DOI: 10.5748/19CONTECSI/PSE/SEC/7115

A SYSTEM FOR DETECTING OAUTH 2.0 VULNERABILITIES RUNNING OVER HTTP/HTTPS IN SMART ENVIRONMENTS

UM SISTEMA PARA DETECÇÃO DE VULNERABILIDADES OAUTH 2.0 EXECUÇÃO SOBRE HTTP/HTTPS EM AMBIENTES INTELIGENTES

Anderson Aparecido Alves da Silva ; https://orcid.org/0000-0001-5426-6478
IPT/USP/UNIP/SENAC

Roberto Aparecido Ferreira
UNOESTE

Roberto César Oliveira ; https://orcid.org/0000-0003-1055-9310
IPT

José Claudio Simão
IPT

A SYSTEM FOR DETECTING OAUTH 2.0 VULNERABILITIES RUNNING OVER HTTP/HTTPS IN SMART ENVIRONMENTS

ABSTRACT

With the growth in the spread of ransomware, this malware has become a major threat to businesses and computer users. Ransomware is a different kind of malware that can block the screen of infected computers and/or encrypt the files, and only release them for payment. Due to the evolution of the techniques of obfuscation of ransomware, it becomes more difficult to detect by antivirus software among others. Because of the financial return it provides, because in most attacks users make the payment because they do not have an information security policy and together with the lack of regular backups. The present work uses an approach in which it identifies and classifies types of ransomware using machine learning algorithms such as Naive Bayes, Support Vector Machines - SVM, and K-nearest neighbors KNN. In the end, it is expected that the samples presented can be correctly identified and classified, and that which algorithm has obtained the best result.

Keywords: ransomware, malware, SVM, Naive Bayes, KNN

UM SISTEMA PARA DETECÇÃO DE VULNERABILIDADES OAUTH 2.0 EXECUÇÃO SOBRE HTTP/HTTPS EM AMBIENTES INTELIGENTES

RESUMO

Com o crescimento da disseminação do ransomware, esse malware tornou-se uma grande ameaça para empresas e usuários de computador. Ransomware é um tipo diferente de malware que pode bloquear a tela dos computadores infectados e/ou criptografar os arquivos, e apenas liberá-los mediante pagamento. Devido à evolução das técnicas de ofuscação do ransomware, torna-se mais difícil de detectar por software antivírus entre outros. Pelo retorno financeiro que proporciona, pois na maioria dos ataques os usuários efetuam o pagamento por não terem uma política de segurança da informação e juntamente com a falta de backups regulares. O presente trabalho utiliza uma abordagem na qual identifica e classifica tipos de ransomware utilizando algoritmos de aprendizado de máquina como Naive Bayes, Support Vector Machines - SVM, e K-vizinhos mais próximos KNN. Ao final, espera-se que as amostras apresentadas possam ser corretamente identificadas e classificadas, e qual algoritmo obteve o melhor resultado.

Palavras-chave: ransomware, malware, SVM, Naive Bayes, KNN

1. INTRODUÇÃO

Atualmente existem várias formas de detecção de ransomware[1], que utilizam diferentes maneiras de análise, e as principais são: dinâmica, estática, estatística e de contexto. Essas técnicas são utilizadas por softwares de proteção contra softwares maliciosos que, instalados em computadores corporativos e pessoais, tem como objetivo detectar a ameaça antes que ela se instale ou se propague.

O desafio dos profissionais de Tecnologia da Informação, seja nas corporações ou dentro das empresas que entregam para o mercado soluções para detectar e corrigir software mal-intencionado tem sido cada vez mais complexo e laborioso, essa tese inclusive é suportada por diversos relatórios da indústria de proteção contra vírus, gestão de vulnerabilidades e detecção de falhas, dentre esses relatórios destacam-se o da SonicWall [2] importante companhia de firewall e proteção de rede, que indica que os eventos de Ransomware tiveram um aumento de 151% no primeiro trimestre de 2021 em comparação ao mesmo período de 2020, tendo em primeiro lugar como país alvo os Estados Unidos e o Brasil estando na quarta posição desse relatório e dentre os países que mais espalham software maliciosos estão Vietnã, Sri Lanka, Polônia e Brasil, sucessivamente, outro importante relatório da Fortinet [3] aponta o crescimento de 10,7 vezes em junho de 2021 comparado a junho de 2020 as detecções de software malicioso nas empresas que utilizam sua tecnologia.

Ransomware é um tipo de código malicioso que torna um arquivo inacessível por meio de criptografia e esse bloqueio é possível pois apenas o atacante possui a chave para decifrá-lo, isso aplica-se aos dados armazenados em dispositivos de conexão local ou remota [4].
O crescimento desse tipo de ataque, tem como motivador a alta possibilidade de retorno financeiro, uma vez que após ter os arquivos criptografados pelo Malware, fica o usuário, instituição ou sistema impedido de utilizar-lo e para que essa criptografia seja revertida e os arquivos decifrado é pedido o pagamento de cifras elevadíssimas com criptomoedas, usualmente Bitcoin e assim possibilitar a reversão do dano causado pelo Ransomware, estima-se[5] que até o ano de 2031 os custos globais na prevenção, recuperação e controle do dano causado por esses ataques, ultrapassem 265 bilhões de dólares, hoje segundo o Centro de Reclamações de clientes de internet (Internet Crime Complaint Center – IC3) do Escritório Federal de investigação (Federal Bureau of Investigation – FBI) [4] do governo dos Estados Unidos, contabiliza perdas de mais de 29 milhões de dólares, de cidadãos, empresas ou instituições em 2.474 incidentes reportados apenas em 2020, no Brasil o montante despendido também não deve ficar muito menor já que nosso nível de conexão com a internet e digitalização de negócios e instituições é crescente.

O pagamento de resgate é desestimulado e desaconselhado por várias instituições sejam governamentais como IC3 [4] ou publicações [5] por não haver garantia de reversão do dano causado e também por estimular a prática do agressor que em última análise acaba financiando o aprimoramento de sua técnica e sofisticação de suas ferramentas.

Ter backup estruturado, protegido e aplicando melhores práticas dos fabricantes de software de mercado são capazes de proteger qualquer sistema de ataques e associados a guias de boas práticas como do Centro de Cibersegurança Nacional (National Cyber Security Centre – NCSC) do Reino Unido [6] podem abreviar o restabelecimento de ambientes tecnológicos além de garantir a integridade dos dados, entretanto como medida alternativa ao citado, existem maneiras de reverter a criptografia desses arquivos, uma delas utilizando ferramentas listadas na iniciativa chamada No More Ransom um website da Unidade de Crime de Alta Tecnologia da Polícia Holandesa, do Centro Europeu de Cibercrime da polícia Européia (European Cybercrime Centre – EC3), das empresas Kaspersky e McAfee que tem por

objetivo ajudar vítimas de Ransomware a recuperar seus arquivos codificados sem a necessidade de pagamento ou não havendo backup, este website disponibiliza uma grande quantidade de informações sobre vários tipos de Ransomware juntamente com ferramentas para decifrar os arquivos.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção são apresentados os principais conceitos para fundamentar e conceituar este artigo.

2.1 *Netflow*

É uma técnica de análise de tráfego de rede passiva, isso significa que não executa injeção de tráfego no canal como a análise ativa, mas apenas capturando a transmissão de dados no meio e podendo ser online, que consiste em captura e observação ou offline, captura, armazenamento e posterior investigação[8].

A tecnologia de fluxo de pacotes ou Netflow (RFC 3954) foi desenvolvida pela empresa Cisco Systems para um melhor monitoramento de redes de grande porte. De acordo com (WAGNER et al., 2011b) Um argumento para usar os registros do Netflow é que eles incluem todas as informações de tráfego de rede relevantes em uma versão compactada.

Atualmente na versão 9 o NetFlow suporta IPv6 assim dando suporte a Firewall, Roteadores e outros equipamentos de redes da nova geração.
Computadores conectados na rede ficam escutando, ou seja, recebendo os fluxos e armazenado para posteriores análises. As formas de armazenamento podem ser em banco de dados relacionais ou em arquivos.csv dependendo do dispositivo que está gerando os fluxos.

Dentre os campos do Netflow podemos destacar:

- Endereço de IP de origem;
- Endereço de IP de destino;
- Porta de origem;
- Porta de destino;
- Protocolo;
- Tipo de serviço ToS.

Com essas informações será possível detectar anomalias no tráfego da rede, ou seja, assinaturas de códigos maliciosos em específico para detecção de ransomwares.

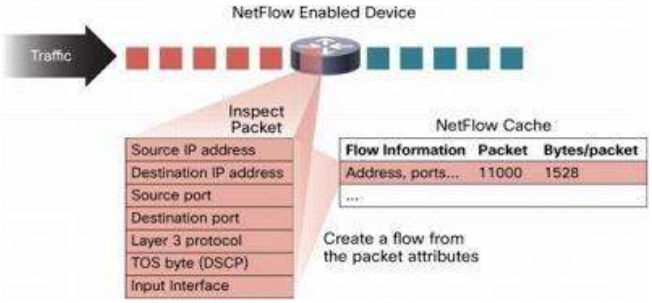

Figura 1 - Estrutura de NetFlow versão 9.

2.2 Aprendizado de máquina

Aprendizado de máquina é uma área de Inteligência Artificial, cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática. Um sistema que utiliza treinamento é um programa de computador que toma decisões baseado em experiências acumuladas através da solução bem-sucedida de problemas anteriores [Monard e Baranauskas (2003)].

No aprendizado de máquina utiliza-se a indução como forma de inferência lógica para obter conclusões genéricas sobre um conjunto de amostras particular. Esta técnica de inferência indutiva pode ser realizada de duas formas: supervisionada e não supervisionada. O presente trabalho vai utilizar o aprendizado supervisionado como linha de pesquisa para solucionar o problema proposto a ser resolvido.

Para isso empregam-se vários algoritmos de aprendizagem de máquinas. Esta pesquisa vai utilizar o aprendizado supervisionado, e em decorrência dessa característica existem alguns algoritmos que poderão ser utilizados em busca de melhores resultados para a solução do problema proposto. De acordo com (VINAYAKUMAR et al., 2017) existem várias abordagens no aprendizado supervisionado, como Árvore de Decisão (DT), Naive Bayes (NB), Máquina de Vetor de Suporte (SVM), Floresta Aleatória (RF), Ada Boost (Ada) e K-vizinho mais próximo (KNN). Também podem ser utilizados modelos de redes neurais artificiais como as Redes Multi Layer Perceptron (MLP) para identificação e classificação.

Os algoritmos utilizados no presente trabalho foram: Naive Bayes, Máquinas de Vetor de Suporte (SVM) e K-vizinhos mais próximos (KNN) ao qual vai ser feita uma breve explanação nas subseções a seguir.

O algoritmo Naive Bayes ou classificador Bayesiano, é um algoritmo inspirado no Teorema de Bayes. O raciocínio do método Bayesiano fornece uma abordagem probabilística para aprendizagem, e está baseada na suposição de que as quantidades de interesse são reguladas por distribuições de probabilidade (KOERICH, 2012). Novas instâncias podem ser classificadas combinando a probabilidade de múltiplas hipóteses ponderadas pelas suas probabilidades.

Por exemplo, objetos que podem ser classificados em uma classe entre m classes possíveis C1,...,Cm. Considere um objeto com p características descritas por variáveis categóricas xi,i=1,p, em um vetor X=(x1,...,xp), cuja classe não é conhecida. A proposta do Naive Bayes é classificar o objeto na classe com a maior probabilidade a posteriori condicionada a X. O classificador associa o objeto com características X à classe Ck se, e somente se, $P(Ck|X) \geq P(Cj|X)$ j,k=1,m e $j \neq k$ a X. O classificador associa o objeto com características X à classe Ck se, e somente se, $P(Ck|X) \geq P(Cj|X)$ j,k=1,m e $j \neq k$ (ALMEIDA; PESSANHA; CALOBA, 2018).

Máquinas de Vetor de Suporte SVM: Fundamentada na Teoria da Aprendizagem Estatística, a Máquina de Vetor de Suporte, do inglês Support Vectors Machine - SVM, foi desenvolvida por (VAPNIK, 1995), com o intuito de resolver problemas de classificação de padrões. As SVMs, em sua essência, fundamentam-se nas Teorias Estatísticas e resolvem problemas de classificações de padrões, prevendo que é possível separar classes diferentes no espaço euclidiano, possibilitando que os dados que não são linearmente separáveis, ou seja, as amostras podem ser separadas em grupos, onde a superfície de decisão irá usar linhas e curvas para delimitar as regiões limites (VAPNIC, 1999).

K-Vizinho mais próximo KNN: O algoritmo K-Nearest Neighbor ou K-Vizinho mais próximo (KNN), é o mais adotado para reconhecimento de padrões. Vários pesquisadores descobriram que o algoritmo KNN tem um desempenho muito superior em seus experimentos em vários conjuntos de dados (RAWAT; CHOUBEY, 2016).

O K-vizinho mais próximo descobre o número de amostras nos dados de treinamento que estão mais próximos da amostra de teste e, em seguida, atribui o rótulo de classe mais frequente entre as amostras de treinamento consideradas para a amostra de teste. Para classificar as amostras, o K-vizinho mais próximo é conhecido como uma abordagem que é a mais simples e não-paramétrica. K-vizinho mais próximo pode ser mencionado como um aprendiz baseado em instância, não um indutivo com base (HAQ et al., 2015).

Os exemplos de treinamento são vetores em um espaço de recurso multidimensional, cada um deles possui rótulo de classe. A fase de treinamento deste algoritmo consiste apenas em armazenar os rótulos de classes e os vetores de características das amostras de treinamento (RAJU; SUBRAHMANIAN; SIVAKUMAR, 2017).

3. TRABALHOS RELACIONADOS

Nessa seção é realizada a análise de alguns trabalhos que abordam assuntos relacionados à ataques Ransomware e técnica para sua detecção em armazenamento de dados local ou remota.

4. ARQUITETURA PARA ANÁLISE DO TRÁFEGO

Este artigo utiliza arquivos de tráfego de rede, previamente identificados como ransomware obtidos no site malware-traffic-anaslysis-net e já estão no formato .pcap, ao acessar o site e baixar os arquivos, é necessário que sejam descompactados, posteriormente para analisar o arquivo o software Caploader, foi utilizado na leitura e conversão do tráfego normal para fluxo de pacotes e posterior envio para uma planilha que pôde ser aberta no Microsoft Excel.

Este trabalho foi realizado para os três tipos de ransomwares que foram utilizados nesta pesquisa, ficando todos os fluxos de pacotes na mesma planilha gerando um arquivo chamado dataset.xlsx.

Tendo extraídos os fluxos e inseridos em uma tabela do Microsoft Excel, algumas ações se fizeram necessárias no tratamento dos dados. Algumas colunas foram retiradas, pois não são relevantes para presente pesquisa, e outras colunas foram convertidas, e ou, substituídas de letras para números, no intuito de que os algoritmos pudessem lê-las.

O conjunto de dados tem um total de 2584 linhas e 11 colunas somadas das amostras dos 3 tipos de ransomware. Primeiramente foram convertidos para números os endereços IP das colunas Client_Ip e Server_Ip para Cli_Ip_Conv e Server_Ip_Conv. A conversão dos endereços IP em número foi realizada por um programa desenvolvido em Python que utilizou o módulo ip address da biblioteca padrão do Python que realizou as conversões. A Tabela abaixo exemplifica como era e como ficaram os campos.

Tabela 1 - Exemplos de conversão de endereços IP para números inteiros.

Client_IP	Cli_Ip_Conv	Server_IP	Server_Ip_Conv
10.6.25.102	168171878	223.27.21.46	374309462
10.6.25.102	168171878	85.25.95.39	142772619
10.6.25.102	168171878	46.30.42.236	773729004
10.6.25.102	168171878	46.30.42.236	773729004
10.6.25.102	168171878	199.59.149.233	334257985
10.6.25.102	168171878	199.59.149.233	334257985
10.6.25.102	168171878	52.3.78.30	872631838
10.6.25.102	168171878	115.28.36.224	193122288

Fonte: Elaborado pelo autor.

A coluna duração do fluxo denominada Duration também foi convertida, e ficou como Duration_Miless. Esta coluna estava em formato hora, minuto, segundo milésimo de segundo. Os dados contidos nesta coluna foram todos padronizados para o formato milésimo de segundo, como está descrito na tabela 2.

Tabela 2 - Exemplo de conversão de hora, minuto, segundo e milissegundo para milissegundo.

Duration	Duration_Miless
00:00:01.284	1284
00:01:16.382	76382
00:01:13.849	73849
00:00:01.646	1646
00:01:13.488	73488
00:01:13.419	73419
00:01:25.018	85018
00:01:39.148	99148
00:00:03.804	3804

Fonte: Elaborado pelo autor.

Outra coluna que também recebeu uma atenção especial foi a TCP_Flags que estava em tipo string (onde constavam como um grupo de palavras) também fora convertida em números. Como pode ser visto a Tabela 3 ilustra a tabela após as devidas conversões.

Tabela 3 – Tabela Flags TCP_Flags.

Tipo	Classificação
AP	3
AP F	4
AP S	5
AP SF	6
APR	7
APR F	8
APRRS	9
APRSF	10

Fonte: Elaborado pelo autor.
Legenda A = Ack; P = Push; R = Reset; S = Syn; F = Fin.

Foi criada uma coluna Class para identificar e classificar os tipos de ransomware que estava no conjunto de dados. Na tabela 4 está demonstrando a classificação dos ransomwares.

Tabela 4 – Classificação dos ransomwares.

Tipos	Classificação
Cerber	0
Criptomix	1
Criptowall	2

Os tratamentos realizados nas referidas colunas se fizeram necessários porque os algoritmos de aprendizagem de máquinas necessitam que os dados estejam de forma padronizada, para que possam efetivamente fazer a leitura e consequentemente a classificação correta das amostras.

Foram utilizados os algoritmos Naive Bayes, Máquina de Vetor de Suporte SVM e K-vizinhos mais próximos – KNN como mencionado em seções anteriores.

5. RESULTADOS

Aqui serão apresentados os resultados da execução dos algoritmos de aprendizagem de máquina ao arquivo dataset.xlsx. Após a execução, é feita uma comparação entre os algoritmos baseados nos métodos de avaliação de precisão como acurácia e também foi utilizado um método de concordância denominado coeficiente Kappa. O coeficiente Kappa é

uma medida estatística que leva em consideração as probabilidades de as concordâncias de itens categóricos terem ocorrido ao acaso (COHEN, 1960).

O experimento teve início utilizando o algoritmo de aprendizado de máquinas supervisionado na Naive Bayes. A tabela 5 mostra uma matriz de confusão com o resultado do algoritmo Naive Bayes.

Tabela 5 - Matriz de Confusão Naive Bayes.

	Cerber	Criptomix	Criptowall	Precisão
Cerber	**187**	20	0	0.90
Criptomix	15	**51**	0	0.77
Criptowall	0	0	**503**	1.0

Fonte: Elaborado pelo autor.

O resultado foi o seguinte:

- 187 amostras do tipo Cerber classificadas VP;
- 20 amostras do tipo Cerber classificadas FP do tipo Criptomix;
- 0 amostras do tipo Cerber classificadas como FP do tipo Criptowall;
- 15 amostras do tipo Criptomix classificadas FP do tipo Cerber;
- 51 amostras do tipo Criptomix classificadas VP;
- 0 amostras do tipo Criptomix classificadas como FP do tipo
- Criptowall;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Cerber;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Criptomix;
- 503 amostras do tipo Criptowall classificadas como VP do tipo Criptowall.

No segundo experimento foi utilizado o algoritmo KNN, que como mostra a tabela 6 os resultados obtidos:

Tabela 6 Matriz de Confusão algoritmo KNN.

	Cerber	Criptomix	Criptowall	Precisão
Cerber	**188**	18	1	0.91
Criptomix	16	**50**	0	0.76
Criptowall	0	0	**503**	1.0

- 188 amostras do tipo Cerber classificadas VP;
- 18 amostras do tipo Cerber classificadas como FP Criptomix;
- 1 amostra do tipo Cerber classificadas como FP do tipo Criptowall;
- 16 amostras do tipo Criptomix classificadas FP do tipo Cerber;
- 50 amostras do tipo Criptomix classificadas VP;
- 0 amostras do tipo Criptomix classificadas como FP do tipo
- Criptowall;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Cerber;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Criptomix;
- 503 amostras do tipo Criptowall classificadas VP.

O terceiro e último experimento foi realizado com o algoritmo SVM, segue abaixo os resultados que podem ser vistos na tabela 7.

Tabela 7 Matriz de Confusão Teste SVM.

	Cerber	Criptomix	Criptowall	Precisão
Cerber	**207**	0	0	1.0
Criptomix	0	**66**	0	1.0
Criptowall	0	0	**503**	1.0

Fonte: Elaborado pelo autor.

- 207 amostras do tipo Cerber classificadas como VP;
- 0 amostras do tipo Cerber classificadas FP do tipo Criptomix;
- 0 amostras do tipo Cerber classificadas FP do tipo Criptowall;
- 0 amostras do tipo Criptomix classificadas FP do tipo Cerber;
- 66 amostras do tipo Criptomix classificadas VP;
- 0 amostras do tipo Criptomix classificadas FP do tipo Criptowall;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Cerber;
- 0 amostras do tipo Criptowall classificadas como FP do tipo Criptomix;
- 503 amostras do tipo Criptowall classificadas como VP.

Na tabela 5 pode ser observada os números com as porcentagens de acerto de cada algoritmo e o método utilizado.

Tabela 8 – Resultado dos métodos de precisão aplicados.

	Acurácia	Kappa
Naive Bayes	95,49%	91,02%
KNN	95,49%	91,00%
SVM	100%	100%

Fonte: Elaborado pelo autor.

Tendo em vista os resultados apresentados pode-se constatar que o algoritmo Naive Bayes e KNN apresentaram um desempenho igual no método de acurácia, e no método kappa com uma pequena vantagem para o algoritmo Naive Bayes.

Ambos apresentam uma taxa de acerto muito satisfatória, pois atingem mais de 95 % na acurácia e 91 % no modelo kappa respectivamente. Já o SVM foi o melhor entre os três. Obteve uma taxa de acerto de 100% em ambos os métodos.

5. CONCLUSÃO

O trabalho demonstrou como melhor algoritmo para resolução do problema proposto o SVM, que teve uma taxa de acerto de 100% na classificação das amostras nos referidos métodos acurácia e kappa demonstrada na de matriz de confusão.

Com o intuito de propor trabalhos futuros, sugere-se a aplicação de outros tipos de algoritmos de aprendizagem de máquina supervisionado, com a intenção de verificar qual o que possui melhor desempenho na classificação de ransomware. Pode-se também utilizar como outra sugestão algoritmos não supervisionados, e comparar com algoritmos supervisionados para ver qual tem o melhor desempenho

REFERÊNCIA BIBLIOGRÁFICA

Rupali Komatwar & Manesh Kokare (2021) A Survey on Malware Detection and Classification, Journal of Applied Security Research, 16:3, 390-420, DOI: 10.1080/19361610.2020.1796162

SonicWall Exposes Soaring Threat Levels, Historic Power Shifts In New Report em: https://blog.sonicwall.com/en-us/2021/03/sonicwall-exposes-soaring-threats-historic-power-shifts-in-new-report . Acesso em: 16 de outubro de 2021.

Global Threat Landscape Report em: https://informationisbeautiful.net/visualizations/ransomware-attacks/ Acesso em: 16 de outubro de 2021.

Internet Crime Report em: https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf Acessado em 17 de outubro de 2021.

Global Ransomware damage costs predicted to exceed $265 Billion by 2031 em: https://cybersecurityventures.com/global-ransomware-damage-costs-predicted-to-reach-250-billion-usd-by-2031/ Acessado em 17 de outubro de 2021.

Mitigating malware and ransomware attacks em: https://www.ncsc.gov.uk/guidance/mitigating-malware-and-ransomware-attacks#stepsifinfected Acessado em 17 de outubro de 2021.

No More Ransom em: https://www.nomoreransom.org/pt/decryption-tools.html Acessado em: 17 de outubro de 2021.

R. Hofstede et al., "Flow Monitoring Explained: From Packet Capture to Data Analysis With NetFlow and IPFIX," in IEEE Communications Surveys & Tutorials, vol. 16, no. 4, pp. 2037-2064, Fourthquarter 2014, doi: 10.1109/COMST.2014.2321898.