Intrusion Detection System in Ad Hoc Networks with Neural Networks Artificial and K-Means Algorithm

D. Canêdo, and A. Romariz

***Abstract*—There has been a great technological advance in the infrastructure of mobile technologies. The increase in the use of wireless local area networks and the use of services from satellites is also noticeable. The high rate of use of mobile devices for various purposes brings the need to monitor the wireless networks to ensure the integrity and confidentiality of the information. Therefore it is necessary to quickly and efficiently identify the normal and abnormal traffic of these wireless networks so that their administrators can take action. This paper presents a proposal for a Ad Hoc Wireless Intrusion Detection System composed of two stages, based on data grouping through the algorithm K-Means and Artificial Neural Networks through the Multilayer Perceptron algorithm, for the detection and classification of anomalies caused by attacks on the networks of Computers.**

***Index Terms*—Ad Hoc Wireless Networks, Multilayer Perceptron, K-Means, Intrusion Detection System.**

I. Introdução

REDES de Computadores tem visto um aumento significativo em sua infraestrutura de conexão, tornando a segurança da informação um desafio. Atualmente, há um aumento no uso de Redes de Computadores Sem Fio, tanto em ambientes residenciais quanto no ambiente corporativo. Dados recentes da Anatel (Agência Internacional de Telecomunicações) mostram que cerca de 260 milhões de brasileiros usam dispositivos móveis para acessar a Internet para criar, transmitir ou consumir informações [1].

Redes Ad Hoc segue o padrão 802.11, sendo definida como Redes de Computadores Sem Fio sem a presença de um componente concentrador, tornando cada nó da Rede responsável pelo roteamento e controle de acesso ao meio e gerenciando algumas características da Rede como: Baixa taxa de transmissão; Probabilidade de erro; Variações no meio de transmissão. Essas redes são formadas em ambientes onde há necessidade de comunicação, mas há uma inoperabilidade de Redes Sem Fio com estrutura, tornando as Redes Ad Hoc de natureza temporária e complexa [2].

No entanto, Redes Ad Hoc estão sujeitas a ataques que podem ter origem interna e externa, alguns deles com o objetivo de paralisar alguns serviços dos nós da própria Rede, enquanto outros têm o objetivo de capturar informações que trafegam entre os nós de Redes Ad Hoc.

D. R. Canêdo, Universidade de Brasília, Brasília, Distrito Federal, Brasil e Instituto Federal de Goiás, Luziânia, Goiás, Brasil, daniel.canedo@ifg.edu.br.

A. R. S. Romariz, Universidade de Brasília, Brasília, Distrito Federal, Brasil, alromariz@gmail.com.

A confidencialidade, integridade e disponibilidade dos recursos das Redes são fundamentais para prover a segurança da informação, sendo que um processo de anomalia em Redes de Computadores incluindo as Redes Ad Hoc podem comprometer sistemas, caracterizando uma intrusão. O IDS (*Intrusion Detection System*) tem o propósito de identificar intrusões em Redes de Computadores, sem comprometer o funcionamento normal da Rede. O Sistema de Detecção de Intrusão é considerado uma ferramenta de segurança de Redes, que em conjunto com outras ferramentas de segurança são organizadas para reforçar a segurança da informação em sistemas de comunicação [3].

A análise do tráfego de rede nas Redes AdHoc é dificultada pela falta de gerenciamento central. Outra característica importante a considerar é a alta mobilidade dos componentes da Rede Ad Hoc, já que pode-se entrar e sair da Rede sem restrições. Outra característica é que os componentes da Rede Ad Hoc são na maioria das vezes dispositivos móveis, que possuem restrições em seu estado ativo, pois dependem da energia de seus recursos. Estas características das Redes Ad Hoc remetem que os Sistemas de Detecção de Intrusão tradicionais não são usados diretamente.

Este artigo apresenta uma proposta de Sistema de Detecção de Intrusão para Redes Ad Hoc através de duas etapas. A primeira etapa destina-se a agrupar todos os tráfegos da Rede de um determinado nó através da utilização do Algoritmo K-Médias, enquanto a segunda etapa é classificar as anomalias levando em conta as informações do grupo, através da aplicação de Redes Neurais Artificiais com o algoritmo *Multilayer Perceptron*.

A estrutura deste artigo esta organizada em seções. Na seção dois serão apresentadas propostas para Sistemas de Detecção de Intrusão. Na seção 3 apresenta-se a fundamentação teórica abordando Redes Sem Fio Ad Hoc, enquanto que na seção 4 aborda-se Sistemas de Detecção de Intrusão em Redes Ad Hoc. Na seção 5 será apresentada a abordagem proposta do Sistema de Detecção de Intrusão, bem como os resultados da simulação da mesma. Na seção 6 apresenta-se a conclusão do trabalho e a apresentação de trabalhos futuros.

II. Trabalhos Relacionados

Na literatura existem trabalhos de classificação de tráfego de Redes Wireless, os quais podem ser aplicados em Sistemas de Detecção de Intrusão. Estas propostas utilizam métodos de aprendizagem supervisionados e não supervisionados.

A proposta de Chandrashekar(2014) [4] fornece uma abordagem geral dos vários métodos de classificação, usando dados de alta dimensão e uma técnica de seleção de variáveis com o objetivo de reduzir o tempo computacional e a velocidade de aprendizagem.

Govindarajan apresenta uma proposta [5] de dois métodos de classificação envolvendo perceptron multicamada e função de base radial. Propõe-se neste trabalho uma arquitetura híbrida envolvendo ambos os classificadores para sistemas de detecção de intrusão.

Cervantes apresenta uma proposta [6] de sistema de detecção de intrusão contra ataques *sinkhole* e *selective forwarding* sobre o roteamento na IoT densa e móvel. Utiliza agrupamento para lidar com a densidade e a mobilidade, e combina estratégias de *watchdog*, reputação e confiança na detecção de atacantes, a fim de garantir confiabilidade aos dispositivos.

EdWilson apresenta uma proposta [7] de Sistema de Detecção de Intrusão híbrido, em que realiza-se um processamento de sinais através da utilização de transformações Wavelets e posteriormente a classificação das anomalias utilizando Redes Neurais Artificais.

EdWilson apresenta uma proposta [8] que propõe a elaboração de uma base de dados reais de tráfego de Redes Wireless, a qual será utilizada na avaliação do Sistema de Detecção de Intrusão - IDS - proposto. Estes dados por sua vez sofrem um pré-processamento para posteriormente serem classificados por técnicas de reconhecimento de padrões, como por exemplo Redes Neurais Artificiais.

III. Redes Sem Fio Ad Hoc

Uma Rede Ad Hoc é formada em situações onde há necessidade de comunicação e uma infraestrutura fixa não está disponível ou não é desejável [9]. Nesse caso, os nós móveis formam uma rede para uso temporário, a fim de atender às necessidades de comunicação naquele momento, ou ad hoc. Uma Rede Sem Fio Ad Hoc, também denominada de MANET, é um sistema de rede sem fio com nós móveis que podem mover-se livremente e são auto-organizáveis com topologia dinâmica e permite que equipamentos possam utilizar a rede sem comunicação preexistente, diferente de uma rede com infraestrutura fixa [9].

A Figura 1 [10] apresenta um modelo de Rede Sem Fio Ad Hoc que permite a comunicação diretamente entre os nós, e estes por sua vez podem realizar o repasse de pacotes através de múltiplos saltos. Cada elemento da rede é responsável pelo encaminhamento de pacotes de seus vizinhos. Cada nó é equipado com uma ou mais interface de rádio, e a cobertura da rede depende diretamente do alcance destes enlaces. Até certo ponto, é possível adicionar mais nós na rede e, consequentemente, aumentar sua cobertura.

Os nós da rede também podem funcionar como roteadores para outros nós, com o encaminhamento de pacotes para o destinatário final. Esta rede pode possuir conexão com rede com infraestrutura, através de *gateways*. Alguns exemplos de aplicações deste modelo de rede são campos de batalhas militares, locais onde existe necessidade de formação rápida de redes, missões de resgate, redes de sensores para automação e aplicações em eventos [11].

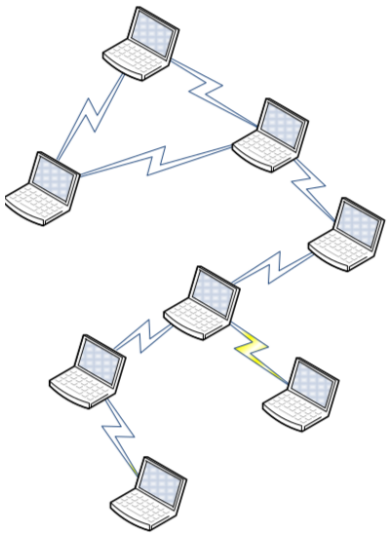

Fig. 1. Redes Sem Fio Ad Hoc [10].

IV. Sistema de Detecção de Intrusão em Redes Ad Hoc

Uma intrusão é definida como certas ações cuja finalidade é comprometer as propriedades de confidencialidade, integridade e disponibilidade dos recursos da Rede de Computadores. Um Sistema de Detecção de Intrusão - IDS - deve ser capaz de identificar ações maliciosas, no entanto, não deve comprometer a operação da Rede de Computadores. O IDS, por outro lado, deve consumir poucos recursos computacionais, para não prejudicar usuários legítimos.

Confidencialidade, integridade e disponibilidade de recursos representam fatores vitais para a segurança da informação, onde uma ação maléfica ou não intencional pode comprometer o sistema, caracterizando uma intrusão. O sistema de detecção deve conseguir identificar essa ação, mas sem comprometer o funcionamento normal da rede. Um sistema de detecção é uma ferramenta de segurança que, como outras medidas, a exemplo de antivírus e *firewalls*, destinam-se a reforçar a segurança da informação em sistemas de comunicação [3].

Os IDS's são usados para monitorar, avaliar e informar violações de segurança que podem ser intencionais ou não. No entanto, as técnicas de detecção e prevenção não avançam no mesmo ritmo, o que dificulta sua aproximação.

De uma forma geral os Sistemas de Detecção de Intrusão tradicionais não são empregados diretamente nas Redes Ad Hoc devido à particularidade de sua infraestrutura, pois apresenta influência direta no funcionamento do IDS. Atualmente existem propostas envolvendo Sistemas de Detecção de Intrusão em Redes Ad Hoc, sendo algumas delas apresentadas na seção de Trabalhos Relacionados.

V. Abordagem Proposta

A proposta deste artigo caracteriza-se por um Sistema de Detecção de Intrusão em Redes Ad Hoc através de duas etapas,

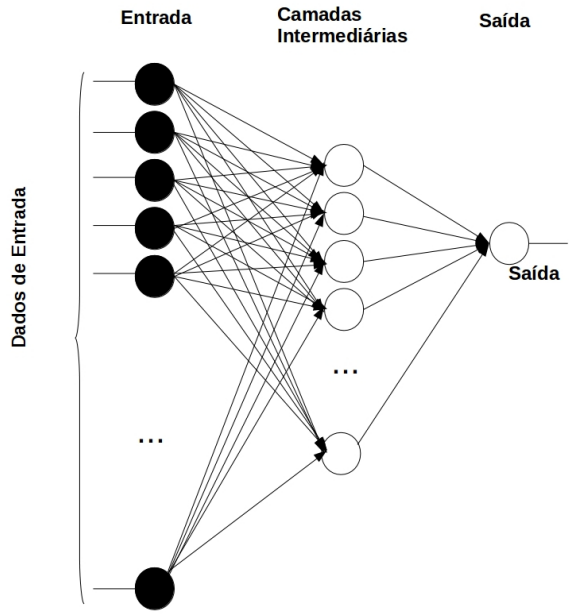

Fig. 2. Redes Neurais de Várias Camadas.

sendo a primeira dedicada ao agrupamento dos dados em grupos provenientes de uma Rede Sem Fio Ad Hoc, enquanto a segunda etapa é responsável pela classificação de anomalias pré-definidas.

O Sistema de Detecção de Intrusão proposto faz uso das Técnicas de Inteligência Computacional para executar as duas etapas relatadas. Para a primeira etapa, o agrupamento de dados do tráfego da Rede Ad Hoc é realizado pelo algoritmo K-Médias, que é um algoritmo de aprendizado não supervisionado. O algoritmo K-Médias separa certos objetos em grupos, chamados *clusters*. Estes *clusters* são formados através da aplicação de técnicas de medição de distância ou técnicas de similaridade entre objetos [12]. Este algoritmo é escolhido principalmente pela simplicidade computacional. O algoritmo K-Médias é capaz de processar grandes volumes de dados, cuja complexidade de armazenamento é O((m + K)n), onde *m* é o número de pontos e *n* é o número de atributos [12].

Após o agrupamento de dados da Rede Ad Hoc, o segundo passo é realizado para classificar anomalias pré-definidas utilizando a técnica de inteligência computacional Redes Neurais Artificiais através do algoritmo *Multilayer Perceptron*. O algoritmo *Multilayer Perceptron* será composto por pelo menos uma camada oculta entre a entrada e a saída. As camadas ocultas não possuem conexões com o mundo exterior, como mostra a Figura 2. Esse tipo de Rede Neural está sendo utilizado em larga escala para resolver problemas complexos, pois tem como característica o treinamento supervisionado com o processo de Correção de Erros, como o algoritmo de retropropagação [13]. A Figura 3 mostra o comportamento de um neurônio no processo de Aprendizagem por Correção de Erros, com os elementos fundamentais sendo o vetor de entrada, camada de neurônio oculto, neurônio de saída, função de soma.

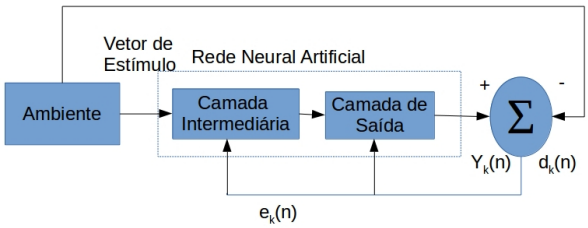

Fig. 3. Aprendizagem por Correção de Erro.

A. Base de Dados

A base de dados utilizada neste trabalho é uma coleção real de tráfegos de rede capturados na arquitetura Ad Hoc. Estes dados por sua vez representam o comportamento de usuários que frequentemente utilizam a Rede Sem Fio Ad Hoc para acessar diversas informações, bem como para a utilização da Internet. Segundo Ferreira(2015) [8] a base de dados é construída a partir do tráfego obtido pela comunidade acadêmica da instituição na qual o experimento é realizado, sendo que esta base ainda não se encontra disponível publicamente.

Para a coleta dos dados utiliza-se dois cenários distintos, pois tem o objetivo de aumentar as possibilidades de tráfego da Rede. Os cenários abordados possuem configurações e topologias próprias, sendo um cenário representando um ambiente doméstico típico de Redes Wireless, enquanto o outro cenário é um ambiente um pouco mais complexo, corporativo.

Esta base de dados é composta por um total de 616047 registros, sendo que cada registro é composto por 16 variáveis que são características do próprio tráfego de rede. Além disso, em cada registro do banco de dados, a classe à qual pertence determinado registro é definida. A classificação é realizada nas seguintes classes:

- Normal: Dados que possuem características de tráfego de Redes Wireless aceitáveis;
- *EAPOLStart*: Uso do protocolo *Extensible Authentication Protocol*(EAP), cujo o objetivo é realizar um método de autenticação tanto na utilização do protocolo *Wired Equivalent Provacy*(WEP), tanto para o protocolo *Wi-Fi Protected Access*(WPA), em suas versões comerciais para acesso a Rede Sem Fio. Esta anomalia se caracteriza por uma carga excessiva de solicitação EAPOL - Start, que em um sobrecarregamento do *Access Point*, responsável pela interconexão dos dispositivos da Rede Wireless;
- *BeaconFlood*: Solicitações de tipo de gerenciamento, destinadas a transmitir milhões de Beacons inválidos, dificultando que um componente sem fio específico identifique um ponto de acesso legítimo. Esses *beacons* ajudam a identificar a localização do BSS (*Basic Set Service*) de uma Rede Wireless [14];
- *Deauthentication*: Solicitações do tipo gerenciamento, que são injetados na Rede Wireless. Os quadros pertencentes a esta anomalia são transmitidos como pedidos imaginários, os quais solicitão a desautenticação de um dispositivo que se encontra autorizado na Rede Wireless;
- *RTSFlood*: Denominado *Request-to-Send Flood* é um quadro do tipo controle. Esta anomalia se baseia na

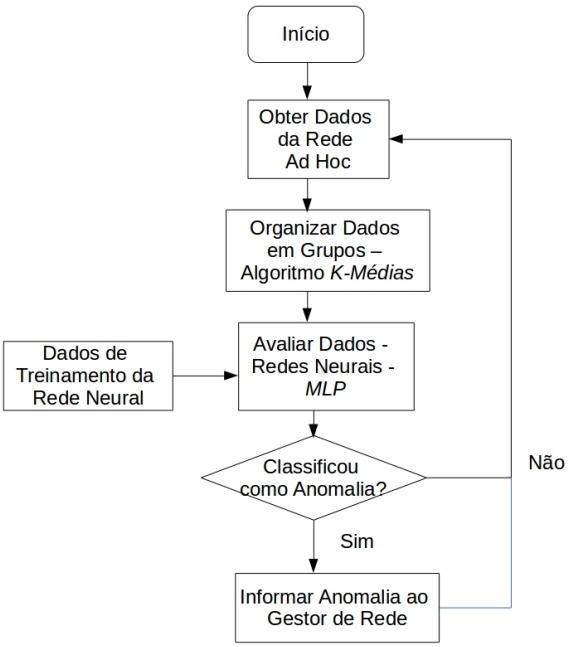

Fig. 4. Algoritmo Proposto.

transmissão em grande escala de pacotes ou frames RTS por um curto período de tempo. A inundação de frames RTS na Rede Wireless proporcionará o congestionamento na reserva do canal Wireless, resultando no processo de negação de serviço aos nós da Rede Wireless [14].

B. Modelo Proposto

Nesse trabalho é proposto o uso de um Sistema de Detecção de Intrusão local com detecção em duas etapas. A primeira agrupa os dados capturados da Rede Sem Fio Ad Hoc, enquanto que a segunda, classifica os ataques. A primeira etapa agrupa e classifica os dados em clusters, que são compostos por dados que possuem uma proximidade.

A segunda etapa é composta por uma Rede Neural Artificial *MultiLayer Perceptron*, com cinco neurônios na saída. Essa Rede Neural é treinada para reconhecer cinco classes, sendo quatro classes de ataques e uma classe de tráfego normal, como é apresentado na Figura 6.

A Rede Neural é formada por um MLP (*Multilayer Perceptron*) treinado com o algoritmo *backpropagation* [15]. A entrada é composta por 17 neurônios, sendo 16 deles referentes às variáveis da Rede Sem Fio Ad Hoc e 1 neurônio referente a informação do *cluster* resultante da etapa anterior. A camada oculta é formada por 10 neurônios.

O algoritmo para o Sistema de Detecção de Intrusão proposto é apresentado na Figura 4. Os dados da Rede, são obtidos através da captura do tráfego da rede. Note que o algoritmo fica continuamente executando para analisar os dados e gerar seus *clusters*, através do algoritmo K-Médias. Em seguida, o algoritmo *Multilayer Perceptron* classifica as anomalias reconhecidas para posterior comunicação com o gestor da Rede.

A proposta deste Sistema de Detecção de Intrusão, conforme apresentado ns Figura 4, deve ficar em execução continuamente, para obter os dados da rede. É realizado um pré-processamento dos dados que contenham características importantes da arquitetura de Redes Sem Fio Ad Hoc, como por exemplo quadros da camada de enlace, a exemplo dos quadros de solicitação de associação em pontos de acesso. Com isso, será possível identificar anomalias exclusivas dessa arquitetura. Como os dados são obtidos diretamente na rede, não é esperado aumento de *overhead*, independente da arquitetura utilizada.

O próximo passo após realizar o pré-processamento dos dados obtidos da Rede Sem Fio Ad Hoc é a geração dos *clusters* através da execução do algoritmo K-Médias. Para a execução do algoritmo K-Médias define-se a construção de 10 *clusters* e a utilização da função de distância Euclidiana para medir a similaridade entre os dados de cada grupo.

A Figura 5 mostra o fluxograma de funcionamento do algoritmo K-Médias, que é composto de seis passos fundamentais. O primeiro, abrange o valor preliminar dos centróides, ou seja, (C1, C2, ...) representa os centróides que se harmonizam. O segundo, a distância dos objetos dos centróides, sendo a distância entre o centróide do cluster e todos os objetos calculados. A distância euclidiana é usada e depois a matriz de distância na iteração 0 é calculada. Cada coluna na matriz de distância significa um objeto. A distância da matriz na primeira linha corresponde à distância de cada objeto à segunda linha e o primeiro centróide representa a distância de cada objeto no segundo centróide. No terceiro passo, têm-se o agrupamento de objetos, alocando todos os objetos baseados na menor distância. O quarto passo, determina os centróides, identificando os componentes de todos os grupos, sendo o novo centróide de cada conjunto determinado com base nessas novas associações. No quinto passo, repetindo a partir do segundo passo. Por fim, o último passo realiza a comparação do último agrupamento de iteração e essa iteração indica que os grupos não são movidos pelos objetos. Assim, realiza-se a classificação dos dados da Rede Sem Fio Ad Hoc, sendo acrescentado a informação do *cluster* a qual cada quadro de dados pertence [16].

O terceiro passo do algoritmo proposto é classificar as anomalias existentes na base de teste, através da execução do algoritmo *Multilayer Perceptron*. Para a execução do algoritmo *Multilayer Perceptron* é definido uma Rede Neural com 17 neurônios na entrada(16 referentes à rede sem fio Ad Hoc e 1 referente ao *cluster*) e 10 neurônios para a camada oculta, como mostra a Figura 6. A utilização de 10 neurônios na camada oculta se justifica pelo fato da Rede Neural convergir para a melhor saída, sendo que é reduzido ao número máximo de neurônio. O algoritmo *Multilayer Perceptron* de acordo com o algoritmo proposto é treinado com dados preexistentes, contendo algumas anomalias definidas. Se uma nova anomalia for encontrada, o algoritmo proposto deve ser iniciado novamente para dar eficiência aos administradores da rede.

Se uma anomalia for detectada pelo algoritmo *Multilayer Perceptron*, deve-se relatar a intrusão ao administrador da rede, atualizando os registros em um arquivo de log. Caso contrário, retoma o fluxo normal de processamento.

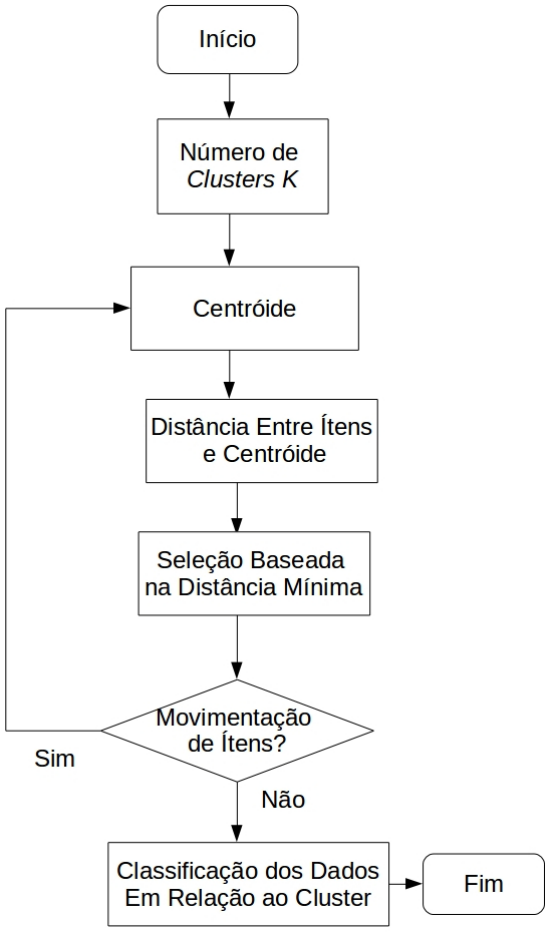

Fig. 5. Fluxograma de Funcionamento do Algoritmo K-Médias.

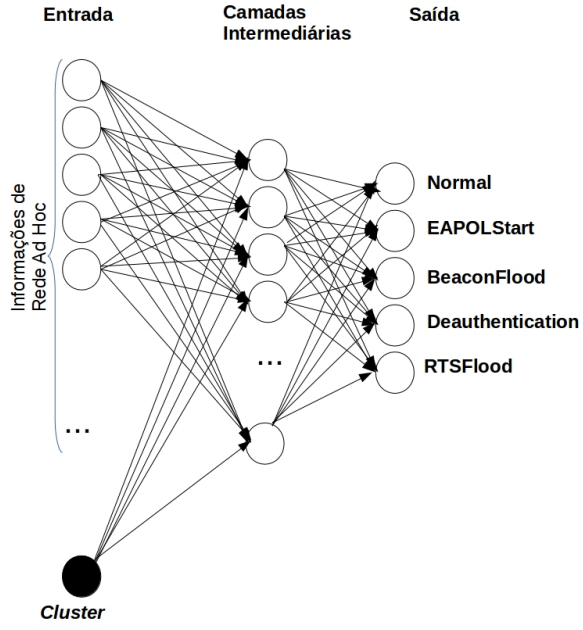

Fig. 6. Proposta de Arquitetura da Rede Neural.

C. Simulações e Resultados

Para realizar a validação do Sistema de Detecção de Intrusão proposto, o algoritmo K-Médias primeiro agrupa os dados em dez *clusters*. O algoritmo *Multilayer Perceptron*, por sua vez, realizará a classificação, fazendo uso das informações do *cluster*, resultante da etapa anterior, como entrada da Rede Neural, conforme Figura 6. Para a classificação a Rede Neural é treinada com 90% dados completo e 10% são utilizados para teste.

Para a realização do agrupamento dos dados define-se um total de 500 iterações e 10 *clusters* ou grupos para o algoritmo K-Médias. A rotulação das classes de anomalias para cada *cluster* é apresentado na Tabela I. A Figura 7 apresenta a discriminação dos dados em cada *cluster*, ou seja, apresenta a quantidade de registros de cada classe em cada *cluster*.

Para a classificação é utilizado um Conjunto de Testes, formado por 10% dos dados após o agrupamento dos mesmos. O algoritmo *Multilayer Perceptron* irá treinar a Rede Neural através da base de treinamento, que é composta por 90% dos dados da Rede.

A Figura 8 apresenta o percentual de classificação em relação a cada classe através do algoritmo *Multilayer Perceptron*, que representa o percentual de dados classificados corretamente em cada classe. Enquanto que a Tabela II mostra a taxa de classificação total para o sistema proposto.

A matriz de confusão da base de dados aplicada à Rede Neural *Multilayer Perceptron* é apresentada na Figura 9. A Tabela III mostra os valores obtidos para as métricas de avaliação do Sistema proposto que são: Falso Positivos, Falso Negativos, Verdadeiro Positivos, Verdadeiro Negativos.

Os testes realizados no Sistema de Detecção de Intrusão

TABELA I
ROTULAÇÃO DOS *Clusters*

Rotulação	Cluster
Normal	0,1,3,4,5,6,7,8,9
EAPOLStart	2
BeaconFlood	—
Deauthentication	—
RTSFlood	—

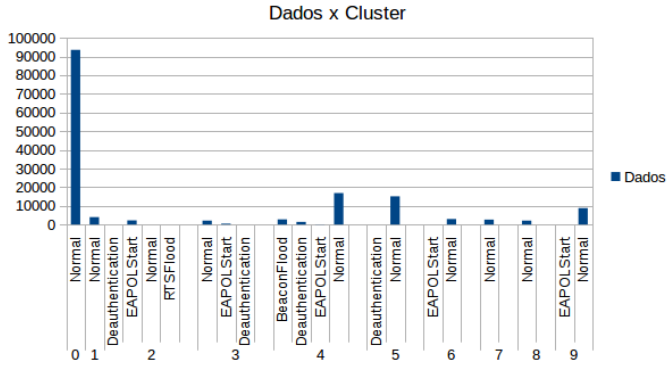

Fig. 7. Dados por *Cluster*.

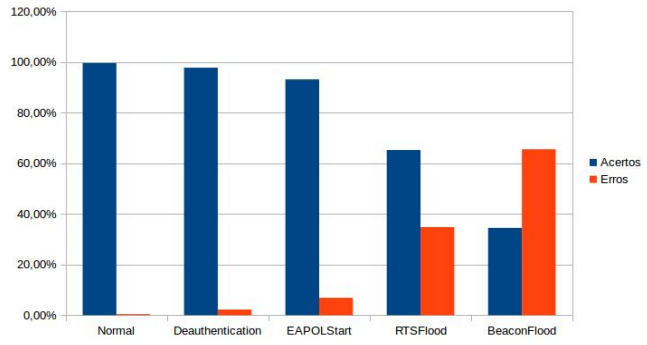

Fig. 8. Classificação - *Multilayer Perceptron*.

TABELA II
CLASSIFICAÇÃO DO SISTEMA PROPOSTO

Acertos	Erros
97%	3%

```
=== Confusion Matrix ===

     a      b      c      d      e   <-- classified as
541466     30    194   2496      0 |      a = Normal
  8229  19205      0      0      0 |      b = BeaconFlood
   538      0  27511     22      0 |      c = EAPOLStart
    78      0    246  15893      0 |      d = Deauthentication
    44      0      0      0    105|      e = RTSFlood
```

Fig. 9. Matriz Confusão - Sistema Proposto

TABELA III
MÉTRICAS DE AVALIAÇÃO - SISTEMA PROPOSTO

Falso Positivo (FP)	Falso Negativo (FN)	Verdadeiro Positivo (VP)	Verdadeiro Negativo (VN)
0,50%	12,36%	87,25%	99,50%

proposto apontam para resultados relevantes. Para a validação do grupo de dados, são utilizados o banco de dados local [8] e o algoritmo K-Médias, que organiza os dados em 10 *clusters* em 97 iterações do algoritmo em 406,94 segundos. A organização dos dados pode ser verificada na Figura 7. Para a classificação dos dados já agrupados, é utilizado o algoritmo *Multilayer Perceptron*, que possui uma taxa de precisão de 97% dos dados, com erro médio em torno de 1,79% e erro quadrático médio em torno de 9,31%.

O sistema proposto é eficaz para o processo de classificação de dados de Redes Sem Fio Ad Hoc, através do uso dos algoritmos K-Médias e *Multilayer Perceptron*, pois permite a redução de falso positivos, quando comparado ao uso isolado de estratégias de Inteligência Computacional adotadas.

A contribuição deste trabalho é apresentar um Sistema de Detecção e Classificação de Intrusão para Redes Sem Fio Ad Hoc. Este sistema por sua vez contribui com a utilização de uma abordagem combinando estratégias de técnicas de inteligência computacional com aprendizagem supervisionada e não supervisionada. Essa proposta é aplicada em cada componente da Rede Sem Fio Ad Hoc, sendo possível realizar o agrupamento do tráfego da rede sem o uso de algum evento externo, ou seja, utiliza-se os dados de forma fiel sem restrições. Após este agrupamento, um segundo método é utilizado para classificar as ações maléficas, caso existam, exigindo um pouco mais de recurso computacional. Outra contribuição importante neste trabalho é a utilização deste Sistema de Detecção e Classificação de Intrusão na política de segurança de ambientes de Redes Ad Hoc.

TABELA IV
COMPARAÇÃO ENTRE DIVERSOS TRABALHOS

Proposta	Classificação
IDS Wavelet [7]	99%
IDS Híbrido [5]	98%
Nossa Proposta	97%
IDS Thatachi [6]	96%

Este trabalho difere da proposta de Ferreira [7], pois a proposta de Ferreira realiza uma filtragem de dados da rede através do cálculo de um limiar em relação ao sinal capturado para posterior classificação. Ferreira [7] utiliza a base de dados KDD 99 [10] para o treinamento da Rede Neural utilizada no processo de classificação dos dados filtrados pela camada *Wavelet*.

Govindarajan, apresenta o IDS Híbrido apontado na Tabela IV que utiliza dois métodos de classificação envolvendo *Perceptron Multicamada* e função de base radial. Propõe-se neste trabalho uma arquitetura **híbrida** envolvendo ambos os classificadores para sistemas de detecção de intrusão. Os dados usados neste estudo baseiam-se em um sistema imunológico desenvolvido na Universidade do Novo México [5]. É um serviço privilegiado para enviar *e-mails*. Os dados incluem tanto traços normais quanto anormais.

Cervantes, no entanto, apresenta um sistema de detecção de intrusão contra ataques *sinkhole e selective forwarding* utilizando agrupamento para lidar com a densidade e a mobilidade, e combina estratégias de *watchdog*, reputação e confiança na detecção de atacantes. O Thatachi foi comparado ao sistema INTI, desenvolvido para mitigar principalmente ataques *sinkhole* [6]. Ambos os sistemas foram implementados no simulador Contiki-Cooja, um sistema operacional de código aberto.

Na Tabela IV, é possível perceber que a proposta apresentada proporcionou uma taxa aceitável de sucesso, demonstrando que esta é uma abordagem viável para a construção do IDS.

VI. CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho apresenta uma proposta de Sistema de Detecção de Intrusão em Redes Sem Fio Ad Hoc composto por duas etapas, baseado em agrupamento de dados através do algoritmo K-Médias e Redes Neurais Artificiais, para detecção

e classificação de anomalias, causados por ataques às Redes de Computadores.

O algoritmo K-Médias permite agrupar os dados em grupos denominados *clusters*, que comporão dados com similaridade, através da distância euclidiana até o centróide mais próximo. Após este passo, o reconhecimento de padrões, que indica a anomalia, é simples e rápido. Para a detecção e classificação de novas anomalias, é necessário treinar novamente a Rede Neural.

O Sistema de Detecção de Intrusão proposto permite compartilhar as melhores características de cada método. A utilização em conjunto permite a redução de falso positivos, se comparado com a utilização isolada de ambas as técnicas.

A validação da proposta deste artigo fundamenta-se em dados obtidos de ambientes de Redes Sem Fio Ad Hoc doméstico e de organização. Para a classificação dos dados já agrupados utiliza-se o algoritmo *Multilayer Perceptron*, que possui taxa de acerto em 97% dos dados possuindo erro médio em torno de 1,79% e erro médio quadrático em torno de 9,31% em cada *cluster*. Os resultados obtidos aqui permitem concluir que a abordagem proposta é promissora, e um bom nível de detecção é conseguido nas avaliações realizadas.

Os trabalhos futuros podem ser: aplicação em Redes Sem Fio, avaliando o sistema proposto em uma rede sem fio real em seus diversos ambientes, desde corporativos até em ambientes de desastres; estudo da adaptação deste sistema proposto para avaliações instantâneas na rede, ou seja, possibilidade de detecção on line; estudo e análise da utilização de outras técnicas de inteligência computacional, com o objetivo de acelerar o processamento do sistema proposto, sem grandes prejuízos para os dispositivos das Redes Ad Hoc; integração com Sistema de prevenção de ataques, pois o sistema proposto indica se determinado dispositivo está sob condição normal ou de anomalia; realização da análise semântica das informações pertencentes aos quadros classificados de maneira correta, podendo o administrador ter acesso a informações mais específicas do tráfego anômalo da Rede.

REFERÊNCIAS

[1] D. R. Canedo and A. R. S. Romariz, "Data analysis of wireless networks using computational intelligence," *Journal of Communications*, vol. 13, no. 11.

[2] J. Loo, J. L. Mauri, and J. H. Ortiz, *Mobile ad hoc networks: current status and future trends*. CRC Press, 2016.

[3] J. Amudhavel, V. Brindha, B. Anantharaj, P. Karthikeyan, B. Bhuvaneswari, M. Vasanthi, D. Nivetha, and D. Vinodha, "A survey on intrusion detection system: State of the art review," *Indian Journal of Science and Technology*, vol. 9, no. 11, pp. 1–9, 2016.

[4] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014.

[5] M. Govindarajan and R. Chandrasekaran, "Intrusion detection using neural based hybrid classification methods," *Computer Networks*, vol. 55, no. 8, pp. 1662–1671, 2011.

[6] C. Cervantes, M. Nogueira, and A. Santos, "Mitigação de ataques no roteamento em iot densa e móvel baseada em agrupamento e confiabilidade dos dispositivos," in *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC, 2018.

[7] E. W. T. Ferreira, G. A. Carrijo, R. de Oliveira, and N. V. de Souza Araujo, "Intrusion detection system with wavelet and neural artifical network approach for networks computers," *IEEE Latin America Transactions*, vol. 9, no. 5, pp. 832–837, 2011.

[8] E. W. T. Ferreira, A. A. Shinoda, R. D. Oliveira, V. E. Nascimento, and N. V. D. S. Araújo, "A Methodology for building a Dataset to Assess Intrusion Detection Systems in Wireless Networks," *WSEAS Transactions on Communications*, vol. 14, pp. 113–120, 2015.

[9] A. Dorri, S. R. Kamel, and E. Kheirkhah, "Security challenges in mobile ad hoc networks: A survey," *arXiv preprint arXiv:1503.03233*, 2015.

[10] E. W. T. Ferreira, "Proposta de um sistema de detecção e classificação de intrusão em redes de computadores baseado em transformadas wavelets e redes neurais artificiais," Tese, Universidade Federal de Uberlândia, Uberlandia, Brasil, Dezembro 2009.

[11] A. X. d. Marins, "Protocolos de roteamento para redes móveis comparativo : Olsr x aodv," Projeto Final, Universidade Federal Fluminense, Niterói, Brasil, 2017.

[12] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14. Oakland, CA, USA, 1967, pp. 281–297.

[13] S. O. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Ontario Canada: Pearson, 2009.

[14] R. F. De Moraes, N. V. de Souza Araújo, and C. Maciel, "Avaliação de um conjunto de dados quanto à sua qualidade na especificação de perfis de ataque e não-ataque numa rede ieee 802.11 w," *Anais da Escola Regional de Informática da Sociedade Brasileira de Computação (SBC)–Regional de Mato Grosso*, vol. 6, pp. 145–150, 2015.

[15] N. F. Haq, A. R. Onik, M. A. K. Hridoy, M. Rafni, F. M. Shah, and D. M. Farid, "Application of machine learning approaches in intrusion detection system: a survey," *IJARAI-International Journal of Advanced Research in Artificial Intelligence*, vol. 4, no. 3, pp. 9–18, 2015.

[16] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. M. Zeebaree, "Combination of k-means clustering with genetic algorithm: A review," *International Journal of Applied Engineering Research*, vol. 12, pp. 14 238–14 245, 2017.

Daniel Rosa Canêdo possui graduação em Engenharia de Computação pela Pontifícia Universidade Católica de Goiás (2003) e com mestrado em Engenharia Elétrica pela Universidade de Brasília (2006). Atualmente é professor exclusivo do Instituto Federal de Goiás - Campus Luziânia. Atualmente é aluno de doutorado do Programa de Pós-Graduação em Engenharia de Sistemas Eletrônicos e Automação do Departamento de Engenharia Elétrica da Universidade de Brasília (UnB).

Alexandre Ricardo Soares Romariz possui graduação em Engenharia Elétrica pela Universidade de Brasília (1992), Mestre em Engenharia Elétrica pela Universidade Estadual de Campinas (1995) e Doutor em Engenharia Elétrica pela Universidade do Colorado em Boulder (2003). Atualmente é professor associado na Universidade de Brasília. Tem experiência na área de Inteligência Computacional, Circuitos Integrados, Optoeletrônica e Processamento Digital de Sinais.