabalho apresentado por Veras e Ruggiero (2005), discute-se a análise contínua da utilização da aplicação como parte essencial para complementar a autenticação inicial do sistema, assim como o trabalho de Carmo et al. (2007), também discute o índice de confiança como um fator para avaliar o usuário.

Ambos os trabalhos basearam-se no trabalho de Platzer (2004), no que diz respeito à representação do julgamento humano, em relação ao comportamento

referenciado pelos índices de confiança, onde, 0 (zero) para totalmente não confiável e 1 (um) para confiável. Quanto mais próximo do zero estiver o índice de confiança, menos liberdades o usuário terá no sistema.

Este índice é incrementado com o uso de autenticações contínuas. Sempre que for necessário maior confiança, será enviado uma autenticação para o usuário, que, sendo positiva, aumentará o valor do índice de confiança, liberando funcionabilidades do sistema .

No trabalho de Lane e Brodley (1999), é feita uma avaliação do problema de detecção de anomalia de comportamento, focada no uso de comandos *Shell* e depois generalizado para outras áreas do sistema, com o uso de um modelo IBL (*Instance based Learning*) (Wangenheim & Wangenheim, 2003; Senger et al, 2007).

Para tanto, foi criado uma base de conhecimento com casos "normais" e, a partir destes, detectado os "anormais". A medida utilizada para a detecção destas anomalias de comportamento, é a distância de similaridade entre as entradas, onde utilizou-se a distância temporal entre os eventos do sistema (*keystroke*, eventos de interface, chamadas de sistema) (Lane & Brodley, 1997).

2.2 MÉTRICA ESTATÍSTICA

A distância de Mahalanobis é uma medida estatística introduzida pelo matemático indiano Prasanta Chandra Mahalanobis, publicado em 1936 (Mahalanobis, 1936). Tem como base a correlação entre variáveis com as quais diferentes padrões podem ser identificados e comparados. É uma medida útil para determinar a similaridade entre um conjunto de dados conhecidos e uma amostra desconhecida. Distintamente da distância euclidiana, considera as correlações do conjunto de dados e é independente da escala das variáveis que compõem uma amostra.

Formalmente, a distância de Mahalanobis entre um grupo de valores conhecidos e um conjunto de valores desconhecidos, onde a média dos valores conhecidos é: $\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_p)^T$ e os desconhecidos representados por $x = (x_1, x_2, x_3, \ldots, x_p)^T$ e ainda com matriz de covariância Σ é definida como: $D_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)}$.

Se tomarmos como exemplo duas observações, de um conjunto variáveis (x, y). E que essas variáveis possuíssem as seguintes características: x: média = 250 e y: média = 250, e matriz de covariância:

Matriz de Variância/Covariância		
	X	Y
X	6291.55737	3754.32851
Y	3754.32851	6280.77066

E em um única observação obtivéssimos os valores x = 225 e y = 216, calcularíamos a distância de Mahalanobis como:

Dado a equação da distância de Mahalanobis, teremos:

$$(x - \mu) = \begin{pmatrix} 225-250 \\ 216-250 \end{pmatrix} = \begin{pmatrix} -25 \\ -34 \end{pmatrix}$$

$$\Sigma^{-1} = \begin{pmatrix} 6291,55737 & 3754,32851 \\ 3754,32851 & 6280,77066 \end{pmatrix}^{-1} = \begin{pmatrix} 0,00025 & -0,00015 \\ -0,00015 & 0,00025 \end{pmatrix}$$

Então, a $D^2 = (-25 \quad -34) \times \begin{pmatrix} 0,00025 & -0,00015 \\ -0,00015 & 0,00025 \end{pmatrix} \times \begin{pmatrix} -25 \\ -34 \end{pmatrix} = 0,19025$.

Logo, uma única amostra possui distância de 0,19025 da média x = 225 e y = 216.

Se pegarmos muitas observações e colocarmos graficamente sua distribuição no espaço teremos algo como:

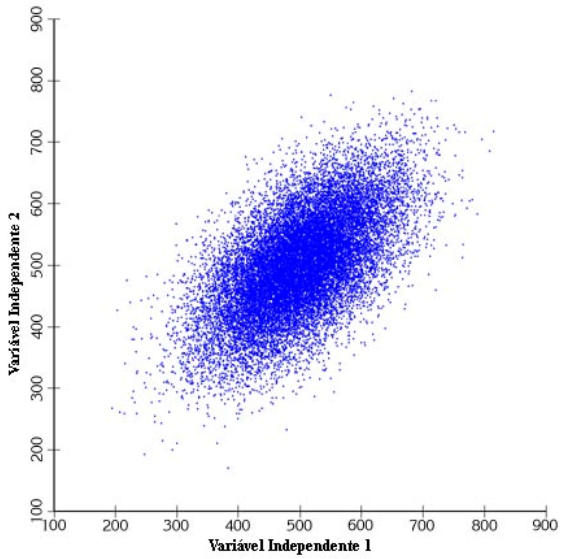

Figura 1. Exemplo de plotagem 20.000 pontos aleatórios contendo duas variáveis independentes.

Fonte: Adaptado de Jenness Enterprises (2008).

Ao calcularmos a distância de Mahalanobis para cada um dos pontos e colorimos de acordo com o valor da distância teremos:

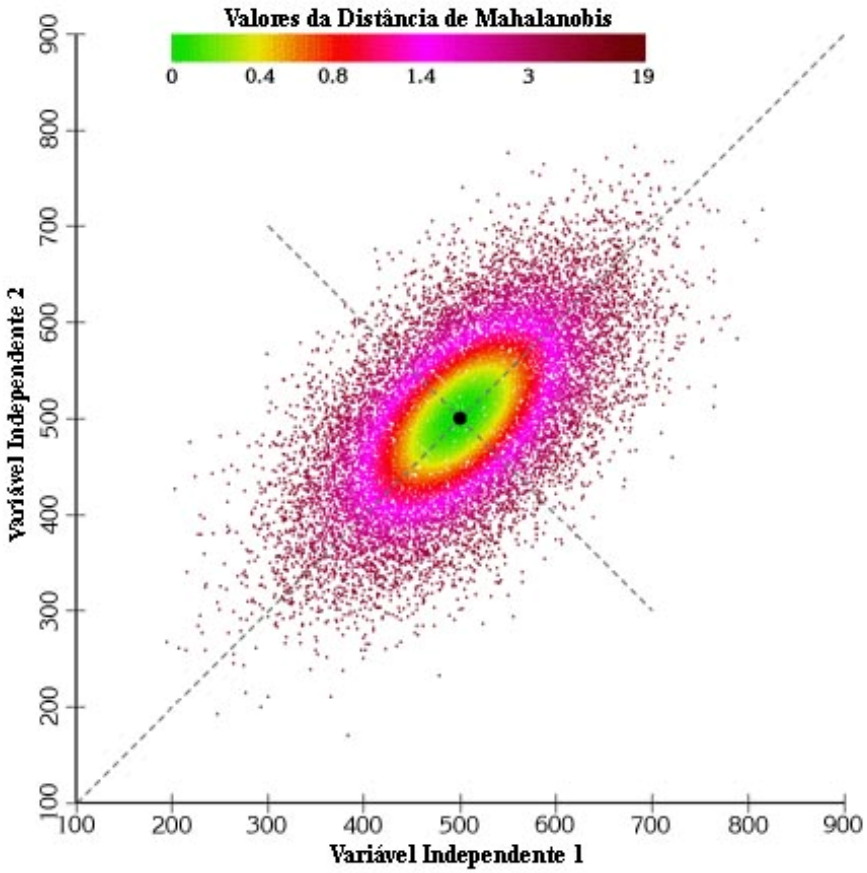

Figura 2. Coloração da distribuição de dados com base na distância de Mahalanobis.
Fonte: Adaptado de Jenness Enterprises (2008).

Com a coloração das diferentes faixas com valores da distância de Mahalanobis, como mostrado na Figura 2, podemos claramente notar os padrões elípticos surgindo.

As teorias de detecção de comportamento e métrica estatística de Mahalanobis aqui descritas neste capítulo definem a base teórica necessária para o desenvolvimento desde trabalho, detalhado a seguir.

Capítulo 3

METODOLOGIA E DESENVOLVIMENTO

Neste capitulo é apresentada uma revisão detalhada dos métodos e tecnologia empregada para o desenvolvimento deste trabalho. O enfoque geral é a criação de assinaturas de usuários e detecção de anomalias no uso do sistema online com base nessas assinaturas.

3.1 BASE PARA O DESENVOLVIMENTO

Devido ao fato de que os sistemas alvo para a utilização de nossa metodologia serem sistemas online, procuramos desenvolver a esta utilizando tecnologias similares as encontradas na maioria dos servidores web. Tal escolha fundamenta-se em uma das características básicas do nosso método, nomeadamente, para facilitar a integração do mesmo com os sistemas a que este se presta a proteger. Desta forma listamos os artefatos de software e linguagem de programação utilizadas:

- PHP – as rotinas necessárias ao processo foram todas desenvolvidas utilizando-se esta linguagem, por tratar-se de uma linguagem interpretada, bastante difundida em ambientes *WEB* (PHP, 2008), conhecida por seu alto desempenho e rápido desenvolvimento.
- Zend Studio – esta ferramenta de desenvolvimento da Zend (Zend, 2008) foi utilizada para a confecção do sistema final, pelo fato do amplo suporte a linguagem PHP, suporte a teste de unidade, *refactoring* do código e ter sido originalmente a plataforma de desenvolvimento do sistema em produção que serviu como base para os testes do sistema proposto serem realizados.
- PostgreSQL – foi o banco de dados escolhido. Tal escolha se ocorreu devido a sua segurança, robustez e desempenho. Ele ainda é um sistema de banco de dados mundialmente aceito, compatível com SQL, e finalmente por ser amplamente utilizado nos sistemas em produção.

3.2 ESTRUTURA DO PROCESSO

O modelo proposto requer que seja realizada uma análise sobre o conjunto de dados relativos a sua utilização para que se possa determinar o perfil de cada usuário de um sistema. Para tal, esse trabalho propõe a criação e verificação de assinatura de usuário em sistemas web. O modelo pode ser subdividido em quatro passos bem definidos:

A. Mapeamento
B. Aquisição
C. Pré-processamento
D. Geração da Assinatura

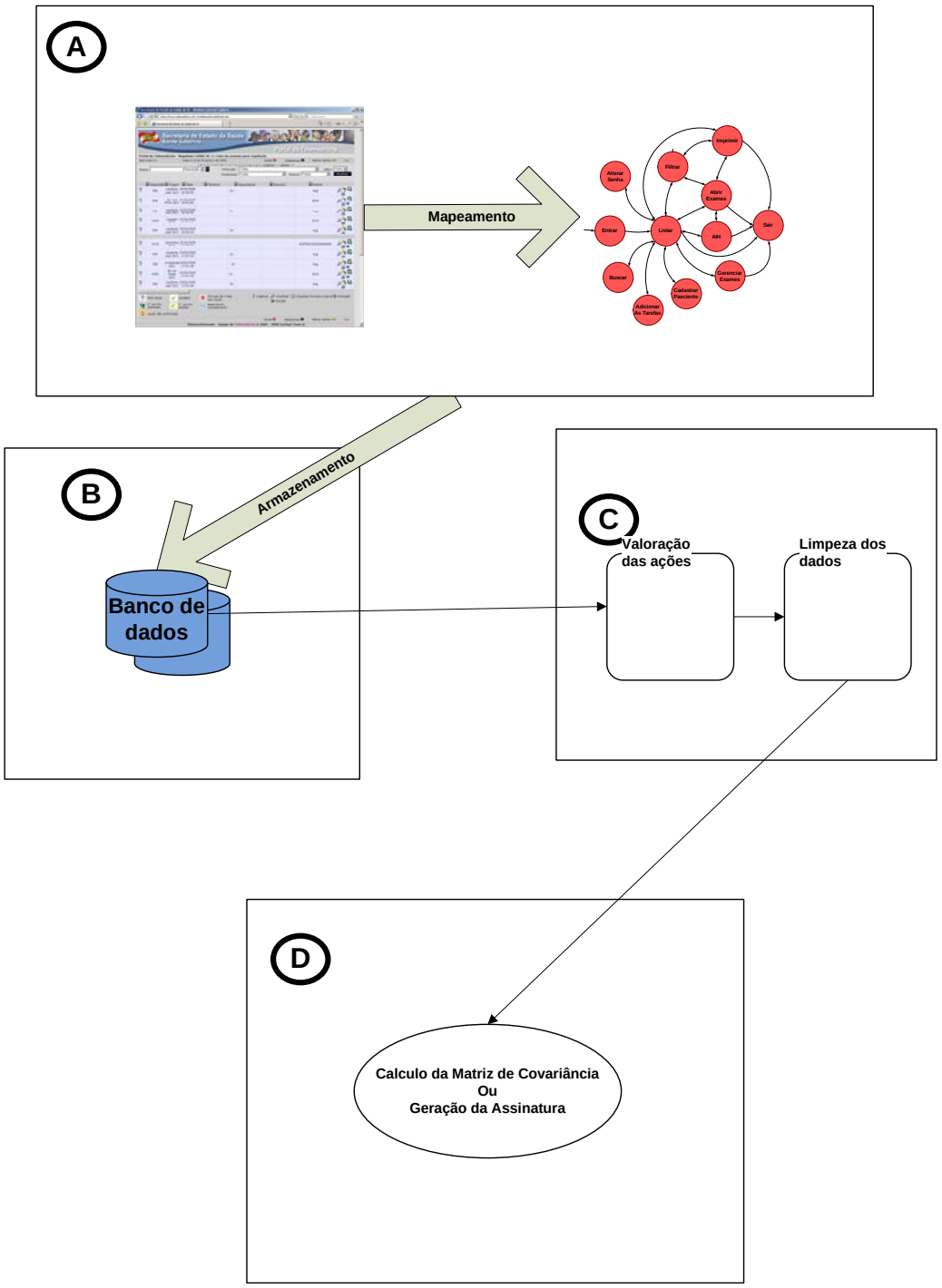

Figura 3. Modelo esquemático da metodologia de detecção de comportamento

A Figura 3 representa um modelo esquemático de cada uma das etapas do processo. Como já mencionado, o processo pode ser desmembrado em quatro sub-processos, que serão descritos adiante.

O primeiro sub-processo, refere-se ao mapeamento das ações do sistema, para a geração da assinatura do usuário. Este mapeamento é uma alternativa para a tradicional análise de arquivos de auditoria a fim de determinar, não somente quais as ações, mas também quem seria o possível responsável pela ação. Uma vez mapeado o sistema, passa a ser possível armazenar apenas as informações relevantes para criar a assinatura de um usuário.

Na segunda etapa, aquisição, foi criado um módulo para ser acoplado em todas as possíveis ações do sistema, anteriormente mapeados. Este módulo é composto de um leitor de sessão, que coleta um conjunto de variáveis, variáveis estas que podem ser definidas pelo administrador do sistema para cada ação que a ser armazenada. O leitor de sessão é composto de um conjunto e funções que a própria linguagem de programação proporciona, para saber mais www.php.net (PHP, 2008).

O modo de armazenamento das mesmas também pode ser definido pelo administrador. A opção feita por banco de dados relacional se dá devido a facilidade de manipulação dos dados, visto que, os banco de dados mais modernos possuem um conjunto de ferramentas que auxiliam a mineração dos dados, afetando diretamente o desempenho de uma aplicação como esta. Como descrito no capítulo 2, a análise de arquivos de auditoria que possuam uma conjunto de dados muito grande, pode comprometer a resposta de um sistema de detecção de comportamento.

Entre as informações coletadas em uma sessão, o módulo de aquisição também armazena o endereço IP de cada usuário, porque, em alguns casos e/ou sistemas, ele poderá ser utilizado para informar posição geográfica do indivíduo (*IANA, 2008*). O endereço IP, na versão 4 (IPv4), é um número de 32 bits escrito com quatro octetos e no formato decimal, separados por ponto (exemplo: 140.3.7.1)[1]. Devemos notar que um endereço IP não identifica uma máquina individual, mas uma conexão. Assim, um gateway conectando à n redes tem 'n' endereços IP diferentes, um para cada conexão (*IANA, 2008*). Nos casos onde é possível utilizar o endereço IP, torna mais fácil determinar acesso de um mesmo usuário por diferentes pessoas, em tempos diferentes, já que fisicamente impossível uma pessoa acessar o sistema a partir de São Paulo e 15

[1] A primeira parte do endereço identifica uma rede específica, chamado de classe A, a segunda parte identifica um host dentro dessa rede, classe B, a terceira e a quarta classe C e D respectivamente.

minutos depois acessar novamente a partir de Florianópolis. Além disso, ao armazenar as ações de forma individualizada e previamente identificadas, o sistema pode verificar se o usuário pedindo acesso, já não estava previamente acessando o sistema. Este simples processo, impossibilita o acesso simultâneo utilizando o mesmo par: usuário e senha.

O passo C refere-se ao pré-processamento, onde as diversas entradas que compõem o conjunto de passos armazenados, são reorganizados de forma a obter um novo conjunto, que contenha apenas os dados analisáveis pela distância de Mahalanobis.

No pré-processamento, os dados são analisados tanto em conjunto quanto separadamente. Ou seja, uma sessão de utilização inicia-se no momento da entrada no sistema, e será apenas finalizada na saída. Porém algumas sessões ficam incompletas e por meio de varredura das ações de cada usuário, é possível identificar a ocorrência de uma tentativa de entrada antes da finalização da ultima sessão. Geralmente as falhas de armazenamento de ações se dá por falha temporária de conexão entre o cliente e o sistema, resultando em uma desconexão prematura. Por meio do algoritmo desenvolvido especificamente para filtragem de dados, todas as sessões incompletas são removidas. Notou-se experimentalmente que tais sessões correspondem a menos de 1% do conjunto total de ações. O passo, seguinte refere à geração de uma matriz de referência que sumariza o comportamento do usuário. Tal matriz conterá n linhas e m colunas, onde n é o número de ações que serão utilizadas para gerar a assinatura e m é o número de variáveis que compõe uma entrada de uma ação (IP, ação, horário em segundos do dia). De posse desta matriz

Na etapa D, uma vez que a matriz supracitada esteja disponível, é realizada a geração da matriz de covariância e vetor médio, únicos para cada indivíduo, e que servem como base para calculo da distância entre novos eventos e os previamente armazenados.

Detalhes da implementação e tópicos relevantes a serem considerados em cada um dos 4 passos descritos acima serão apresentados nas próximas 4 subseções.

3.3 MAPEAMENTO

Como frisado na sessão anterior, o mapeamento é uma alternativa para que não seja necessária a varredura de arquivos de auditoria, coleta das ações e identificação dos usuários responsáveis. O mapeamento é utilizado como referência para o acoplamento do módulo de coleta de informações e o de análise de comportamento (detecção de anomalia).

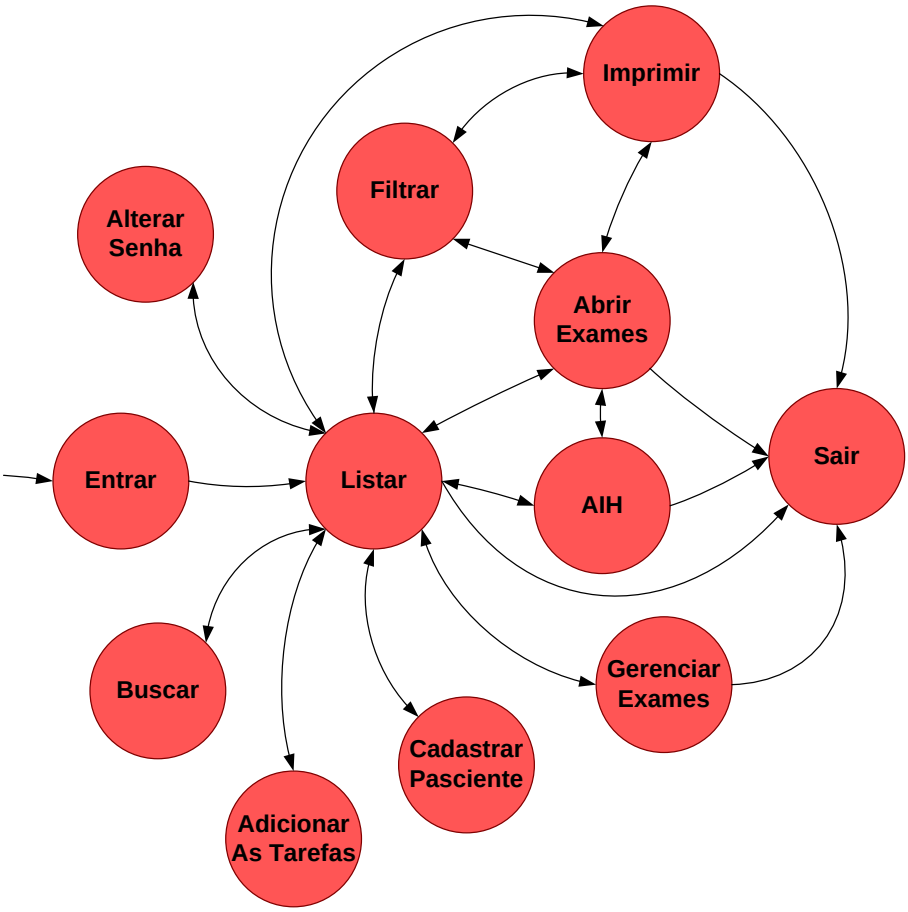

Figura 4. Exemplo do mapeamento das ações do Portal de Telemedicina utilizando-se um grafo dirigido

Como cada ação de um usuário no sistema pode ser descrita como uma transição de um estado do sistema para outro, segundo a teoria de grafos, tal problema pode ser modelado como um grafo dirigido, como exemplificado na Figura 4. Esta abordagem trás algumas vantagens. A ligação entre uma ação e outra, ou seja, um usuário não

poderia passar de um estado inicial “Entrar” para o estado final “Sair” sem antes passar, pelo estado “listar” no menor caminho. Esta figura exemplifica o mapeamento do Portal de Telemedicina, sistema utilizado como plataforma de testes. Uma descrição detalhada de todas as possíveis ações executáveis pelo sistema de Telemedicina pode ser encontrada no Apêndice 2.

Para mapear um sistema web qualquer em formato de grafo dirigido, é possível utilizar duas abordagens. A primeira, no caso de um sistema bem documentado, onde ele teria referência para todas as possíveis ações do sistema. Seria operar o sistema seguindo o que foi descrito pela documentação, gerando material suficiente para a criação do grafo. Outra abordagem, para sistemas com documentação incompleta ou inexistente seria inteiramente manual. O problema no caso de mapeamento manual, é que existe a possibilidade de ações não mapeadas, como descrito no trabalho de Carmo e Costa (2007). Porém é possível adicionar nodos ao grafo a qualquer momento em que uma nova função é identificada ou adicionada ao sistema.

A ligação entre os nodos do grafo nada mais é do que a passagem de um estado do sistema para outro após a execução de uma ação.

3.4 AQUISIÇÃO

Para esta etapa foi criado um algoritmo, descrito na Figura 5, que é responsável pela aquisição e envio dos dados para o banco de dados, no seu formato ainda não criticado, ou seja, ainda em formato de seqüência de caracteres, para que, se necessário, verificação visual de ocorrências de intrusão ou falha no sistema. Este módulo será acoplado em todos os passos (ações) mapeados, para que possam ser analisados na etapa seguinte.

```
Conecta no banco de dados;
Se existe ação faça
        Coleta nome do usuário;
        Coleta data e hora da ação;
        Coleta IP do usuário;
        Monta a consulta para o banco de dados;
        Armazena as informações no banco de dados;
Fim faça;
Desconecta do banco de dados;
```

Figura 5. Algoritmo de armazenamento das ações do usuário

No algoritmo da Figura 5, são utilizadas apenas as variáveis necessárias para criação de uma assinatura, ou seja, o nome do usuário, a fim de facilitar e direcionar a busca das informações armazenadas no banco de dados, o momento em que ocorreu a ação e o endereço IP da conexão que estava sendo utilizada.

3.5 PRÉ-PROCESSAMENTO

Para que a distância de Mahalanobis possa ser utilizada, os dados devem possuir determinadas características, como serem numéricos e ter variação em relação a cada componente do conjunto que corresponde a uma entrada ou ação. Campos que não variam quando analisado o conjunto de dados, devem ser eliminados, pois no momento do calculo da matriz de covariância, ocorrerá divisão por 0, inviabilizado a utilização desta métrica (Mahalanobis, 1936).

A modificação dos dados inicia-se com a correspondência, mapeamento ou valoração de cada ação armazenada em formato de cadeia de caracteres, para um valor numérico equivalente como exemplificado a seguir:

- Entrar → 0;
- Listar → 1:
- Filtrar → 2:
- Buscar → 3;
- Visualizar → 4;
- Imprimir → 5;
- Altera Senha → 6

Maiores detalhes sobre os mapeamentos do Sistema exemplo de Telemedicina se encontram disponíveis no Anexo 1.

Além da valoração de cada ação é também necessário transformar o endereço IP em um conjunto de quatro variáveis, cada uma delas uma parte do endereço IP, já que as métricas de existentes apenas trabalham com números e não seqüência de caracteres.

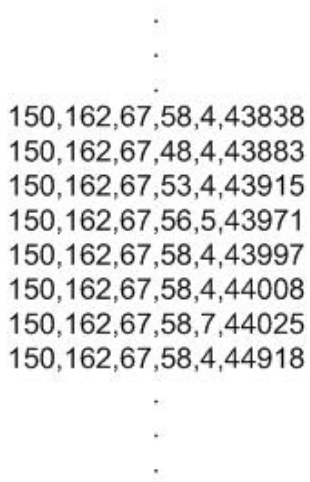

Figura 6 Amostra de dados armazenados, onde cada variável está separada por vírgula

Nos dados apresentados na Figura 6, é possível notar a repetição de valores nas três primeiras variáveis que compõem uma ação, como citado anteriormente, ambos os três campos deverão ser eliminados.

Após a valoração das ações e preparação dos dados é realizada a normalização dos dados, já que cada componente de uma entrada possui uma escala própria, logo, cada valor armazenado será mapeado para o intervalo [0, 1].

Ao excluir os campos sem variação será gerado um conjunto Y - (campo analisado), resultando em uma matriz n X m, onde n corresponde a cardinalidade de ações do usuário e m é a quantidade de variáveis contidas em uma ação.

3.6 GERAÇÃO DA ASSINATURA

A geração da assinatura é composta de matriz de covariância e vetor médio. Tem como base para sua geração, todas as ações armazenadas anteriormente e pré-processadas, como descrito no tópico anterior.

De forma sintética, a matriz de covariância é uma matriz quadrada, simétrica, cuja diagonal principal contém a variância da variável e em cada interseção linha (i) coluna (j) a covariância das variáveis Xi e Xj.

A geração da matriz de covariância de cada indivíduo é realizado conforme o algoritmo descrito na Figura 7.

```
para cada linha da matriz
    para cada coluna da matriz
        v1=0
        v2=0
        acumulado=0
        para cada lilnha da matriz
            v1=matriz[linha][linha]
            v2=matriz[linha][coluna]
            acumulado=acumulado+(v1-vetormedio[linha])*(v2-vertormedio[coluna])
        fim para
        resultado[linha][coluna]=acumulado/(total de linhas-1)
    fim para
fim para
```

Figura 7. Algoritmo para criação da matriz de covariância

Um conjunto de dados, como o mostrado na Figura 6, tem como resultado a seguinte matriz de covariância:

$$\begin{bmatrix} 0{,}00356098682431 & 0{,}0001470178650 83 & 0{,}00144945437771 \\ 0{,}0001470178650 83 & 0{,}07355884285 44 & -0{,}0004338168155 86 \\ 0{,}00144945437771 & -0{,}0004338168155 86 & 0{,}0167063156548 \end{bmatrix}$$

Figura 8. Exemplo de Matriz de Covariância

O vetor médio nada mais é do que a média aritmética de cada variável que compõem uma ação.

Como descrito na seção 2.2, para o cálculo da distância de Mahalanobis se faz necessário ambos, matriz de covariância e vetor médio, para a detecção de anomalia, ou detecção de comportamento.

3.7 VERIFICAÇÃO DE COMPORTAMENTO

A verificação da identidade de um usuário que acessa o sistema é realizada de forma tecnicamente simples. Após a geração do espaço representativo, como descrito na seção anterior, qualquer nova ação pode ser diretamente comparada com a assinatura do usuário, bastando, apenas, que seja calculada a distância desta ação em relação à assinatura desse indivíduo.

O algoritmo para verificação de comportamento faz uso da distância de Mahalanobis para o cálculo da distância da ação do usuário e a sua assinatura. Já que trata-se de uma métrica de fácil aplicação, e por ser estatística, recomendada pelos trabalhos existentes como sendo a de melhor resultado na área de detecção de anomalia (Micarelli & Sansonetti, 2007).

Para a comparação de uma nova ação com a assinatura do usuário, se faz necessário observar os campos eliminados na geração da assinatura, confrontando-os com os mesmos campos no vetor correspondente a ação. Se algum dos campos eliminados possuir valor diferente dos campos contidos na ação, indicaria um desvio de comportamento (kemmerer & Vigna, 2002). No caso em que os valores são iguais, devem ser eliminados para calcular a distância, já que tanto o vetor correspondente a ação quanto a matriz de covariância e o vetor médio devem possuir a mesma cardinalidade para que seja aplicada a distância de Mahalanobis (Mahalanobis, 1936).

A detecção de comportamento é realizada através de um módulo que calcula, como descrito anteriormente, a distância de Mahalanobis para todas as ações do usuário em relação à assinatura. Este módulo é acoplado diretamente a cada ação previamente mapeada conforme descrito na Figura 9.

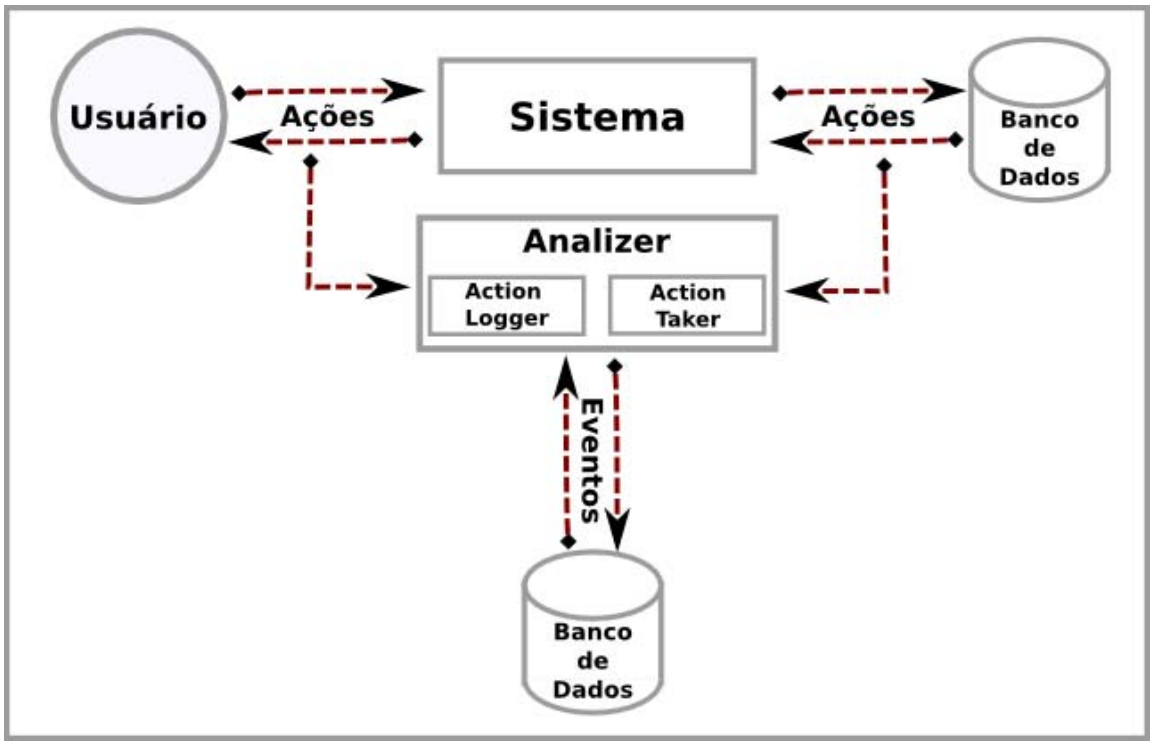

Figura 9. Modelo de ligação dos módulos de aquisição e detecção de anomalia

Para validar esta proposta de detecção foram realizadas avaliações quantitativas e qualitativas da metodologia descrita, utilizando a base de dados do Portal de Telemedicina como base de testes, estas avaliações estão descritas no próximo capítulo.

Capítulo 4

RESULTADOS

De acordo com kemmerer e Vigna (2002), Carmo e Costa (2007), Micarelli e Sansonetti (2007), entre outros, a qualidade de um método para determinar detecção de anomalia depende diretamente da quantidade de Falsos Negativos (FN) e Falsos Positivos (FP) gerados. A ocorrência de FN é devido à classificação de uma ação legítima como anomalia, resultando em um falso alerta. E FP é o seu complemento, ou seja, a classificação de uma ação de possível invasão como sendo uma ação normal ou legítima. Quão menor for o valor de FN e FP, melhor será a qualidade da técnica empregada.

Com o intuito de detectar o comportamento do *framework* proposto em ambiente de produção, e coletar índices de FN, foi executado uma avaliação com o conjunto de ações armazenadas no banco de dados. Dividindo o conjunto de dados em dados para criação da assinatura e dados para comparação e cálculo da distância.

Para coletar o índice de FP foi realizada um espelhamento do Portal de Telemedicina com seu conteúdo, a fim de permitir testes com dados reais de produção sem afetar o sistema em questão. Assim, foi liberado o acesso para um grupo, devidamente credenciado para manipular dados médicos, para operar o sistema utilizando usuário e senha de outro profissional.

As avaliações partem de duas hipóteses principais, a primeira, que a partir de um conjunto de dados devidamente identificado como pertencente a um determinado indivíduo, podemos criar a assinatura que servirá como base de comparação. E ações futuras deste usuário devem ser possíveis de identificar como sendo normais.

E como segunda hipótese, que é possível detectar anomalia de comportamento, gerando a assinatura com base no histórico de atividades de um indivíduo, comparando com ações de outra pessoa com o mesmo usuário do sistema e senha do sistema.

O Portal de Telemedicina utilizado pelos testes contém atualmente 94.143 pacientes, 969 profissionais de saúde e 105.758 exames e gerou 598.296 ações do período de janeiro de 2008 à dezembro de 2008.

A bibliografia sugere que quanto maior o número de casos para utilizar Mahalanobis, melhor o resultado (Takeshita et al., 1993). Logo, selecionamos todos os usuários que continham 10.000 ações ou mais, como fonte de comparação, resultando em um grupo de 15 profissionais.

A validação dessas hipóteses, bem como seus desdobramentos, estão demonstrados a seguir.

4.1 VALIDAÇÃO COM DADOS NORMAIS

O Presente experimento visou calcular a média de Falsos Negativos utilizando a detecção de comportamento com base na distância de Mahalanobis.

A primeira parte do experimento tem o intuito de determinar qual a quantidade ideal de ações a serem utilizadas para geração da assinatura, ou seja, obter um baixo índice de FN utilizando o menor número de casos possíveis para a geração da assinatura, já que quanto maior o número de casos, maior será a complexidade computacional. O que, dependendo da quantidade de casos e variáveis que compõem um caso, pode tornar o sistema inviável do ponto de vista de tempo de resposta para uma determinada análise.

Os testes foram realizados com geração de assinatura variando de 2.000 à 9.000 ações, e comparado com outros 1.000 casos normais. E ainda buscando demonstrar a eficiência da métrica de Mahalanobis, comparamos os resultados com as distâncias euclidiana e Minkowski, duas distâncias bem difundidas e de fácil aplicação.

4.1.1 Resultados da Validação com Dados Normais

Inicialmente variamos a quantidade de ações para gerar a assinatura de 2.000 à 5.000, levando em conta o tempo de resposta do algoritmo principal, Mahalanobis, com o intuito de obter o menor valor necessário para aplicar esta distância, com tempo de

resposta satisfatório. Observando a Tabela 1, podemos notar que a partir de 5.000 ações já obtivemos um bom resultado, com baixo tempo de resposta.

Este teste de foi realizado com um conjunto de 1.000 ações comparadas com assinatura, logo, 14,65 segundos para a execução do teste, representa em média, 0,01465 segundos por comparação.

Tabela 1 - Tabela comparativa de total de ações em relação ao tempo de teste e FN

Total de Ações	Tempo	FN
2000	1,81s	5,87%
2500	2,35s	5,31%
3000	2,85s	4,89%
3500	4,24s	4,34%
4000	7,65s	3,3%
4500	10,25s	2,38%
5000	14,65s	0,99%

Com base na Tabela 2 demonstramos a superioridade da distância de Mahalanobis, frente a distância euclidiana e a distância de Minskowski. Foi utilizado um valor de distância de 0,09 como fator de comparação, o que ainda é dez vezes menor que o limiar utilizado por Micarelli e Sansonetti (2007).

Tabela 2 - Tabela resultados da hipótese de Falsos Negativos

Usuário	Total de Ações	Mahalanobis	Euclidiana	Minkowski
Usuário1	1.000	0,50%	90,815%	85,759%
Usuário2	1.000	0,00%	98,956%	98,330%
Usuário3	1.000	0,00%	87,716%	87,208%
Usuário4	1.000	0,00%	93,798%	93,254%
Usuário5	1.000	4,60%	100,000%	100,000%
Usuário6	1.000	0,00%	100,000%	100,000%
Usuário7	1.000	5,90%	100,000%	100,000%
Usuário8	1.000	0,00%	100,000%	100,000%
Usuário9	1.000	0,00%	83,953%	78,321%
Usuário10	1.000	0,00%	100,000%	100,000%
Usuário11	1.000	0,00%	100,000%	100,000%
Usuário12	1.000	3,80%	82,305%	82,204%
Usuário13	1.000	0,00%	96,548%	91,694%
Usuário14	1.000	0,00%	92,191%	91,988%
Usuário15	1.000	0,00%	90,461%	90,252%
Média		**0,99%**	**94,450%**	**93,267%**

Os usuários foram anonimizados por se tratar de dados médicos e sigilosos. As porcentagens mostram o índice de Falso Negativo em 1.000 ações de comparação, para as diferentes distâncias utilizadas.

4.2 VALIDAÇÃO DE DETECÇÃO COM DADOS ANÔMALOS

O Presente experimento visou calcular a média de Falsos Positivos utilizando a detecção de comportamento com base na distância de Mahalanobis.

Os 15 usuários que serviram de base para a hipótese anterior também foram utilizados neste teste já que possuíam as quantidade satisfatória de ações armazenadas. De um grupo devidamente credenciado para cessar dados sigilosos como os presentes no Portal (dados médicos), foram selecionados 15 pessoas para participar do experimento,. Foram divididos em três subgrupos, que realizaram acessos a cópia do Portal de Telemedicina, seguindo determinadas características, que serão descritas nos itens subseqüentes:

1. Acesso como *Hacker*: desempenha papel de invasor com fins de alteração de conteúdo simulando atitude maliciosa dentro do sistema.
2. Acesso como Seguradora de Saúde: realiza pesquisas por determinado paciente com o intuito de levantar seus dados clínicos e laboratoriais, que seriam interessantes no momento de determinar o valor de um seguro saúde para este paciente, por exemplo.
3. Acesso Livre: Usuário comum tentando acessar dados restritos ou apenas realizando consultas de rotina por mera curiosidade.

Após, foram enviados mensagens para uma determinada pessoa deste grupo, com os devidos dados de acesso ao Portal e requisitando atuação com um tos perfis descritos nos itens 1, 2 ou 3, durante o período de sete dias. Só então, foi possível analisar estatisticamente os dados coletados.

4.2.1 Resultados dos Experimento Para Detecção de Dados Anômalos

A Tabela 3**Erro! Fonte de referência não encontrada.** demonstra o grau de detecção das ações simuladas por um invasor, que, de posse de nome do usuário e senha, invadiu e utilizou o sistema das formas descritas no experimento anterior, com média de **0,18**% de falha na detecção de anomalia utilizando métrica de Mahalanobis, superior, quando comparado a métrica euclidiana com **6,44**% e Minkowski com **6,83**%.

Tabela 3 - Tabela resultado da hipótese de Falso Positivo

Usuário	Total de Ações	Mahalanobis	Euclidiana	Minkowski
Usuário1	148	0,00%	64,19%	64,16%
Usuário2	12	0,00%	0,00%	0,00%
Usuário3	20	0,00%	0,00%	0,00%
Usuário4	115	0,00%	0,00%	0,00%
Usuário5	134	1,74%	0,00%	0,00%
Usuário6	125	0,00%	18,40%	18,40%
Usuário7	138	0,00%	0,00%	0,00%
Usuário8	111	1,80%	0,00%	0,00%
Usuário9	144	0,00%	0,00%	2,72%
Usuário10	142	0,00%	0,00%	0,00%
Usuário11	125	0,00%	0,00%	3,20%
Usuário12	111	0,00%	9,02%	9,02%
Usuário13	126	0,00%	0,00%	0,00%
Usuário14	103	0,00%	0,00%	0,00%
Usuário15	80	0,00%	5,00%	5,00%
Média		**0,24%**	**6,44%**	**6,83%**

Pelos mesmos motivos citados anteriormente os usuários foram anonimizados na tabela anterior. As porcentagens mostram o índice de Falso Positivo de acordo com a quantidade de ações que foram consideradas pertencentes ao usuário em questão. Em alguns casos o número reduzido de interações da pessoa simulando a invasão, como no caso do usuário 2 e 3, pode prejudicar o resultado. Porém mesmo desconsiderando os dois casos, a técnica empregada ainda obteve um baixo índice de FP.

É interessante notar que Mahalanobis possui maior estabilidade em relação as outras métricas aplicadas, comparando os diferentes usuários. Métricas como a euclidiana e a Minkowski são bem mais simples de calcular porém em determinados casos como os dos usuários 1, 6 e 12, retornam um valor muito alto de FP.

Capítulo 5
CONCLUSÕES

Neste trabalho apresentamos um método para detecção de comportamento com base na métrica de Mahalanobis, para interpretação de dados armazenados a partir da navegação e operação do Portal de Telemedicina do Estado de Santa Catarina. Estes dados em conjunto com a técnica proposta nos permitiu obter baixo índice de Falsos Negativos (0,99%) e Falsos Positivos (0,24%), como demonstrado pelos experimentos descritos no capítulo anterior. Levando em conta todos os casos utilizados para testes, 16.531 de comparação, onde 15.000 de casos reais para teste de FN e 1.531 de casos de simulação de invasão, ainda obtivemos apenas três falhas de detecção de comportamento anômalo e 148 detecções de anomalia que era casos normais.

O uso de um índice para distância de 0,09 permite que o sistema ainda possa discernir de ataques de mimetização, onde o intruso finge ser um usuário credenciado para operar o sistema, como descrito no trabalho de Tan et al. (2002).

Com base no primeiro experimento, observou-se que existe um compromisso entre o conjunto de ações a serem utilizadas pelo algoritmo gerador da assinatura e o desempenho do sistema. Executou-se um treinamento do sistema usando um número incremental de casos, variando de 2.000 à 5.000 com o passo de 500 em 500. Resultados apresentados na sessão 4.1.1 mostram que 5.000 casos obteve uma boa relação entre tempo de operação da detecção e número de ações utilizados para geração da assinatura, isso em tempo real. As operações realizadas sobre matrizes, como descrito na sessão **Erro! Fonte de referência não encontrada.**, fazem com que quanto maior o número de variáveis que compõem um caso e o volume de casos, utilizados para a técnica, maior o tempo gasto com a geração da assinatura e calculo da distância para um novo caso.

É válido notar que entre o conjunto de variáveis foi utilizado o endereço IP, que como citado por Carmo e Costa (2006), "não é uma variável confiável", já que existem formas de mascarar. Porém, no caso do Portal de Telemedicina, é totalmente viável, já que os hospitais e postos de saúde utilizam rede do Centro de Informação e Automação de Santa Catarina (CIASC, 2008), como forma de acesso a internet, sendo assim,

possuem endereço IP rastreável e fixo. Apenas os casos de profissionais que acessam de forma externa essa rede é que podemos considerar duvidosa a utilização dessa variável.

Um problema neste tipo de abordagem de reconhecimento de comportamento, é a necessidade de um grande número de casos necessários para gerar a assinatura do usuário, ou seja, entre a primeira entrada e o valor ideal para a criação da assinatura, a técnica é ineficiente.

Um dos fatores de maior estresse computacional citado em muitos dos artigos referenciados ao longo deste trabalho, é a varredura ou análise de trilhas de navegação utilizando os arquivos de auditoria dos servidores de páginas. Para contornar este problema, utilizamos banco de dados relacional para armazenamento da informação, já que sistemas mais modernos como PostgreSQL e outros, possuem um conjunto completo de ferramentas que auxiliam nestas tarefas. Em contrapartida, existe a necessidade de mapeamento e utilização de um artefato de software externo ao sistema acoplado a cada ação, armazenando as informações na base de análise.

Por fim, a técnica utilizada possibilitou o acompanhamento das atividades dos usuários, bem como sua evolução, podendo ser empregada em sistemas que possuam características similares ao sistema em produção testado, para fins de incremento de segurança.

5.1 OBJETIVOS ALCANÇADOS

Foi analisada boa parte da literatura e dos métodos relativos à detecção de comportamento e sistemas similares, nas sessões **Erro! Fonte de referência não encontrada.** e **Erro! Fonte de referência não encontrada.**, com especial ênfase na distância de Mahalanobis , na sessão **Erro! Fonte de referência não encontrada.**, como forma de criação de assinatura e para cálculo de distância de similaridade.

A análise no conjunto de dados armazenados, mostraram a existência de pequenas deturpações provenientes de falhas temporárias na comunicação entre o servidor de aplicação e o cliente ou entre o servidor de aplicação e o de banco de dados

onde foram armazenadas as ações, resultando na criação de um algoritmo para detecção e eliminação do conjunto de dados comprometidos.

Com base nos dados armazenados, foi possível criar uma assinatura para cada usuário onde o número de casos seria 5.000 obteve uma boa relação entre precisão e tempo de execução, como verificado na sessão 4.1.1.

No conceito da técnica aplicada, a distância de Mahalanobis obteve um desempenho satisfatório, levando em conta o número de casos disponíveis, como demonstrado nos experimentos executados e descritos no capítulo anterior.

5.2 DIFICULDADES ENCONTRADAS

Um dos principais problemas neste trabalho foi determinar o melhor número de casos a serem utilizados para geração da assinatura, necessitando de dois diferentes passos, primeiro, avaliar qual o melhor índice de distância a ser utilizado e depois, verificar sua precisão junto ao grupo de testes, executando exaustivamente o algoritmo responsável por medir os FP e os FN.

A necessidade de um grande número de casos também é um fator complicador, pois o sistema encontra-se vulnerável enquanto o número de casos não alcançar o volume ideal para a criação da assinatura do usuário.

5.3 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Um possível trabalho futuro poderia ter ligação direta com as dificuldades encontradas. A adoção de uma metodologia híbrida utilizando técnicas lineares ou estatísticas nos primeiros casos armazenados, possivelmente seria uma solução para o momento considerado crítico na técnica apresentada, onde o número de casos ainda não é suficiente para seu emprego. Após alcançar um volume suficiente para análise, utilizar métrica estatística de Mahalanobis.

Outro trabalho futuro interessante, seria a utilização de índice de confiança como base para tomada de decisões do sistema, em conjunto com a distância de Mahalanobis para medida de distância entre a assinatura e as ações dos usuários, ou seja, um módulo

que avaliasse a confiança do sistema no usuário, e com base nos indicativos providos pela técnica descrita neste trabalho, variasse o grau de confiança, liberando ou restringindo operações no sistema.

Uma proposta, seria o uso de uma técnica que operasse simultaneamente com a detecção de comportamento, realizando a tarefa de resposta do sistema, que poderia ser simples como um alerta enviado diretamente ao usuário, ou ao administrador do sistema, bem como medidas mais drásticas tais como bloqueio temporário ou permanente das suas funções. Assim, poder-se-ia obter um sistema completo, com detecção e resposta automática ou semi-automática, necessitando avaliar qual das duas abordagens seria mais interessante.

Além disso, existe a necessidade de detectar outras possíveis variáveis do sistema que possam ser utilizadas em conjunto com as já mapeadas, lembrando que um conjunto maior de variáveis será igual a maior processamento e tempo de reposta da detecção.

REFERÊNCIAS

Ames, W., **Understanding spyware: risk and response**, IT Professional, v.6, n.5, p. 25-29, 2004.

Balajinath, B. and Raghavan, S.V., **Intrusion detection through learning behavior model**. Computer Communications. v. 24 cap. 12, p. 1202-1212, 2001.

Brocardo, M. L. et al., **Um Modelo de Segurança e Controle de Acesso ao Registro Eletrônico de Pacientes na Web**, XI Congresso Brasileiro de Informática em Saúde - CBIS'2008**,** Campos do Jordão – SP, 2008.

Carmo, L. F. R. C. Costa, D., **Reconhecimento de padrões de comportamento individual baseado no histórico de navegação em um Web Site**, SBSEG-VII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais, 2007, Rio de Janeiro. Anais do SBSEG 2007, p. 18-29, 2007.

Carmo, L.F.R.C. et al., **Trust Evaluation for Web Applications based on behavioral Analyses**. : 22th International Information Security Conference... New approaches for security privacy and trust in complex environments (IFIP 07), Sandton, Africa do Sul, 2007.

CIASC, **Centro de Informação e Automação de Santa Catarina**, Disponível em: < http://www.ciasc.gov.br/>, Acessado em: 26 de Agosto de 2008.

Franz, M., **Containing the Ultimate Trojan Horse**, Security & Privacy, IEEE v. 5, cap. 4, pp. 52-56, Julho – Agosto, 2007.

Gilleland, M., **Levenshtein Distance**, in Three Flavors, disponível em: <http://www.merriampark.com/ld.htm>, Acessado em: 30 de Julho de 2008.

IANA, **Internet Assigned Numbers Authority**, disponível em: <http://www.iana.org/numbers/>, Acessado em: 26 de outubro de 2008.

ITL Security Bulletin, **Computer Attacks: What They Are and How to Defend Against Them**. Disponível em: <http://csrc.nist.gov/publications/nistbul/html-archive/may-99.html>, Acessado em: 28 de Julho de 2008.

Javitz, H.S., Valdes, A., **The SRI IDES Statistical Anomaly Detector**, Proceedings of the IEEE Symposium on Security and Privacy, Oakland, CA, IEEE Computer Society Press, Los Alamitos, Maio,1991.

Jenness Enterprises, **Mahalanobis Description**, Disponível em: < http://www.jennessent.com/arcview/mahalanobis_description.htm>, Acessado em: 27 de outubro de 2008.

Kemmerer, R.A.; Vigna, G., **Intrusion detection: a brief history and overview**, Computer, v. 35, cap. 4, p. 27-30, Abril, 2002.

Kock, J., **Frobenius Algebras and 2D Topological Quantum Field Theories**, London Mathematical Society student texts, Cambridge: Cambridge University Press, 2003. ISBN 0-521-83267-5.

Lane, T. and Brodley, C., **Temporal Sequence Learning and Data Reduction for Anomaly Detection**, ACM Transactions on Information and System Security, New York, v.2, p. 295–331,1999.

Liu, A.X. et al., **A secure cookie protocol**, 14th International Conference Computer Communications and Networks, ICCCN 2005, p. 333–338, Outubro, 2005.

Mahalanobis, P.C, **On the generalized distance in statistics**, Proceedings of the National Institute of Science of India 12, p.49-55, 1936.

Maia, R.S. et al., **A Statewide Telemedicine Network for Public Health in Brazil**, Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on Computer Based Medical Systems, P.95-500, 2006.

Micarelli, A., Sansonetti, G., **Case-Based Anomaly Detection**, International Conference of Case Based Reasoning, ICCBR 2007, Belfast, Irlanda, p.269-283, Agosto, 2007.

Mohamed, M.A., Gader, P., **Generalized hidden Markov models. I. Theoretical frameworks**, Fuzzy Systems, IEEE Transactions, v.8, cap. 1, p.67-81, Fevereiro , 2000.

Noel, S., et al., **Applications of Data Mining in Computer Security**, chapter Modern Intrusion Detection, Data Mining, and Degrees of Attack Guilt, Boston, MA, Kluwer Academic Publisher, 2002, p. 2–25.

Patcha, A., **An overview of anomaly detection techniques**: Existing solutions and latest technological trends, v. 51, cap. 12, p. 3448-3470, 2007.

Perez, C.A, et al., **Real-Time Iris Detection on Coronal-Axis-Rotated Faces**, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions, v. 37, cap. 5, p. 971-978, Setembro, 2007.

pgAdmin, Disponível em: <http://www.pgadmin.org/>, Acessado em: 26 de Agosto de 2008.

FireFox, Disponível em: < http://pt-br.www.mozilla.com/pt-BR/firefox/>, Acessado em: 26 de Agosto de 2008.

Internet Explorer, Disponível em: < http://www.microsoft.com/brasil/windows/products/winfamily/ie/default.mspx>, Acessado em: 26 de Agosto de 2008.

PHP, **Personal Home Page**, Disponível em: <http://www.php.net/>, Acessado em: 26 de Agosto de 2008.

Platzer, C., **Trust- Based security in web services**, Masther's Thesis, Information Institute, Tecnical University of Vienna, Austria, 2004.

Porras, P., Neumann, P., **EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances**, Proceedings of the 20th National Information Systems Security Conference, Baltimore, MA, Outubro, 1997.

Senger, L. J. et al, **Aprendizado baseado em instâncias aplicado à predição de características de execução de aplicações paralelas,** Revista de Informática Teórica e Aplicada, v. 14, 2007.

Takeshita, T.; Nozawa, S.; Kimura, F., **On the bias of Mahalanobis distance due to limited sample sizeeffect**, Document Analysis and Recognition, 1993, Proceedings of the Second International Conference on, p.171-174, Outubro, 1993.

Tan, K., Maxion, R., **"Why 6?" Defining the Operational Limits of Stide, an Anomaly-Based Intrusion Detector**, Proceedings of the IEEE Symposium on Security and Privacy, Berkeley, CA, p.188–202, Maio, 2002.

Véras, L.M.A e Ruggiero, W.V., **Autenticação Contínua de Usuários em Aplicações Seguras na Web**. : V Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais (SBC 2005), Florianópolis, p.40-53, 2005.

Wallauer, J. et al, **Building a National Telemedicine Network**, IT Professional, vol. 10, no. 2, p. 12-17, Mar-Abr, 2008.

WALLAUER, J., **Sistema de Gerencia de Exames via Web**, Trabalho de conclusão de Curso, Universidade Federal de Santa Catarina, Departamento de Informática e Estatística – INE, 2005.

Wangenheim, C. G. von, Wangenheim, A. von, **Raciocínio Baseado em Casos**, Barueri-SP, Brasil, Manole , 2003, 1ª edição.

Weaver, A.C. **Secure Sockets Layer**, Computer v. 39, cap. 4, p. 88-90, Abril, 2006.

Younhee, G. et al., **Access control system with high level security using fingerprints**, Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 32nd, p.238-243, Outubro, 2003.

Zend, **The PHP Company**, Disponível em: <http://www.zend.com/>, Acessado em: 26 de Agosto de 2008.

ANEXO 1 - Funções do Portal de Telemedicina

Descrição de cada ação do Portal de Telemedicina, que serviu como base de pesquisa para este trabalho.

Entrar: ação de entrada no sistema, onde o usuário acessa a primeira pagina do Portal de Telemedicina, que contem campos para preenchimento de nome do usuário e senha correspondente, que serão validados em comparação ao conteúdo armazenado, de forma criptografada, no banco de dados relacional.

Listar: após a validação do usuário no Portal, a primeira tarefa executada é a listagem de todos os exames armazenados, de forma paginada contendo 20 exames, e é ordenado do mais novo exame para o mais velho.

Filtrar: opção de restrição do conteúdo mostrado, a fim de mostrar somente exames com características definidas pelo campo de filtros de pesquisa.

Buscar: opção para refinamento de pesquisa, em relação a código de exame ou nome de paciente.

Visualizar: mostra em uma tela o conteúdo completo de um determinado exame.

Imprimir: formata o conteúdo completo de um exame para impressão.

Altera Senha: realiza a alteração da senha do usuário.

Segunda Opinião: ação de requisição de segunda opinião, em relação ao laudo de um exame.

AIH: ação responsável pela requisição de Autorização de Internação Hospitalar.

Cadastra Paciente: Ação de cadastro de um novo paciente na base de dados.

Insere Exame: ação de inserção de exame na lista de exames a serem listados no sistema.

Laudar: ação de avaliação e resposta ao conteúdo de um exame sem laudo no sistema.

Encaminhar Exame: opção de encaminhamento de um exame para outras regiões para avaliação de casos específicos.

Sair: saída do sistema, libera todos os exames escolhidos para laudo, que não foram laudados, para outros usuários dos sistema.

ANEXO 2 – Portal de Telemedicina do Estado de Santa Catarina

Esta sessão contém parte do Trabalho de Conclusão de Curso defendido em 2005, que descreve o funcionamento do Portal de Telemedicina (Wallauer, 2005).

3 Sistema Gerenciamento de Exames

O sistema de gerenciamento de exames proposto neste trabalho, tem como principais objetivos a marcação de exames e alocação de pessoal, especializado e disponível, ligado ao Sistema Único de Saúde do Estado de Santa Catarina sendo acessível a partir da Internet e de fácil utilização para pessoal não técnico, garantindo os requisitos de segurança necessários a um sistema desta qualidade e integrado com o MIB (Medical Image Browser).

Para tal, idealizou-se um sistema computacional concebido de acordo com os seguintes critérios:

- O sistema deve possuir três formas distintas de visualização: como requisitante, executor ou moderador.
- O requisitante deve ver somente os exames requisitados por ele, com a possibilidade de visualizar as imagens referentes ao exame e o Laudo assim que disponível.
- O executor pode visualizar somente os exames que ele selecionou para dar laudo e os que estão sem um responsável por um laudo, podendo dar laudo ou anexar um novo laudo ao exame, bem como escolher um exame para sua responsabilidade, ou desistir de um exame previamente escolhido por ele.
- O regulador pode visualizar o exame, bem como as imagens, podendo apenas abrir no MIB (Formato original) sem poder de alteração.

Pode-se imaginar o sistema do ponto de vista estrutural tal como apresentado na figura 4.

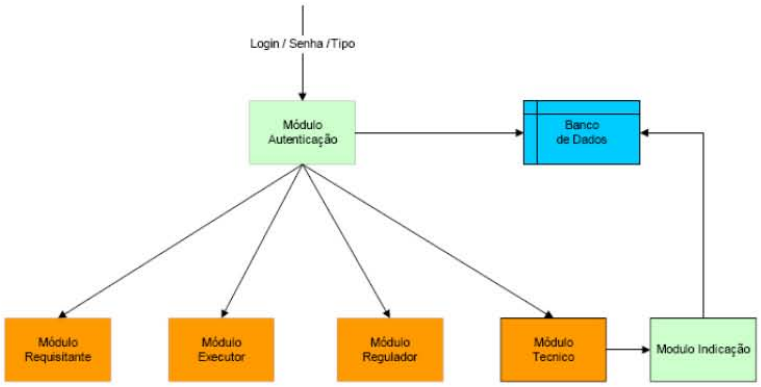

Figura 4 - Com o modelo esquemático do sistema em forma de módulos

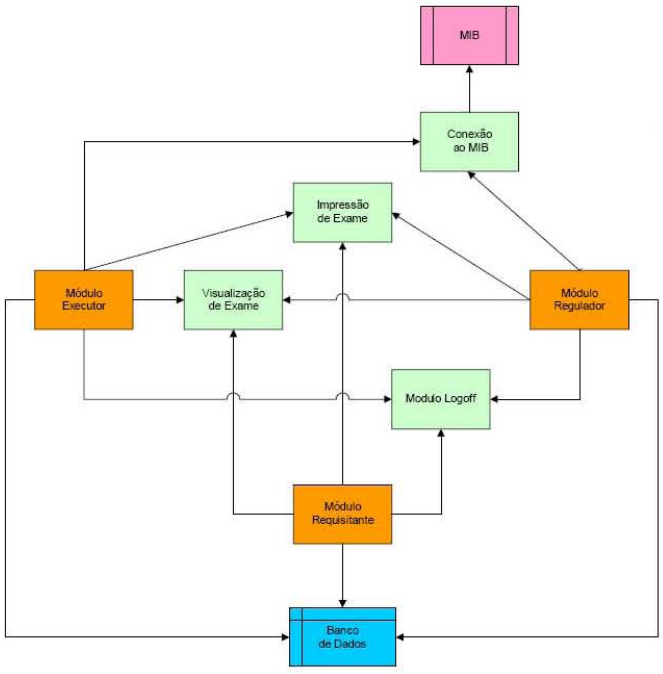

Figura 5 - Detalhamento dos Módulos da Figura 4.

13

3.1 Módulos do sistema

Neste tópico será feita uma descrição de cada módulo pertencente ao sistema de moderação e marcação de consulta via WEB.

3.1.1 Modulo 1

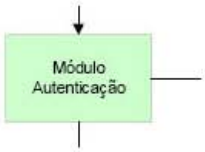

Figura 6 – Módulo Autenticação

Descrição simplificada

Nesse módulo o sistema utiliza o login, senha e o tipo (executor, requisitante, regulador) para requisitar as permissões armazenadas ao módulo de conexão com o banco de dados, autenticando o usuário e o direcionando para o módulo requisitado.

Atributos

Login, Senha e Tipo (executor, requisitante e regulador).

Funcionalidades

Ao conectar-se ao módulo de banco de dados passando os valores de login e senha, verifica-se se o usuário pertence ao sistema, e o módulo de banco de dados por sua vez retorna o valor de permissões que ele possui caso pertença, assim, a autenticação verifica se o usuário possui a referencia correta para o tipo de conexão requisitada e o direciona para o módulo seguinte, que pode ser o módulo de executor, requisitante ou de regulador.

3.1.2 Modulo 2

Figura 7 – Módulo de Banco de Dados

Descrição simplificada

Nesse módulo o sistema utiliza o login e senha para conexão com o banco de dados e retornar as permissões e valores referentes ao conjunto validado pelos valores de login e senha. Assim como recebe dos módulos executor, requisitante e regulador os pedidos ao banco de dados, que são executados e retorna um conjunto de valores.

Atributos

Host (local onde a base está armazenada), Login, Senha, Porta (porta no computador para conexão da base de dados), Tabela (tabela a ser consultada), Query (Consulta a ser realizada na base de dados).

Funcionalidades

O módulo de banco de dados é o responsável direto por todas as conexões realizadas pelo sistema, tanto para verificação das permissões dos usuários do sistema, quanto pelas consultas feitas pelos módulos executor, requisitante e regulador.

3.1.3 Modulo 3

Figura 8 – Módulo executor

Descrição simplificada

Nesse módulo tem como funções principais o tratamento dos dados referentes às permissões de usuário do tipo executor, enviando consultas para o módulo de banco de dados e retornando um conjunto formatado de dados contendo uma referência para o módulo de conexão com o MIB (do acrônimo Medical Image Browser, ou visualizador de imagens médicas), assim como uma referência para visualização utilizando o módulo visualização.

Atributos

Login, Senha, Ordem (Ordem ascendente ou descendente de amostragem dos dados), Filtro, Ação, Referencia, Restrição.

Funcionalidades

Envia o login, senha e query para o módulo de banco de dados, recebendo como retorno um conjunto de dados referentes aos valores do Filtro, Ordem, Referencia e Restrição, mostrando esse conjunto de dados para o usuário num formato pré-definido, conforme a figura 9, que contém uma referencia para o módulo de conexão com o MIB e com o módulo de visualização, além de referência para o módulo de logoff.

Figura 9 – Formato dos dados na interface do Executor.

3.1.4 Modulo 4

Figura 10 – Módulo Requisitante

Descrição simplificada

Esse módulo tem como funções principais o tratamento dos dados referentes às permissões de usuário do tipo requisitante, enviando consultas para o módulo de banco de dados e retornando um conjunto formatado de dados contendo uma referência para visualização utilizando o módulo visualização.

Atributos

Login, Senha, Ordem (Ordem ascendente ou descendente de amostragem dos dados), Filtro, Ordem, Referencia e Restrição.

Funcionalidades

Envia o login, senha e query para o módulo de banco de dados, recebendo como retorno um conjunto de dados referentes aos valores do Filtro, Ordem, Referencia e Restrição, mostrando esse conjunto de dados para o usuário num formato pré-definido, conforme a figura 11, que contém uma referencia para o módulo de visualização e para o módulo de logoff.

Figura 11 - Formato dos dados na interface do Requisitante.

3.1.5 Modulo 5

Figura 12 – Módulo Regulador

Descrição simplificada

Nesse módulo tem como funções principais o tratamento dos dados referentes às permissões de usuário do tipo regulador, enviando consultas para o módulo de banco de dados e retornando um conjunto formatado de dados contendo uma referência para o módulo de conexão com o MIB (do acrônimo Medical Image Browser, ou visualizador de imagens médicas), assim como uma referência para visualização utilizando o módulo visualização e uma referencia para o módulo de conexão ao SisREG.

Atributos

Login, Senha, Ordem (Ordem ascendente ou descendente de amostragem dos dados), Filtro, Referencia, Restrição.

Funcionalidades

Envia o login, senha e query para o módulo de banco de dados, recebendo como retorno um conjunto de dados referentes aos valores do Filtro, Ordem, Referencia e Restrição, mostrando esse conjunto de dados para o usuário num formato pré-definido, conforme a figura 13, que contém uma referencia para o módulo de conexão com o MIB (Medical Image Browser), uma referencia pra o módulo de conexão ao SisREG, com o módulo de visualização, além de referência para o módulo de logoff.

Figura 13 - Formato dos dados na interface do Regulador.

3.1.6 Modulo 6

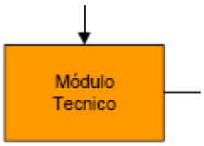

Figura 14 – Módulo Técnico

Descrição simplificada

Nesse módulo tem como funções principais o tratamento dos dados referentes às permissões de usuário do tipo técnico, enviando consultas para o módulo de banco de dados e retornando um conjunto formatado de dados contendo uma referência para o módulo de Indicação.

Atributos

Login, Senha, Ordem (Ordem ascendente ou descendente de amostragem dos dados), Filtro, Referencia e Restrição.

Funcionalidades

Envia o login, senha e query para o módulo de banco de dados, recebendo como retorno um conjunto de dados referentes aos valores do Filtro, Ordem, Referencia e Restrição, mostrando esse conjunto de dados para o usuário num formato pré-definido, conforme a figura 15, que contém uma referencia para o módulo de indicação.

Figura 15 - Formato dos dados na interface do Técnico.

3.1.7 Modulo 6

Figura 16 – Módulo Visualização.

Descrição simplificada

No módulo de visualização a referência passada, pelos módulos de executor, requisitante e regulação, é formatada.

Atributos

Identificador (referencia do exame na base de dados).

Funcionalidades

No módulo de visualização a referência passada pelos módulos de executor, requisitante e regulação é formatada segundo o padrão mostrado na figura 17, onde na parte superior contém um cabeçalho, logo após pode vir um texto de laudo e por ultimo as imagens.

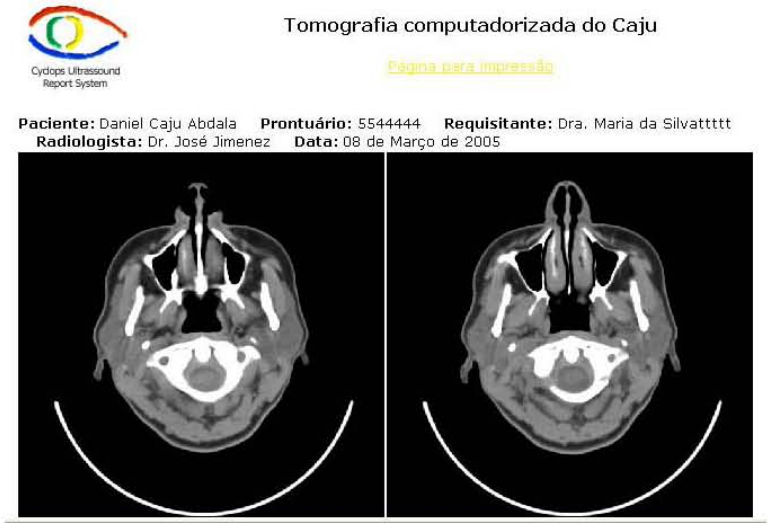

Figura 17 – Formato dos dados no módulo de Visualização.

3.1.8 Modulo 7

Figura 18 – Módulo de Visualização de Exames.

Descrição simplificada

No módulo de visualização a referência passada, pelos módulos de executor, requisitante e regulação, é formatada.

Atributos

Identificador (referencia do exame na base de dados).

Funcionalidades

No módulo de visualização a referência passada pelos módulos de executor, requisitante e regulação é formatada segundo o padrão mostrado na figura 19, onde na parte superior contém um cabeçalho, logo após pode vir um texto de laudo e por ultimo as imagens.

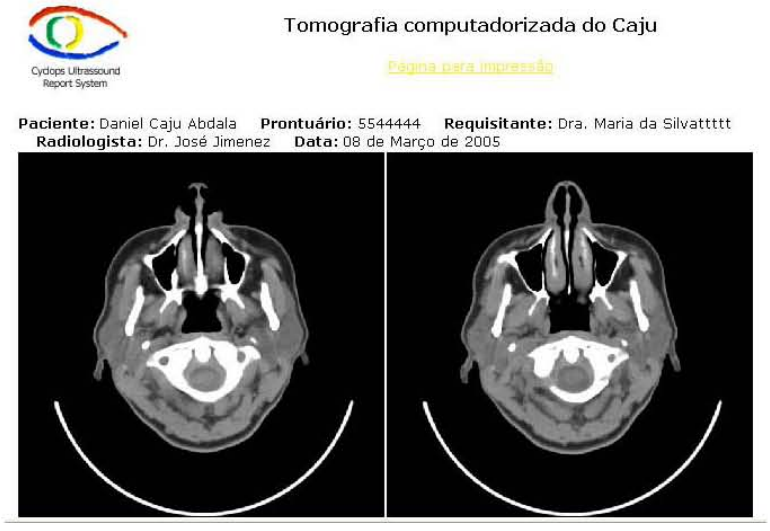

Figura 19 – Formato dos dados no módulo de Visualização.

3.1.9 Modulo 8

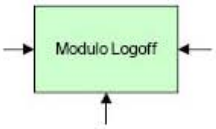

Figura 20 – Módulo Logoff.

Descrição simplificada

No módulo Logoff é feito o processo de saída do sistema.

Atributos

Identificador (identificador da sessão).

Funcionalidades

No módulo Logoff é onde são destruídos todos os dados de sessão, como senha, login e referencias de utilização do sistema que possam ter ficadas armazenadas temporariamente.

3.1.10 Modulo 9

Figura 21 – Módulo de conexão com o MIB

Descrição simplificada

No módulo Conexão ao MIB é enviado um conjunto de dados para o artefato de software MIB.

Atributos

Data.

Funcionalidades

No módulo, Conexão ao MIB, é utilizado o conjunto de dados enviados pelos módulos executor e regulador, para gerar um meta dado em formato XML para enviar informações, referente ao exame a ser manipulado, para o MIB.

3.1.11 Modulo 10

Figura 22 – Módulo de Indicação

Descrição simplificada

No módulo de indicação é feito o processo de indicação de um exame.

Atributos

Indicação e query.

Funcionalidades

O módulo de indicação monta uma consulta de atualização com a indicação e requisitante, e envia para o banco de dados na posição referente ao exame a ser indicado.

3.2 *Administração do sistema*

Para obter acesso à área administrativa, é necessário possuir um login e senha de administrador. A tela de login para administradores é exibida abaixo.

Figura 23 – Tela de login para administradores.

Após efetuar o login com sucesso, o administrador terá acesso à tela principal da área administrativa, a partir de onde as demais funcionalidades estarão acessíveis através de um menu, do lado esquerdo da tela.

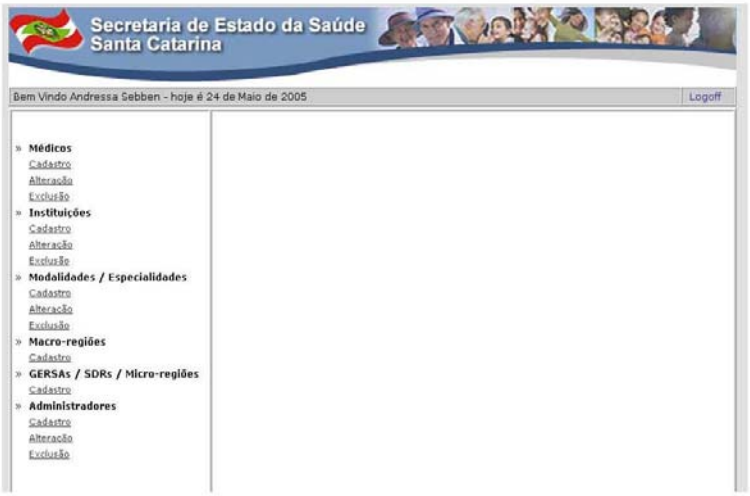

Figura 24 – Tela pós login na área de administradores do sistema

As tarefas que podem ser acessadas a partir do menu são descritas a seguir.

3.2.1 Médicos

Manutenção de cadastro de médicos, composto das seguintes opções:

3.2.1.1 Cadastro

Para que um médico tenha acesso ao sistema, deverá ser previamente cadastrado. O cadastro é constituído dos seguintes campos, sendo que todos são obrigatórios:

- *Nome*: nome do médico.
- *Sexo*: M/F.
- *Login*: nome de acesso ao sistema para o médico que está sendo cadastrado. O login deve possuir no mínimo 6 caracteres e deve ser único, ou seja, não é possível cadastrar dois médicos com o mesmo login.
- *Senha*: juntamente com o login, permite ao médico acessar o sistema. A senha deve conter no mínimo 6 caracteres.
- *Confirmação de senha*: por motivo de segurança, a senha deve ser digitada novamente. Dessa forma, evita-se que erros de digitação no momento do cadastro impeçam o acesso.
- *Instituição*: este campo traz uma lista de todas as instituições já cadastradas. Deve-se selecionar a instituição na qual o médico atua.
- *Tipo*: determina as permissões do usuário no sistema. Deve-se selecionar ao menos um dos seguintes tipos:
    - Requisitante: // descrever cada um....
    - Executor:
    - Regulador:
    - Técnico:

Após a digitação das informações, o botão **Cadastrar** deve ser pressionado para confirmar os dados. Se alguma restrição no preenchimento

dos campos não foi respeitada, um erro é exibido na parte superior do formulário, e os dados não serão confirmados até que o erro seja corrigido.

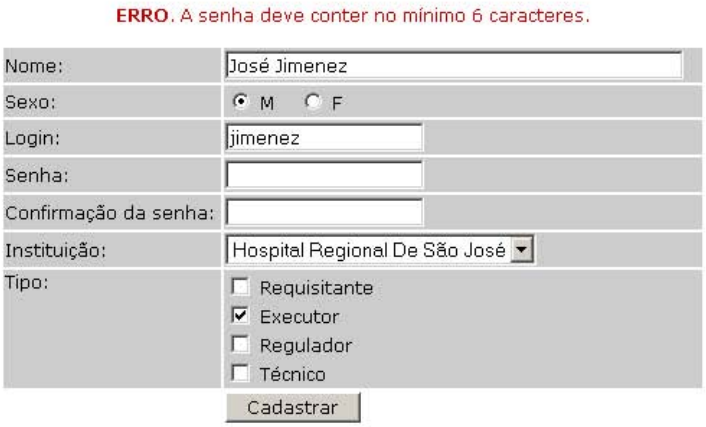

Figura 25 – Demonstração de erro no cadastro de médico.

3.2.1.2 Alteração

As informações do médico podem ser alteradas através desta opção. Primeiro, é necessário escolher o médico que se deseja alterar:

Figura 26 – Seleção do médico a ter o cadastro alterado.

Ao selecionar o nome do médico o clicar **OK**, o cadastro do médico será aberto, permitindo que as informações sejam editadas. As restrições para a validade dos campos são as mesmas descritas no tópico cadastro de médico.

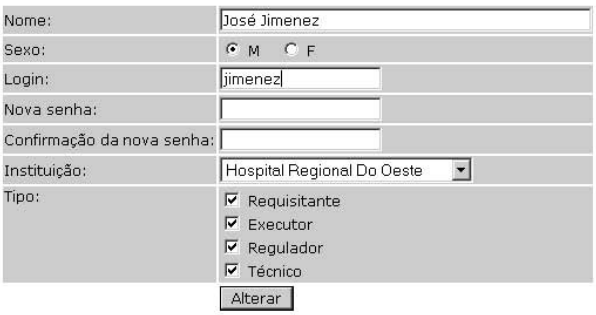

Figura 27 – Tela de alteração de cadastro médico.

Se o campo *Nova senha* for deixado em branco, o usuário permanecerá com sua senha antiga. Se este campo for preenchido e confirmado (campo *Confirmar Nova Senha*), a senha do usuário será alterada.
Para confirmar as alterações, pressione o botão **Alterar**.

3.2.1.3 Exclusão

Para excluir um médico, é necessário digitar efetuar uma busca, digitando seu nome, ou as primeiras letras de seu nome.

Após digitar e pressionar o botão **OK**, serão listados logo abaixo os registros de médicos que coincidem com o nome digitado.

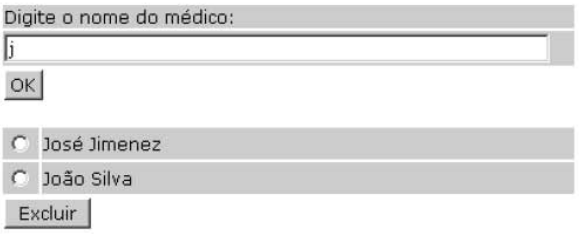

Figura 28 – tela de listagem e exclusão de médico.

Deve-se escolher o nome do médico a ser excluído e clicar sobre o botão Excluir. Será solicitada uma confirmação para esta operação.

Figura 29 – Confirmação da exclusão de um médico do sistema.

Clicando em **Cancelar** a ação será ignorada e o médico não será excluído. Pressionar o botão **OK** fará com que a ação seja confirmada e a exclusão do médico efetivada.

3.2.2 Instituições

Manutenção do cadastro de hospitais. As ações permitidas são Cadastro, Alteração e Exclusão.

3.2.2.1 Cadastro

Para cadastrar uma instituição, todos os dados do formulário devem ser preenchidos. A validade do CNPJ digitado será verificada, e não será possível efetuar o cadastro se alguma informação estiver incorreta ou em branco. Também não é possível cadastrar duas instituições com o mesmo CNPJ.

Figura 30 – Tela de cadastro de instituição.

Ao cadastrar uma instituição deverá também ser selecionada a GERSA à qual ela encontra-se vinculada.

3.2.2.2 Alteração

A alteração de um cadastro de instituição pode ser feita selecionando-se o nome da instituição e clicando em **OK**.

Figura 31 – listagem das instituições para alteração de cadastro.

Em seguida, o cadastro da instituição será aberto para edição.

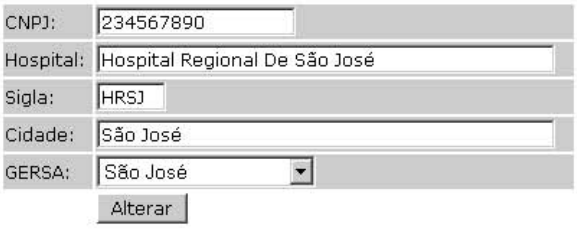

Figura 32 – Edição do cadastro de instituição.

As restrições sobre a validade das informações são as mesmas já apresentadas no tópico 2.1. Não é possível confirmar a alteração se os dados não estiverem no formato correto.

3.2.2.3 Exclusão

Para excluir uma instituição, digite o nome ou as primeiras letras do nome na caixa de busca. Após clicar em OK, serão listados logo abaixo todos os nomes das instituições que coincidem com o nome digitado.

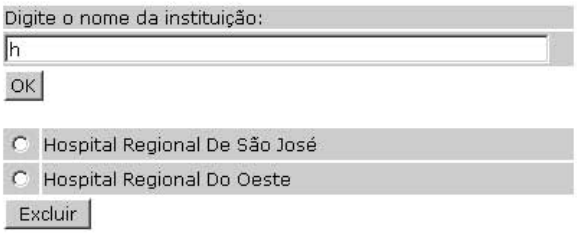

Figura 33 – tela de pesquisa e exclusão de instituição.

Selecione a instituição a ser excluída e clique em **Excluir**. Uma confirmação de exclusão será solicitada.

Figura 34 – Confirmação da exclusão de um a instituição.

3.2.3 Modalidades / Especialidades

Manutenção do cadastro dos tipos de exame.

3.2.3.1 Cadastro

O cadastro de modalidades é composto de apenas um campo, o qual contém o nome da modalidade / especialidade. Basta digitar o nome e clicar sobre o botão **Cadastrar**

Figura 35 – Tela de cadastro de modalidade de exame.

3.2.3.2 Alteração

Para alterar uma modalidade, é necessário selecioná-la primeiro na lista de modalidades e clicar em **OK**.

Figura 36 – seleção de modalidade para alteração.

Após selecionar a modalidade, será possível alterá-la em uma caixa de edição padrão, idêntica ao formulário de cadastro de modalidades (3.2.3.1).

3.2.3.3 Exclusão

Para excluir uma modalidade deve-se digitar seu nome (ou as primeiras letras do nome) na caixa de busca e, em seguida, pressionar o botão **OK**. A lista de modalidades que coincidem com o valor digitado na caixa de busca é mostrada logo abaixo dela.

Figura 37 – Tela de pesquisa e exclusão de modalidade.

Ao selecionar uma modalidade e clicar em **Excluir**, uma confirmação será solicitada antes que a modalidade seja efetivamente excluída.

Figura 38 – Confirmação de exclusão de modalidade.

3.2.4 Macro regiões

Cadastro de novas macro regiões do estado de Santa Catarina. Para cadastrar uma macro região, basta informar o nome e clicar em **Cadastrar**.

Figura 39 – Tela de cadastro de macro regiões.

3.2.5 GERSAs / SDRs / Micro-regiões

Este módulo permite que sejam cadastradas novas GERSAs (Gerência Regional de Saúde) em cada macro região. Basta selecionar a macro região à qual a GERSA pertence, informar seu nome e clicar em **Cadastrar**:

Figura 40 – Cadastro de GERSAS para uma referida Macro Região.

3.2.6 Administradores

Manutenção do cadastro de administradores do sistema, onde é possível executar as seguintes ações:

3.2.6.1 Cadastro

É possível cadastrar novos administradores para o sistema. Todos os campos devem ser preenchidos, observando a restrição de 6 caracteres no mínimo para login e senha. O CPF será verificado quanto à sua validade, e não será permitido cadastrar dois administradores com o mesmo CPF.

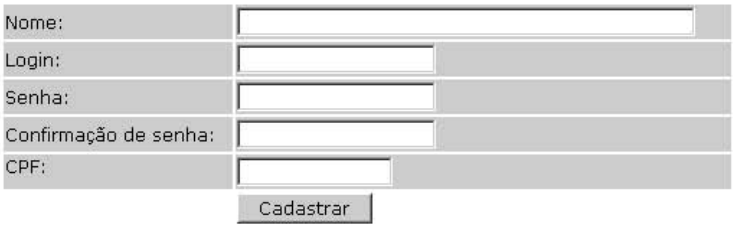

Figura 41 – Tela de Cadastro de Administradores do sistema.

Caso algum campo não seja preenchido ou alguma restrição não seja respeitada, o sistema informará o erro na parte superior do formulário, e não será possível confirmar o cadastro até que o erro seja contornado.

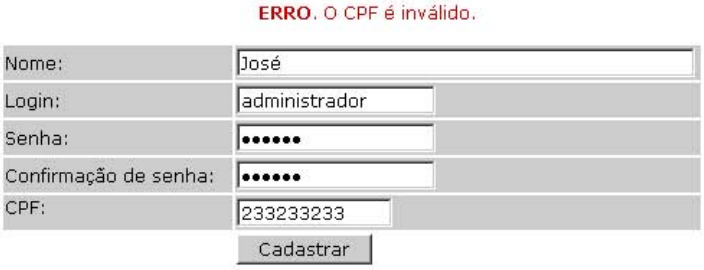

Figura 42 – Demonstração de erro no cadastro de um administrador.

3.2.6.2 Alteração

Para alterar as informações de um administrador, é necessário selecionar primeiro seu nome na lista:

Figura 43 – seleção de um administrador para alteração de informações.

Após selecionar um nome e clicar em **OK**, o cadastro do administrador será aberto para edição.

Figura 44 – Tela de alteração de cadastro de um usuário.

Os dados estão sujeitos às mesmas restrições citadas n cadastro de um administrador.

3.2.6.3 Exclusão

Para excluir um administrador, digite seu nome, ou as primeiras letras do nome e clique em **OK**. Serão mostrados logo abaixo os nomes dos administradores que coincidem com o texto digitado na caixa de busca.

Figura 45 – Tela de listagem e exclusão de um administrador.

Após selecionar um nome e clicar em **Excluir**, uma confirmação será solicitada. O administrador será excluído apenas se a ação for confirmada, clicando no botão **OK**.

Figura 46 – Confirmação para a exclusão de um administrador.

36