



**X**

## Referências

Acesso fornecido por:


► [Mostrar no contexto Google Scholar](#)

► **Mostrar no contexto** **Ver artigo**  
Google Scholar

► [Mostrar no contexto](#)   [CrossRef](#)  

► [Mostrar no contexto](#) [Financiar a publicação deste conteúdo](#) [Ver artigo](#) 

[Google Scholar](#)  [Adicionar aos alertas de citação](#)

► [Mostrar no contexto](#)   [CrossRef](#)  

[Mostrar no contexto](#) [Ver artigo](#) 

Navegar▼ Minhas configurações▼ Ajuda▼ PUC-MG

## Todos

### Arquitetura reconfigurável de K-Means/K-Modes baseada em FPGA para detecção de intrusão de rede

**Cite isto**

 PDF

Todos os autores ...

704  
Completo  
Visualizações de  
texto

## Seções do documento

## II. Fundo

 Trabalho relacionado

#### 4. Arquitetura K-Means/K-Modes baseada em FPGA

## V. Metodologia de Avaliação

Mostrar esboço completo ▼

## Autores

## Figuras

## Referências

## Citações

### Palavras-chave

## Métricas

Mais como isso


 baixar PDF

PDF

Ajuda

**Resumo:** Ao longo dos anos, a quantidade de dados compartilhados entre usuários de diversas áreas cresceu consideravelmente. Consequentemente, o mesmo aconteceu com os ataques à segurança de informações. Neste trabalho, apresentamos uma nova abordagem para a detecção de anomalias em dados de segurança. A abordagem proposta utiliza uma técnica de aprendizado de máquina, especificamente, uma rede neural profunda (DNN), para detectar anomalias em dados de segurança. Os resultados experimentais demonstram que a abordagem proposta apresenta uma performance superior em relação aos métodos tradicionais de detecção de anomalias. **Ver mais**

### ► Metadatos

**Abstrato:**

Ao longo dos anos, a quantidade de dados compartilhados entre usuários de diversas áreas cresceu consideravelmente. Consequentemente, o mesmo aconteceu com os ataques à rede. As estratégias de monitoramento de segurança devem classificar os tipos de informação nas redes de forma rápida e eficaz. Sistemas de Detecção de Intrusão têm sido propostos com técnicas de Aprendizado de Máquina e Computação de Alto Desempenho para evitar anomalias de segurança. Assim, os dispositivos FPGA são bons candidatos para melhorar o desempenho e a eficiência energética. Neste resumo, propomos uma arquitetura K-means/K-modes reconfigurável baseada em FPGA para acelerar o agrupamento de dados para detecção de intrusão na rede. Avaliamos nossa abordagem no conjunto de dados NSL-KDD e os resultados mostraram que K-means e K-modes podem alcançar até 15× e 994× mais operações por Watt do que versões de software paralelas.

**Publicado em:** Transações IEEE em Circuitos e Sistemas II: Express Briefs ( Volume: 67 , Edição: 8 , agosto de 2020 )

**Página(s):** 1459 - 1463

DOI: 10.1109/TCSII.2019.2939826

Data de publicação: 06 de setembro de 2019?

Editora: IEEE

► **Informações ISSN:**

► Agência Financiadora:

Conteúdo

Referências

SEÇÃO I.

Introdução

A Internet tornou-se uma ótima maneira de compartilhar informações entre muitos usuários e máquinas. Grandes volumes de dados de diversas áreas são transferidos em redes locais e globais a cada segundo, aumentando as preocupações com a segurança dos dados. Isso leva a investimentos em estratégias de monitoramento e segurança. Para reduzir os ataques cibernéticos, foram propostos Sistemas de Detecção de Intrusão (IDSs) [1]. Esses sistemas protegem as informações nas redes, identificando possíveis ataques. Os administradores do sistema sobre possíveis invasões ou até mesmo evitam a chegada de intrusões. A detecção de intrusões pode ser realizada de duas maneiras: (i) A detecção baseada em assinaturas de intrusões comparando os dados avaliados e os armazenados previamente conhecidos. (ii) A detecção baseada em anomalias identifica registros que possuem características diferentes do comportamento normal.

Para melhorar a eficiência do IDS, foram aplicadas técnicas de Machine Learning. O aprendizado de máquina revelou uma ampla gama de possibilidades para detecção de invasões [2]. Nesse contexto, métodos de agrupamento K-means e K-modes são constantemente utilizados. Os clusters formados por estes algoritmos podem representar tipos de anomalias e acessos normais [3]. Isto permitiu que os IDSs avaliassem novos registros de acordo com semelhanças entre os dados.

A grande quantidade de dados numéricos ou categóricos provenientes do tráfego de rede exige alto desempenho para serem processados. Arquiteturas paralelas poderosas, como unidades de processamento de Intel Xeon Phi têm sido amplamente utilizadas para acelerar algoritmos de cluster. Porém, estas arquiteturas são compostas por circuitos complexos para oferecer alto desempenho em ponto flutuante, consumindo energia considerável [4]. Assim, arquiteturas eficientes baseadas em Field Programmable Gate Arrays (FPGAs) devem ser consideradas [5]. Neste resumo, nosso objetivo é propor uma arquitetura K-means/K-modes reconfigurável baseada em FPGA de alto desempenho e eficiência energética para detecção de intrusão de rede.

Em resumo, nossas contribuições para o estado da arte são:

- Uma arquitetura K-means/K-modes que suporta dados de entrada numéricos ou categóricos com 64 bits;
- Flexibilidade para alterar os parâmetros de ambos os algoritmos (número de clusters e tipo de algoritmo) em tempo de execução;
- Uma arquitetura reconfigurável com eficiência energética e alto desempenho em comparação com as versões de software paralelas.

Este resumo está organizado da seguinte forma. A Seção II fornece uma explicação dos algoritmos. A Seção III apresenta trabalhos relacionados. Na Seção IV, mostramos nossa arquitetura de hardware proposta. A metodologia de avaliação é apresentada na Seção V. Os resultados obtidos são apresentados na Seção VI e na Seção VII apresentamos nossas considerações finais.

SEÇÃO II.

Fundo

K-means é um algoritmo de agrupamento que agrupa dados em  $k$  clusters, com base em suas semelhanças [6]. A ideia básica é, a partir de um conjunto de  $n$  pontos de dados com  $d$  atributos, associam iterativamente esses pontos aos centros de cada cluster, denominados centróides, com base em sua distância euclidiana. O algoritmo começa a inicializar  $k$  centróides com valores estratégicos. Posteriormente, a distância entre cada ponto e um centróide é calculada, e o ponto é atribuído ao seu centróide mais próximo [7]. Na última etapa, os centróides são atualizados, com base na média aritmética dos atributos dos pontos de dados a eles associados. O algoritmo executa essas etapas enquanto os centróides continuam mudando seus valores ou até atingir um determinado número de iterações. Assim, para cada iteração  $t$ , todas essas etapas são executadas com complexidade computacional de  $O(nkd)$ . Ao final, a saída é um mapeamento de centróides e seus pontos de dados atribuídos, representando os clusters gerados.

Os modos K funcionam de forma semelhante aos K-means, mas usando dados categóricos, ou seja, dados não numéricos, como entrada [8]. Contudo, as distâncias entre pontos e centróides são definidas por técnicas de dissimilaridade (por exemplo, correspondência simples). Além disso, o modo aritmético é usado para atualizar os atributos dos centróides.

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018. [Mostrar no contexto](#) [Ver artigo](#) [Google Scholar](#)

2. Desai, M. Z. J. Fei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusões em redes de nuvem", *IEEE Access*, vol. 6, pp. 21954-21961, 2017. [Mostrar no contexto](#) [Ver artigo](#) [Google Scholar](#)

3. A. Bohara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando agrupando diversos dados de monitor", *Proc. 9 Ciência do Bootcamp. Segurança (HotSoS)*, p. 16, 2016. [Mostrar no contexto](#) [CrossRef](#) [Google Scholar](#)

4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rur a computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eui (DATA)*, pp. 398-403, março de 2017. [Mostrar no contexto](#) [Ver artigo](#) [Google Scholar](#)

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial basead IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017. [Mostrar no contexto](#) [CrossRef](#) [Google Scholar](#)

6. S. Lloyd, "Quantização de mínimos quadrados PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. pp. 129-137, março de 1982. [Mostrar no contexto](#) [Ver artigo](#) [Google Scholar](#)

PDF

Ajuda

https://ieeexplore.ieee.org/document/8826329

2/10

A Figura 1 mostra um exemplo de clustering com  $k = 3$ . Os algoritmos K-means e K-modes são capazes de agrupar dados com características semelhantes gerados por diversos contextos. As informações da rede podem ser utilizadas como entrada para eles, os grupos gerados podem ser marcados como acessos maliciosos ou normais. Desta forma, Sistemas de Detecção de Intrusões baseados em métodos de agrupamento são amplamente utilizados para encontrar novos padrões de anomalias para classificar e detectar intrusões desconhecidas.

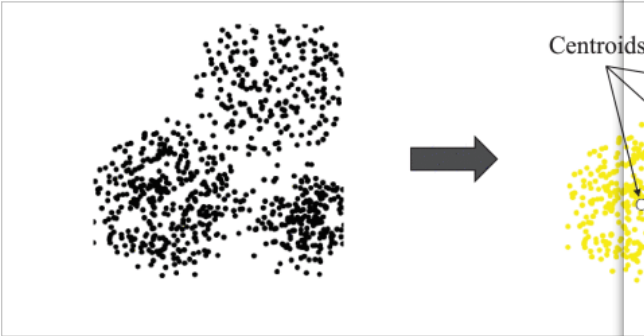


Figura 1.  
Exemplo de agrupamento com  $k = 3$ .

SEÇÃO III.

## Trabalho relatado

Nesta seção, apresentamos trabalhos relacionados e uma breve comparação na Tabela I com trabalhos que escalam com arquiteturas intrusões em redes.

TABELA I Visão Geral dos Trabalhos Relacionados

Work	Device	Methods	Characteristics
[9]	CPU	K-means	Evaluate Detection Rate, NSL-KDD data set
[10]	CPU	K-means SVM and ELM	Evaluate Detection Rate, KDDCup'99 data set
[11]	CPU	K-modes Active Learning	Evaluate 7 clusters, Synthetic data set
[12]	CPU	K-means	Evaluate 7 clusters, KDDCup'99 data set
[13]	CPU	K-means PCA	Evaluate Time, KDDCup'99 data set
[14]	CPU	K-means Naive Bayes	Evaluate 2 clusters, KDDCup'99 data set
[15]	CPU	K-means Random Forest	Evaluate 2 clusters, NSL-KDD data set
[16]	FPGA	K-means	32-bit data, FPGA with 40 MHz, Evaluate Time, DARPA data set
[17]	FPGA	K-means	16-bit data, 82mW and 162 MHz, Evaluate Energy, Not related to IDS
[18]	CPU + FPGA	K-means	128-bit data, 126W (CPU+FPGA), Evaluate Time/Energy, Not related to IDS
[19]	FPGA	K-means	FPGA with 250 MHz, Evaluate Time/Energy, Not related to IDS
[20]	65nm CMOS	K-means	16-bit data, 41mW and 250 MHz, Evaluate Energy, Not related to IDS
Our Work	FPGA	K-means K-modes	64-bit data, Reconfiguration Inputs, 1.5W and 50 MHz, Evaluate Cycles/Energy, Evaluate up to 8 clusters, NSL-KDD data set

Um método chamado MinMax aumenta a qualidade dos centróides iniciais em K-means [9] . Isso melhorou a taxa de detecção com uma redução nos falsos positivos. Um modelo multinível [10] para um IDS baseado em clustering K-means, Support Vector Machine (SVM) e Extreme Learning Machine (ELM), foi proposto para agrupar dados de treinamento usando K-means, produzindo novos conjuntos de dados para aprimorar o SVM. e ELMO. Outro trabalho [11] realizou detecção de anomalias em conjuntos de dados de alta dimensão usando redes Bayes, modelo de mistura gaussiana (GMM), modos K e aprendizado ativo. Os autores identificaram subgrupos de dimensões relevantes, que poderiam ser classificados como ataques.

## Referências

1. A. W. Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018.  
▶ Mostrar no contexto Ver artigo Google Scholar
2. C. Yin, Y. Zhu, J. Fei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusões usando redes neurais recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.  
▶ Mostrar no contexto Ver artigo Google Scholar
3. A. Bohara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando e agrupando diversos dados de monitor", *Proc. 9º Simpósio Brasileiro de Segurança da Informação e da Comunicação (SBSEG) - Segurança (HotSoS)*, p. 16, 2016.  
▶ Mostrar no contexto CrossRef Google Scholar
4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rural network intrusion detection using a multi-scale approach", *Proc. Projeto Autom. Teste e Validação (DATA)*, pp. 398-403, março de 2017.  
▶ Mostrar no contexto Ver artigo Google Scholar
5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial baseada em IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.  
▶ Mostrar no contexto CrossRef Google Scholar
6. S. Lloyd, "Quantização de mínimos quadrados por mínimos quadrados", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 1, pp. 129-137, março de 1982.  
▶ Mostrar no contexto Ver artigo

Uma abordagem modificada usando K-means teve como objetivo aplicar o método de agrupamento para IDS [12]. Após testar sua abordagem no conjunto de dados KDDCup'99, os autores mostraram uma eficiência de até 95% na detecção de intrusões, apresentando distâncias interclusters maiores que a versão tradicional, formando assim grupos mais distintos e de maior qualidade. Em outro experimento [13], foi utilizado um método de agrupamento para IDS usando K-means com Mini Batch e Análise de Componentes Principais (PCA). A abordagem desenvolvida foi comparada com versões tradicionais de K-means e PCA, utilizando também o conjunto de dados KDDCup'99. O modelo foi eficiente no cenário de big data, apresentando tempos de execução mais rápidos que os clássicos. Um classificador Naive Bayes [14] também foi usado junto com K-means e testado no KDDCup'99. O objetivo era categorizar os dados em instâncias normais e ataques potenciais. Da mesma forma, também foi utilizado um classificador Random Forest [15] com K-means, apresentando taxa de precisão de 99,86% na detecção de intrusões no conjunto de dados NSL-KDD.

Embora esses trabalhos tenham obtido bons resultados, consideraram cenários com baixo poder computacional e eficiência energética, que não podem ser negligenciados atualmente. Pensando nisso, a arquitetura proposta para detecção de intrusão de arquitetura de hardware do algoritmo K-means [16]. Este hardware foi capaz de usar os dados de entrada para identificar características de quatro tipos de ataques. Utilizando um modelo FPGA de 64 bits e uma entrada de 32 bits e um número fixo de iterações, os autores mostraram que a arquitetura proposta está à altura de 300 × mais rápido do que uma versão de software ao avaliar registros gerados sintaticamente existentes no conjunto de dados DARPA. Este resultado representa uma redução no tempo de processamento de 99%, semelhante a este briefing que obteve uma redução de 91%. Em termos de elementos lógicos, este trabalho relacionado mostra uma ocupação FPGA de 58.484 elementos (apenas K-means); valor pouco mais da metade dos 106.824 elementos alcançados por nossa implementação híbrida K-means/K-modes.

Existem outras propostas relacionadas a K-means implementadas em FPGA [17] – [19] e CMOS [20]. Elas não estão relacionados ao IDS, mas destacam a importância dos K-means baseados em hardware como aceleradores para diferentes aplicações.

Este resumo difere dos outros, pois propomos uma arquitetura K-means/K-modes reconfigurável baseada em FPGA com dados de 64 bits. Nossa abordagem oferece maior desempenho e menor consumo de energia no cenário de detecção de intrusões do que as abordagens de software. Além disso, é capaz de avaliar até 8 clusters com ambos os algoritmos, mostrando uma solução FPGA mais robusta que as apresentadas nos trabalhos anteriores.

SEÇÃO IV.

Arquitetura K-Means/K-Modes baseada em FPGA

Nossa arquitetura reconfigurável de K-means/K-modes é baseada em blocos interconectados, gerenciados por uma unidade de controle central, conforme mostrado na Figura 2. Esta arquitetura é descrita em VHDL (VHSIC Hardware Description Language) e seus blocos básicos são: (i) Banco de Registradores que armazena em registradores internos as operações disponíveis; (ii) Contagem de Iterações que determina a condição de parada dos algoritmos; (iii) Resultado de saída que recebe os mapeamentos.

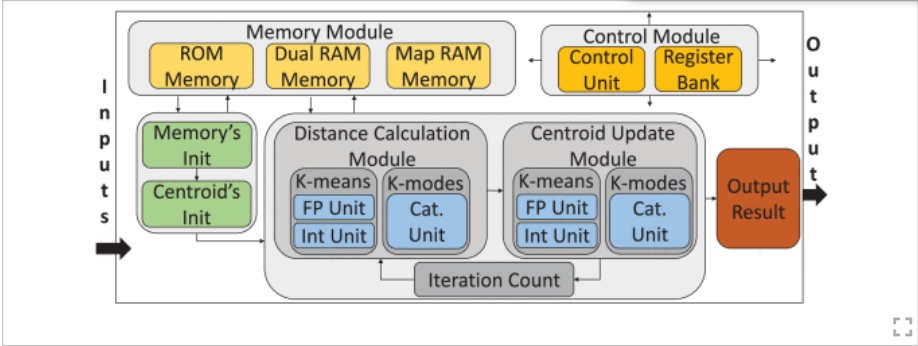


Figura 2. Arquitetura reconfigurável de K-means/K-modes.

As entradas são relógio, reset, algoritmo escolhido (K-means-Integer, K-means-FloatingPoint ou K-modes), número de pontos, centróides, recursos e iterações. A saída é o mapeamento dos centróides e seus pontos. Usamos entradas de dados de 64 bits com significados diferentes dependendo do algoritmo escolhido. Para os modos K, são usados atributos categóricos, portanto a palavra de 64 bits contém oito atributos de oito bits cada.

Referências

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018.
2. C. Xin, Y. Zhou, D. Pei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusão usando redes neurais recorrentes", *IEEE Access*, vol. 6, pp. 21954-21961, 2017.
3. A. Phara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando agrupando diversos dados de monitor", *Proc. 9. Ciência do Bootcamp, Segurança (HotSoS)*, p. 16, 2016.
4. K. O'Brien, LD Fucci, G. Durelli e M. Blott, "Rural a computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eui (DATA)*, pp. 398-403, março de 2017.
5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial basead IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.
6. S. Lloyd, "Quantização de mínimos quadrados PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. pp. 129-137, março de 1982.

Por outro lado, com K-means os atributos são numéricos, portanto definimos 2 atributos por palavras de 64 bits. Os atributos de ponto flutuante são codificados no padrão de precisão simples IEEE 754.

Para manter os dados e acelerar o acesso durante a execução dos algoritmos, foram utilizadas memórias internas. Nosso protótipo primeiro usa uma ROM para armazenar o conjunto de dados em tempo de compilação. Esta memória simplifica o protótipo de hardware no FPGA. Apesar disso, o objetivo futuro é utilizar memórias externas para armazenar o conjunto de dados. Num segundo momento, o protótipo utiliza uma RAM Dual para armazenar dinamicamente os pontos e centróides. Esta Dual RAM possui dois canais de entrada e dois canais de saída, possibilitando o processamento de duas informações no mesmo ciclo de clock. Finalmente, um Map RAM é usado para manter o mapeamento final dos pontos de dados.

A cada ciclo de clock, a unidade de controle seleciona o próximo bloco que será ativado, seguindo o fluxo de execução representado na Figura 3. A inicialização dos dados é realizada por um módulo de controle. Memória (1) e Inicialização Centróide (2). O primeiro lê um dado existente na ROM e o segundo inicializa o endereço disponível na Dual RAM. Este último atribui os valores de  $k$  pontos iniciais e a estratégia reduz a complexidade e a latência do hardware.

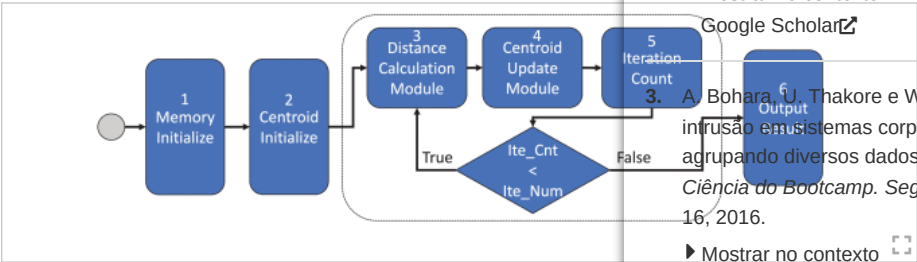


Figura 3. Diagrama de estado de hardware.

Além disso, a arquitetura K-means/K-modes é dividida em dois módulos principais: cálculo de distância (3) e atualização de centróide (4). O primeiro calcula a distância  $d$ , no qual  $n$  é o número de atributos e  $c_{eu}$  são os atributos do ponto e centróide. Este bloco consiste em três componentes, um para dados fixo, outro para dados de ponto flutuante e outro para dados categóricos.

$$d(p, c)^2 = \sum_{eu=1}^n |p_{eu} - c_{eu}|^2$$

Ver fonte

PDF

Ajuda

Nossa implementação permite a execução paralela de dois atributos quando K-means é escolhido e de um atributo para K-modes, reduzindo a ordem de complexidade dos algoritmos para  $O(n)$ . Eles possuem uma unidade comparadora que adiciona a diferença entre os atributos dos pontos e centróides a um acumulador. Em seguida, o resultado é armazenado no Map RAM. Adaptamos a distância euclidiana, de modo a retirar o cálculo da raiz quadrada, conforme mostrado na Equação 1. Este cálculo adiciona maior complexidade ao hardware e removê-lo não altera o mapeamento final. Com dados categóricos, medições de distância baseadas em números não podem ser usadas. Assim, usamos uma métrica de dissimilaridade de correspondência simples da Equação 2 para encontrar a distância entre tais atributos. A métrica atribui 0 quando os atributos são iguais e 1 quando são diferentes, conforme Equação 3.

$$d(p, c) = \sum_{eu=0}^n \delta(p_{eu}c_{eu}) \tag{2}$$

Ver fonte

onde

$$\delta(p_{eu}c_{eu}) = \begin{cases} 0, & \text{se } (p_{eu} = c_{eu}) \\ 1, & \text{se } (p_{eu} \neq c_{eu}) \end{cases} \tag{3}$$

Ver fonte

Referências

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018.

Mostrar no contexto Ver artigo

Google Scholar

2. M. Li, Y. Zhu, e P. He, "Uma abordagem de aprendizagem profunda para detecção de intrusão em redes de dados recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.

Mostrar no contexto Ver artigo

Google Scholar

3. A. Bohara, J. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando agrupando diversos dados de monitor", *Proc. 5ª Ciência do Bootcamp. Segurança (HotSoS)*, p. 16, 2016.

Mostrar no contexto Ver artigo

CrossRef

Google Scholar

4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rural à computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eui (DATA)*, pp. 398-403, março de 2017.

Mostrar no contexto Ver artigo

Google Scholar

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial basead IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.

Mostrar no contexto Ver artigo

CrossRef

Google Scholar

6. S. Lloyd, "Quantização de mínimos quadrados por um único representante", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 1, pp. 129-137, março de 1982.

Mostrar no contexto Ver artigo

CrossRef

Google Scholar



O módulo de atualização centróide possui dois blocos, um para K-means e outro para K-modes. Para calcular a média, o bloco K-means usa arrays para armazenar a soma dos atributos dos pontos atribuídos a um determinado centróide. O bloco K-modes usa oito buffers de 256 posições para armazenar a frequência de cada valor possível de atributos. As frequências são usadas para encontrar o maior elemento de cada buffer e calcular a moda. Após sua execução, este módulo armazena centróides atualizados na região mais baixa da Dual RAM. Esses cálculos são realizados até que os centróides não mudem mais ou até que um determinado número de iterações seja alcançado (5). Por fim, o mapeamento final é enviado para o bloco de resultados de saída (6), que mostra os clusters e seus pontos atribuídos.

## SEÇÃO V.

## Metodologia de Avaliação

Comparamos nossa arquitetura com versões de software paralelo OpenMP (Open Multi-Processing) com os mesmos tamanhos de entrada e definições de operação. As versões OpenMP foram executadas em um computador com 6 threads em um processador Intel Xeon E5-2620 de 6 núcleos e 2,40 GHz. Nossa arquitetura foi implantada em uma placa DE2-115 com Intel Cyclone IV-E FPGA EP4CE115F29C7, descrita em VHDL e sintetizada pelo software Quartus Prime Lite 16.1. OpenMP possui uma boa abstração para programadores e é um caminho fácil para uma solução de software focada em memória compartilhada, mas de acordo com nossos resultados na Seção VI, há uma melhoria considerável de desempenho e energia com o uso de um FPGA como acelerador.

O NSL-KDD [21] foi utilizado em nossos experimentos. É um conjunto de dados amplamente utilizado para testes em cenários de detecção de intrusão de rede. Este conjunto de dados é uma atualização da base KDD'99 resolvendo os problemas de redundância do conjunto original. Contém exemplos de 22 tipos de ataques e registros de acesso normal, com um conjunto de 125.973 dados para treinamento e 22.543 dados para teste. Cada registro possui 41 atributos numéricos e categóricos, classificados em recursos básicos, de conteúdo e de tráfego.

Nossa arquitetura recebe como entrada os dados do conjunto de dados NSL-KDD e os processa de acordo com a *Proc. Projeto Autom. Teste Euclid* configurados. O primeiro contém atributos inteiros: o número de conexões no mesmo destino (Data), pp. 398-403, março de 2017. de bytes transferidos da fonte de dados para o destino; e do destino à origem. O segundo grupo consiste em atributos de ponto flutuante: os percentuais de conexões no mesmo serviço; e esse percentual em diferentes serviços. O terceiro grupo é composto por valores categóricos: o protocolo utilizado; o serviço de rede de destino; Conecte-se; e status da conexão.

Para cada grupo, definimos um número máximo de iterações igual a 5; um número de centróides variando entre 2, 4 e 8; e total de pontos entre  $2^8$ ,  $2^{10}$  e  $2^{12}$ . Definimos o número de atributos como 4. Assim, selecionamos os atributos mais significativos dos registros com representação mais robusta, boa variedade e melhores valores diferentes de zero [22].

Para avaliar o desempenho, comparamos os ciclos de clock gastos na execução de K-means-Int (Inteiro), K-means-FP (Ponto Flutuante) e K-modes-Cat (Dados Categóricos). O consumo de energia do FPGA foi avaliado utilizando o Intel PowerPlay EPE. Para o processador Xeon, utilizamos o framework PAPEL [23]. A eficiência energética foi medida usando milhões de operações por segundo (MOPS) por consumo de energia (Watts), ou seja, MOPS/Watt. O número de operações foi calculado através do produto entre número de centroides, pontos, atributos e iterações.

## SEÇÃO VI.

## Resultados

A Tabela II mostra a ocupação do FPGA, considerando o número total de registradores, multiplicadores, elementos lógicos e bits de memória do hardware sintetizado. O grande número de operações necessárias para executar os K-means (Int/FP) e K-modes (Cat) no FPGA escolhido explica os altos números de ocupação nos elementos lógicos.

**TABELA II** Ocupação de FPGA

## Referências

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos: um número de iterações seja em sistemas de controle de rede sem fio", *IEE Trans. Sistema de Circuitos II Exp. Resumos*, 65, não. 8, pp. 1049-1053, agosto de 2018.

► [Mostrar no contexto](#)    [Ver artigo](#)

Google Scholar

2. C. Yin, Y. Zhu, J. Fei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusão usando redes neurais recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.

[Mostrar no contexto](#)
[Ver artigo](#)

Google Scholar

- gramadores e é um caminho fácil
3. A. Böhara, U. Thakore e WH Sanders, "Detecção de novos resultados na Seção Intrusão em sistemas corporativos combinando FPGA como acelerador agrupando diversos dados de monitor", *Proc. 5ª Ciência do Bootcamp. Segurança (HotSoS)*, p. 18, 2016.

atualização da base KDD'99 [▶ Mostrar no contexto](#) [CrossRef](#)

22 tipos de ataques e  
- Google Scholar

4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rur à computação em exaescala com arquiteturas DD com heterogeneidade", *Proc. Projeto Autom. Teste Eui* (DAEP), pp. 399-403, maio de 2017.

► [Mostrar no contexto](#) [Ver artigo](#)

Google Scholar

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial baseada em IEC 61850" *JCT Exp.*, vol. 4, não. 1, pp. 1-5, 2019.

Mostrar no contexto CrossRef

Google Scholar

6. S. Lloyd, "Quantização de mínimos quadrados de energia do EPGA foi avaliado PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 1, pp. 129-137, março de 1982.

► [Mostrar no contexto](#) [Ver artigo](#)

Elements	Available	Used	Used (%)
Multipliers	532	65	12%
Registers	114480	31864	28%
Logics	114480	106824	93%
Memory bits	3981312	2270602	57%

Mostrar no contexto

Ver artigo

Google Scholar

As colunas K-means da Tabela III apresentam os ciclos de execução da arquitetura proposta no FPGA em comparação com a implementação de software paralela no processador Xeon. Os resultados mostram que o FPGA tem desempenho melhor que o Xeon para aplicação K-means, realizando operações com 72% a 91% menos ciclos, com redução entre 72% e 91% para dados inteiros (Int) e entre 62% e 85% para dados reais (FP) em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018.

Mostrar no contexto

Ver artigo

Google Scholar

2. S. Min, Y. Zhu, J. Li e X. He, "Uma abordagem de detecção de intrusão baseada em algoritmo de agrupamento de dados fluídos recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.

Mostrar no contexto

Ver artigo

Google Scholar

3. A. Bohara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando agrupando diversos dados de monitor", *Proc. 5ª Ciência do Bootcamp. Segurança (HotSoS)*, p. 16, 2016.

Mostrar no contexto

Ver artigo

Google Scholar

4. K. O'Brien, L.D. Tucci, G. Durelli e M. Blott, "Rur a computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eur (DATA)*, pp. 398-403, março de 2017.

Mostrar no contexto

Ver artigo

Google Scholar

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial baseada IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.

Mostrar no contexto

Ver artigo

Google Scholar

6. S. Lloyd, "Quantização de mínimos quadrados PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 1, pp. 129-137, março de 1982.

Mostrar no contexto

Ver artigo

Google Scholar

TABELA III Número de Ciclos (Milhões) por Número de Pontos (Pts) e Centróides (Ctr)

Ctr	Pts	K-means				K-modes	
		FPGA		Xeon		FPGA	Xeon
		Int	FP	Int	FP	Cat	Cat
2	2 <sup>8</sup>	3.546	5.567	39.904	38.031	0.046	0.047
	2 <sup>10</sup>	2.884	6.724	35.755	43.754	0.056	0.056
	2 <sup>12</sup>	7.417	9.079	43.672	45.355	0.057	0.057
4	2 <sup>8</sup>	5.011	6.685	22.774	23.600	0.047	0.047
	2 <sup>10</sup>	3.158	8.679	34.730	39.706	0.056	0.056
	2 <sup>12</sup>	11.537	15.917	40.519	44.422	0.057	0.057
8	2 <sup>8</sup>	5.217	7.274	35.914	41.448	0.047	0.047
	2 <sup>10</sup>	9.055	14.851	43.926	39.212	0.057	0.057
	2 <sup>12</sup>	12.791	16.201	50.693	54.901	0.058	0.058

Mostrar no contexto

Ver artigo

Google Scholar

O consumo de energia, em milijoules, foi avaliado, utilizando o FPGA em 50MHz. O FPGA-Int consome entre 78% e 94% menos energia que o Xeon-Int, e o FPGA-FP consome entre 72% e 88% menos energia que o Xeon-FP, conforme mostrado na Tabela IV. A mesma avaliação foi realizada para o Xeon-Cat de que o FPGA consome até 99% menos energia que o Xeon. Isso se deve a dois fatores: a diferença na frequência de operação e o número de ciclos gastos por cada plataforma. Dessa forma, vale lembrar que a versão do software utiliza unidades computacionais mais complexas e com frequência superior ao FPGA. Entre eles, os sistemas mais robustos de memória, suporte multicore e multithread aumentam o consumo de energia.

Mostrar no contexto

Ver artigo

Google Scholar

TABELA IV Energia Consumida (mJ) por Número de Pontos (Pts) e Centróides (Ctr)

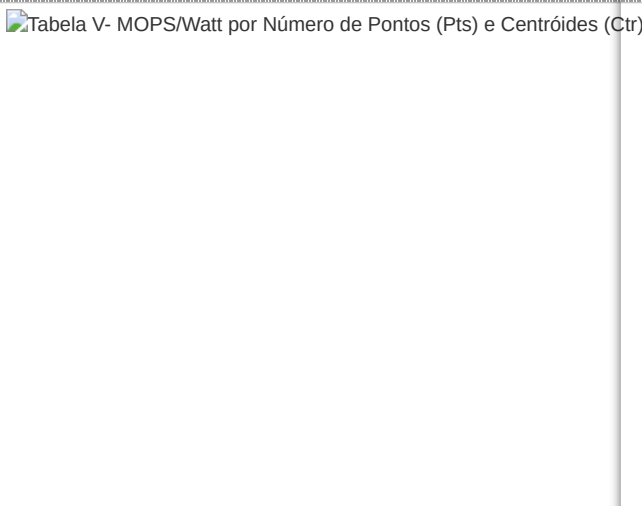
Tabela IV- Energia Consumida (mJ) por Número de Pontos (Pts) e Centróides (Ctr)

https://ieeexplore.ieee.org/document/8826329

7/10

energia de 1,50 Watts, enquanto a potência de design térmico (TDP) do Xeon E5-2620 é em média de 59 Watts. Para o pequeno conjunto de dados, o K-means baseado em FPGA executa até 15× mais rápido (custo flutuante), mais operações por Watt do que o processador Xeon. Com um conjunto de dados maior esse ganho chega a até 8× e 7×. Os resultados para modos K (dados categóricos) mostram o FPGA rodando 994× mais MOPS/Watt que o Xeon graças à paralelização das operações durante o cálculo da correspondência simples. Finalmente, dados categóricos podem dar à clusterização baseada em FPGA uma vantagem de desempenho. Nossos resultados mostram que K-modes-Cat é de até 220×, 225× e 185× melhor que K-means-Int, que é mais eficiente que K-means-Float, para ciclos, energia e MOPS/Watt, respectivamente.

TABELA V MOPS/Watt por número de pontos (Pts) e centróides (Ctr)

	
--	--

SEÇÃO VII.  
Conclusão

Apresentamos o projeto de uma arquitetura para algoritmos de agrupamento K-means e K-modes em um dispositivo FPGA. Suporta dados de entrada de 64 bits, sendo 32 bits para atributos numéricos e atributos categóricos de 8 bits, permitindo alterar seus principais parâmetros de entrada em tempo de execução. Descrevemos uma avaliação do desempenho, consumo de energia e eficiência energética do nosso hardware, quando comparado com um processador Intel Xeon. Nossas propostas baseadas em PBF e LHM são as melhores que as versões baseadas em OpenMP com um aumento de MOPS/Watt de até 225× e 185× para modos K. Os resultados mostraram que nosso hardware apresentou menor número de ciclos (até 91%) e menor consumo de energia (até 99%) do que as abordagens comparadas. Assim, nossa arquitetura pode aumentar a eficiência exigida pelos sistemas de detecção de intrusão. Para trabalhos futuros pretendemos utilizar dispositivos FPGA mais robustos (ex. Intel Arria 10); use memórias externas em vez de ROM; e implementar um modelo híbrido no FPGA capaz de classificar dados em tempo real.

Referências

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, 65, não. 8, pp. 1049-1053, agosto de 2018.  
Mostrar no contexto Ver artigo Google Scholar

2. C. Yin, Y. Zhu, J. Fei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusão usando redes neurais recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.  
Mostrar no contexto Ver artigo Google Scholar

3. A. Bohara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando e agrupando diversos dados de monitor", *Proc. 9ª Ciência do Bootcamp. Segurança (HotSoS)*, p. 16, 2016.  
Mostrar no contexto CrossRef Google Scholar

4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rumo à computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eui (DATA)*, pp. 398-403, março de 2017.  
Mostrar no contexto Ver artigo Google Scholar

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial baseada IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.  
Mostrar no contexto CrossRef Google Scholar

6. S. Lloyd, "Quantização de mínimos quadrados PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 1, pp. 129-137, março de 1982.  
Mostrar no contexto Ver artigo

Autores	▼
Figuras	▼
Referências	▼
Citações	▼
Palavras-chave	▼
Métricas	▼



Publicado: 2019

Análise de um novo projeto de controlador de tabela de consulta não volátil (NV LUT) com memórias de acesso aleatório resistivas (RRAM) para matrizes de portas programáveis em campo (FPGA)

Simpósio Regional IEEE 2019 sobre Micro e Nanoeletrônica (RSM)

Publicado: 2019

Referências

1. AW Al-Dabbagh, Y. Li e T. Chen, "Um sistema de detecção de intrusão para ataques cibernéticos em sistemas de controle de rede sem fio", *IEEE Trans. Sistema de Circuitos II Exp. Resumos*, vol. 65, não. 8, pp. 1049-1053, agosto de 2018.

Mostrar no contexto

Ver artigo

Google Scholar

Conta Pessoal IEEE

Detalhes da compra

Informação do Perfil

ALTERAR NOME DE USUÁRIO/SENHA

OPÇÕES DE PAGAMENTO

VER DOCUMENTOS ADQUIRIDOS

PREFERÊNCIAS DE COMUNICAÇÃO

PROFISSÃO E EDUCAÇÃO

INTERESSES TÉCNICOS

Sobre o IEEE Xplore | Contate-nos | Ajuda | Acessibilidade | Termos de Uso | Política de Não Discriminação | Mapa do site | Política de privacidade do IEEE

Uma organização sem fins lucrativos, o IEEE é a maior organização profissional técnica do mundo dedicada ao benefício da humanidade.

© Copyright 2024 IEEE – Todos os direitos reservados.

Conta IEEE

Detalhes da compra

Informação do Perfil

Preciso de ajuda?

» Alterar nome de usuário/senha

» Atualizar endereço

» Opções de pagamento

» Histórico de pedidos

» Ver documentos adquiridos

» Preferências de comunicação

» Profissão e Educação

» Interesses Técnicos

» EUA e Canadá: +1 800 678 4333

» Em todo o mundo: +1 732 981 0060

» Contato e suporte

PDF

Ajuda

2. C. Yin, Y. Zhu, J. Fei e X. He, "Uma abordagem de aprendizagem profunda para detecção de intrusão usando redes neurais recorrentes", *IEEE Access*, vol. 5, pp. 21954-21961, 2017.

Mostrar no contexto

Ver artigo

Google Scholar

3. A. Bohara, U. Thakore e WH Sanders, "Detecção de intrusão em sistemas corporativos combinando agrupando diversos dados de monitor", *Proc. 9ª Ciência do Bootcamp, Segurança (HotSec)*, pp. 16, 2016.

Mostrar no contexto

CrossRef

Google Scholar

4. K. O'Brien, LD Tucci, G. Durelli e M. Blott, "Rur à computação em exaescala com arquiteturas heterogêneas", *Proc. Projeto Autom. Teste Eur (DATA)*, pp. 398-403, março de 2017.

Mostrar no contexto

Ver artigo

Google Scholar

5. J. Kim e J. Park, "Detecção de intrusão de rede baseada em FPGA para rede industrial basead IEC 61850", *ICT Exp.*, vol. 4, não. 1, pp. 1-5, 2017.

Mostrar no contexto

CrossRef

Google Scholar

6. S. Lloyd, "Quantização de mínimos quadrados PCM", *IEEE Trans. Inf. Teoria*, vol. IT-28, não. 3, pp. 129-137, março de 1982.

Mostrar no contexto

Ver artigo

Google Scholar

