Monitoramento Colaborativo de Redes Sem-fio: Acurácia do Sistema e Denúncia de Farejadores Maliciosos

Miguel Elias M. Campista[1], Matteo Sammarco[2], Marcelo D. de Amorim[2] e Tahiry Razafindralambo[3] *

[1]GTA/PEE-COPPE/DEL-Poli – UFRJ – Brasil
[2]LIP6/CNRS – UPMC Sorbonne Universités – França
[3]INRIA Lille Nord Europe – França

miguel@gta.ufrj.br,{matteo.sammarco, marcelo.amorim}@lip6.fr
tahiry.razafindralambo@inria.fr

Resumo. *A concepção de estratégias adaptativas baseadas em medidas de tráfego em redes sem-fio é desafiadora. Neste trabalho, demonstra-se através de experimentos que a caracterização acurada dessas redes é viável com o aumento do número de traços coletados. Isso significa que a colaboração dos usuários é fundamental. Uma consequência dessa participação, entretanto, é a possível incidência de ações maliciosas no sistema de medidas. Assim, este trabalho avalia a influência de nós maliciosos bem como analisa uma estratégia de detecção baseada na relevância de cada traço coletado. Os resultados mostram que a eficiência da detecção é ligada à densidade dos nós monitores e à quantidade de informações falsas inseridas.*

Abstract. *Conceiving measurement-based adaptive strategies based on wireless networking traffic is challenging. In this paper, we show via experiments that accurate characterization of these networks is feasible by increasing the number of traces collected. This means that the collaboration of users is fundamental. A consequence of such a participation, however, is the possible incidence of malicious actions in the measurement system. Thus, this paper assesses the influence of malicious participants and analyzes a detection strategy based on the relevance of each trace. Results show that the detection efficiency is related to the density of sensing nodes and to the amount of fake information introduced.*

1. Introdução

Estatísticas revelam que em 2014 o número de *smartphones* aumentou 26% no mundo e 27% no Brasil [Cisco, 2015]. Esses dados indicam o enorme potencial que os usuários atuais possuem para aplicações em rede e, consequentemente, o crescente nível de exigência relacionado à qualidade do serviço. Nesse sentido, o aumento de desempenho das redes sem-fio é fundamental, tornando o monitoramento do tráfego um importante recurso de avaliação do sistema. O objetivo é reagir prontamente a problemas de dimensionamento que possam ocorrer durante a operação da rede [Acharya et al., 2010].

Dentre as tecnologias de acesso sem-fio, o WiFi é uma das mais populares. Como consequência, alguns sistemas de monitoramento vêm sendo propostos para

*Este trabalho contou com suporte financeiro de agências de fomento brasileiras (CAPES, CNPq, Faperj, FINEP e RNP) e francesa (ANR) sob o projeto ANR VERSO RESCUE (ANR-10-VERS-003).

obtenção de medidas mais realistas dessas redes [Mahajan et al., 2006, Li et al., 2015]. Essas ferramentas capturam o tráfego da rede e traduz as informações em arquivos de traços de dados [Robinson et al., 2008, Srinivasan et al., 2008, Kone et al., 2011, Claveirole e de Amorim, 2012]. Em ambientes controlados, esse monitoramento é relativamente simples, se tornando mais complexo em cenários reais, como nas redes de sensoriamento urbano [Antoniadis et al., 2008, Yi, 2009]. Para contornar a complexidade das medidas, uma alternativa é aumentar o número de monitores e, assim, aumentar a quantidade de dados disponível. Isso pode ser alcançado a baixo custo com a colaboração dos próprios usuários que participariam em troca de benefícios. Os resultados das medidas poderiam, p.ex., permitir a redistribuição da infraestrutura da rede para que os usuários sejam mais bem atendidos. Tal alternativa ainda não foi avaliada, até onde se sabe.

Um inegável problema da participação dos usuários é a possibilidade de ações maliciosas. Em sistemas de monitoramento, os usuários poderiam inserir traços falsos para maximizar seus benefícios ou simplesmente para deteriorar o serviço. Esses resultados poderiam ser alcançados, respectivamente, com a atração de recursos da infraestrutura de rede para regiões próximas ao atacante (*ataque de atração*) ou com a repulsão de recursos da infraestrutura próxima ao atacante para regiões afastadas (*ataque de repulsão*). Tendo em vista esses possíveis ataques, este trabalho também propõe um sistema de detecção capaz de identificar traços falsos que possam ter sido adicionados ao sistema por nós maliciosos. Para tal, um modelo em grafo das relações entre os traços é proposto, onde cada vértice representa um traço e cada enlace representa a relação entre o par de traços. Tal relação é quantificada a partir de uma métrica de acurácia proposta. Desse modelo, pode-se inferir o comportamento de um monitor baseado na soma dos pesos das arestas de seu traço (força do traço). Se a força do traço for alta ou baixa comparada aos outros, significa que o nó gerador é possivelmente malicioso.

Este trabalho conduz testes experimentais usando dois cenários com diferentes densidades de nós monitores, chamados de *colocalizado* e *esparso*. Em ambos os cenários, demonstra-se que cada um dos traços é relevante para a melhoria da acurácia do sistema e que a distribuição geográfica dos monitores impacta no resultado final. Adicionalmente, o desempenho do sistema de detecção é avaliado, demonstrando ser possível reconhecer os nós maliciosos considerando tanto o ataque de atração quanto o de repulsão. Em resumo, as principais contribuições são as seguintes: (i) proposta de um sistema de monitoramento colaborativo; (ii) identificação de dois possíveis ataques baseados na contribuição maliciosa de um monitor; e (iii) proposta de uma metodologia baseada em grafo para a detecção de possíveis nós maliciosos.

Este trabalho está organizado da seguinte forma. A Seção 2 motiva o emprego de sistemas de monitoramento colaborativos, enquanto a Seção 3 detalha o problema abordado. A Seção 4 propõe a métrica de acurácia e o sistema de detecção. A Seção 5 descreve os cenários utilizados; enquanto as Seções 6, 7 e 8 apresentam, respectivamente, os resultados da métrica proposta, o impacto do ataque e o desempenho do sistema de detecção. A Seção 9 descreve os trabalhos relacionados e a Seção 10 conclui este trabalho.

2. Sistema de Monitoramento Colaborativo

O objetivo principal do sistema de monitoramento considerado neste trabalho é auxiliar o planejamento da infraestrutura de rede por parte dos provedores. Esse plane-

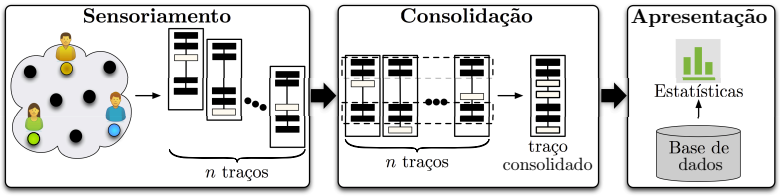


Figura 1. Módulos da arquitetura de um sistema de monitoramento colaborativo.

jamento é realizado *deslocando-se a infraestrutura disponível para regiões menos favorecidas*. A arquitetura do sistema, ilustrada na Figura 1, é composta de três módulos principais: um módulo de sensoriamento composto por farejadores de pacotes de redes sem-fio (*packet sniffers*); um módulo de consolidação de traços, responsável por construir um traço final consolidado; e um módulo de apresentação que armazena e entrega estatísticas de uso da rede aos administradores. Os farejadores, ou monitores, são distribuídos e produzem individualmente um traço de dados composto por uma sequência de quadros em ordem cronológica de recepção. Esses quadros podem ser de diferentes comunicações. O traço consolidado é produzido por um elemento central, que recebe de tempos em tempos os traços individuais. Pode-se notar que há um compromisso entre frequência de envio dos traços e sobrecarga na rede que não é tratado neste trabalho.

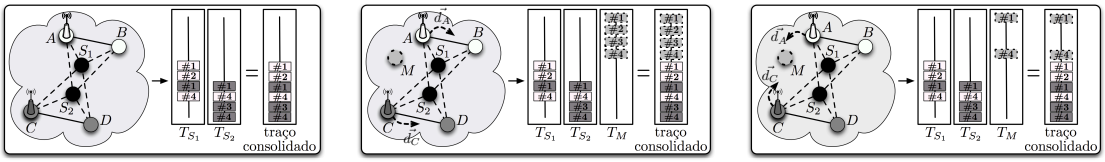
Os farejadores também podem ser executados em dispositivos de usuários. Sabe-se que quanto maior o número de usuários, menor é o custo de implantação e menor é o problema de posicionamento dos monitores. O único requisito para os usuários colaborarem é executar um farejador de pacotes. Assume-se que esse requisito não seja um entrave já que um farejador de pacotes em plano de fundo consume menos que 1,5 kB de memória virtual, incluindo código, dados, bibliotecas compartilhadas e páginas de memória.[1] A participação dos usuários pode ser estimulada com incentivos como uma melhor qualidade de serviço, recompensas financeiras e até mesmo melhor reputação. No sistema proposto, assume-se que o módulo de consolidação está sob o controle de uma entidade central confiável, enquanto o de sensoriamento está distribuído e sob o controle de diferentes entidades, dentre as quais *algumas podem ser maliciosas*.

3. Definição do Problema

A participação dos usuários pode resultar em dois ataques: *ataque de repulsão*, no qual um traço falso é inserido no sistema de monitoramento com sequências completas de quadros (isto é, contém fluxos com todos os quadros); e *ataque de atração*, no qual um traço falso é inserido no sistema de monitoramento com sequências vazias de quadros (isto é, contém fluxos com apenas o primeiro e último quadro).

Dependendo da recompensa da colaboração, um usuário malicioso poderia disparar um dos dois ataques. Por exemplo, um usuário que colaborasse com traços contendo fluxos de dados com sequências completas de quadros poderia provocar um "ataque de repulsão", já que o sistema de monitoramento inferiria uma alta acurácia na região próxima ao atacante e, portanto, assumiria que essa área pode ceder infraestrutura de rede. Em oposição, um usuário que colaborasse com traços contendo fluxos com apenas o primeiro e o último quadro poderia provocar um "ataque de atração", já que o sistema inferiria

[1] Esse teste foi conduzido com o `tcpdump` em uma máquina Debian Linux.



(a) Operação normal da rede. (b) Ataque de repulsão. (c) Ataque de atração.

Figura 2. Operação da rede em condições normais e sob ataque.

baixa acurácia na região. Como consequência, o sistema assumiria que a região próxima ao atacante precisa de mais infraestrutura de rede para melhorar o serviço.

A Figura 2(a) ilustra a operação normal do sistema. Nessa figura, os nós monitores S_1 e S_2 monitoram áreas distintas da mesma rede sem-fio e produzem traços de dados T_{S_1} e T_{S_2}, respectivamente. Já os nós A e C são pontos de acesso e os nós B e D são usuários da rede. Nenhum desses nós (A, B, C, D) participam do monitoramento. Na figura, as linhas pontilhadas representam os enlaces entre monitores e usuários e as linhas contínuas entre usuários apenas. Note que o nó monitor S_1 captura mais quadros da comunicação $A{\leftrightarrow}B$ (quadros brancos), enquanto S_2 captura mais quadros de $C{\leftrightarrow}D$ (quadros cinza) por questões de proximidade. Assumindo que em cada comunicação os nós trocam quatro quadros, os nós S_1 e S_2 capturam 50% do total de quadros, sendo que S_1 e S_2 capturam mais quadros de $A{\leftrightarrow}B$ e $C{\leftrightarrow}D$, respectivamente. Após a consolidação, o resultado final tem 75% do total de quadros, melhorando o desempenho do monitoramento.

Considerando a presença de um nó M controlado por um usuário malicioso, um traço falso (T_M) é inserido no sistema. Assim, no ataque de repulsão, o nó M forja um traço com um fluxo contendo todos os quatro quadros da comunicação entre um par de nós não existente. Após a consolidação, a fração de quadros capturados é de aproximadamente 83% (3 quadros de 4 da comunicação $A{\leftrightarrow}B$, mais 3 de 4 de $C{\leftrightarrow}D$ e mais 4 de 4 do par falso), que é maior que os 75% da operação normal. Como consequência, a infraestrutura disponível (nós A e C) pode ser afastada de M, como visto na Figura 2(b) a partir dos deslocamentos $\vec{d_A}$ e $\vec{d_C}$. No ataque de atração, por outro lado, depois da consolidação, 66% dos quadros são considerados capturados (3 quadros de 4 de $A{\leftrightarrow}B$, mais 3 de 4 de $C{\leftrightarrow}D$ e mais 2 de 4 do par falso). Logo, a fração de quadros se torna mais baixa do que a normal, aproximando de M possível infraestrutura extra (Figura 2(c)). Neste trabalho, cada monitor contribui com um traço. Ainda, cada traço tem o mesmo formato e é obtido ou forjado no mesmo período legítimo de monitoramento.

4. Acurácia do Sistema de Monitoramento e Detecção de Vulnerabilidades

Este trabalho aborda dois problemas complementares do monitoramento colaborativo. O primeiro é consequência dos desafios impostos pelo ambiente sem-fio. Já o segundo é consequência da proposta deste trabalho baseada na participação dos usuários.

4.1. Acurácia do monitoramento

O aumento da acurácia do monitoramento permite uma melhor caracterização da rede e, consequentemente, evita ações equivocadas baseadas em medidas incorretas. No contexto deste trabalho, a acurácia é formalmente definida baseada em quatro premissas: (i) cada nó da rede envia quadros em ordem crescente de número de sequência para

qualquer destino (tempo não é suficientemente grande para que os números de sequência se repitam); (ii) os traços são consolidados em um ponto central; (iii) não há requisitos de tempo real no sistema; e (iv) há uma probabilidade não negligenciável de perda de quadros, mesmo considerando todos os traços capturados. Baseado nessas premissas, considera-se que o número máximo de quadros que um nó pode enviar para outro seja igual à diferença entre o último e o primeiro quadro do mesmo fluxo mais o total de retransmissões. Note que o último e o primeiro quadro de um fluxo são, respectivamente, aquele com o maior e o menor número de sequência capturados. Uma vez que em experimentos reais é difícil obter traços completos, utiliza-se esse recurso para estimar a acurácia dos traços [Yeo et al., 2004, Schulman et al., 2008]. Assim, denota-se como $\mathcal{N}$ o conjunto de nós na rede, onde i e j são dois nós em $\mathcal{N}$, e como $v_i(t)$ o conjunto de nós na vizinhança de i no intervalo de tempo t. Ainda, considerando que quadros consecutivos têm número de sequência incrementais e que as retransmissões são estimadas tanto por números de sequência repetidos quanto por sinais de retransmissão ligados, o número máximo de quadros $s_i(t)$ que i pode enviar em t é:

$$s_i(t) = \sum_{j=0}^{|v_i(t)|} (n_{ij}^{\max}(t) + s'_{ij}(t)), \tag{1}$$

onde $n_{ij}^{\max}(t)$ e $s'_{ij}(t)$ são, respectivamente, o maior número de sequência e o número de quadros retransmitidos do fluxo de i para j. Note que i pode não enviar quadros para um dado vizinho ($n_{ij}^{\max}(t) = s'_{ij}(t) = 0$). Além disso, como pode haver quadros perdidos, o número total de quadros recebidos do nó i em t, $r_i(t)$, tem como limiar superior $s_i(t)$. Logo, o número de quadros perdidos do nó i é a diferença entre o número máximo de quadros e o número de quadros recebidos de i, ou seja $m_i(t) = s_i(t) - r_i(t)$. Generalizando em $\mathcal{N}$, tem-se que o número de quadros perdidos, $m(t)$, é igual à diferença entre o número de quadros enviados e recebidos por todos os nós da rede. Assim,

$$m(t) = s(t) - r(t) = \sum_{i=0}^{|\mathcal{N}|} (s_i(t) - r_i(t)). \tag{2}$$

Baseado na Equação 2, pode-se definir a métrica de acurácia a como se segue.

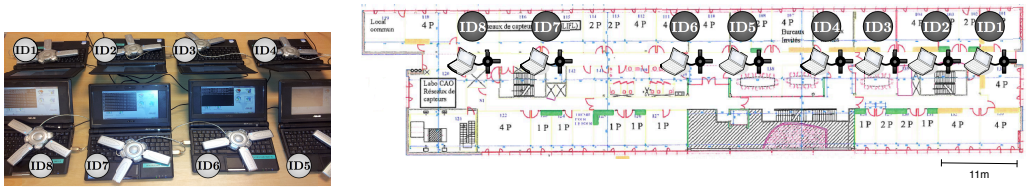
Definição 1 (Métrica de acurácia a) *É definida como a porcentagem de quadros enviados na rede que são capturados por pelo menos um nó monitor:*

$$a = 1 - (m(t)/s(t))\cdot \tag{3}$$

O cálculo de a requer que na única leitura do traço consolidado, o número de sequência de cada quadro seja comparado ao anterior do mesmo fluxo para verificar perdas.

4.2. Sistema de detecção

O sistema de detecção proposto é modelado a partir de um grafo não direcionado com pesos do tipo $\mathcal{G}(\mathcal{V}, \mathcal{E})$, onde $\mathcal{V}$ e $\mathcal{E}$ representam, respectivamente, o conjunto de vértices e enlaces. Em $\mathcal{G}$, cada vértice representa um traço capturado T_i e cada enlace T_iT_j descreve a acurácia relativa entre os traços conectados. Assim, no modelo proposto, o grafo $\mathcal{G}$ é totalmente conectado já que a acurácia de cada um dos traços pode ser calculada em relação a todos os outros. Considerando $a(T_iT_j)$ como sendo a acurácia calculada



(a) Cenário colocalizado. (b) Cenário esparso.

Figura 3. Cenários experimentais.

do traço consolidado entre T_i e T_j, tem-se que o peso da aresta T_iT_j é igual a $a(T_iT_j)$. Baseado no peso das arestas, calcula-se a força de cada vértice, $\sigma(T_i)$, como se segue:

$$\sigma(T_i) = \sum_{j=1}^{|\mathcal{V}|} a(T_iT_j), \tag{4}$$

onde o número de traços é equivalente ao número de nós monitores, ou seja, $|\mathcal{V}| = |\mathcal{N}|$. Neste artigo, assume-se que o peso do enlace contendo um nó malicioso tem um valor discrepante comparado aos outros já que a adição de informações falsas nos traços tem como efeito o aumento ou a redução da acurácia, dependendo do ataque. A partir dessa observação, o traço com força diferente dos outros pode ser identificado usando um teste de detecção de *outlier*. Neste trabalho, assume-se no máximo um único nó malicioso.

Apesar do modelo em grafo proposto ser genérico, ele é usado neste trabalho apenas para detecção de nós maliciosos. Consequentemente, não há nenhuma premissa sobre a movimentação da infraestrutura de rede monitorada. Além disso, não são considerados outros ataques além dos descritos na Seção 3.

5. Redes de Testes e Conjuntos de Dados

Os experimentos deste trabalho foram conduzidos em dois cenários, chamados de *colocalizado* e *esparso*, durante 90 minutos. O cenário colocalizado foi montado em uma sala do laboratório de informática (LIP6) da UPMC Sorbonne Universités em Paris. Todos os monitores foram posicionados lado-a-lado em uma mesa (Figura 3(a)). Os traços coletados têm em média 253 MB, enquanto o traço consolidado tem 450 MB. Já o cenário esparso foi montado no laboratório de computação IRCICA/LIFL em Lille. Nesse cenário, todos os nós monitores foram espalhados no segundo andar do prédio de acordo com a disponibilidade de tomadas elétricas do corredor (Figura 3(b)). Os traços coletados têm em média 205 MB, enquanto o traço consolidado tem 1,2 GB. Os tamanhos dos arquivos já indicam que quanto mais esparso for o cenário, maior é a quantidade de dados agregados com a consolidação. Nos dois cenários, 8 netbooks Asus EEEPC-4G com 512 MB de memória RAM, 3 interfaces USB WiFi Netgear WG111v3 configuradas em canais de rádio diferentes, sistema operacional Xandros com kernel modificado e software WiPal para farejamento de tráfego foram usados. Esses netbooks cumpriram o papel de monitores e, consequentemente, coletaram os traços da rede sem-fio.

Apesar dos cenários serem internos (*indoor*), eles demonstram naturezas diferentes que permitem compreender a operação de um nó malicioso. A densidade de nós na rede é variada, visto que uma maior distância entre monitores pode aumentar a discrepância entre os traços capturados. Em ambos os experimentos, os nós monitores capturam qualquer quadro transmitido em sua área de cobertura, já que estão em modo monitor.

Tabela 1. Resultados de acurácia ($\times 10^{-3}$) obtidos com traços individuais e consolidados em ambos os cenários avaliados.

Cenários	Traços								
	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	**Consolidado**
Colocalizado	1,57	1,58	1,70	1,57	1,99	1,82	2,37	1,71	**3,27**
Esparso	0,34	0,29	0,25	1,35	1,01	0,65	0,17	0,12	**3,07**

Neste trabalho, assume-se que qualquer tipo de tráfego deve ser capturado pelo sistema de monitoramento, independente de sua natureza ou duração. Vale mencionar ainda que as medidas deste trabalho capturaram tráfego sem-fio de controle, de gerenciamento e de dados originados ou destinados para pontos de acesso. Essas medidas não apontam diferenças entre os canais não sobrepostos do IEEE 802.11g (canais 1, 6 e 11). Sendo assim, optou-se pelas medidas do canal 1.

Software WiPal: Este trabalho utiliza a ferramenta WiPal para consolidação de traços. O WiPal consolida um número arbitrário de arquivos PCAP de traços do IEEE 802.11 de diferentes nós monitores e calcula o traço PCAP consolidado. O procedimento de consolidação, entretanto, é realizado em pares de traços de forma recursiva. Logo, o resultado da rodada anterior é usado como um dos traços da rodada seguinte. De forma simplificada, para cada par de traços consolidados, o WiPal identifica quadros de referência para cálculo da defasagem entre os relógios dos monitores, sincroniza todos os quadros e escreve o resultado em um arquivo de saída. O mesmo procedimento é repetido até que todos os traços de entrada do WiPal sejam consolidados [Claveirole e de Amorim, 2012].

6. Acurácia das Medidas

Antes das questões de segurança, o comportamento da acurácia das medidas é investigado em uma rede sem-fio operacional livre de ataques.

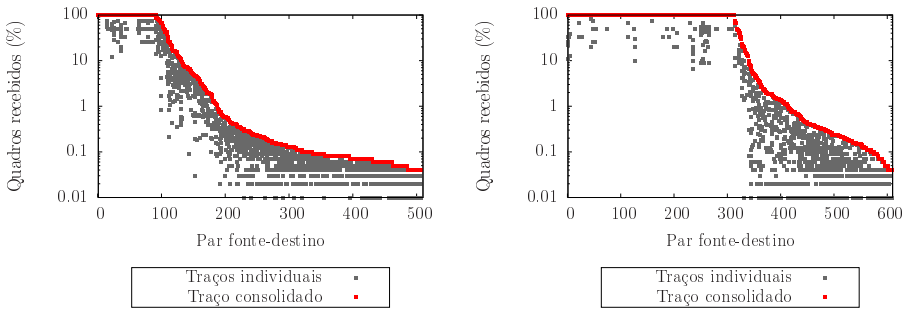
6.1. Traços individuais e consolidados

A Tabela 1 mostra a acurácia de cada traço em comparação ao traço final consolidado. Os parâmetros necessários para a computação da fração de quadros recebidos, ou seja, o número máximo de sequência e o número de retransmissões por par fonte-destino ($s(t)$), são extraídos do traço consolidado. Assim, pode-se dizer que a acurácia computada de todos os traços usa o traço consolidado como referência. Note que a diferença de acurácia entre o traço consolidado e o traço com a menor acurácia varia de $2\times$ a $25\times$ no cenário colocalizado e esparso, respectivamente. O impacto da acurácia em função da distância entre os nós monitores pode ser observado mesmo considerando que o valor da métrica de acurácia seja sempre bem menor do que 1. Então, quanto mais próximos forem os monitores, menor é a diferença da acurácia entre eles. Isso acontece porque quanto mais próximos estiverem os monitores, maior é a interseção de quadros capturados.

Nas duas próximas seções, a fração de quadros capturados de cada par fonte-destino é usada ao invés da métrica acurácia. O objetivo é analisar o impacto do número de traços e da posição dos monitores nas medidas colaborativas.

6.2. Número de traços

A Figura 4 ilustra a fração de quadros capturados de todos os pares fonte-destino escutados. No eixo-x, os pares fonte-destino são ordenados de acordo com a fração de

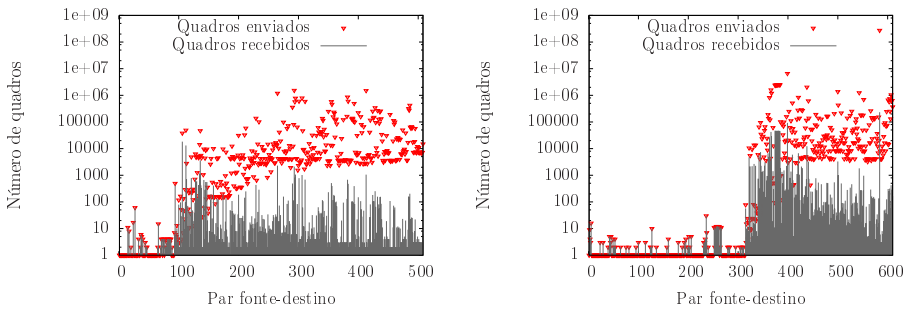


(a) Cenário colocalizado. (b) Cenário esparso.

Figura 4. Fração de quadros capturados de todos os pares fonte-destino da rede.

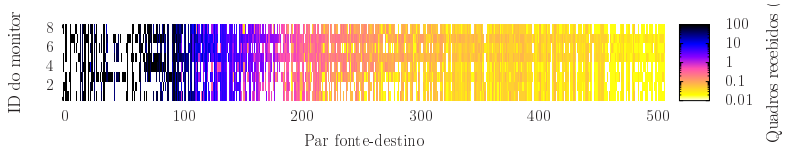
quadros capturados. Note que, em ambos os cenários, há uma grande discrepância entre a fração de quadros capturados pelos diferentes traços individuais e pelo traço consolidado para qualquer par origem-destino. Logo, decisões baseadas em um único traço podem levar a ações equivocadas. A comparação entre os dois cenários permite verificar um comportamento similar. Há a presença de um degrau em 100% devido à presença de quadros de controle e de gerenciamento, assim como pequenas sequências de quadros de dados. Os quadros de gerenciamento e de controle são mais robustos a perdas, já que são enviados na taxa básica. Já as pequenas sequência de quadros de dados podem ser totalmente capturadas por apenas um único monitor.

A Figura 5 mostra a diferença entre o número de quadros enviados e capturados de cada par fonte-destino na rede. Nessa figura, considera-se apenas o traço consolidado. O eixo-x é o mesmo da Figura 4. Novamente, o número de quadros enviados e capturados é semelhante até certo par de nós. Esse par é o mesmo onde a acurácia cai de 100% na Figura 4, ou seja, o 100^o e o 310^o pares no cenário colocalizado e esparso, respectivamente. Adicionalmente, observa-se que o número de sequência é pequeno até esses pares nos dois cenários. Isso reforça o argumento no qual se dizia que os quadros ou seriam de gerenciamento, ou de controle ou de pequenas sequências de dados. Caso as sequências de quadros sejam grandes e enviadas a uma taxa de transmissão maior, as perdas são mais significativas, independente do monitor.

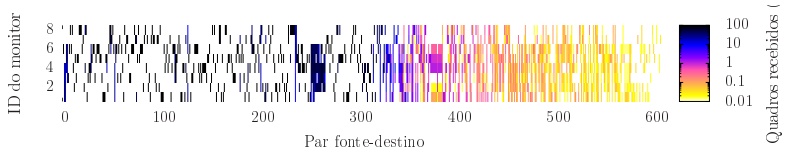


(a) Cenário colocalizado. (b) Cenário esparso.

Figura 5. Variação do número de sequência considerando todos os pares fonte-destino em ambos os cenários.



(a) Cenário colocalizado.



(b) Cenário esparso.

Figura 6. Distribuição geográfica de acurácia considerando todos os pares fonte-destino em ambos os cenários.

6.3. Posição dos nós monitores

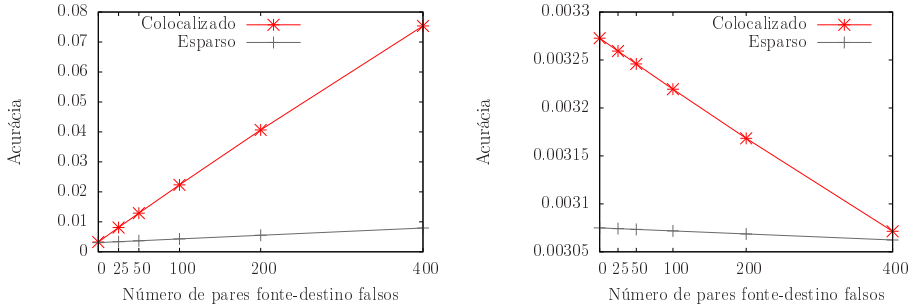
A Figura 6 mostra a correlação entre a fração de quadros capturados e a posição dos nós monitores. Para tal, o eixo-y representa o identificador (ID) do monitor, enquanto o eixo-x é o mesmo usado nas figuras anteriores. Os IDs dos monitores correspondem àqueles apresentados na Figura 3. Como terceira magnitude, tem-se a fração de quadros capturados por cada monitor, representada a partir da intensidade da cor. Novamente, o traço consolidado é usado como referência.

As Figuras 6(a) e 6(b) mostram que até o 100^o e 310^o par, respectivamente, há pelo menos um monitor que captura quadros. Isso pode ser confirmado pelos pontos em preto na figura. Entretanto, mesmo assumindo que esses quadros sejam de gerenciamento ou de controle, eles frequentemente não são capturados por todos os monitores. Isso ocorre porque a posição do monitor afeta até mesmo os quadros enviados em taxas de transmissão mais baixas. Por outro lado, com o aumento do número de sequência, a porcentagem de quadros perdidos aumenta para todos os monitores (representado pelas cores mais claras), indicando que os nós estão enviando dados em taxas de transmissão mais elevadas. Um resultado importante da Figura 6 é a diferença entre a fração de quadros capturados pelos monitores, independente do cenário, devido às posições diferentes. Essa conclusão é possível verificando a variação da intensidade de cor de qualquer linha vertical no gráfico, ou seja, de qualquer par fonte-destino.

7. Impacto dos Ataques

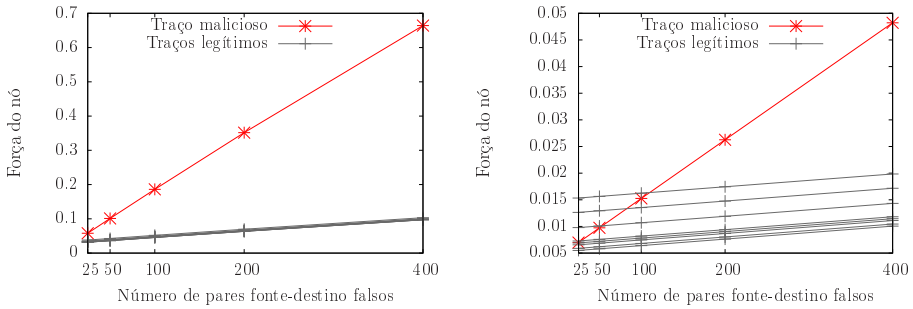
Na seção anterior, foi observado que: a acurácia do monitoramento aumenta com o número de traços e que mesmo pequenas distâncias entre monitores podem alterar a visão da rede. Essas observações têm uma influência direta neste trabalho, já que o aumento do número de monitores pode melhorar a acurácia das medidas, mas a possibilidade da participação dos usuários pode culminar na inserção de traços falsos. Nesta seção, o impacto dos ataques de atração e de repulsão na acurácia do traço consolidado é avaliado. Em ambos os ataques, um usuário malicioso forja um traço contendo sequências falsas de quadros entre pares de nós inexistentes. Para alcançar os efeitos desejados, os ataques de atração e de repulsão devem, respectivamente, reduzir e aumentar a acurácia do sistema.

A Figura 7 apresenta o impacto dos dois ataques em ambos os cenários. Para



(a) Ataque de repulsão. (b) Ataque de atração.

Figura 7. Impacto da inserção de traços com pares comunicantes falsos na acurácia do traço consolidado.



(a) Cenário colocalizado. (b) Cenário esparso.

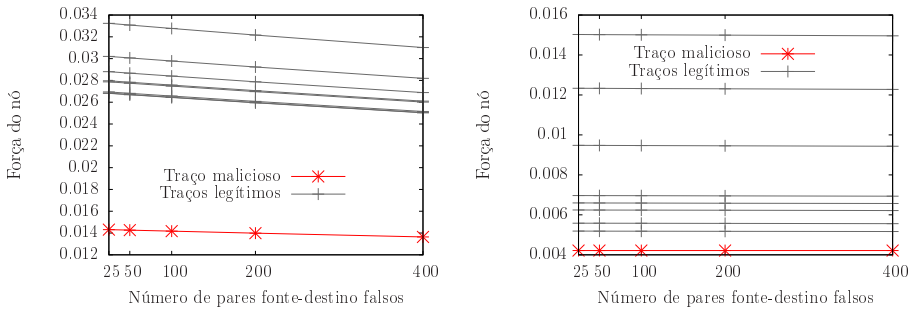
Figura 8. Variação da força do nó no ataque de repulsão.

isso, o número de pares de nós falsos é variado. Note que os ataques em cenários densos são mais efetivos que os em cenários esparsos. Isso ocorre porque a semelhança entre os traços individuais coletados por nós próximos é maior. Assim, o impacto de uma maior quantidade de pares falsos é mais relevante. Apesar das figuras terem um comportamento linear, o ataque de repulsão é mais eficiente que o de atração do ponto de vista da mudança de acurácia. Enquanto no ataque de repulsão a mudança é perto de 2.200% e 150% considerando os cenários colocalizados e esparsos, respectivamente; no ataque de atração, essa diferença é próxima de 6% e 0,4%, respectivamente. Isso acontece porque a acurácia das medidas é naturalmente baixa sem atacantes, como mostrado na Tabela 1. Portanto, reduzir ainda mais a acurácia requer um número elevado de pares de nós falsos.

8. Detecção de Atacantes Potenciais

O sistema de detecção proposto é baseado na acurácia entre pares de traços. Isso significa que primeiro deve-se consolidar todos os pares de traços para, em seguida, calcular a acurácia entre eles. A Figura 8 mostra a variação da força de cada vértice com o aumento do número de pares fonte-destino falsos, conforme o ataque de repulsão. Note que a força relativa ao traço malicioso cresce principalmente no cenário colocalizado já que as distâncias menores entre os monitores contribuem com uma maior semelhança entre os traços legítimos. Logo, pode-se concluir que o aumento da densidade de monitores diminui as chances de sucesso de um ataque de repulsão.

A Figura 9 apresenta resultados similares para o ataque de atração. Nesse caso,



(a) Cenário colocalizado. (b) Cenário esparso.

Figura 9. Variação da força do nó no ataque de atração.

o efeito é oposto em termos de força dos nós se comparado ao ataque de repulsão, ou seja, a menor força é relativa ao nó malicioso. Note que o impacto não é tão significativo quanto no ataque de repulsão. No cenário colocalizado, há uma redução na força de todos os vértices como consequência da presença do traço do monitor malicioso. Já no cenário esparso, a redução da força é pequena, o que demonstra que esse ataque requer mais fluxos de dados falsos. Vale mencionar que se o impacto do ataque não for facilmente observável, o sistema proposto não é capaz de identificar o usuário malicioso. Esse problema não é grave já que a eficiência do sistema de detecção é proporcional ao impacto do ataque. Em outras palavras, se o ataque não for efetivo, a detecção também não é.

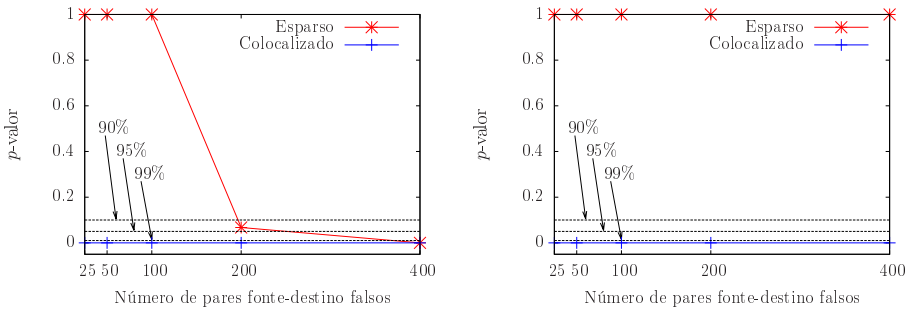
Como visto, a modelagem proposta tem como consequência a diferenciação da força do nó malicioso em comparação à força dos outros nós. Essa constatação, é utilizada na identificação do nó malicioso a partir de um teste de detecção de *outlier*. Para tal, o teste de Dixon é usado já que ele pode ser executado sobre um pequeno conjunto de dados [Dean e Dixon, 1951]. Antes de usar o teste de Dixon, é necessário verificar se a força dos nós legítimos segue uma distribuição normal, já que o teste de Dixon é somente válido nesse caso. Executando o teste de Cramér-von Mises [Darling, 1957] para todas as forças dos nós legítimos, não é possível rejeitar a hipótese de que os dados sigam uma distribuição normal. Em ambos os casos, o p-valor está acima de 0,05, o que garante um intervalo de confiança de 95%. No modelo proposto, todos os vértices têm o mesmo grau e a força do vértice é uma função da acurácia par-a-par entre os traços.

Note que um teste preliminar poderia ser conduzido para verificar a distribuição das forças dos vértices. Se a distribuição desviar da normal, pode-se ter um indicativo de um possível ataque. A Tabela 2 apresenta os p-valores obtidos com o teste de Cramér-von Mises, considerando as forças dos vértices legítimos com e sem o traço malicioso. Os resultados para 400 pares origem-destino falsos mostram que já se pode ter uma ideia de que um ataque está ocorrendo, uma vez que a hipótese de distribuição normal pode ser rejeitada. No cenário esparso com ataque de atração, entretanto, a hipótese não pode ser rejeitada, já que as diferenças entre as forças dos vértices não são suficientemente grandes mesmo com a presença do traço malicioso.

A Figura 10 mostra o resultado do teste de Dixon considerando intervalos de confiança de 90%, 95% e 99%. Semelhante ao teste de normalidade, é avaliada a hipótese do traço malicioso *não ser* um *outlier*. Observa-se que a hipótese é sempre rejeitada no cenário colocalizado, ou seja, o traço *é* malicioso, independente do ataque. Por outro

Tabela 2. Resultados do teste de hipótese de normalidade.

Cenários	Legítimos	Ataque de Repulsão	Ataque de Atração
Colocalizado	0,09199	$4{,}224\times10^{-7}$	0,0006139
Esparso	0,08413	0,005008	0,08194



(a) Ataque de repulsão. (b) Ataque de atração.

Figura 10. Detecção de atacante baseado no teste de Dixon.

lado, no cenário esparso, a hipótese só não é confirmada no ataque de repulsão para mais de 200 e 400 pares falsos, com intervalo de confiança de 90% e 99%, respectivamente. Já que os testes realizados são estatísticos, novas métricas para avaliação dos traços obtidos podem ser adicionadas. Por exemplo, a combinação dos testes propostos com métricas de confiança poderia reforçar os resultados obtidos. Isso evitaria a presença de falsos positivos, que eventualmente podem surgir, principalmente considerando as variações do meio sem-fio. Outra forma de evitar os efeitos do meio sem-fio é monitorar a rede por tempo suficientemente grande para suavizar efeitos imprevistos ou para traçar perfis conhecidos (p.ex., um nó monitor que esteja isolado).

9. Trabalhos Relacionados

Medidas em redes sem-fio e ferramentas de consolidação de traços são frequentemente abordados na literatura. Srinivasan et al. propuseram o SWAT, que é uma ferramenta para execução de experimentos, coleta de dados e obtenção de métricas de camada física [Srinivasan et al., 2008]. AirLab é outra iniciativa que tem como objetivo evitar falhas possivelmente inseridas por hardware ou software utilizados por nós monitores [Kone et al., 2011]. O objetivo é propor uma metodologia única para coletar traços, e consequentemente, permitir avaliações mais consistentes. Já o Wit é uma ferramenta projetada para consolidar deferentes traços, incluindo a possibilidade de reconstrução de pacotes baseado na máquina de estados de protocolos conhecidos [Mahajan et al., 2006]. Na mesma direção, o Jigsaw é outra ferramenta para sincronização e consolidação de múltiplos traços [Cheng et al., 2006]. O Jigsaw pode também reconstruir pacotes possivelmente perdidos das camadas de enlace e transporte. Tanto o Wit quanto o Jigsaw poderiam ter sido usados neste trabalho. Entretanto, decidiu-se pelo WiPal pela facilidade de uso já previamente existente.

Tipicamente, os trabalhos da literatura não consideram a participação de usuário nas medidas. Quando o fazem, ou o usuário participante é considerado idôneo e as questões de segurança são deixadas de lado ou o usuário malicioso não faz parte do sis-

tema e o monitoramento é usado para encontrá-lo. Ravindranath et al. aumentaram o conjunto típico de dispositivos monitores, normalmente limitados a interfaces de redes sem-fio, incluindo outros sensores como o GPS e os acelerômetros [Ravindranath et al., 2011]. Ao utilizar essas informações adicionais, eles podem melhorar o desempenho da rede. O trabalho de Kanuparthy et al. é outro exemplo que considera a participação de usuários confiáveis. Eles propõem um ferramenta para monitorar as condições do meio físico sem a necessidade de qualquer equipamento específico de rede [Kanuparthy et al., 2012]. Todas as medidas são conduzidas pelos próprios usuários e requerem o uso de sondas e de um servidor em um computador conectado a um ponto de acesso. Baseado nas informações das sondas, é possível conhecer os componentes de atraso como a espera pela disponibilidade do canal, espera na janela de contenção, atraso de transmissão e certos intervalos entre-quadros do IEEE 802.11. Como o atraso não considera a espera em filas na fonte, é possível estimar propriedades da camada de enlace e condições físicas. Paul et al. coletam múltiplos traços para analisar a interferência entre nós da rede sem-fio e, além disso, detectar comportamentos egoístas [Paul et al., 2013]. Os autores argumentam que a partir da análise de traços é possível detectar usuários que poderiam ganhar acesso ao meio de forma maliciosa através da manipulação dos parâmetros do protocolo MAC, como o tamanho da janela de contenção. O Map [Sheng et al., 2008] e o DOMINO [Raya et al., 2006] são outros sistemas de monitoramento especializados na captura, consolidação e avaliação de múltiplos traços de redes sem-fio que têm como foco a avaliação da presença de usuários maliciosos.

10. Conclusão

Este trabalho apresentou experimentalmente o impacto de um nó malicioso em um sistema de monitoramento colaborativo para redes sem-fio. Foram demonstrados que a acurácia do sistema aumenta com o número de traços de dados e que a participação dos usuários é bem-vinda. O desafio, porém, é lidar com possíveis usuários maliciosos, interessados em influenciar a organização da infraestrutura da rede. Dentre os possíveis ataques, foram identificadas duas possibilidades onde o usuário malicioso provoca ou o afastamento ou a atração da infraestrutura. Desses dois ataques, foi mostrado que o de repulsão tem mais chances de ser bem-sucedido, especialmente em cenários onde os nós monitores estão mais próximos uns dos outros. Para evitar tais ataques, este trabalho ainda propôs um sistema de detecção que modela a rede como um grafo totalmente conectado, onde os vértices são os traços e as arestas representam a acurácia entre eles. Desse modelo, pode-se detectar a presença de um nó malicioso através do cálculo da força de cada traço, já que os nós maliciosos possuem valores discrepantes em relação aos outros legítimos. Como trabalho futuro, planeja-se conduzir mais experimentos e aprimorar o sistema de detecção para que ele consiga identificar ataques em conluio.

Referências

Acharya, P. A. K., Sharma, A., Belding, E. M., Almeroth, K. C., Member, S. e Papagiannaki, D. (2010). Rate adaptation in congested wireless networks through real-time measurements. *IEEE TMC*, 9(11):1535–1550.

Antoniadis, P., Grand, B. L., Satsiou, A., Tassiulas, L., Aguiar, R., Barraca, J. P. e Sargento, S. (2008). Community building over neighborhood wireless mesh networks. *IEEE Technology and Society*, 27(1):48–56.

Cheng, Y.-C., Bellardo, J., Benkö, P., Snoeren, A. C., Voelker, G. M. e Savage, S. (2006). Jigsaw: solving the puzzle of enterprise 802.11 analysis. Em *ACM SIGCOMM*, p. 39–50.

Cisco (2015). VNI mobile forecast highlights, 2014 - 2019.

Claveirole, T. e de Amorim, M. D. (2012). Manipulating Wi-Fi packet traces with WiPal: design and experience. *Software Practice & Experience*, 42(5):585–599.

Darling, D. A. (1957). The Kolmogorov-Smirnov, Cramér-von Mises Tests. *The Annals of Mathematical Statistics*, 28(4):823–838.

Dean, R. B. e Dixon, W. J. (1951). Simplified statistics for small numbers of observations. *Analytical Chemistry*, 23(4):636–638.

Kanuparthy, P., Dovrolis, C., Papagiannaki, K., Seshan, S. e Steenkiste, P. (2012). Can user-level probing detect and diagnose common home-WLAN pathologies? *SIGCOMM Comput. Commun. Rev.*, 42(1):7–15.

Kone, V., Zheleva, M., Wittie, M., Zhao, B. Y., Belding, E. M., Zheng, H. e Almeroth, K. (2011). AirLab: consistency, fidelity and privacy in wireless measurements. *SIGCOMM Comput. Commun. Rev.*, 41(1):60–65.

Li, W., Mok, K. P., Wu, D. e Chang, R. (2015). On the accuracy of smartphone-based mobile network measurement. Em *IEEE INFOCOM*, p. 1–9.

Mahajan, R., Rodrig, M., Wetherall, D. e Zahorjan, J. (2006). Analyzing the MAC-level behavior of wireless networks in the wild. Em *ACM SIGCOMM*, p. 75–86.

Paul, U. K., Buddhikot, M. M. e Das, S. R. (2013). Passive measurement of interference in WiFi networks with application in misbehavior detection. *IEEE TMC*, 12(3):434–446.

Ravindranath, L., Newport, C., Balakrishnan, H. e Madden, S. (2011). Improving wireless network performance using sensor hints. Em *USENIX NSDI*, p. 1–14.

Raya, M., Aad, I., Hubaux, J.-P. e Fawal, A. E. (2006). DOMINO: detecting MAC layer greedy behavior in IEEE 802.11 hotspots. *IEEE TMC*, 5(12):1691–1705.

Robinson, J., Swaminathan, R. e Knightly, E. W. (2008). Assessment of urban-scale wireless networks with a small number of measurements. Em *ACM MobiCom*, p. 187–198.

Schulman, A., Levin, D. e Spring, N. (2008). On the fidelity of 802.11 packets traces. Em *PAM*, p. 132–141.

Sheng, Y., Chen, G., Yin, H., Tan, K., Deshpande, U., Vance, B., Kotz, D., Campbell, A., Mcdonald, C., Henderson, T. e Wright, J. (2008). Map: a scalable monitoring system for dependable 802.11 wireless networks. *Wireless Communications*, 15(5):10–18.

Srinivasan, K., Kazandjieva, M. A., Jain, M., Kim, E. e Levis, P. (2008). SWAT: enabling wireless network measurements. Em *ACM SenSys*, p. 395–396.

Yeo, J., Youssef, M. e Agrawala, A. (2004). A framework for wireless LAN monitoring and its applications. Em *ACM WiSe*, p. 70–79.

Yi, C.-W. (2009). A unified analytic framework based on minimum scan statistics for wireless ad hoc and sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 20(9):1233–1245.